



Pontificia Universidad
Católica del Ecuador | Sede
Ambato

CENTRO DE POSGRADOS

Tema:

**SIMULACIONES DIGITALES PARA FACILITAR LA COMPRENSIÓN DOCENTE
EN ANÁLISIS ESTADÍSTICO IMPLICATIVO Y EL *LEARNING ANALYTICS***

**Proyecto de investigación y desarrollo previo a la obtención del título de
Magister en Ciencias de la Educación**

Línea de investigación:

Desarrollo e innovación curricular, procesos didácticos y
recursos educativos

Autor:

José Luis Mullo Gualán

Director:

Dr. Rubén Antonio Pazmiño Maji. PhD

Ambato – Ecuador

Abril 2023

PONTIFICIA UNIVERSIDAD CATÓLICA DEL ECUADOR SEDE AMBATO

HOJA DE APROBACIÓN

Tema:

**SIMULACIONES DIGITALES PARA FACILITAR LA COMPRENSIÓN DOCENTE
EN ANÁLISIS ESTADÍSTICO IMPLICATIVO Y EL *LEARNING ANALYTICS***

Línea de investigación:

Desarrollo e innovación curricular, procesos didácticos y
recursos educativos

Autor:

José Luis Mullo Gualán

Rubén Antonio Pazmiño Maji, Dr. PhD.

f.



Firmado electrónicamente por:
RUBEN ANTONIO
PAZMINO MAJI

CALIFICADOR

Carlos Rodrigo Moreta Herrera, Psc. Ms.

f.

CALIFICADOR

Juan José Ramos Paredes, Mg.

f.

CALIFICADOR

Juan Carlos Acosta Teneda, Ms. PhD.

f.

COORDINADOR DE LA OFICINA DE POSGRADOS

Hugo Rogelio Altamirano Villarroel, Dr.

f.

SECRETARIO GENERAL PUCESA



SECRETARIA GENERAL
PROCURADURIA

Ambato – Ecuador

Abril 2023

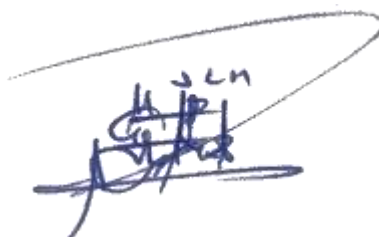
Pontificia Universidad
Católica del Ecuador
BIBLIOTECA

DECLARACIÓN DE AUTENTICIDAD Y RESPONSABILIDAD

Yo: **JOSÉ LUIS MULLO GUALÁN**, con CC. **060298688-7**, autor del trabajo de graduación intitulado: “SIMULACIONES DIGITALES PARA FACILITAR LA COMPRENSIÓN DOCENTE EN ANÁLISIS ESTADÍSTICO IMPLICATIVO Y EL *LEARNING ANALYTICS*”, previa a la obtención de título profesional de **MAGISTER EN CIENCIAS DE LA EDUCACIÓN**, en la escuela de **POSGRADOS**.

1. Declaro tener pleno conocimiento de la obligación que tiene la Pontificia Universidad Católica del Ecuador, de conformidad con el artículo 144 de la Ley Orgánica de Educación Superior, de entregar a la SENESCYT en formato digital una copia del referido trabajo de graduación para que sea integrado al Sistema Nacional de Información de la Educación Superior del Ecuador para su difusión pública respetando los derechos de autor.
2. Autorizo a la Pontificia Universidad Católica del Ecuador a difundir a través de sitio web de la Biblioteca de la PUCE Ambato, el referido trabajo de graduación, respetando las políticas de propiedad intelectual de la Universidad.

Ambato, abril 2023



José Luis Mullo Gualán

C.C. 060298688-7

DEDICATORIA

“Porqué Jehová da la sabiduría, y de su boca viene el conocimiento y la inteligencia”. Proverbios. 2: 6.

A mis Padres: Pedro y Rosario, por los cuidados y consejos en todo tiempo.

A mi esposa Bertha Zenaida, por su paciencia, apoyo, motivación, amor y cariño.

José Luis Mullo Gualán

AGRADECIMIENTO

Gratitud a Dios por la salud y la vida que me regala cada día, por la oportunidad de ingresar en el mundo del conocimiento. Fue factible la culminación de este proyecto gracias a la guía y la motivación por parte del Dr. Rubén Antonio Pazmiño Maji, por el asesoramiento en cada una de las etapas de desarrollo de este trabajo, también, merece reconocer el aporte de Raphael Couturier PhD por la actualización continua de la página de *Rchic*, que facilitó la comprensión de las técnicas comunes entre el Análisis Estadístico Implicativo y el *Learning Analytics*.

De la misma manera mi gratitud al cuerpo docente de esta prestigiosa Universidad, por los conocimientos impartidos que me han facilitado la consecución de este trabajo.

RESUMEN

La investigación tiene por objetivo facilitar la comprensión docente de los conceptos y algoritmos: similaridad, cohesión, implicación, *hclust.vector*, *dendro.variables*, *dendro.diana*, *apriori*, *weclat* y *eclat*, técnicas comunes del Análisis Estadístico Implicativo (ASI) y el *Learning Analytics* (LA). Esta investigación se desarrolló entre los años 2017 y 2022, se trabajó con los profesores del nivel de bachillerato de la Unidad Educativa del Milenio Intercultural Bilingüe Chibuleo, de la provincia de Tungurahua, Ecuador. Se emplearon metodologías de tipo inductivo, cualitativo y mixto transversal, el proceso de elaboración de las simulaciones digitales fue el “Building valid models: how to build valid and credible simulation models” de Law y McComas (2001). En base al diagnóstico de las aptitudes docentes en las herramientas TIC, el principal aporte fue la elaboración de las simulaciones digitales basados en el proceso antes mencionado, diseñadas en *Office* y *R* (disponible en: <https://bit.ly/2PUWrPw>) que facilita que los docentes de bachillerato comprendan los conceptos y algoritmos ya citados y los usen en su labor de enseñanza y aprendizaje.

Palabras claves: análisis estadístico implicativo, *learning analytics*, *rchic*.

ABSTRACT

The objective of the present research is to facilitate the teacher's comprehension of the concepts and algorithms such as: similarities, cohesion, implication, hclust,vector, dendro.variables, dendro.diana, apriori, weclat and eclat, that are common techniques of the Statistical Implicative Analysis (SIA) and the Learning Analytics (LA). The research was carried out between the 2017 and 2022, where high school teachers from the "Unidad Educativa del Milenio Intercultural Bilingüe Chibuleo", located in the Tungurahua Province in Ecuador collaborated with the research. The simulation method Building Valid Models: How to Build Valid and Credible Simulation Models by Law and McComas (2001), has been used, being it the main contribution to the creation of digital simulations based on the aforementioned methodology which was done in Office and R programs, the same that would help teachers from high school to understand the mentioned concepts and algorithms and use them in the teaching-learning process.

Keywords: statistical implicative analysis, learning analytics, rchic.

ÍNDICE GENERAL DE CONTENIDOS

DECLARACIÓN DE AUTENTICIDAD Y RESPONSABILIDAD	iii
DEDICATORIA.....	iv
AGRADECIMIENTO.....	v
RESUMEN	vi
ABSTRACT	vii
INTRODUCCIÓN	1
CAPÍTULO I. ESTADO DEL ARTE Y LA PRÁCTICA	5
1.1. Los simuladores en la educación.....	5
1.2. Análisis estadístico implicativo.....	9
1.3. Análisis estadístico implicativo y la educación.....	18
1.4. <i>Learning analytics</i> y la educación	19
CAPÍTULO II. DISEÑO METODOLÓGICO	34
2.1. Metodología de la investigación.....	34
2.2. Recolección de la información	37
2.3. Propuesta.....	40
CAPÍTULO III. ANÁLISIS DE LOS RESULTADOS DE LA INVESTIGACIÓN.....	63
3.1. Procesamiento y análisis de la información	63
3.2. Evaluación de expertos de la propuesta	74
3.3. Resultados del análisis cuasi experimental.....	74
CONCLUSIONES.....	78
RECOMENDACIONES	80
BIBLIOGRAFÍA	81
ANEXOS	89

ÍNDICE DE TABLAS

Tabla 1. Publicaciones en relación a las técnicas del LA.....	31
Tabla 2. Estadístico de confiabilidad.....	39
Tabla 3. Especificaciones del hardware necesario para las simulaciones	49
Tabla 4. Aplicaciones a ser instaladas	49
Tabla 5. Tipos de datos presentes en las técnicas	56
Tabla 6. Estadísticos descriptivos pruebas pre test y post test	75
Tabla 7. Prueba de normalidad pre-test y post-test.....	76
Tabla 8. Estadísticos de contraste de Wilcoxon.....	76

ÍNDICE DE GRÁFICOS

Gráfico 1. El sexo del encuestado.....	64
Gráfico 2. El rango de edad del encuestado	64
Gráfico 3. Años de servicio del docente encuestado.....	65
Gráfico 4. Número de estudiantes a quienes imparten asignaturas.....	65
Gráfico 5. Asignaturas de mayor carga horaria con las que trabajan los docentes.....	66
Gráfico 6. Las herramientas TIC con las que está familiarizado el docente.....	67
Gráfico 7. Las herramientas TIC que utiliza el docente para la enseñanza-aprendizaje ...	67
Gráfico 8. Uso docente de la herramienta excel	68
Gráfico 9. Uso docente del software estadístico R	69
Gráfico 10. Uso docente de medios digitales en clases.....	69
Gráfico 11. Análisis estadístico de aprendizajes.....	70
Gráfico 12. Datos estadísticos en promociones estudiantiles.....	71
Gráfico 13. Conocimiento de las técnicas ASI	71
Gráfico 14. Conocimiento de las técnicas LA.....	72
Gráfico 15. Conocimiento de las técnicas del ASI y LA en datos educativos.....	73

ÍNDICE DE CUADROS

Cuadro 1. Técnicas comunes entre el ASI y el LA.....	33
Cuadro 2. Diseño de investigación.....	37
Cuadro 3. Detalle de las técnicas del ASI y LA con enfoque educativo	43
Cuadro 4. Enlaces de las hojas de simulaciones.....	52

ÍNDICE DE FIGURAS

Figura 1. Conjunto de eventos.....	11
Figura 2. La intensidad entrópica	13
Figura 3. Grafo de implicación	14
Figura 4. Grafo implicativo	15
Figura 5. Técnicas más utilizadas del LA en educación.....	21
Figura 6. ¿qué ocurre en 60 segundos en las redes?	22
Figura 7. El entorno del learning analytics.....	22
Figura 8. El entorno del big data.....	23
Figura 9. Las seis universidades con publicaciones LA de autores ecuatorianos	32
Figura 10. Técnicas comunes del ASI y el LA.....	42
Figura 11. Flujograma del proceso de elaboración de las simulaciones digitales	47
Figura 12. Simulación en excel para el índice de similaridad	52
Figura 13. Flujograma del uso de las simulaciones digitales ASI-LA.....	56
Figura 14. Análisis de similaridad en excel.....	57
Figura 15. Análisis estadístico implicativo en excel	57
Figura 16. Análisis cohesivo en excel	58
Figura 17. Uso educativo de las técnicas comunes del ASI y el LA en R	58
Figura 18. El cálculo del coeficiente V de Aiken de la propuesta	74

ÍNDICE DE FÓRMULAS

Fórmula 1. Índice de similaridad.....	10
---------------------------------------	----

INTRODUCCIÓN

Este trabajo, se titula “Simulaciones digitales para facilitar la comprensión docente en: Análisis Estadístico Implicativo (ASI) y el *Learning Analytics* (LA)”. A ser empleada en una primera fase por los docentes del bachillerato de la Unidad Educativa del Milenio Intercultural Bilingüe Chibuleo (en adelante, UEMIB Chibuleo), centro educativo que tiene en sus procesos de enseñanza aprendizaje, medios digitales. Se responde a la pregunta ¿por qué los docentes de bachillerato de la UEMIB Chibuleo, no aplican las técnicas comunes del Análisis Estadístico Implicativo y el *Learning Analytics*, sobre los datos generados en el uso de las TIC en los procesos educativos?

El vacío que evidencia la pregunta de investigación, corresponde a las Analíticas de Aprendizajes (en inglés, *Learning Analytics* (LA)) que, combinadas con las técnicas del Análisis Estadístico Implicativo (ASI), favorecen al docente a mejorar sus procesos de enseñanza y aprendizaje. Como se muestra en este trabajo, en la mencionada Unidad Educativa, en el nivel de bachillerato, no existen experiencias del uso de estas técnicas, sin embargo, los docentes están con la predisposición de aprender cosas nuevas y útiles para su labor. La población beneficiaria de esta investigación, son directamente los docentes de la UEMIB Chibuleo, ubicada en Ambato-Tungurahua-Ecuador.

El trabajo de investigación tiende a responder la proposición el uso de las simulaciones digitales en la comunidad de docentes facilita la comprensión de los conceptos y algoritmos de la tendencia ASI y LA.

El uso de las técnicas del *Learning Analytics* (LA), son una de las tendencias en auge en los países anglosajones, y ahora en Latinoamérica, como mencionan Chiappe and Rodríguez (2017):

Las técnicas del *Learning Analytics* ha cobrado un interés en la comunidad de investigadores educativos, proporcionan herramientas importantes para la toma de decisiones, y avance de los procesos

educacionales, su implementación no debe limitarse solo a los niveles superiores (universidades), debe ser abierto a toda la comunidad académica y al público en general. (p. 971)

Objetivo general

- Diseñar un sistema de simulaciones digitales para que el profesor de nivel de bachillerato familiarice con las técnicas comunes entre el Análisis Estadístico Implicativo (ASI) y el *Learning Analytics* (LA).

Objetivos específicos

1. Fundamentar los aspectos teóricos asociados al desarrollo de las simulaciones digitales y la teoría del Análisis Estadístico Implicativo y el *Learning Analytics*.
2. Diagnosticar el nivel de conocimiento de estas bases teóricas en la comunidad de los docentes de la UEMIB Chibuleo.
3. Elaborar un sistema de simulaciones digitales encaminada a la adquisición de conocimientos referentes a las técnicas comunes del Análisis Estadístico Implicativo y el *Learning Analytics*.
4. Realizar un análisis de validación de expertos.

Con estas técnicas, el docente logra reconocer las reglas de interacción digital de sus aprendices y las utiliza de forma permanente para mejorar los entornos de enseñanza aprendizaje y lograr así, estudiantes exitosos. Como mencionan (Larrea de Granados & Granados Boza, 2013):

Es imprescindible un alineamiento de los actores de la educación ecuatoriana, con políticas estatales que propicien una revolución paradigmática y del pensamiento, con prácticas pertinentes, colaborativas, distribuidas, que de manera compleja contribuyan a una sociedad más justa, equitativa y solidaria, basada en el bio-conocimiento, en el marco de la innovación social; por, lo cual, hay

que avanzar hacia la generación de modos y formas posibles de aplicación del paradigma conectivista, en la educación formal de las instituciones educativas, ampliando las oportunidades de aprendizaje sin violentar los fundamentos teóricos y metodológicos de la educación. (p. 141)

Este trabajo cubre cuatro etapas, la identificación de las técnicas comunes entre el Análisis Estadístico Implicativo (ASI) y el *Learning Analytics* (LA), basadas en las sugerencias de investigadores Pazmiño R, Naranjo M y sus colaboradores. A continuación, seleccionan los conceptos y las definiciones que sustentan las técnicas, se elaboran ejemplos prácticos asociados a las técnicas, para la difusión, utilizan las simulaciones digitales presentadas en una página web y, por último, se procede a la evaluación por expertos.

El trabajo propuesto cobra mucha importancia para el docente de la UEMIB Chibuleo que utiliza las diferentes herramientas TIC con sus aprendices, los estudiantes y los docentes por el empleo en los procesos de aprendizajes herramientas digitales, aplicaciones y páginas web, dejan huellas de su quehacer educacional, refleja así la permanencia estudiantil en centros o comunidades de aprendizaje. Esta huella es factible el análisis con las técnicas del ASI y LA y descubrir conocimientos como: reglas de comportamiento digital junto a las formas de organización de las temáticas y su relación con los aprendizajes.

Este trabajo es una investigación aplicada, de tipo descriptiva y de enfoque cualitativo: Investigación aplicada, se pretende encontrar una estrategia (las simulaciones digitales), hacia los docentes de bachillerato de la UEMIB Chibuleo, para que utilicen las técnicas ASI y LA en sus procesos de enseñanza aprendizaje. Es investigación descriptiva, se va a diseñar las simulaciones en el uso de la metodología "*How to build valid and credible simulation model*" de (Law & McComas, 2001), y presentar a los profesores. La investigación es de enfoque cualitativo, (Daza, 2018) menciona: "Permite profundizar con mayor amplitud la práctica, el fenómeno educativo (...), posibilita analizar y explicar la comprensión

de la complejidad” (p. 105). En este caso, desde la detección de la problemática hasta presentar la propuesta a los docentes.

De la misma forma, con respecto al enfoque cualitativo (Sampieri, Collado, & Lucio, 2014) mencionan: “con frecuencia, estas actividades sirven, primero, para descubrir cuáles son las preguntas de investigación más importantes, y después, para refinarlas y responderlas” (p. 7). Lo que conduce hasta la ejecución de la propuesta de investigación hacia la comunidad de docentes, quienes en estos tiempos necesitan una vanguardia basada en los datos generados por los aprendices en medios digitales.

CAPÍTULO I. ESTADO DEL ARTE Y LA PRÁCTICA

En este capítulo, se presentan algunos de los conceptos del análisis estadístico implicative y el *learning analytics* encaminada al ámbito educativo que, en la comunidad científica, se ha generado en el transcurso de los últimos años. Una primera parte, es información acerca de las simulaciones digitales, su congruencia con el enfoque constructivista, seguidamente los aspectos relacionados al aprendizaje docente, para después en la parte central de la propuesta, responder la pregunta: ¿Cuáles son las técnicas del análisis estadístico implicative (ASI) empleadas en el ámbito educativo?, seguidamente, se han tratado los conceptos del *learning analytics* y las técnicas comunes al análisis estadístico implicative y el *learning analytics*.

Son técnicas que, han innovado el ambiente de las investigaciones educativas, existe una tendencia de uso en países extranjeros, en Ecuador utilizan en el ámbito de la educación superior. En esta propuesta, se lleva estos contenidos hacia la población docente de bachillerato de la UEMIB Chibuleo del cantón Ambato, existe un gran interés de los docentes de estos niveles, en aprender nuevas herramientas de análisis de datos, pero la alta mayoría de ellos no tiene formación de estadísticos o analista de datos, de allí surge, la propuesta de las simulaciones digitales para facilitar su aprendizaje y el uso de estas herramientas.

1.1. Los simuladores en la educación

En el entorno de los simuladores en la educación, está la programación informática, de acuerdo con Ceballos, (2004): “La programación informática tradicional, se compone de procedimientos y de datos”, por lo tanto, este simulador se corresponde a un conjunto de procesos que el docente tiene que aproximar, a través de ejercicios en herramientas como el *excel*.

Las técnicas de simulación

Entre las principales técnicas de simulaciones se tiene: la simulación física, la simulación analógica, la simulación numérica, las simulaciones específicas y las simulaciones digitales.

Simulación física

Utilizados en un entorno de simulación dentro de la enseñanza de la asignatura de física, las simulaciones que suelen organizarse por parte del profesor y estudiantes son limitadas, las simulaciones son siempre visuales, el estudiante suele obtener datos numéricos o gráficos de un buen número de variables implicadas.

La simulación analógica

Se refiere al uso de dispositivos electrónicos con la finalidad de presentar, explicar, comprender conceptos, se aplica en el campo de la Electrónica, por ejemplo, los circuitos electrónicos.

La simulación numérica

Es el uso de métodos numéricos en la resolución de problemas cotidianos, ejemplo, algoritmo matemático para la solución de un modelo del enfriamiento de un cuerpo.

La simulación estadística

Es el uso de las principales herramientas con las que, se pretende explicar conceptos estadísticos, he ahí que, se disponen de *Software* Estadístico como: SPSS, SAS y SAS/STAT, en particular, herramientas *Excel*, con respecto al Análisis Estadístico Implicativo y el *Learning Analytics*, se las lleva a cabo a través del *Software R*.

Simulaciones en educación

Son procesos que permiten a los docentes y estudiantes el aprendizaje por medio de exploraciones, la inferencia y el aprendizaje por descubrimientos, mediante procesos investigativos. Sitúa al aprendiz en un ambiente en que imite aspectos de la vida real y la búsqueda a las soluciones de las problemáticas o productivas. Según Contreras Gelves, García Torres, & Ramírez Montoya, (2010) “una de las funciones principales de los simuladores en educación es el apoyo a docentes en la transferencia de conocimiento” (p. 3). De la misma forma Bender & Fish, (2000) mencionan:

Una jerarquía de conocimiento cuando abordan la transferencia, y detallan los siguientes niveles: datos (unidad mínima de información), información (al añadir significado a los datos), conocimiento (al dar la aprehensión de hechos, verdades o principios), hasta la destreza (estadio superior al dar respuesta al porqué de las cosas y se generan habilidades y métodos de aplicación). (p. 128)

En esta era digital, los docentes y los estudiantes utilizan las técnicas de las simulaciones para la comprensión de los fenómenos estudiados con la finalidad de minimizar operaciones y llegar al conocimiento, desde el uso de las calculadoras hasta programar en el uso de los simuladores y entender así los fenómenos educativos.

Las simulaciones digitales

Las simulaciones digitales son medios de uso multimedia con la que se puede gestionar el conocimiento, según Contreras Gelves et al., (2010):

Los simuladores constituyen un procedimiento tanto para la formación de conceptos y construcción de conocimiento, en general, como para la aplicación de éstos a nuevos contextos, a los que, por diversas razones, el estudiante no puede acceder desde el contexto metodológico donde, se desarrolla su aprendizaje; de hecho, buena

parte de la ciencia de frontera, se basa cada vez más en el paradigma de la simulación, más que en el experimento en sí. (p. 4)

Es por ello que, a través de las simulaciones digitales y la programación, es posible el aprendizaje de fenómenos educativos en el uso de medios digitales, ser así efectividad en la educación.

¿Cómo aprende la gente adulta?

Para responder esta interrogante tal como la Pedagogía, se encarga de la educación de los niños, la ciencia que encarga de la educación de la gente adulta, es la llamada la Andragogía, Fasce (2006) menciona “la Pedagogía viene del griego *paidos*, que significa niño y *agogus* que indica guía de, la Andragogía proviene de la acepción *andros* que significa hombre”, complementa Caraballo Colmenares, (2007) con que “la Andragogía ha sido descrita como una ciencia, es un conjunto de supuestos, es un método, es una serie de lineamientos, es una filosofía, es un cuerpo, campo de conocimiento, disciplina, es un modelo educativo, es una teoría, y como un proceso de desarrollo integral del ser humano” (p. 191).

Está el rol tanto del aprendiz y el que guía, en este caso, se emplea tareas autónomas para que el adulto, aproxime al conocimiento. En este sentido, el término Andragogía “se trata del arte de enseñar a los adultos a aprender, el aprendizaje del adulto, se sustenta en dos atributos principales: (1) los aprendices poseen autonomía y la capacidad de dirigir el sentido del aprendizaje y (2) el profesor pasa a ocupar un rol de facilitador del aprendizaje en vez de ser un transmisor de información” (Knowles, 1984).

En el mismo sentido, la función principal del docente es la de facilitar la información a sus aprendices adultos, tal como Benavides Herrera, (2014) menciona:

En la educación de jóvenes y adultos, hay que tener presente que la enseñanza, no es transferir conocimientos y aprendizajes, no es

memorizar; en consecuencia, el rol del facilitador es vital considerarlo, porque deja de ser transmisor de conocimientos, para convertirse en orientador de mecanismos que ha de adoptar el adulto en las situaciones de aprendizaje; también, hay que considerar que el adulto no llega en cero a una situación de aprendizaje. (p. 49)

En este contexto y debido a que la sociedad, se encuentra en un proceso eminente de cambio, por ejemplo, el empleo más exponencial de las tecnologías de la información y comunicación (TIC) entre otros cambios significativos para el hombre, es muy importante difundir el conocimiento, a través de medios más fiables y en el menor tiempo posible.

1.2. Análisis estadístico implicativo

Está en desarrollo, por su aplicación en ámbitos de la didáctica específicamente de la matemática, pero no excluye otras ramas de la didáctica a más de otras ramas de conocimiento. Menciona Orús et al. (2009) “El Análisis Estadístico Implicativo (*Analyse Statistique Implicative*, ASI, en francés) es un *método de análisis no simétrico de datos*” (p. 77).

Árbol de similaridad

Orús et al., (2009) parte de la consideración de un conjunto I formado por n individuos y un conjunto A formado por p características, $A = \{a_1, a_2, \dots, a_p\}$, donde:

$A_i = \{x \in I : a_i(x) = 1\}$, $\text{Card}(I) = n$ y $\text{Card}(A_i) = n_{a_i}$. CHIC calcula los índices de similaridad en términos de una probabilidad, la misma que obedece la variable aleatoria $\text{Card}(X_i \cap X_j) (\dots)$ las leyes de distribución para $\text{Card}(X_i \cap X_j)$ son: Hipergeométrica, Binomial y Poisson, en CHIC, Binomial y Poisson únicamente. Además, en este

Software informático, se puede determinar los nodos significativos. (p. 80-81)

Es un proceso de cálculo de probabilidad, la misma que indica cuanto un grupo de acciones educativas son similares en un grupo de estudiantes, tal como mencionan Orús et al., (2009):

Para la pareja (a_i, a_j) , se calculan los índices de proximidad o similaridad cita a Lerman. (1970): $s(a_i, a_j) = \text{Prob}[\text{Card}(X_i \cap X_j) \leq K]$, donde $K = \text{Card}(A_i \cap A_j)$, es decir, el valor observado de la cantidad de co-presencias entre a_i y a_j . El cálculo de estas probabilidades dependerá de la ley de probabilidad seleccionada (en CHIC Poisson o Binomial), sin embargo, se puede aproximar esta probabilidad a la Normal. (p. 81)

Este proceso de cálculo es posible, gracias a las definiciones y el uso básico de la herramienta *Excel*, en el *software R*, se emplea el paquete *Rchic*.

Fórmula 1. Índice de similaridad

$$s(a_i, a_j) = \text{Prob} \left[\frac{\text{Card}(X_i \cap X_j) - \frac{n_{a_i} * n_{a_j}}{n}}{\sqrt{\frac{n_{a_i} * n_{a_j}}{n}}} \leq K_C \right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{K_C} e^{-\frac{1}{2}x^2} dx.$$

$$\text{Donde } K_C = \frac{K - \frac{n_{a_i} * n_{a_j}}{n}}{\sqrt{\frac{n_{a_i} * n_{a_j}}{n}}}.$$

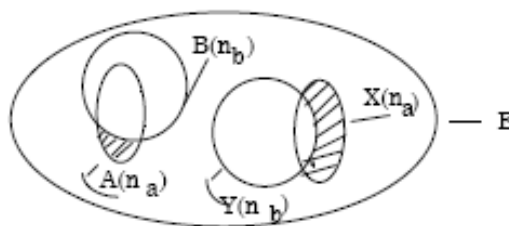
El número $s(a_i, a_j)$, indica las similaridades de cada pareja (a_i, a_j) al nivel cero de la jerarquía. Este proceso es continuo, hasta completar el cálculo para todas las parejas de eventos que representan eventos.

Intensidad de implicación

En el caso binario, según R. Gras y P. Kuntz, para la formalización de este concepto, Gras & Kuntz, (2009) mencionan:

Sean X e Y dos partes cualesquiera de un conjunto E , elegidas de forma aleatoria e independiente (ausencia del vínculo a priori entre las dos partes) y de cardinales iguales a los de A y B respectivamente, además, al considerar \bar{Y} y \bar{B} los complementos de Y y de B respectivamente en E de igual cardinal $n_{\bar{b}} = n - n_b$, se dice que: Definición 1.- $a \Rightarrow b$ es admisible al nivel de confianza $1 - \alpha$ sí y solo sí: $P[\text{Card}(X \cap \bar{Y}) \leq \text{Card}(A \cap \bar{B})] \leq \alpha$

Figura 1. Conjunto de eventos



Fuente: tomado a partir de (Gras & Kuntz, 2009)

Las partes subrayadas corresponden a los contraejemplos de la implicación. (p. 6)

Por otra parte, al calcular las probabilidades, (Lerman) citada por Gras y Kuntz (2009), menciona: “la variable aleatoria $\text{Card}(X \cap \bar{Y})$ sigue la distribución de Poisson, con parámetro $\frac{n_a n_{\bar{b}}}{n}$, por lo tanto, se da el cumplimiento de: $P[\text{Card}(X \cap \bar{Y}) = s] = e^{-\lambda} \frac{\lambda^s}{s!}$.”

(p. 6).

Definición 2.- El número: $q(a, \bar{b}) = \frac{n_{a \wedge \bar{b}} - \frac{n_a n_{\bar{b}}}{n}}{\sqrt{\frac{n_a n_{\bar{b}}}{n}}}$ es llamado **índice de implicación**, es

un indicador de la no implicación de a sobre b.

Definición 3.- El parámetro denominado **intensidad de implicación**, que calcula la **calidad inductiva** de a sobre b es:

$$\varphi(a, b) = \begin{cases} 1 - P[Q(a, \bar{b}) \leq q(a, \bar{b})] = \frac{1}{\sqrt{2\pi}} \int_{q(a, \bar{b})}^{\infty} e^{-\frac{t^2}{2}} dt & \text{si } n_b \neq n \\ 0 & \text{si } n_b = n \end{cases}$$

Es un cálculo de probabilidades que va de entre cero y el uno. Si aproxima al uno, existe mayor fuerza de implicación.

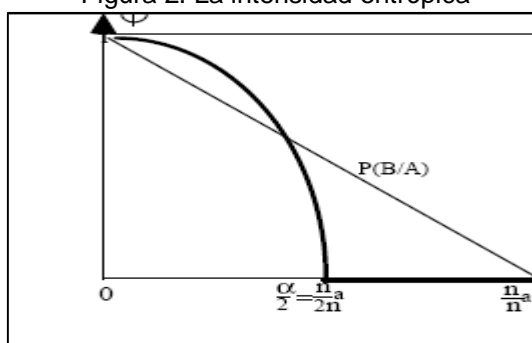
Definición 4.- Según, Gras & Kuntz, (2009), “la implicación $a \Rightarrow b$ es admisible al nivel de confianza $1 - \alpha$ si y solo si: $\varphi(a, b) \geq 1 - \alpha$ ” (p. 9). Es un concepto de cálculo basado en probabilidades.

Definición 5.- El **índice de inclusión** A, soporte de a, en B, soporte de b, es el número: $i(a, b) = [(1 - h_1^2(t))(1 - h_2^2(t))]^{\frac{1}{4}}$ que integra la información a partir de la realización de un escaso número de contraejemplos, de una parte por la regla $a \Rightarrow b$ y por otra parte la regla $\text{no } b \Rightarrow \text{no } a$. (p. 21)

Es uno de los cálculos básicos de la entropía de un sistema, que según Gras y Kuntz (2009), la **intensidad de implicación – inclusión** (o la intensidad entrópica), nueva medida de la calidad inductiva, es el número:

$\psi(a, b) = [i(a, b) \cdot \varphi(a, b)]^{\frac{1}{2}}$ que integra a la vez el asombro estadístico y la calidad inclusiva, la función Ψ según la variable t admite una representación que tiene la forma de la Figura 2. siguiente: (p. 21)

Figura 2. La intensidad entrópica



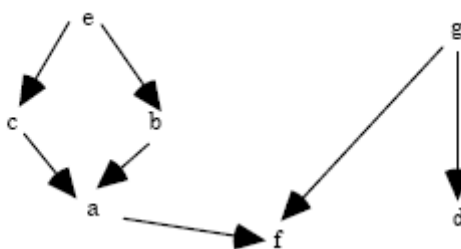
Fuente: tomado a partir de (Gras & Kuntz, 2009)

Es decir, para n_a y n_b fijos, la Figura 2. muestra la diferencia de comportamiento de la función conforme a la probabilidad condicional $P(B/A)$, índice fundamental de las otras modelizaciones de la medida de las reglas, (...) además, de su carácter lineal, esta probabilidad conduce a una medida que decrece muy rápido para los primeros contraejemplos, resiste entonces durante mucho tiempo, cuando éstos, se vuelven importantes.

Grafo de implicación.- Al comprender la estructura de datos, en la salida de los cálculos de intensidad de implicación, el usuario queda con incógnitas de su estructura, no es posible la determinación de las reglas de las estructuras presentes, pero existe una técnica asociada a estas tablas de resultados, llamado grafo orientado, ponderado (hacia las intensidades de implicación) (Gras & Kuntz, 2009). Es un esquema de representación (de las relaciones o implicaciones) de los índices de implicación, asociada a eventos.

Según Gras & Kuntz, (2009): Hay eventos que no obedecen a la relación de transitividad, por convenio, se dice que: si $a \Rightarrow b$ y si $b \Rightarrow c$, se acepta el cierre transitivo $a \Rightarrow c$, solamente si: $\psi(a,b) \geq 0,5$, ejemplo: suponer que entre las siete variables a, b, c, d, e, f y g existen al límite máximo superior 0.5, las siguientes reglas: $e \Rightarrow c, a, f, b$; $c \Rightarrow a, f$; $b \Rightarrow a, f$; $g \Rightarrow d, f$; $a \Rightarrow f$. Se podrá entonces traducir este conjunto de relaciones por el grafo siguiente: (p. 24)

Figura 3. Grafo de implicación



Fuente: tomado a partir de (Gras & Kuntz, 2009)

El gráfico, se reduce o amplía si se eleva o si se reduce el límite máximo de consideraciones de las reglas.

Definición 6.- El cálculo denominado **la cohesión de la clase (a, b)** es el número calculado por $c(a,b)$, dado por:

- Si $p = \psi(a,b)$ y $H = -p \log_2 p - (1-p) \log_2 (1-p)$, entonces $\text{coh}(a, b) = (1 - H^2)^{1/2}$,
- Si $p=1$, entonces $\text{coh}(a,b) = 1$,
- Si $p=0.5$, entonces $\text{coh}(a,b) = 0$.

La **cohesión es una medida del desequilibrio** de las frecuencias de los eventos $\bar{a} \vee b$ y $a \wedge \bar{b}$ a favor del primero. Es un método clásico del cálculo de la entropía de un sistema.

Definición 7.- El concepto denominado **cohesión de la clase ordenada de variables** $\underline{A} = (a_1, \dots, a_r)$ está definida por la fórmula:

$$\text{Coh}(\underline{A}) = \left[\prod_{\substack{i \in \{1, \dots, r-1\} \\ j \in \{2, \dots, r\}, j > i}} c(a_i, b_j) \right]^{\frac{2}{r(r-1)}}$$

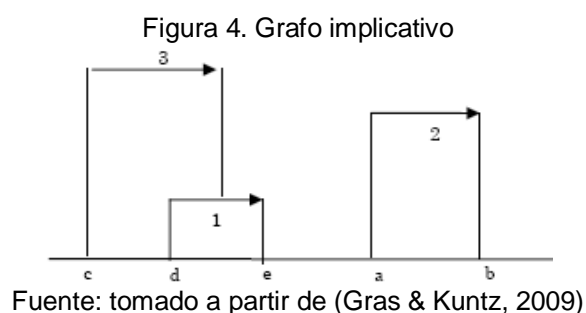
Es la media geométrica de las cohesiones de parejas, se anulan debido a que una de las cohesiones en juego, se hace cero. Se nota por \underline{A} y \underline{B} las clases de

variables de elementos genéricos a_i y b_i , $\text{Coh}(\underline{A})$ y $\text{Coh}(\underline{B})$ sus cohesiones respectivas.

Definición 8.- La **intensidad de implicación** de \underline{A} sobre \underline{B} es:

$$\psi(\underline{A}, \underline{B}) = \left[\begin{array}{l} \text{Sup}_{\substack{i \in \{1, \dots, r\} \\ i \in \{1, \dots, S\}}} \psi(a_i, b_i) \end{array} \right]^{rs} \cdot [\text{Coh}(\underline{A})\text{Coh}(\underline{B})]^{1/2}$$

Este índice satisface las dificultades semánticas. Define un método de clasificación descendente clásico por el criterio de cohesión decreciente, se puede obtener un gráfico (árbol).



Este grafo (árbol), se interpreta como: La regla $c \Rightarrow (d \Rightarrow e)$, la segunda parte, la regla $a \Rightarrow b$, la expresión $c \Rightarrow (d \Rightarrow e)$, se representa lógicamente como: $c \Rightarrow d \Rightarrow e$.

Niveles significativos de una jerarquía cohesitiva, según R. Gras y P. Kuntz

Pre orden cohesivo.- “Sea el conjunto V de $\{a_1, a_2, \dots, a_m\}$ variables y el conjunto de las parejas (a, b) de $V \times V$ con $a \neq b$. Existe $m(m-1)$ de estas parejas, a las cuales, les corresponden sus cohesiones $c(a, b)$ ” (p. 32).

Definición 9.- Con la finalidad de cimentar la noción del pre ordenamiento, mencionan Gras & Kuntz, (2009): “se llama pre orden inicial y global cohesivo sobre $V \times V$ (o pre ordenamiento), al pero orden Ω inducido por la aplicación cohesión c sobre $V \times V$ ” (p. 33).

De la misma forma, para generar el conjunto de parejas desde su inicio, Gras & Kuntz, (2009) consideran:

Sea $G(\Omega)$ su grafo en $V \times V$. Se sigue que: por una parte, la clase de pre orden correspondiente a $c=0$ contiene todas las parejas tales que $\psi(a,b) \leq 0,5$, de la misma forma, si $\psi(A,B)$ entonces $c(b,a) \leq c(a,b)$, por el contrario, si $c(a,b) \leq c(c,d)$, no se tiene necesariamente $c(b,a) \leq c(d,c)$ o. $c(b,a) \geq c(d,c)$. (p. 33)

Al respecto de la determinación de los niveles significativos, los autores Gras & Kuntz, (2009) consideran:

Colocando a un nivel k cualesquiera de la jerarquía. A este nivel, se forma una clase de m_i variables, con $2 \leq m_i \leq m$ cuya cohesión es menos buena que la de las clases anteriores formadas, de acuerdo con el algoritmo elegido, y mejor que la de las clases que deben venir. Y sea Π_k la partición de V definida a este nivel, construida por las clases que ya, se formaron allí y eventualmente, por conjuntos unitarios aun no asociados. Π_k es más fina que Π_{k+1} . Sea $S\Pi_k$ el conjunto de las parejas separadas a este nivel y $R\Pi_k$ el conjunto de las parejas que, se reúnen allí por primera vez, quedando entendido que, se dirá que el par (a,b) , se reúne si a y b pertenecen a la misma clase del tipo $(\dots((\dots a, \dots) \dots) \dots b) \dots)$. El conjunto $G(\Omega) \cap [S\Pi_k \times R\Pi_k]$, se constituye de los pares de parejas que al nivel k respetan al pre orden inicial. Por ejemplo, si $c(e,f) < c(a,b)$ [pues $((e,f)$ y $(a,b)) \in G(\Omega)$] y si al nivel k , e y f son separados mientras que a y b , se reúnen en la clase que, se forman el par $((e,f), (a,b))$

pertenece a $G(\Omega) \cap [S\Pi_k \times R\Pi_k]$, para el cardinal $A \cap \bar{B}$, se asocia al cardinal de $G(\Omega) \cap [S\Pi_k \times R\Pi_k]$, el índice aleatorio $\text{Card}[G(\Omega^*) \cap [S\Pi_k \times R\Pi_k]]$ donde Ω^* es un pre ordenamiento aleatorio en el conjunto, provisto de una probabilidad uniforme, de todos los reordenamientos del mismo tipo cardinal que Ω . Este índice tiene por esperanza $0.5 * \text{Card}[S\Pi_k \times R\Pi_k]$ y por varianza $\text{Card}[S\Pi_k \times R\Pi_k] \times \text{Card}[G(\Omega)]$. De allí, la obtención del $s(\Omega, k)$ denominado índice centrado reducido, bajo la expresión:

$$\frac{\text{Card}[G(\Omega^*) \cap [S\Pi_k \times R\Pi_k]] - 0.5 \times \text{Card}[S\Pi_k \times R\Pi_k]}{\sqrt{\text{Card}[S\Pi_k \times R\Pi_k] \times \text{Card}[G(\Omega)]}}. \quad (\text{p. 33-34})$$

Las mismas que tal como se va a exponer en este trabajo, es factible su comprensión en el ámbito de los docentes de nivel medio, esto gracias al uso de los simuladores, en ello, se considera.

Definición 10.- Se denomina **nodo significativo** a todo nodo correspondiente a un máximo local de $s(\Omega, k)$ durante la creación de la jerarquía implicativa, se dice en este caso que la partición Π_k está en **resonancia parcial** con Ω , si, además, $G(\Omega) \cap [S\Pi_k \times R\Pi_k] = S\Pi_k \times R\Pi_k$, se dirá que la partición Π_k está en **resonancia total** con Ω .

Este trabajo, para llegar a la población docente, se ajusta las fórmulas de las definiciones antes citadas, en el uso de las hojas de excel y los comandos R. Se acompaña a través de ejercicios del ámbito educativo, para su comprensión y aplicación en el ámbito educativo.

Es así que, hasta estas líneas, se ha recogido las versiones que dieron origen a las técnicas del Análisis Estadístico Implicativo (ASI), conceptos como: implicación estadística, inducción, implicación-inclusión, entropía, cohesión, grafo

implicativo, árbol jerárquico, nivel significativo, dialéctica, variable binaria. Conceptos, aparentemente difíciles para la mayoría de los docentes del bachillerato ecuatoriano que, gracias a las simulaciones digitales, es factible la aproximación a su uso educativo.

1.3. Análisis estadístico implicativo y la educación

En ambientes educativos que utilizan medios de comunicación digitales (presencia de tecnologías) entre el profesor y los estudiantes, se genera una cantidad incontable de datos (frecuencias de uso de recursos, tiempos de respuestas a actividades, frecuencia de interacción en foros, chats, uso de elementos de internet, entre otras actividades) que hacen a los aprendices más efectivos en el logro de los objetivos y estándares educativos. Las técnicas comunes del Análisis Estadístico Implicativo (ASI) y *Learning Analytics (LA)*, son técnicas de análisis de datos generados en éstos medios, con la finalidad de alertar a tiempo, la actuación eficaz de los administradores, docentes y estudiantes.

Esta tendencia motiva a los educadores a utilizar esta nueva técnica estadística multivariada de datos masivos, a los investigadores a fortalecer el grupo internacional del ASI y por qué no decir a los programadores a apoyar a Raphael Couturier, Rubén Pazmiño y Pablo Gregori con el desarrollo del actual y futuro *Rchic*, parte del ASI. R. Pazmiño (2014) clasifica el avance del uso del ASI por los diferentes congresos (desde el año 2000 hasta el año 2012), concluye que: el ASI históricamente tuvo su primera aplicación en el ámbito educativo, en el área de la matemática, de todos los artículos científicos publicadas en estos congresos, de entre el 48% y el 83% son relacionados al ámbito educativo (hasta el ASI8, 71 de 122).

1.4. *Learning analytics* y la educación

Uno de los primeros aportes acerca del *Learning Analytics (LA)*, se observa en LAK11 (<https://tekri.athabascau.ca/analytics/>), quienes mencionan: “El *Learning Analytics* es la medición, recopilación, análisis e informe de datos sobre los alumnos y sus contextos, con el fin de comprender y optimizar el aprendizaje y los entornos en los que ocurre”. Concepto que ha revolucionado (en aspectos académicos), la industria de la matemática, de la informática y la gestión de los procesos de esta índole.

Learning Analytics es una de las multi-herramientas que está en auge en esta década, y la tendencia es inmensurable, la revista *NMC Horizon Report 2012 Higher Education Edition*, considera: “*Learning Analytics* combina libremente una variedad de herramientas de recopilación de datos y las técnicas de análisis, para estudiar la participación, el rendimiento y el progreso de los estudiantes en la práctica, con el objetivo de utilizar lo aprendido para revisar el currículo, la enseñanza y la evaluación, en tiempo real”.

Estos son los conceptos que dieron el génesis de estas herramientas novedosas. Utilizan diferentes técnicas, entre ellas la de la minería de datos masivos, direccionado en la inteligencia de gestión y mejora, es que *Learning Analytics* se ha convertido no como una tendencia más, sino como una de las disciplinas que los docentes tienen que emplear, para perfeccionar continuamente los procesos de enseñanza y aprendizajes.

En este sentido mismo, los autores Montero and Suhonen (2014) mencionan del *Learning Analytics*:

Generalmente, la analítica del aprendizaje, se ocupa del desarrollo de métodos que aprovechan los conjuntos de datos educativos, para respaldar el proceso de aprendizaje. El *Learning Analytics (LA)* es un campo multidisciplinario que involucra el aprendizaje automático

(*machine learning*), la inteligencia artificial, la recuperación de la información, las estadísticas y la visualización. (p. 2)

De la misma forma, al hablar del mundo de los algoritmos en desarrollo, Softic *et al.* (2014) mencionan con respecto al *Learning Analytics*:

El campo recientemente establecido de la Analítica de Aprendizajes (*Learning Analytics-LA*), presenta un interés inherente en el método computacional o algorítmico de análisis de datos. En esta perspectiva, la analítica es más que el análisis empírico de las interacciones de aprendizaje en entornos ricos en tecnología, en realidad, también, exige enfoques computacionales y matemáticos específicos como parte del análisis. (p. 2)

Por lo cual, es necesaria la interacción entre los diferentes profesionales, en áreas como la educación, computación, estadística y matemática para desarrollar nuevas metodologías, con la intención siempre de optimizar los procesos de enseñanza-aprendizajes. Además, en la gran interacción en medios computacionales, es necesario el surgimiento de algoritmos de análisis, procesos netamente matemáticos, lo cual hace tan rico estas herramientas novedosas de análisis de datos masivos.

Por otra parte, de acuerdo a los avances recientes, y en base a los congresos LAK, el autor Peña-Ayala (2017), con respecto a *learning analytics* menciona:

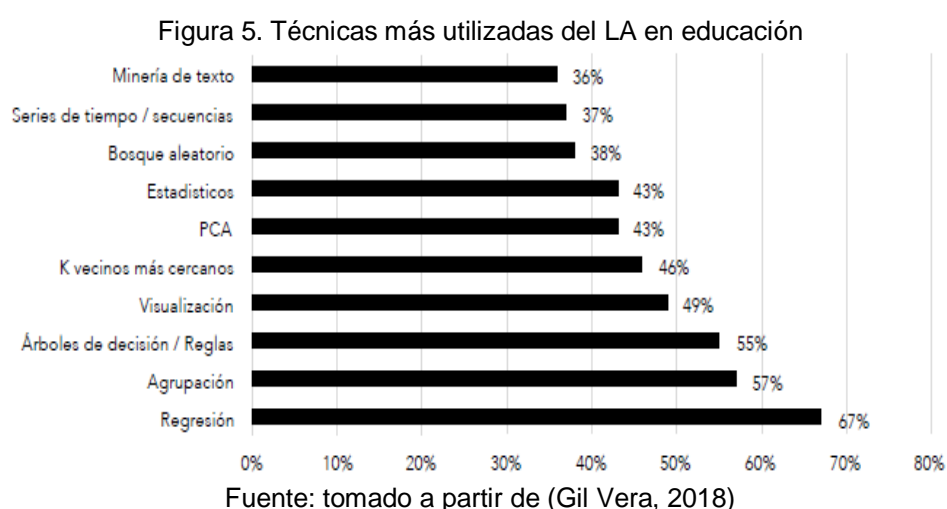
Surge como una disciplina emergente que busca mejorar la enseñanza y el aprendizaje, por medio de una evaluación crítica de datos; *Learning Analytics* apoya la toma de decisiones, adapta el contenido legible, facilita evaluaciones realistas y proporciona supervisión personal del progreso del estudiante, (...) es una disciplina que monitorea el desarrollo del proceso de la enseñanza aprendizaje, constituye de metodologías que realizan una evaluación de los datos y el descubrimiento de reglas del desenvolvimiento de los estudiantes, puede realizar las predicciones, para continuar con

la retroalimentaciones respectiva y así evitar problemas educativos.
(p. 5)

Por ello, si una institución educativa ya sea de nivel medio o superior en, la cual, una de sus políticas es la optimización de recursos apuntados al éxito estudiantil, empezarían a trabajar o a minar sus datos educativos (huella digital educativo) y sacar el máximo provecho, gracias al aporte de las técnicas de *Learning Analytics*.

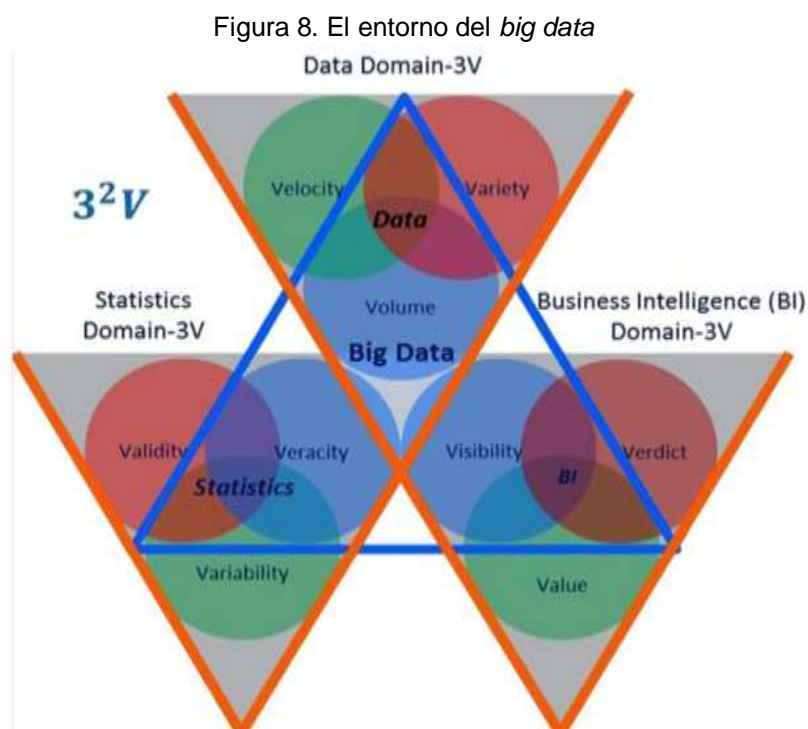
Learning analytics y la educación

Se adjunta las diferentes técnicas del *LA* sugeridas por Gil Vera, (2018) para el ámbito educativo: Predicción, agrupamiento, detección de valores atípicos, minería de relación, análisis de redes sociales, proceso de minería, minería de texto, destilación de datos para el juicio humano, descubrimiento de modelos, gamificación, aprendizaje de máquinas, estadística.



En la Figura 5. se muestra la posición de Gil Vera Víctor Daniel acerca de los diez de los métodos o técnicas del *LA* más utilizados.

La tendencia *LA*, aporta a la ciencia de los datos mediante: las analíticas, la minería de datos educativos, analítica académica y tecnologías del entorno (algoritmos, aplicaciones, entre otros).



Nota: adaptación del autor³.

La tecnología *big data*, se trata de las dimensiones: Dominio, manejo o tratamiento óptimo de los datos, el análisis estadístico, las mismas que están orientados a la inteligencia de negocios, en el ámbito educativo, a la mejora continua de los procesos.

En la gran mayoría de los centros educativos a nivel mundial, los docentes, cada vez utilizan medios tecnológicos en sus procesos de enseñanza aprendizaje, y al interactuar los aprendices y profesores por estos medios, generan datos de uso

² Figura para el entorno del proceso del *Learning Analytics*, del Estado del Arte: *Learning Analytics* (LA). Seminario de LA para la Educación Mediada por Tecnologías de la Información, 2018, (<https://slideplayer.es/slide/13143140/79>).

³ Figura para el entorno del *Big Data*, del Estado del arte: *Learning Analytics* (LA). Seminario de LA para la Educación Mediada por Tecnologías de la Información, 2018, (<https://slideplayer.es/slide/13143140/79>).

de éstos medios, las mismas que por medio de analíticas, se aprovecha en beneficio de la comunidad educativa.

Contexto de las técnicas *learning analytics*

Las Instituciones Educativas, en la actualidad, están en el mundo de las redes, trabajan con las herramientas TIC en sus procesos educativos, en Ecuador, según los datos del INEC (2019), el analfabetismo digital corresponde al 11.4%, con una tendencia a la baja, es decir, el resto, 88,6%, por lo menos utilizan internet, un computador y tiene un celular activo. En esta realidad, el docente, administradores o la comunidad educativa, proponen la mejora de sus procesos de enseñanza-aprendizaje, basados en la analítica de datos generados por los aprendices en éstos medios, las técnicas del LA es la mejor opción, son una tendencia, es un complemento necesario para todo docente.

Al respecto, al tratar de las técnicas y algoritmos presentes en plataformas como el R, hay una variada literatura, en este trabajo, se acoge las sugeridas por Baker & Siemens, (2014), quienes sostienen que *learning analytics* son:

Métodos de predicción (clasificación; regresión; estimación de conocimiento latente), descubrimiento de estructuras (*clustering* (agrupación), análisis de factores, detección de valores atípicos, descubrimiento de estructuras de dominio, análisis de redes sociales), minería de relación (minería de reglas de asociación, minería correlacional, minería secuencial, minería de datos casuales), descubrimiento con modelos, otros métodos (procesos de minería, análisis del discurso, enfoques multimodales, entre otros).
(pp. 4-5)

De ahí que, las diferentes herramientas informáticas (plataformas del LA vigentes), se resume en la parte A Anexo 2. En el mismo sentido, existen, también, una variedad de literatura de la forma de uso de éstas herramientas, mientras que, para fines de este trabajo investigativo, se basa en las técnicas

comunes tanto al ASI como al LA, las mismas que en el capítulo 2, se realiza, una aproximación de su uso.

Limitaciones de las técnicas del *learning analytics*

Interesa el aporte de Gil Vera (2018), con la pregunta: “¿Cuál es el principal problema de la implementación de algoritmos para el análisis de datos en educación?” (p. 17). En la que cita a F. J. García-Peñalvo, Griffiths, Johnson, Sharples, and Sherlock (2014); Greller and Drachsler (2012) y Romero & Ventura, (2010) con:

Algoritmos para el análisis del aprendizaje aún no han sido desarrollados y los existentes (árboles y reglas de decisión, regresión, agrupación, estadísticos descriptivos, visualización, series de tiempo, análisis de secuencias, reglas de asociación, métodos de ensamble, minería de texto, vectores de soporte, entre otros), pueden resultar útiles únicamente a desarrolladores, debido a ello, no hay herramientas generales que puedan ser aplicadas a cualquier sistema educativo (...), se necesitan algoritmos, modelos, herramientas prediseñadas para el contexto educativo (...). (p. 17)

Por lo que, por una parte, al momento de aplicar técnicas de LA en ambientes educativos, se observa la necesidad de creación de ciertos algoritmos y las existentes ajustar a los diferentes ambientes, por ello este conjunto de técnicas está en constante desarrollo.

Al momento de hablar de las limitaciones (o dimensiones) del *del learning analytics* Chatti, Dyckhoff, Schroeder, & Thüs, (2012) menciona: tiene que basar en un modelo de referencia del LA, al responder preguntas de éstas cuatro dimensiones; el quién (analistas), el por qué (motivación u objetivos del análisis), el qué (tipo de información utilizada), y el cómo (métodos y técnicas de análisis) de un proyecto del LA. Son pasos que superarían a la hora de implementar las herramientas del LA.

También merece exponer acerca de las limitaciones externas e internas de la tendencia learning analytics, tal como sostiene Greller and Drachsler (2012) acerca de las dos limitaciones (dimensiones) en el marco de las analíticas de aprendizajes.

Limitaciones externas. - Esta dimensión, se refiere a convenciones (ética, privacidad personal, limitaciones motivadas por la sociabilidad) y las normas (restricciones legales y organizacionales).

Limitaciones internas. - Esta dimensión, se refiere a los factores humanos relevantes, como la competencia, por ejemplo: la interpretación y el pensamiento crítico, la aceptación que podrían entrar en conflicto al compilar las técnicas del LA.

Obstáculos para las técnicas del *learning analytics*

Al hablar de los obstáculos que presenta la tendencia LA, al momento de ser implementada en los diferentes centros de educación, se tiene en cuenta lo siguiente:

El especialista, tiene que ser un profesional por lo menos de datos, ciencias matemáticas y de la computación. Al hablar de la parte pedagógica, los datos tienen ser utilizados con la finalidad de optimizar los ambientes de aprendizajes y lograr los estándares de calidad. Otro de los obstáculos es la velocidad del procesamiento de los datos (Chiappe & Rodríguez, 2017), la latencia, también, conocida como el estado latente del dato. El costo, es otro de los inconvenientes a la hora de implementar LA en una Institución Educativa. Aun así, hay retos como: muchos de los estudiantes aun no superan las expectativas, no están mejorando el rendimiento (por ejemplo, las pruebas estandarizadas PISA).

Otro de los problemas, es la generación o la creación de datos educativos, Wolpers, Najjar, Verbert, & duval, (2007), citado por Ferguson, (2012). En este

sentido Greller & Drachsler, (2012), mencionan: “los usuarios tienen que proporcionar los valores apropiados a los parámetros” (p. 51), de la misma manera Peña-Ayala (2017): “para obtener buenos resultados y modelos, por lo tanto, deben poseer una cierta cantidad de conocimientos especializados para poder encontrar la configuración correcta” (p. 17). Es un reto, disponer una información certera para aplicar los diferentes algoritmos y sus análisis.

Adicionalmente, con respecto a los algoritmos en desarrollo para la ejecución de las técnicas del learning analytics menciona Gil Vera, (2018) que:

Tienen que seleccionar el método o algoritmo específico que quieren utilizar de una amplia gama de métodos y algoritmos disponibles de análisis de datos. La mayoría de los algoritmos para el análisis de datos necesitan ser configurados antes de ser ejecutados, y no están diseñados para usuarios con pocos conocimientos en programación. (p. 16)

En su trabajo, responde las siguientes preguntas: ¿Cuáles son los principales problemas de la implementación de algoritmos para el análisis de datos en educación? ¿Cuáles son las posibles soluciones? ¿Qué casos reales de aplicación del análisis del aprendizaje, se han reportado en la educación superior?

Con respecto al reto de la educación en el manejo de datos de gran volumen y el uso de las técnicas eficientes, Gil Vera (2018) menciona que: “el manejo estratégico y sistematizado de grandes y complejas bases de datos está permitiendo solucionar problemáticas que, se presentan en la educación” (p. 24).

Por lo citado, a la hora de la implementación de estas técnicas fabulosas, se va encontrar con ciertos inconvenientes fuertes, pasa por lo económico, del talento humano, de la falta de capacitación docente en temas de análisis de datos, búsqueda de expertos, el tiempo, entre otros. Por lo que, el reto para los nuevos docentes es que, a más de la gestión del conocimiento, tienen que estar

preparados para el uso de las herramientas tecnológicas y del tratamiento (convertir en información válida) de datos generados en estos medios digitales.

La ética y el *learning analytics*

Al hablar de la parte ética del uso de datos y la privacidad de las mismas, hay una inmensidad de contribuciones al respecto, mencionan Drachsler and Greller (2016):

La adopción generalizada del *LA* y la *Educational Data Mining* (EDM), se ha detenido un poco recientemente y en algunos casos estacados, incluso, se ha revertido debido a las preocupaciones de los gobiernos, las partes interesadas y los grupos de derechos civiles sobre la privacidad y la ética aplicada al manejo de datos personales. (p. 1)

También, Fynn (2016) trata de una manera formal el aspecto de la ética, los valores y los argumentos morales que giran en torno al uso de *Big Data*, utiliza una demostración práctica de análisis de aprendizaje aplicada.

De la misma manera, al llegar a la población estudiantil, en el proceso del minado de datos, Garaizar and Guenaga (2014) muestran en su estudio que: “Los estudiantes esperan que los sistemas de análisis de aprendizaje incluyan tableros personalizados y adaptables; además, los estudiantes son bastante conservadores al compartir datos para sistemas de análisis de aprendizaje” (p. 278-279).

Por otra parte al tratar la privacidad de la información y su uso eficiente, Pardo and Siemens (2014) aportan con: “La privacidad y los problemas éticos que surgen en este contexto están estrechamente interconectados con otros aspectos como la confianza, la responsabilidad y la transparencia” (p. 438).

Al hablar de los beneficios y las oportunidades de las técnicas del learning analytics, Vihavainen, Helminen, and Ihantola (2014), habla de los desafíos: “Los enfoques adoptados para comprender las oportunidades y los desafíos éticos del LA dependen necesariamente de muchos supuestos ideológicos y epistemologías” (p. 109).

En el mismo sentido, de los cuidados que se debe tener con la información Nespereira, Dai, Vilas, and Redondo (2014) mencionan que: “Se ve a la ética y la privacidad como habilitadores cruciales en el *Learning Analytics*” (p. 285), estos aportes nos llaman al cuidado del tratamiento y el uso desenfrenado de los datos generados por los alumnos y sus contextos.

Después de implementar un sistema de analíticas de aprendizajes, en vías de hermetismo o ambientes en las que detecten cuestiones éticas del manejo de datos de actores educativos Martin (2014) sugiere que: para resolver este tipo de problemas éticos, podrían aumentar la transparencias del proceso para los estudiantes y obtener el consentimiento en múltiples niveles.

Al hablar de la barrera en la recopilación de la información, Ochoa, Worsley, Chiluzia, and Luz (2014) Consideran que: “cada vez más, las preocupaciones éticas sobre la recopilación y recolección de datos de los estudiantes, se han planteado, pero no hay una indicación clara de cómo abordar o supervisar estas preocupaciones” (p. 531). Por lo que, existe una barrera ética en el instante del manejo adecuado de los datos privados de los estudiantes.

Por lo expuesto, el último de los problemas detectado, es la parte ética y privacidad del uso de datos, en muchos contextos estos datos, se las usan no en función de mejorar los aprendizajes, como que usan en perseguir a los actores educativos, la cual, crea un hermetismo en los ambientes de aprendizajes.

Tendencia del *learning analytics*

El uso creciente de una multitud de recursos de *e-learning*, *Software* Educativo, herramientas de aprendizaje e Internet en la educación produce una gran cantidad de datos relacionados con el aprendizaje. Esta información proporciona una mina de oro de datos educativos que pueden explorarse y explotarse para comprender cómo aprenden los alumnos (Romero & Ventura, 2010).

El marco de análisis de aprendizaje actualmente en desarrollo, constituye un enfoque para evaluar, monitorear, informar y promover conocimiento de manera efectiva y competencia, integra datos educativos provenientes de múltiples actividades y fuentes de datos. Este enfoque, se caracteriza por una combinación de métodos basados en la teoría y los datos como base para *LA* centrada en competencias y ayuda a construir una comprensión más completa y precisa de aprendizaje y progreso que, se podrían informar a los estudiantes e instructores y sirve de apoyo para aprender, optimizar la enseñanza y perfeccionar las metodologías de *LA* en sí. La investigación conceptual es implementada en términos de un conjunto de servicios de *LA* e integrado en el sistema global.

Retos futuros del *learning analytics*

Construir una fuerte conexión con las ciencias del aprendizajes, desarrollar métodos que integre diferentes frentes de trabajo para optimizar los ambientes de aprendizajes (*mobile data*, *biometric data* y *mood data*), enfocarse en las perspectivas de los estudiantes, más que en las necesidades de la institución, desarrollo y aplicación de un claro conjunto de líneas éticas (Ferguson, 2012).

Hay una creciente necesidad de sistemas de indicadores y métricas para medir el desempeño personalizado de los alumnos (Higueras, Campano, Olivera, Suances, & Gómez, 2014). Plantea la interrogante: ¿Se encuentran las instituciones preparadas para adoptar el *LA*?, (Sclater *et al.*, 2017). Por lo que la

correcta implementación del *LA*, con procesos de transparencia, se podría pensar en un futuro en ambientes educativos más tranquilos y efectivos.

Difusión de las técnicas *LA* en el contexto ecuatoriano

Con respecto al uso de las técnicas del *LA* en el Ecuador, R. Pazmiño-Maji, Naranjo-Ordoñez, Conde-González, & García-Peñalvo, (2019) presentan:

Tabla 1. Publicaciones en relación a las técnicas del *LA*

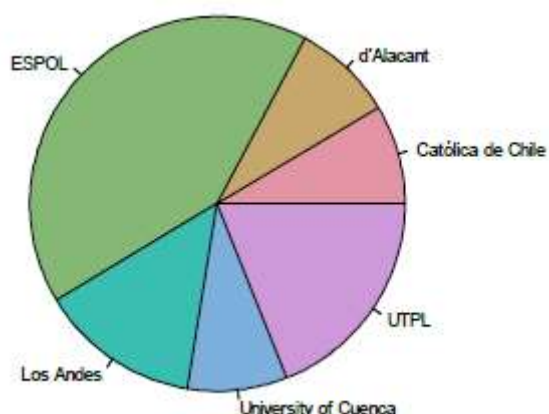
Año	2013	2014	2015	2016	2017	2018	Junio 2019
SCOPUS	1	3	9	10	8	13	8
WOS			7	5	2		
IEEE							
RRAAE				1		1	

Fuente: tomado a partir de (R. Pazmiño-Maji et al., 2019)

La Tabla 1. indica que: desde al año 2013, cerca de 68 artículos referentes al *learning analytics*, se han publicado en el País.

De la misma forma, al hablar de los centros de educación superior, en las que han sido publicadas. En la Figura 9. se observa que el 41.4% de las publicaciones de ecuatorianos son en la Escuela Superior Politécnica del Litoral (ESPOL), seguido por el 19% en la Universidad Técnica Particular de Loja (UTPL) y 8.6% por la Universidad de Cuenca. Por otra parte, el 69% de las publicaciones tienen a ecuatorianos como autores principales, mientras que el 31% de las publicaciones no presentan a los ecuatorianos como autores principales. Las universidades que presentan como autores principales son: Universidad de los Andes de Venezuela, la Universidad Católica de Chile y la Universidad de Alicante de España.

Figura 9. Las seis universidades con publicaciones LA de autores ecuatorianos



Fuente: tomado a partir de (R. Pazmiño-Maji et al., 2019)

En estos aportes, también, se observan que en el año 2018 está un trabajo de tesis de Maestría publicada en la Pontificia Universidad Católica del Ecuador sede Ambato, con el tema: Estudio comparativo del Análisis Estadístico Implicativo y el *Learning Analytics* en relación al uso de las técnicas de exploración de datos educativos, esta predisposición continúa para mediados del año 2019, con 18 nuevas publicaciones. Por otra parte, (Pazmiño Maji, 2021) en su tesis doctoral, menciona: “Las investigaciones de alto impacto en *Learning Analytics* en Ecuador son todavía pocas, pero suficientes, para establecer una línea de base, que permitirá a los investigadores, tener un punto de partida (...) en el análisis de aprendizajes”.

Este efecto abre un campo, para que las técnicas del *Learning Analytics* sean utilizadas el nivel medio del ámbito educativo ecuatoriano, existen trabajos de base, que impulsan la continuación de la presente propuesta de las simulaciones digitales.

Razones del uso de las técnicas ASI y LA en el bachillerato

Por una parte, se ha mencionado, varias problemáticas a la hora de utilizar las diferentes técnicas del LA, entre ellos, la parte económica, la privacidad de los datos y últimamente la parte ética de cómo, se utilizan éstos datos, en ello R. A. Pazmiño-Maji *et al.* (2016), reporta el uso educativo de los métodos del ASI más

próximos a las técnicas del *LA*: La Minería de reglas de asociación (en un 95,8% de los artículos analizados), *Clustering* (en un 37,5%) y Estadísticas (en un 20,8% de los documentos analizados), el autor concluye que existe mucha experiencia en la aplicación en el área educativa, y motiva a los educadores a utilizar esta nueva técnica de estadística multivariada.

De la misma forma como se citó más arriba, en el contexto de la Universidad PUCESA, (Naranjo, Maji, Conde, & Peñalvo, 2018), aporta el uso educativo de las técnicas del ASI y el LA.

Técnicas comunes del ASI y el LA

Interesa el aporte de R. A. Pazmiño-Maji et al., (2016), quienes después de un estudio, describen las técnicas comunes a estas dos tendencias de análisis de datos masivos en ambientes educativos: De los artículos analizados: según la clasificación de Baker e Inventado, el 95,8% de los documentos hablan de la minería de reglas de asociación, de la minería de datos casuales el 4,2%, y el 37,5% de los artículos tratan del *Clustering* (Agrupación), por otra parte, de la clasificación de *Papamitsiou* y *Economides*, el 20,8% de los artículos tratan de las técnicas estadísticas.

De la misma manera, Naranjo et al., (2018), con su diseño cuasi-experimental sugieren: el uso docente de las técnicas comunes del ASI y el LA, resumidas en el siguiente cuadro:

Cuadro 1. Técnicas comunes entre el ASI y el LA

Técnicas	Learning Analytics	Análisis Estadístico Implicativo
Clustering (Agrupación)	hclust.vector, dendro.variables, dendro.diana	Árbol jerárquico (Hierarchy Tree), Árbol de similaridad (Similarity Tree),
Minerías de reglas de asociación	A priori, Weclat, Eclat.	Grafo Implicativo (ImplicativeGraph).
Técnicas estadísticas	Summary del LA	Summary ASI

Fuente: tomado a partir de (Kaufman & Rousseeuw, 2009)

CAPÍTULO II. DISEÑO METODOLÓGICO

2.1. Metodología de la investigación

El paradigma de la presente investigación, se fundamenta en la teórica, crítica y propositiva, debido a la participación activa de los sujetos en el proceso investigativo, su aplicación, basado en conocimientos previos. En este sentido, se pretende crear simulaciones digitales para aproximar a la comprensión docente, los conceptos y algoritmos de la tendencia ASI y el LA. Por lo tanto, es necesario fundamentar la construcción de ideas a través de la recopilación de información oportuna con sustento bibliográfico primario, diagnosticar el uso docente de las herramientas digitales como el *Office* y los simuladores, proponer medios de interacción amigables y estimular el uso de la tendencia en su quehacer educacional.

Para la explicación de los preámbulos del presente trabajo, se utiliza la alternativa del enfoque cualitativo, como menciona con respecto a este enfoque Daza, (2018): “Permite profundizar con mayor amplitud la práctica, el fenómeno educativo y contribuye a la concienciación social, (...), posibilita analizar y explicar la comprensión de la complejidad” (p. 105-106). De la misma forma Sampieri et al., (2014) menciona: “con frecuencia, estas actividades sirven, primero, para descubrir cuáles son las preguntas de investigación más importantes, y después, para refinarlas y responderlas” (p. 7).

Enfoque que nos permite interiorizar en el problema, buscar la alternativa para la ejecución de la propuesta en la comunidad de docentes de la Unidad Educativa, quienes, en estos tiempos al utilizar la propuesta, para analizar los datos generados por los aprendices, en el uso de las TIC.

Tipo de investigación

Conforme al enfoque cualitativo del trabajo de investigación, el estudio, se relaciona a: Investigación descriptiva: Como recomienda Sampieri *et al.* (2014) “con los estudios descriptivos, se busca especificar las propiedades, las características y los perfiles de personas, grupos, comunidades, procesos, objetos o cualquier otro fenómeno que, se someta a un análisis” (p. 92). En efecto, se reconocen las posibles aristas que impiden a la comunidad docente utilizar las técnicas del ASI y LA, y buscar las posibles alternativas de llevar este conocimiento a esta comunidad y su uso en su quehacer educacional.

Investigación analítica-sintética: La investigación analítica-sintética, observa los hechos, trata de descomponer el objeto de estudio en sus partes y analizarlas en forma individual, para su integración posterior. En este sentido, para el caso, se indaga la teoría ASI, las técnicas de las que constituye, los estudios comparativos con los algoritmos del LA, y abstraer lo factible, de acuerdo a la capacidad docente para el uso cotidiano.

La investigación propositiva: Se elaboran simulaciones digitales para cada técnica, relacionadas a contextos educativos. Ambientes amigables, que facilitan la comprensión y aplicación docentes de éstas técnicas.

Investigación documental-bibliográfica: Se sintetiza, la información bibliográfica en relación a las técnicas del ASI y LA, para abstraer bases del marco teórico de la investigación y el proceso de interpretación de dicha información.

Modalidad de investigación

La modalidad de la presente investigación va enfocada al aspecto socioeducativo, la propuesta pretende aportar al análisis de datos presentes, en el uso de las herramientas tecnológicas, entre el docente y el estudiante.

Revisión bibliográfica-documental

El estudio permite explorar las diferentes técnicas del ASI y el LA y sus aplicaciones educativas en el nivel medio, se contrasta y se analiza una cantidad finita de fuentes de información, recopiladas de fuentes como: textos, artículos científicos, congresos ASI (nacional e internacional), repositorios de las universidades, entre otras, con la finalidad de obtener toda la información pertinente, actual y aplicable al contexto a la que, se desea aplicar.

De campo

El estudio, se enfoca en el sitio en que, se presenta la problemática de falta del uso de las técnicas del ASI y LA en procesos de interacción estudiante-docente quienes emplean medio tecnológicos. Se trabaja con la población docente (18) de acuerdo a sus capacidades, como solución, se diseñan simulaciones digitales con la información de la tendencia ASI y LA, junto con ejercicios propuestos y resueltos en la herramienta *Excel* destinado a la comprensión de las técnicas ASI y LA en forma de algoritmos, ésta información, se presenta en una página web.

Diseño de investigación

El diseño de la investigación es cuasiexperimental, se trabajó con dos grupos, el de control y experimental. En diseños cuasiexperimentales, los individuos no, se asignan al azar a los grupos ni, se emparejan, sino que dichos grupos ya están conformados antes del experimento: son grupos intactos (Sampieri, Collado, & Lucio, 2014).

Cuadro 2. Diseño de investigación

Grupo	Pre-test	Intervención	Post-test
Experimental	G1	SIMULACIONES DIGITALES	G3
Control	G2	-----	G4

Fuente: elaboración propia

En un primer momento, en un estudio mixto, se diagnostica la posibilidad de trabajar con las técnicas del ASI y LA a los docentes de la UE de la sección de bachillerato, para en un posterior, llegar a difundir la tendencia a través de simuladores presentes en una página *web*.

En el desarrollo del trabajo de campo, después aplicar la encuesta a los docentes, se crean las simulaciones digitales, se efectúa un pre-test (Anexos C3 y C4) tanto al grupo experimental como al grupo de control; seguidamente, se trabaja con las simulaciones digitales ASI y LA en los docentes del grupo experimental. Para terminar, se ejecuta un post-test a los dos grupos (Anexos C3 y C4).

2.2. Recolección de la información

El presente trabajo de investigación, se fundamenta en la recopilación bibliográfica y estudio de campo. El aspecto técnico de la información, se inicia con el diagnóstico docente de las capacidades del uso de las TIC, técnicas estadísticas y el *Software R*. Seguidamente, de acuerdo a las competencias docentes, se diseñan simulaciones en la herramienta *Excel* y en *R*, las mismas que para el uso docente, se publican en una página de *Google Site*.

Para la comprobación del nivel de comprensión docente, objeto de esta investigación, se plantea un instrumento Pre y Post Test (parte C del Anexo 3), validada por juicio de expertos (parte D del Anexo 3), finaliza con la parte de la validación de la propuesta por juicio de expertos (parte E del Anexo 3).

La evaluación de expertos de la propuesta, se lo realiza por medio de una ficha de validación, este instrumento permite contrastar los criterios de: calidad,

objetividad, actualidad, organización, suficiencia, intencionalidad, consistencia, coherencia, metodología y pertinencia de las simulaciones ASI y LA, hacia la comunidad de docentes de bachillerato.

Población

Para la presente investigación, se ha tomado como referencia a los 18 docentes (9 mujeres y 9 hombres) de bachillerato de la UEMIB Chibuleo, para el grupo experimental, se emplea el criterio de muestra no probabilística intencionada, y colaboran como grupo experimental 7 docentes (2 mujeres y 5 hombres), basado en el diagnóstico y por sus afinidades a la temática propuesta.

Técnicas e instrumentos de recolección de datos

La técnica empleada fue una encuesta (parte A del Anexo 3), como instrumento utiliza un cuestionario (parte B del Anexo 3), la cual, se aplica a la comunidad de docentes de bachillerato con la finalidad de diagnosticar los conocimientos generales en el ASI, el uso de las TIC y el *Software R*.

El cuestionario, se elaboró en dos secciones, una primera parte que solicita los datos generales del docente y la segunda parte que contiene preguntas relacionadas con el conocimiento y uso docente de las TIC, la herramienta *Excel*, los medios digitales y el *Software* estadístico *R*, de la misma forma el uso de las técnicas ASI y LA en contextos educativos.

En base a los resultados y en el empleo de la metodología cuasi-experimental, se desarrolla la prueba pre y post test (parte C del Anexo 3) y la validación respectiva (parte D del Anexo 3).

Validez del instrumento

Para la validación del instrumento de investigación (encuesta), se emplea la metodología de juicio a expertos para la cual, se solicitó a seis profesionales relacionados con el área de la educación y de estadística matemática, que evalúen los criterios de: calidad, objetividad, actualidad, organización, suficiencia, intencionalidad, consistencia, coherencia, metodología y pertinencia en una escala del 0 a 100% (parte B del Anexo 3).

Los resultados que arrojaron el proceso de revisión de expertos indicaron que el cuestionario es afín al tema y el objetivo que persigue. El cálculo de V de Aiken arroja 0.97, indica la factibilidad de aplicación de dicha encuesta en la población de docentes de bachillerato de la UEMIB Chibuleo.

Confiabilidad del instrumento

Para la obtención de la confiabilidad del instrumento, se aplica el cuestionario a un grupo de 15 docentes de la Unidad Educativa, con similares características al del grupo de investigación con la finalidad de verificar la comprensión de las preguntas. Se centra en los ítems 3,4 y del 6 al 10 de la encuesta (parte A del Anexo 3) debido a que estas preguntas, se centran en las capacidades docentes en el uso de las TIC, la herramienta *Excel*, el *Software* estadístico *R*, de la misma forma el uso de las técnicas ASI y LA en contextos educativos. En la Tabla 2. se muestran estos resultados.

Tabla 2. Estadístico de confiabilidad

Alfa de Cronbach	No. Items
0,8375	7

Fuente: elaboración propia

El estadístico de confiabilidad en la Tabla 2. anterior es de 0,8375. Que de acuerdo a Sampieri, Collado, & Lucio, (2014), el instrumento de medición es de consistencia muy elevado, que da apertura a su aplicación.

Ubicación de la institución educativa

La investigación, se realiza en el Distrito Educativo 18D02, ubicado en el Cantón Ambato de la Provincia de Tungurahua. La comunidad Chibuleo, se encuentra al sur-oeste de la provincia, a 12km de distancia de la ciudad de Ambato, ubicada entre los 3 100 y los 3 340 msnm, con un clima de entre 7°C a 15°C.

La misión de la UEMIB Chibuleo, del cantón Ambato de la provincia de Tungurahua, fusionada desde el 9 de octubre del 2014, garantiza el derecho a una educación bilingüe intercultural en los niveles: inicial, educación general básica, bachillerato general unificado y bachillerato técnico en contabilidad; encaminada a la construcción del proyecto de vida como práctica del *kumak kawsay*, fortaleciendo los saberes ancestrales, sustentada en el modelo socio-constructivista, pedagogía crítica-reflexiva, capaces de interactuar en la sociedad globalizada, con conciencia ecológica e igualdad de deberes y derechos.

La visión de la UEMIB Chibuleo del cantón Ambato de la provincia de Tungurahua, ofrecerá educación integral, científica, inclusiva, con calidad, calidez y pertinencia étnica-cultural; contribuirá a la formación de seres humanos críticos, reflexivos, dinámicos, innovadores, productivos, con mentalidad internacional y capaces de ejecutar su proyecto de vida centrado en el desarrollo holístico de la persona.

2.3. Propuesta

Con la finalidad de rejuvenecer la cultura de análisis de datos generados en ambientes educativos que utilizan medios digitales, se propone como estrategia de aprendizaje docente el uso de las simulaciones con las técnicas de Análisis Estadístico Implicativo (ASI) y el *Learning Analytics (LA)*. En el desarrollo de la propuesta, se presentan los procesos de creación de las simulaciones digitales, como una estrategia para aproximar al uso de las técnicas comunes del ASI y LA en el ámbito del bachillerato.

Desarrollo de la propuesta

La propuesta está conformada por esquemas didácticos del uso de las simulaciones, para aproximar al docente de bachillerato, con las técnicas comunes del ASI y LA, aplicables a este medio.

Presentación

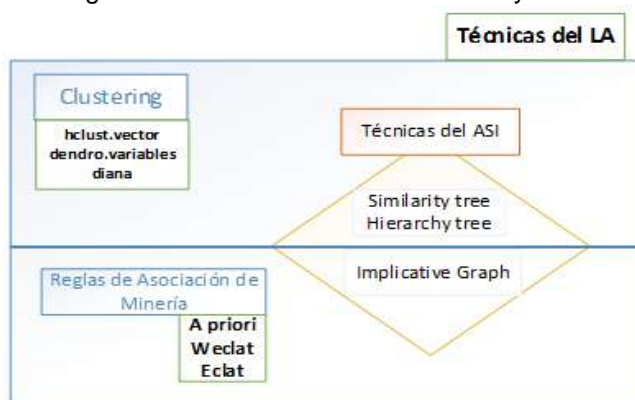
En la UEMIB Chibuleo, por el uso frecuente de medios digitales entre los docentes y estudiantes en el proceso de enseñanza y aprendizaje, se crean datos de esta índole que por el desconocimiento docente no, se procesa para obtener información útil en favor de la mejora continua de los procesos educativos. Las técnicas del ASI y LA, favorecen el descubrimiento en forma de reglas, basado en datos (si la utilización de un tipo de herramienta digital influye en la obtención de estudiantes triunfantes o no).

Con la finalidad de llegar a más docentes y de acuerdo a las capacidades de los usuarios docentes en el uso de las herramientas tecnológicas, se crean simulaciones digitales (procesos de aproximación docente) y se las presentan en una página de *Google Site*.

Contenido y funcionamiento

Las simulaciones digitales, contienen esquemas didácticos para la aproximación a las técnicas del ASI y LA, procesos sugeridos por Pazmiño R, Naranjo M y sus colaboradores, resumidas en: **Técnicas de minerías de reglas de asociación** como: (del ASI: Grafo Implicativo, del LA: *Apriori*, *Weclat* y *Eclat*,) y **técnicas del Clustering** (del ASI: Árboles Jerárquicos '*Hierarchy Tree*' y Similaridad '*Similarity Tree*' y del LA: *hclust.vector*, *dendro.variables* y *dendro.diana*), las mismas que para una mejor comprensión, se esquematizan en la siguiente Figura 10.

Figura 10. Técnicas comunes del ASI y el LA



Fuente: elaboración propia

Para la preparación y presentación, se recurre al método “*Building valid models: how to build valid and credible simulation models*” de (Law & McComas, 2001), las mismas que adaptadas a la realidad comprenden: selección, análisis y descripción en un lenguaje más comprensible, de los fundamentos y de las técnicas comunes del Análisis Estadístico Implicativo y el *Learning Analytics*; delimitación del problema; recolección de la información y construcción del modelo conceptual de la simulación; comprobar el modelo conceptual; programación del modelo de simulación en R; comprobar la validez del modelo programado; diseño, realización y análisis de los experimentos simulados; documentación y presentación de los resultados de la simulación, en efecto:

Técnicas comunes del ASI y el LA en educación

En el proceso de la selección, análisis y descripción para el ámbito educativo de las técnicas comunes del análisis estadístico implicativo y el learning analytics, se acoge los aportes de (Naranjo et al., 2018), resumida en el Cuadro 3. de las técnicas del learning analytics: Los algoritmos de regla de asociación de minería, clustering y minería de datos causales. De las técnicas del análisis estadístico implicativo, se acoge a los algoritmos de cohesión, implicación y similaridad. Esta colección de técnicas para su aproximación se ha reducido su interpretación, y llevadas a las hojas de excel, dominio común en la comunidad de docentes.

Cuadro 3. Detalle de las técnicas del ASI y LA con enfoque educativo

Técnicas	Nombre	Objetivo	Relación con la adquisición y carencias
Learning Analytics	Association Rule Mining. (Regla de Asociación de Minería)	Realizar un proceso de obtención de reglas y asociarlas con el propósito de obtener información que permita la toma de decisiones en el ámbito educativo.	Se representa como un método general para el proceso. Carece de una secuencia de procesos. Es un método que indica los aspectos que, se deben tener en cuenta para el proceso dentro del LA sin el diseño de un proceso específico.
	Clustering	Ofrece <i>cluster</i> o grupo de datos asociados por características comunes, que ayuden en tomar decisiones frente a un problema educativo.	Es una secuencia de pasos generales, basado en agrupación con nivel de complejidad. Es una propuesta importante para el LA. Carece de un enfoque especial para la interpretación de resultados, inserta un nivel de complejidad alto, que dificulta la implementación en pequeños entornos.
	Causal Data Mining, (Minería de Datos Causales)	Establecer causas de un evento y encontrar posibles eventos causales de otros eventos, con la finalidad de obtener resultados de las causas que originan un problema.	Es la ilustración por medio del cual, se centra en obtener los eventos causales, he aquí la importancia dentro del LA. Carece de un enfoque general por lo que dificultaría su adaptación a otros ámbitos, en especial en el área donde es difícil obtener causales.
Análisis estadístico implicativo	Cohesión e implicación (Gras et al., 2008)	Establece cualidades entre variables, con el propósito de obtener una clasificación jerárquica implicativa y cohesiva a través de reglas.	Presenta un método general basándose en las variables ejecutar un proceso, mediante el manejo de reglas que permitan determinar aspectos de cohesión e implicación dentro del ASI, además, con una herramienta (CHIC) muy eficaz a la hora de ejecutar este método.
	Similaridad (Lerman)	Formula cuasi-regla de similaridad, basándose en los niveles de confianza, con el objetivo de obtener similaridad entre las variables, esto basándose en la ley de Poisson.	La ilustración, se da por medio de la formulación de cuasi-reglas de similaridad lo cual, se centra en encontrar niveles de confianza, con la finalidad de encontrar similitudes entre variables, he aquí la importancia dentro del ASI. Posee un enfoque general, por lo que es de gran ayuda dentro de cualquier área.

Fuente: tomado a partir de (Naranjo et al., 2018)

En este contexto, complementa con el aporte de Valls Pla X, (2014), en que: “el *clustering* es utilizada para agrupar los datos, definiendo perfiles para adaptar las soluciones a esos datos, mientras que las reglas de asociación y la clasificación

son utilizadas como método de comprensión de patrones y predicción de resultados” (p. 8).

AL seguir la idea, Valls Pla X, (2014) menciona:

El Análisis Estadístico Implicativo engloba tres procedimientos distintos e independientes: *Análisis de similaridad o clasificadorio*. – *Cluster analysis* con una forma original de medir distancias, las similitudes del Lerman. Se forma un árbol jerárquico similar a un *dendrograma* (*cluster* jerárquico ascendente). *Análisis implicativo*. - Genera una matriz con todas las implicaciones a->b encontradas en los datos y forma un grafo implicativo, con flechas relacionando las variables con las implicaciones más fuertes. *Análisis de cohesiones*. - Construye un árbol jerárquico parecido al árbol de similitudes, en el que, se marcan de nuevo las variables con implicaciones más fuertes mediante flechas. (p. 19)

Por lo tanto, el tipo de dato a ser ejecuta por estas técnicas son del tipo binarios, modales y numéricos, ejemplos de datos binarios en el ámbito educativo (si/no, verdadero/falso, admite/no admite, con propiedad/sin propiedad, aprueba/desaprueba, admitido/no admitido, promovido/no promovido, motivado/desmotivado); ejemplos de datos de tipo modal en el ámbito educativo (bajo/medio/alto, malo/regular/bueno/muy bueno/sobresaliente, nada de acuerdo/poco de acuerdo/ni en acuerdo ni en desacuerdo/muy de acuerdo/completamente de acuerdo); ejemplos de datos de tipo numérico en el ámbito educativo (notas, peso, edad, rendimiento académico, asistencia, número de faltas).

Es decir, las técnicas en mención, permiten observar las asociaciones e implicaciones entre variables como: sexo, edad, nivel de rendimiento, aprovechamiento, entre otros. Por lo tanto, se sigue las recomendaciones de M. Naranjo, basado en los aportes de R. Pazmiño, por el enfoque general, por el enfoque específico, por los procesos, eventos y características como fase inicial,

se acoge para el ASI y el LA: Del *Learning Analytics* (*Association Rule Mining*, (Regla de Asociación de Minería) y *Clustering*) y del Análisis Estadístico Implicativo (Cohesión e implicación y Similitud de (Lerman)), existen ejemplos aplicables al ámbito educativo, ejecutables en *Excel* en línea y en *R*.

Delimitación

En el proceso de la delimitación de la problemática, después de llevar a cabo la encuesta (ver Anexo 3) hacia los 18 docentes del bachillerato de la UEMIB Chibuleo, se detecta que los mismos no tienen bases sólidas en estadística matemática y al parecer como consecuencia, no dominan las técnicas comunes del análisis estadístico implicativo y el *learning analytics*, de la misma forma desconocen del uso y aplicaciones de los *Software: R, RStudio y RStudio Cloud*. Por otra parte, existe experiencia sólida docente en el uso de la herramienta excel, de la misma forma, existe la predisposición del profesorado en adquirir conocimiento en el ámbito del análisis de datos generados en sus ambientes de enseñanza y aprendizaje.

Para dar con una solución a estas dificultades, este trabajo inicia con la identificación de las técnicas comunes entre el ASI y el LA, se detecta las sugerencias de: Kaufman L, Pazmiño R, Naranjo M y sus colaboradores (ver Cuadro 1.), se ejecuta la selección, análisis y descripción en el contexto educativo de dichas técnicas, se crea así las diferentes simulaciones digitales. Seguidamente, con la elaboración del trabajo se las simulaciones digitales, presentadas primero en la herramienta excel hasta el uso de funciones en R. Termina con la presentación de dichas simulaciones en una página de *Google Site* (ver <https://bit.ly/2PUWrPw>).

En el proceso de la recolección de la información para la construcción del modelo conceptual de simulación, que es una fase propuesta por A. Law y M. McComas, se empezó con la investigación bibliográfica en lo referente a las técnicas

comunes al análisis estadístico implicativo y el *learning analytics*, las técnicas del LA son cambiantes, debido a su evolución permanente.

Seguidamente, de acuerdo a la capacidad de los maestros, se elaboran ejemplos prácticos de su quehacer académico (simulaciones en hojas de *excel* en línea) encaminados a un primer acercamiento a la comprensión de las técnicas comunes del análisis estadístico implicativo y el *learning analytics*.

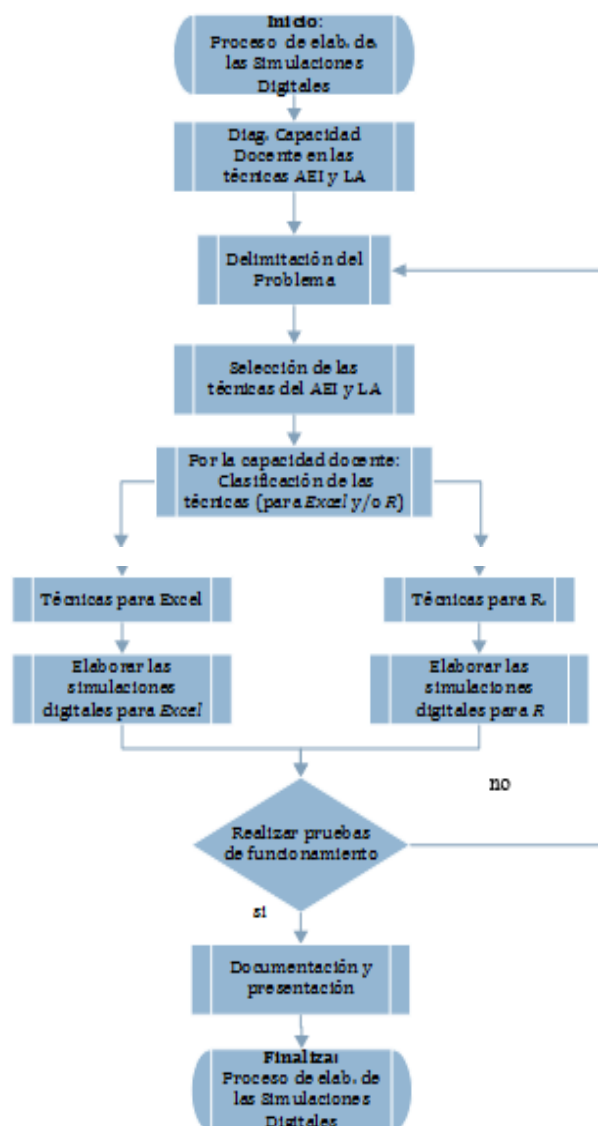
Esta experiencia, se amplía hacia la utilización del *Software Rchic* de *R*, *RStudio* y *RStudio Cloud*, éstos resultados, se presentan en una página de *Google Site* con: (1) Información referente a las técnicas comunes entre el ASI y el LA como son: *similaridad*, *implicación*, *cohesión*, *hclust.vector*, *dendro.variables*, *dendro.diana*, *apriori*, *weclat* y *eclat*, se extiende con la parte explicativa del uso de la herramienta *Rchic* de *R*, *RStudio* y *RStudio Cloud* y (2) la sección de simulaciones en *excel* y en *R* con lo que, se espera que los docentes apropien de estas técnicas, a la hora de extraer conocimiento basados en datos masivos generados por los aprendices en el uso herramientas digitales como medios de aprendizaje.

En el proceso de comprobación del modelo conceptual propuesto por A. Law y M. McComas, los trabajos de simulaciones, se llevan a revisores expertos, se realiza las correcciones del caso. Otra de las formas de comprobación y/o corrección, es la puesta en marcha de los simuladores, en que los docentes aportan con criterios y sugerencias de campo para la mejora continua, es un proceso de corrección y mejoramiento de los simuladores, hasta llegar a satisfacer la mayoría de los criterios u observaciones.

En el proceso de la programación del modelo de simulación, se inicia en dos etapas: (1) en base a las capacidades docentes se elaboran simulaciones digitales en la herramienta *excel*, constituyendo un conjunto de ejercicios en esta herramienta con el uso de datos de uso común educativo y, (2) basado en la experiencia docente del uso de los simuladores en *excel*, se extiende estos

aprendizajes en el uso de fórmulas (funciones o comandos) en el software R. Se utilizó las ideas planteadas en el siguiente flujograma:

Figura 11. Flujograma del proceso de elaboración de las simulaciones digitales



Fuente: elaboración propia

Para la comprobación de la validez del modelo programado, propuesta por A. Law y M. Mc Comas, como se muestra en la Figura 11. existe el proceso de verificación de la validez. En el primer ciclo, se basa en la capacidad de la mayoría de los docentes en el uso de las herramientas *Office*, se emprende a la elaboración de las simulaciones (documentos *excel* en *línea*) para que el docente aproxime a la comprensión de las técnicas comunes del análisis estadístico

implicativo y el learning analytics. Un segundo ciclo, consiste en las simulaciones en *R*, en las que el docente, emplea códigos de *R*, utiliza los datos de su contexto educativo y aproxima al uso de las técnicas en mención. En cada proceso, el docente evidencia el avance, a través de un cuestionario en línea. De la misma forma, otra de las formas de revisar la validez de éstos procesos, es llevar a cabo una validación de expertos.

En el proceso del diseño, realización y análisis de los experimentos simulados, basado en la capacidad de los docentes en el uso de las herramientas *Office*, con la intención de familiarizar a los docentes en las técnicas comunes del análisis estadístico implicativo y el learning analytics, esta experiencia, se ejecuta en dos procesos: (1) las simulaciones en *excel* en línea (las rutinas sugeridas del uso de las técnicas del análisis estadístico implicativo de Zamora y sus colaboradores) y, (2) simulaciones en *R* (de las técnicas comunes del análisis estadístico implicativo y el learning analytics sugeridas por Pazmiño R, Naranjo M y sus colaboradores).

Para la presentación de las simulaciones, se observa las recomendaciones de ((Kuyper, Hoog, & Jong, 2001); citado en (Carvajal, 2010)) quien, para el diseño de una simulación para entornos educativos considera tres pasos principales: “la especificación de los requisitos a partir de los objetivos del diseño; la construcción de bloques basados en la aplicación sobre la base de las necesidades y; la evaluación de la aplicación (...), llamado modelo evolutivo” (p. 12).

En el proceso de la evaluación preliminar de las simulaciones, se presentan los trabajos a la comunidad de docentes a través del enlace: <https://bit.ly/2PUWrPw>, quienes interactúan con la herramienta digital en: *Office Excel*, *RStudio*, *RStudio Cloud* y *Rchic*. Existe gran interés de los docentes en el uso de los simuladores para aproximar su comprensión y aplicación de las técnicas comunes del análisis estadístico implicativo y el learning analytics en contextos educativos. Se recomienda a los usuarios familiarizarse y actualizar de forma continua los diferentes paquetes *R* utilizadas en estas simulaciones debido a la actualización permanente de esta tecnología.

En el proceso de la documentación y presentación de los resultados de las simulaciones, se resume las simulaciones desarrolladas en *Office Excel* y en *R*, que de acuerdo a A. Law y M. McComas comprenden: los supuestos del modelo de simulación, una descripción detallada de las simulaciones y los resultados del estudio actual. Para el uso de las simulaciones digitales, es necesario tener un conocimiento básico de la herramienta *Office-Excel* como del *Software (RStudio, R* y por último, *RStudio Cloud*), la carga de las diferentes librerías (ver la parte C del Anexo 2) en el caso del *Rchic* es necesario contactar con el desarrollador Raphael Couturier quien facilita las últimas versiones.

Las simulaciones, se presentan en una página de *Google Site* (ver enlace <https://bit.ly/2PUWrPw>) con una información referente a las técnicas comunes del análisis estadístico implicative y el learning analytics, acompañado de ejemplos y ejercicios que el docente relacione con el quehacer educativo y así aproximar a su uso.

Materiales y herramientas

A continuación, se detalla las especificaciones técnicas del *hardware* del equipo necesario para la instalación del *R*, *RStudio* y *RStudio Cloud*.

Tabla 3. Especificaciones del *hardware* necesario para las simulaciones

Sistema operativo	Memoria RAM [GB]	Velocidad del CPU [GHz]	Espacio de disco duro [GB]	Tipo de sistema operativo
Windows 8 o superior	4,00 o superior	2,60 o superior	500 o superior	64 bits
Ubuntu 16.04	4,00 o superior	2,60 o superior	500 o superior	64 bits

Fuente: elaboración propia

Tabla 4. Aplicaciones a ser instaladas

Aplicación	Versión
R	4.2.1 (2022-06-23 funcional con Rchic 0,27)
RStudio	2022.02.3+492 (2022-05-20)
Rchic	0,27 (funcional con la versión de R)

Fuente: elaboración propia

Software R

Es un *Software* Estadístico de código libre, gratuito y de libre distribución, se sustenta en una comunidad de desarrolladores muy extenso, la aplicación es descargable de la página: <https://www.r-project.org/>, su instalación, se guía en el enlace: <https://bit.ly/3mu89N9>, (ver la parte A del Anexo 1).

Software RStudio

Según González M (2016), “R Studio es un entorno de desarrollo integrado (IDE) para R, que funciona con la versión estándar de R disponible en CRAN (*The Comprehensive R Archive Network*), es de código libre”. Su instalación, se guía en la página: <https://bit.ly/3mu89N9>, (ver la parte A del Anexo 1).

Plataforma RStudio Cloud

La presentación en forma de plataforma de *RStudio*, se llama *RStudio Cloud* (<https://rstudio.cloud/>), ingresar con una cuenta de *Gmail*. Su instalación y uso guía en el sitio: <https://www.youtube.com/watch?v=o--rUNfzxz8>

Software Rchic

La versión multiplataforma de CHIC de código libre, se llama *Rchic* y se usa desde el año 2015 (R. A. Pazmiño-Maji *et al.*, 2016). Su instalación y uso, se guía en la página: <https://bit.ly/3wExmJb>, (ver Anexo 1).

Diseño y elaboración de las simulaciones digitales

Se diseñan simulaciones digitales para las técnicas: *Clustering* y Reglas de Asociación de Minería comunes del ASI y el LA, es decir, las siguientes técnicas:

De la técnica: Clustering y método: análisis estadístico implicativo. Nombres de las técnicas: árbol jerárquico (en Rchic: callHierarchy Tree), árbol de similaridad (en Rchic: callSimilarity Tree). De la técnica: reglas de asociación de minería y método: análisis estadístico implicativo. Nombre de la técnica: grafo implicativo (en Rchic: ImplicativeGraph)

De la técnica: Clustering y método: learning analytics. Nombres de las técnicas: en R: hclust.vector; dendro.variables y dendro.diana. De la técnica: reglas de asociación de minería y método: learning analytics. Nombres de las técnicas: en R: a priori; weclat y eclat

Para la presentación de las simulaciones, se emplea las recomendaciones de ((Kuyper, Hoog, & Jong, 2001); citado en (Carvajal, 2010)) quienes, para el diseño de una simulación para entornos educativos consideran tres pasos principales: “la especificación de los requisitos a partir de los objetivos del diseño; la construcción de bloques basados en la aplicación sobre la base de las necesidades y; la evaluación de la aplicación en función de los objetivos (...), llamado modelo evolutivo” (p. 12).

Simulaciones digitales en *Excel* - ASI

Por la situación del espacio de difusión de este informe, las tablas de simulaciones, se adjunta en la parte B Anexo 2, de la misma forma, se presenta en la página *Google Site* (<https://bit.ly/2PUWrPw>). Se invita al lector docente a visitar las mismas y robustecer así su experiencia con las técnicas del análisis estadístico implicativo y el learning analytics.

En cada una de las hojas de las simulaciones se consideran, la especificación de los requisitos: Es necesario que el docente demuestre: (1) Conocimiento del uso de las fórmulas de estadística básica en *Excel*, (2) tener un ordenador de acuerdo a las especificaciones (ver Tabla 3.) y (3) descargar el documento *Excel* de acuerdo a cada una de las técnicas como, se muestra en el Cuadro 4. siguiente:

Cuadro 4. Enlaces de las hojas de simulaciones

Técnica	Enlace de la documentación
Árbol de Similaridad	https://bit.ly/32GhyZ8
Grafo Implicativo	https://bit.ly/38luOR0
Cohesión	https://bit.ly/3pmgymR

Fuente: elaboración propia

Al referirse a: bloques basados en la aplicación de acuerdo a las necesidades docentes, el proceso de simulación comprende los pasos: Definiciones de las técnicas, actividades (para aproximar en la comprensión docente de las técnicas) y la evaluación (encaminados a la verificación en la aproximación a la comprensión de las técnicas) proceso propuesto por (Torres Barzabal, 2005). Esta estructura lo encuentra en cada tabla de las simulaciones se adjunta en la parte B del Anexo 2.

A manera de ejemplo, se presenta la siguiente Figura 12. que contiene la simulación de la técnica: índice de similaridad del árbol de similaridad.

Figura 12. Simulación en excel para el índice de similaridad

Simulación en Excel para el Índice de Similaridad

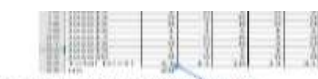
Definiciones: Según Ortiz *et al.* (2009), "la similaridad es una medida de correspondencia o semejanza entre los objetos que van a ser agrupados".
Menciona Ortiz *et al.* (2009), "sea un conjunto I formado por n individuos y un conjunto A formado por p características. $A = \{a_1, a_2, \dots, a_p\}$, donde $A_i = \{n \in I : a_i(n) = 1\}$, $\text{Card}(A_i) = n_i$ y $\text{Card}(A_i \cap A_j) = n_{ij}$ ".
Al número dado por:

$$s_{(A_i, A_j)} = \text{Prob} \left[\frac{\text{Card}(A_i \cap A_j) - \frac{n_i n_j}{n}}{\sqrt{\frac{n_i n_j (n - n_i - n_j + n_i n_j / n)}{n}}} \right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$$

Se le llama índice de similaridad de la pareja $\{A_i, A_j\}$ ".


Proceso de Aproximación a la definición: Ejemplo cotidiano: "A un grupo de n estudiantes se pueden verificar n características".
"La palabra 'Card' quiere decir 'Cantidad de'; en este caso, $\text{Card}(A_i \cap A_j)$ quiere decir cantidad de estudiantes que coinciden en características A_i y A_j la vez con N_i ".
"La expresión 'Prob', calcula la probabilidad de aquellas coincidencias llamadas 'coincidencias'".

Uso de la herramienta digital en rúbrica de aproximación a la comprensión de la técnica: Proceso del cálculo en Excel del elemento denominado "Índice de Similaridad".
Pasos:
1.- "Se inicia con una tabla que refleje los resultados de una encuesta" (quedan los resultados de una evaluación/observación, con datos de tipo binario, es decir datos de la forma: sí/no, verdadero/falso, admite/no admite, con propiedad/sin propiedad, aprueba/desaprueba, admitido/no admitido, promovido/no promovido, motivado/desmotivado), estos datos convertidos a 1 y 0 respectivamente, tal como se muestra en la figura (acompañar con el documento Excel en línea descargado).



2.- Calcular los totales de cada columna (los n_{ij}).

3.- Se forma una nueva tabla.



4.- Se determina las coincidencias (llamadas *coincidencias*) entre A_i y A_j , es decir el: $\text{Card}(A_i \cap A_j)$
Ejemplo: "hay 10 estudiantes que coinciden en sus opiniones entre $A_i = \text{HPD}$ y $A_j = \text{PUN}$, hay 7 estudiantes que coinciden en sus opiniones entre PUN y JAZ , entre otros".

5.- "Se obtiene el valor de $(n_i n_j) / n$, al multiplicar las filas y dividir el resultado entre el total".
Ejemplo: "(11x11)/20=6,05; (11x10)/20=5,5, entre otros".

6.- Se calculan las *coincidencias estandarizadas* K_i : **Ejemplo:** $(10-6,05) = 4,05=1,61$; $(8-7,15) = 0,85=0,32$, entre otros.

7.- "Para la columna $\{a_i, a_j\}$ se calcula el área bajo la curva normal estándar hasta K_i . **Ejemplo:** El área bajo la curva normal estándar hasta 1,61 es 0,9458524; El área bajo la curva normal estándar hasta 1,07 es 0,856789; entre otros".

8.- "Este número hallado, se llama el índice de similaridad".

“Recordando que, si este número se acerca a 1, hay más similitud entre este grupo de datos”. **Ejemplo:** “Hay más similitud entre HIP y PUN, pues $s(\text{HIP}, \text{PUN})=0,9458524$ ”.

Actividades propuestas: En el documento *Excel* descargado, se solicita al lector, revisar la conformación de las siguientes tablas:

3.- Matriz cruzada de todas las Copresencias (K)

Card(A, ∩ A)	HIP	JAZ	HEA	REG	PUN
HIP	11	0	7	0	0
JAZ	0	10	0	0	0
HEA	7	0	10	0	0
REG	0	0	0	10	0
PUN	0	0	0	0	11

4.- Determinar el valor real/*n* de todas las variables.

$D_{ij} = \frac{n_{ij}}{n_i}$	HIP	JAZ	HEA	REG	PUN
HIP	0,09	0,0	0,05	0,0	0,00
JAZ	0,0	0,14	0,0	0,0	0,0
HEA	0,05	0,0	0,14	0,0	0,0
REG	0,0	0,0	0,0	0,14	0,0
PUN	0,00	0,0	0,0	0,0	0,09

5.- Determinar el valor K_c

K_c	HIP	JAZ	HEA	REG	PUN
HIP	2,01	0,2	0,6	1,1	3,01
JAZ	0,21	2,0	1,0	1,0	0,91
HEA	0,61	1,0	2,0	1,1	2,01
REG	1,01	1,0	1,1	2,0	1,07
PUN	3,01	0,9	2,1	1,1	2,01

6.- Determinar la Matriz de Similitud.

S(i,j)	HIP	JAZ	HEA	REG	PUN
HIP	0,9779142	0,5861149	0,7387844	0,856789	0,456789
JAZ	0,5861149	0,9101438	0,9101438	0,7387844	0,7387844
HEA	0,7387844	0,9101438	1	0,9101438	0,856789
REG	0,856789	0,9101438	0,9101438	1	0,856789
PUN	0,456789	0,7387844	0,856789	0,856789	1

Se solicita al lector a completar los siguientes ejercicios de aplicación:

Pasos:

- 1.- “Registrar (cambiar) valores (o 1 o 0, a su gusto), en la tabla 1”.
- 2.- “Observar el cambio de los valores totales en cada columna de la tabla 1 y su registro en la tabla 2”.

3.- “Ahora con estos cambios, podemos responder: ¿A cuántos estudiantes les gusta tanto PUN y REG?, entre otras”.

4.- “Comprobar el cálculo del valor de $(n_{ij} * n_{ij}) / n$, para PUN y JAZ”, y otros.

5.- “Comprobar el cálculo del valor de K_c , para PUN y JAZ”, y otros.

6.- “Calcule el área bajo la curva normal para los nuevos valores de K_c ”.

7.- Se solicita “observar los nuevos valores *índices de similitud*, y como ejemplo, identificar el valor de $s(\text{PUN}, \text{HEA})$ ”.

Evaluación

“Para verificar la aproximación a la comprensión docente de la definición de ésta técnica, se solicita al usuario-docente, contestar una encuesta de los procesos explicados versus las actividades planteadas, la misma que se encuentra en el enlace: <https://bit.ly/2FRx2B0> o haciendo clic en el archivo *Excel* en el botón: ‘Encuesta 1’”.

RECOMENDACIÓN. - Para continuar con las simulaciones, se solicita al lector no guardar los cambios efectuados en el archivo *Excel*, es decir, trabajar con los datos originales, hasta la revisión de toda la documentación, después puede hacer uso de la documentación, como a bien tuviere.

Fuente: elaboración propia

Ver las demás hojas de simulaciones en la parte B del Anexo 2, y en la página de *Google Site*.

Análisis de los simuladores en excel

Con respecto a las simulaciones de las tres técnicas del análisis estadístico implicative presentadas en la parte B Anexo 2: Es necesario que los docentes demuestren competencias en el uso de las herramientas *Excel*, y la interpretación de las fórmulas insertadas. Se ha utilizado el procedimiento y ejemplos del uso de las Técnicas del ASI, sugeridas por Zamora *et al.* (2007). A través de uso de las simulaciones en *Excel*, fue factible un primer acercamiento al uso docente de las mencionadas técnicas. Es necesario estimular a la comunidad docente, en el uso del *Software* como el *R*, que facilita los procesos del cálculo estadístico, en particular, el análisis de datos educativos a través de estas técnicas.

Simulaciones en R para el ASI y el LA

En esta sección, se presentan las simulaciones digitales para facilitar la comprensión docente de las técnicas comunes del ASI y el LA que son: del *Clustering* (del ASI(*callSimilarityTree* y *callHierarchyTree*) y del LA(*hclust.vector*, *dendro.variables* y *dendro.diana*)); de las reglas de asociación de minería (del ASI(*ImplicativeGraph*) y del LA (*A priori*, *weclat* y *eclat*)), de acuerdo a Zamora et al. (2007), R. A. Pazmiño-Maji et al., (2016) y Naranjo et al., (2018).

De la forma similar que en el caso de las simulaciones en *excel*, éstas simulaciones digitales en R siguen los tres procesos sugeridos por (Kuyper et al., 2001), citado en (Carvajal, 2010) que comprenden: “La especificación de los requisitos a partir de los objetivos del diseño; la construcción de bloques basados en la aplicación sobre la base de las necesidades y; la evaluación de la aplicación en función de los objetivos”.

En el proceso de la especificación de los requisitos: es necesario que el docente demuestre: (1) Conocimiento básico del uso del *Software R* y *RStudio*, (2) tener un ordenador de acuerdo a las especificaciones (ver Tabla 3.) y (3) instalar los *Software: R, RStudio* o *RStudio Cloud* y *Rchic* (Ver Anexo 1), junto con los paquetes descritos en la parte C del Anexo 2.

En la fase de: bloques de simulaciones en R, de acuerdo a las necesidades docentes, cada uno de los procesos de simulación describen: Definiciones de las técnicas; actividades (para fortalecer la comprensión docente de las técnicas), este proceso es complementario a las simulaciones en *excel* por lo que, se espera que el docente apropie del uso de las técnicas en su quehacer educacional.

De la misma forma que en el caso anterior, por la situación del espacio de difusión en este informe, las tablas de simulaciones, se adjunta en la parte D del Anexo 2, se adjunta, también, en la página de *Google Site* (<https://bit.ly/2PUWrPw>).

En la fase de la evaluación de la aplicación: en cada uno de los procesos de las simulaciones, se ha desarrollado una encuesta con la finalidad de observar el avance en la comprensión docente de los conceptos citados, lo cual, hasta la presentación de este trabajo, muestra: Herramienta accesible a la mayoría de los docentes. De acuerdo a la capacidad docente, fácil de seguir y comprender los conceptos. Los simuladores están en continua actualización, de acuerdo a las necesidades docentes. Tener en cuenta que *R* está en constante desarrollo, por lo que es necesaria la permanente actualización de los paquetes y los algoritmos propuestos.

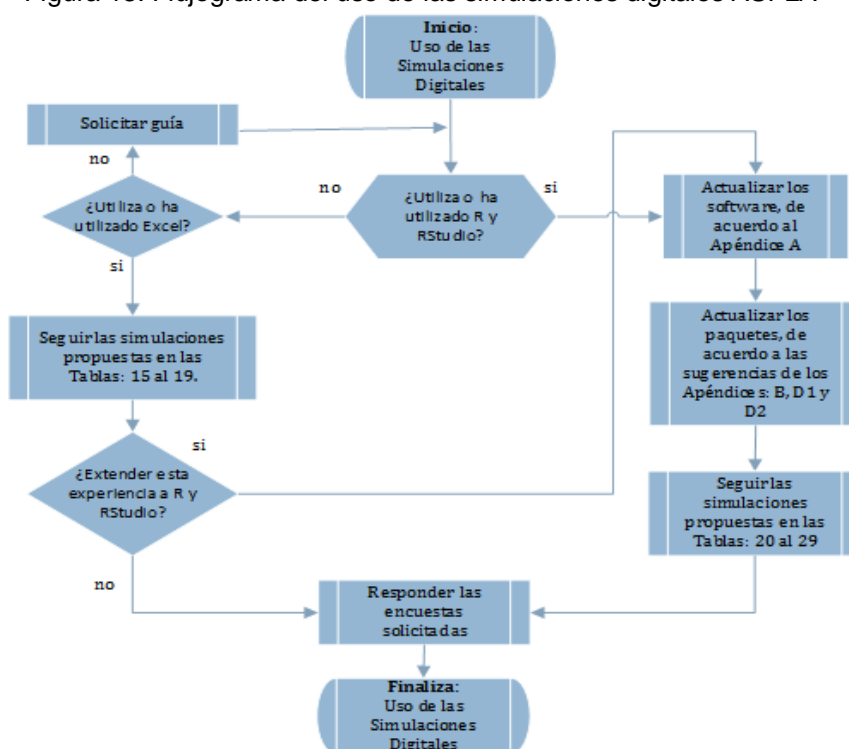
Análisis de los simuladores en *R*

Con respecto a las simulaciones en *R*, de las técnicas comunes del análisis estadístico implicativo y el learning analytics presentadas en esta sección: Se ha utilizado datos del quehacer educativo relacionado a las técnicas. Es necesario que los docentes demuestren competencias en el uso del *Software R*, *RStudio* o *RStudio Cloud* y el paquete *Rchic*. Se ha utilizado el grupo de técnicas comunes del ASI y el LA, sugeridas por Pazmiño R y Naranjo M y sus colaboradores. A través del uso de las simulaciones en *R*, es factible un segundo acercamiento docente en el uso de las mencionadas técnicas.

Diagrama para las simulaciones digitales

Para el uso de las simulaciones digitales de la tendencia análisis estadístico implicativo y el learning analytics, se sugiere el uso del proceso detallado en la Figura 13. que muestra dos rutas: de acuerdo a la capacidad docente en el uso de las herramientas tecnológicas de *office*, en una primera parte acceden a las simulaciones en *excel*, y en una segunda parte, de acuerdo a sus capacidades en el uso del software *R*, acceden a las simulaciones en estas unidades.

Figura 13. Flujograma del uso de las simulaciones digitales ASI-LA



Fuente: elaboración propia

Limitaciones

Las simulaciones presentadas, son de uso general, a más de la predisposición, el docente muestra su capacidad en uso de la herramienta *Excel*, después, familiariza con el *Software R*, los diferentes paquetes como *Rchic*, y experimenta el fabuloso mundo del análisis de datos masivos generados en entornos educativos.

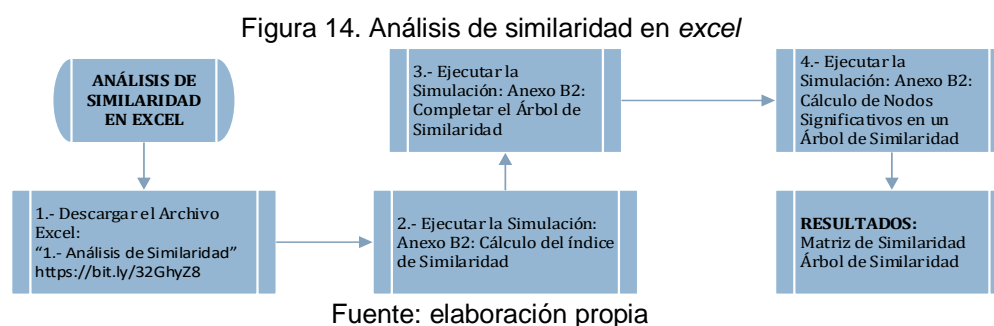
Tabla 5. Tipos de datos presentes en las técnicas

Nombres de las técnicas	Tipos de datos de entrada
Árbol de Similitud, Árbol Jerárquico Cohesivo del ASI (clustering)	Binarios {1, 0}
hclust.vector del LA (clustering)	Vectoriales de datos binarios
dendro.variables del LA (clustering)	Numéricos
dendro.diana del LA (clustering) y Grafo Implicativo del ASI (Reglas de Asociación de Minería)	Binarios {1, 0}
a priori, weclat y eclat del LA (Minería de Asociación)	Caracteres, agrupadas en forma de: transacciones y actividades.

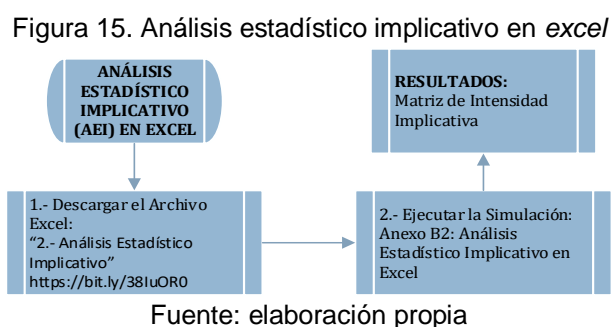
Nota. - Presente en los Anexos B2 y B4

Fuente: elaboración propia

Para el uso de las simulaciones digitales en *excel*, se sugiere a continuar el siguiente proceso:

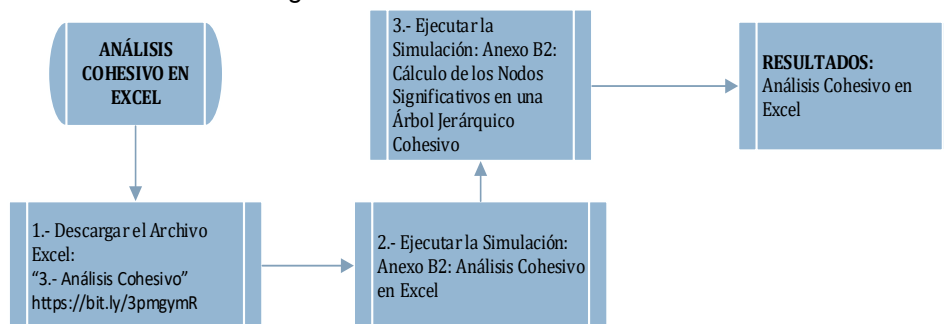


La Figura 14. muestra un proceso que, para aproximar a la comprensión de la técnica: análisis de similaridad en *excel*, después de descargar el archivo digital propuesto, acompaña con las simulaciones y/o actividades propuestas (ver la parte B Anexo 2).



La Figura 15. muestra un proceso que, para aproximar a la comprensión de la técnica: análisis estadístico implicativo en *excel*, después de descargar el archivo digital propuesto, acompaña con las simulaciones y/o actividades propuestas (ver la parte B del Anexo 2).

Figura 16. Análisis cohesivo en excel

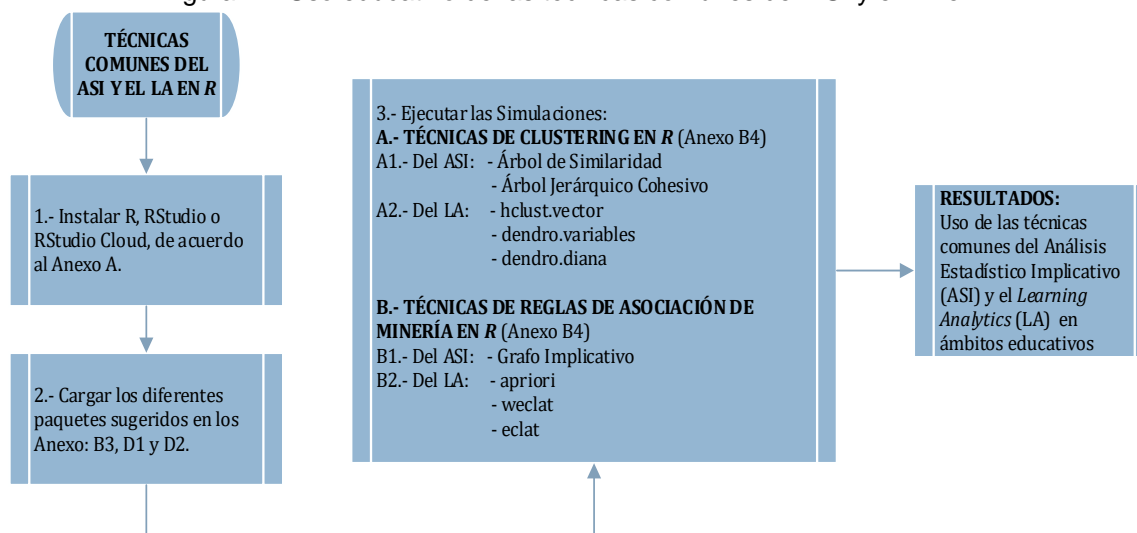


Fuente: elaboración propia

La Figura 16. muestra un proceso que, para aproximar a la comprensión de la técnica: análisis cohesivo en excel, después de descargar el archivo digital propuesto, acompaña con las simulaciones y/o actividades propuestas (ver la parte B del Anexo 2).

Para el uso de las simulaciones digitales en R, se propone el siguiente proceso:

Figura 17. Uso educativo de las técnicas comunes del ASI y el LA en R



Fuente: elaboración propia

La Figura 17. muestra un proceso, para complementar el uso de las técnicas comunes del ASI y el LA en contextos educativos, en la que el docente después de instalar los programas propuestos, junto con los paquetes necesarios,

acompaña con las simulaciones y/o actividades propuestas (ver la parte D del Anexo 2).

Resultados del avance de los simuladores

En el objetivo de crear simulaciones digitales que faciliten la comprensión docente de las técnicas comunes del Análisis Estadístico Implicativo y el *Learning Analytics* (LA), la parte receptora de este trabajo es la población docente de la UEMIB Chibuleo, por la capacidad de los docentes en el uso de las herramientas *Office*, esta experiencia, se ejecutó en dos procesos: (1) las simulaciones en *Excel* (las rutinas sugeridas del uso de las técnicas del ASI de Zamora *et al.* (2007) y (2), simulaciones en *R* (de las técnicas comunes del ASI y el LA sugeridas por Pazmiño R y Naranjo M y sus colaboradores.

Se ha procedido con los siete pasos de un modelo de simulación propuesta por: (Law & McComas, 2001), es decir, después de la selección, análisis y la descripción en un lenguaje más comprensible (para el ámbito educativo) de las técnicas comunes del Análisis Estadístico Implicativo (ASI) y el *Learning Analytics* (LA), se procedió a:

(1) Delimitar el problema; es decir, al observar que: en la UEMIB Chibuleo, por el uso docente de Recursos Educativos Abiertos (REA) en los procesos de enseñanza-aprendizaje, se generan datos como: Número de veces que utiliza el docente cierto REA y su efectividad en su proceso de enseñanza-aprendizaje, entre otros; también, por el desconocimiento del profesorado en el uso de las técnicas comunes del ASI y el LA para el análisis de este tipo de datos, se propone a través de los simuladores digitales llegar con las técnicas en mención, hacia la comunidad de docentes de la UEMIB Chibuleo.

(2) Recolectar la información y construir el modelo conceptual de la simulación; es decir, se abstrae toda la información referente a las técnicas comunes del ASI y el LA, seguidamente, se detectan las competencias docentes en el uso de los

medios tecnológicos y su aspiración en conocer el lenguaje *R*, además, se profundiza en el uso de la metodología de la simulación, para luego plasmar el modelo conceptual de simulación, que de acuerdo a la capacidad docente, constituyó de dos partes (la del *Office Excel* y en *R*).

(3) Para la comprobación del modelo conceptual; la idea básica de los simuladores se compartió con los expertos, se observa las sugerencias y con la premisa de la aprobación, se prosigue con las demás etapas. Al tratar la parte (4), programación del modelo de simulación en *Excel* y en *R*; se basa en las competencias docentes en el dominio de la herramienta *Office Excel*, se recoge ejemplos prácticos para esta herramienta y se ejecuta una primera aproximación, se extiende esta experiencia al lenguaje *R*.

En la fase (5), de la comprobación de la validez del modelo programado; el uso de las fórmulas *Excel* y los algoritmos *R* en los simuladores, están de acuerdo con las definiciones de las técnicas con lo cual, se evidencia la validez de los procesos.

En la fase (6), del diseño, realización y análisis de los experimentos simulados; se procedió a la elaboración de las simulaciones, bajo el proceso sugerido por (Kuyper et al., 2001) citado en (Carvajal, 2010), que contempló: “la especificación de los requisitos a partir de los objetivos del diseño; la construcción de bloques basados en la aplicación sobre la base de las necesidades y; la evaluación de la aplicación en función de los objetivos (...) llamado modelo evolutivo”.

Mientras que, en el proceso didáctico de los bloques de simuladores, se presentaron: las definiciones de las técnicas; actividades (para aproximar a la comprensión docente de las técnicas) y la evaluación (encaminados a la verificación de la aproximación docente en la comprensión de las técnicas), proceso propuesto por (Torres Barzabal, 2005). Estos trabajos, se presentan en la parte B y D del Anexo 2.

En la última fase (7), se documentan y presentan los resultados de las simulaciones, en una página de *Google Site* (ver <https://bit.ly/2PUWrPw>).

Presentación de las simulaciones digitales ASI y LA

El trabajo de las simulaciones digitales, se plasman en una página de *Google Site* (ver <https://bit.ly/2PUWrPw>) que contiene: (1) Información básica referente a las técnicas comunes entre el ASI y el LA como son: *similaridad, implicación, cohesión, hclust.vector, dendro.variables, dendro.diana, apriori, weclat y eclat*, se extiende con la parte explicativa de la instalación y uso de la herramienta *Rchic* de *R*, *RStudio* y *RStudio Cloud* y (2), la sección de simulaciones en *Excel* y en *R*.

En la primera sección de la página, se encuentra publicada contenidos, referentes a las a las siguientes interrogantes: ¿Qué es el Análisis Estadístico Implicativo ASI? ¿Qué es el *Learning Analytics* LA? ¿Cuáles son las técnicas comunes entre el Análisis Estadístico implicativo (ASI) y el *Learning Analytics* (LA)? ¿Cuáles son las definiciones de las técnicas comunes entre el Análisis Estadístico (ASI) y el *Learning Analytics* (LA)?

Mientras que en la segunda sección, se presentan las simulaciones en la herramienta *Excel*, con el objetivo de realizar una primera aproximación docente en la comprensión de las técnicas del ASI, esta sección contiene: Simulaciones en *Excel* para los procesos: Análisis de Similaridad, Análisis Estadístico Implicativo y el Análisis Cohesivo, sugeridos por (Zamora *et al.*, 2007).

Por otra parte en la tercera sección, se presentan las simulaciones en el *Software R*, con el objetivo de realizar una segunda aproximación docente hacia las técnicas del ASI y el LA, esta sección contiene: Información básica y el proceso de instalación de los *Software: R, R Studio, y Rchic*, a más del proceso de carga de los componentes necesarios, de la misma forma, se presentan las Simulaciones en *R* de las técnicas ASI: Árbol de Similaridad del, Árbol Jerárquico Cohesivo, similares al: *hclust.vector, dendro.variables y dendro.diana* del LA,

también, simulaciones para la técnica: Grafo Implicativo del ASI, similar a las técnicas del LA: *a priori*, *weclat* y *eclat*. Termina, la cuarta sección, que facilita el contacto y la recepción de sugerencias de usuarios.

Por lo presentado, es necesario continuar en la difusión de esta herramienta, para el uso docente de las técnicas comunes del ASI y el LA, quien analiza los datos generados en sus procesos de enseñanza-aprendizaje desarrollados a través del uso de medios digitales, con la finalidad de la mejora continua de éstos procesos.

CAPÍTULO III. ANÁLISIS DE LOS RESULTADOS DE LA INVESTIGACIÓN

3.1. Procesamiento y análisis de la información

En la utilización de una página de *Google Site*, un mismo grupo pre-experimental, conformado por los docentes de bachillerato de la UEMIB Chibuleo, son expuestos después de un diagnóstico, a una intervención educativa mediante las *Simulaciones Digitales*, para obtener resultados y verificar si hay una aproximación positiva hacia las técnicas ASI y LA, confirmado a través de pre test y post test.

Se aprecia la participación de 7 (2 mujeres y 5 hombres) docentes del área de matemáticas, emprendimiento, y tecnologías, áreas afines a la propuesta. Se procedió a la realización de la prueba al inicio de la intervención (parte C del Anexo 3). Se comparte la página de *Google Site* con la información de las *Simulaciones Digitales*, se dispuso un tiempo de acompañamiento de 30 minutos, de acuerdo a la disponibilidad de los docentes.

Luego de ello los docentes, se apropian de las técnicas relacionadas al ASI y LA, trabajan con datos educativos tanto en la herramienta *Excel* y R. Al final del uso de éstas herramientas, se realiza la prueba post, la misma prueba que, se emplea al inicio de la intervención, con la finalidad de determinar el impacto del uso de los Simuladores Digitales en la aproximación docente hacia las técnicas ASI y LA, cuyo análisis estadístico, se presenta en la parte final.

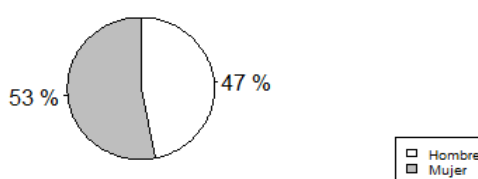
A continuación, se presentan los análisis y las interpretaciones de la encuesta inicial aplicada al grupo de docentes del bachillerato de la Unidad Educativa.

Resultados obtenidos en la encuesta aplicada a los docentes

En el diagnóstico del contexto educativo–docente a la que, se exhibe el presente trabajo investigativo, de acuerdo a la encuesta aplicada a los 18 docentes de la UEMIB Chibuleo (ver Anexo 3), presenta:

1. Con respecto al sexo del personal docente:

Gráfico 1. El sexo del encuestado

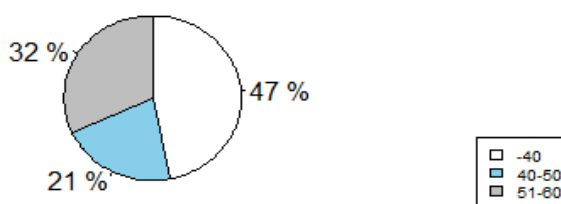


Fuente: elaboración propia

De acuerdo al Gráfico 1. 10 docentes son mujeres (representan el 53%), mientras que la diferencia (9 docentes) son hombres, equivalente al 47%.

2. Con respecto al rango de edad del encuestado.

Gráfico 2. El rango de edad del encuestado

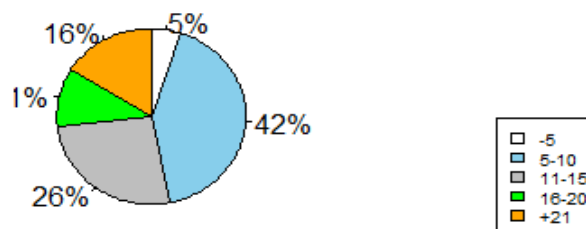


Fuente: elaboración propia

Se observa en el Gráfico 2. que el 47% de los docentes de esta institución, tiene menos de 40 años de edad, el 21% de los profesores tiene entre 40 y 50 años, de la misma forma, un 32% de los docentes están entre los 51 y 60 años. Por otra parte, se observa que: el 68% de los docentes tienen menos de 50 años, que es un indicador favorable para la aplicación de la estrategia.

3. Con respecto a los años de servicio del docente encuestado.

Gráfico 3. Años de servicio del docente encuestado

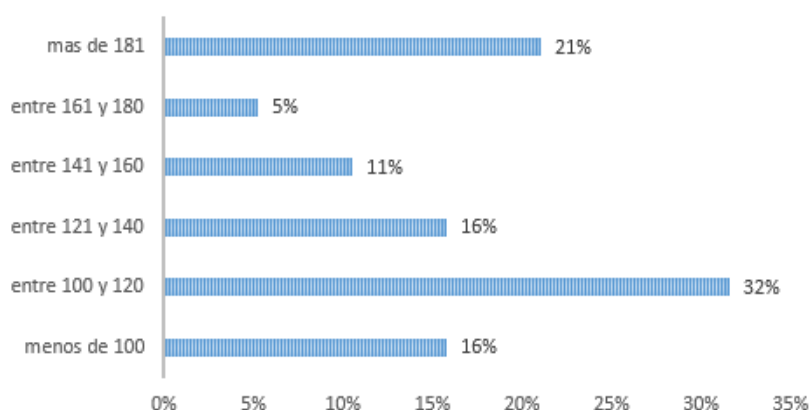


Fuente: elaboración propia

Se observa en el Gráfico 3. que el 5% de los docentes tiene un periodo de servicio de menos de 5 años, el 42% de los profesores están en servicio educativo entre los 5 y 10 años, el 26% de los encuestados están en servicio educativo entre los 11 y 15 años, la diferencia, el 17% están en servicio educativo ya más de los 16 años, esta información, nos hace entender que, se trabaja con personal de experiencia en el ámbito educativo.

4. Con respecto número de estudiantes a quienes imparten asignaturas.

Gráfico 4. Número de estudiantes a quienes imparten asignaturas



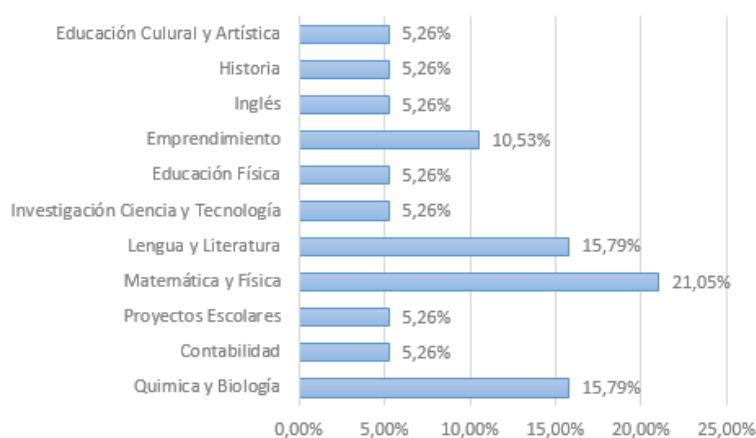
Fuente: elaboración propia

Se observa según el Gráfico 4. anterior, en forma resumida que: el 16% de los docentes encuestados laboran con menos de 100 estudiantes, otro grupo del mismo porcentaje laboran con entre 121 y 140 estudiantes, el mismo porcentaje

de docentes mencionan que trabajan con de entre 141 a 180 estudiantes, el 32% de los docentes laboran con entre 100 y 120 estudiantes, el restante 21% de los profesores de la encuesta laboran con más de 181 estudiantes. Esta información, hace entender que la mayoría de los docentes trabajan con una considerable cantidad de estudiantes.

5. Con respecto a la asignatura con mayor carga horaria del docente.

Gráfico 5. Asignaturas de mayor carga horaria con las que trabajan los docentes

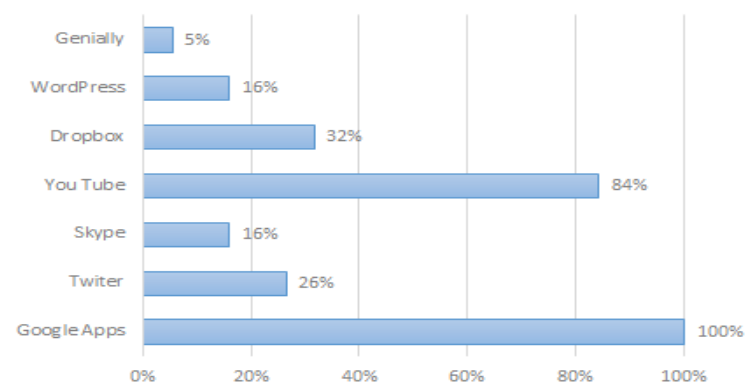


Fuente: elaboración propia

Se observa en el Gráfico 5. que las asignaturas de matemática y física imparten el 21% de los docentes, en el mismo porcentaje el 16% de los docentes están a cargo de las asignaturas de química y biología o lengua y literatura castellana, el 11% de los docentes trabajan en la asignatura de emprendimiento, repartidas en el mismo porcentaje del 5%, están los docentes que imparten asignaturas de educación artística, historia, inglés, educación física, investigación, proyectos escolares y contabilidad. Los resultados apuntan a que un gran porcentaje de docentes trabajan en áreas de ciencias y tecnologías, un indicador más para la ejecución de la propuesta.

6. A la pregunta: ¿Con cuáles herramientas TIC, se encuentra familiarizado? las respuestas de los docentes fueron:

Gráfico 6. Las herramientas TIC con las que está familiarizado el docente

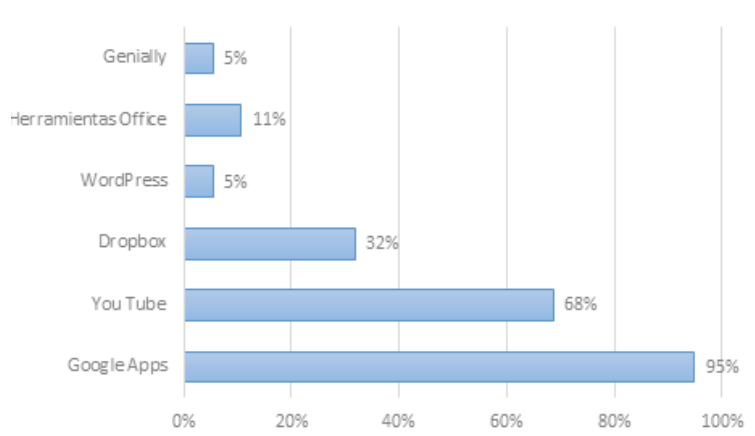


Fuente: elaboración propia

De acuerdo al Gráfico 6. el 100% de los docentes, se encuentran familiarizados con las herramientas de *Google Apps* para sus procesos de enseñanza-aprendizaje, de la misma forma, el 84%, están adaptados a la herramienta *You Tube*, el 32% y el 26% respectiva, los docentes, se encuentran habituados con las herramientas *Dropbox* y *Twiter*. La diferencia de docentes, están familiarizados entre las herramientas *WordPress*, *Skype* y *Genialy*, desconocen herramientas como: *Evernote*, *Edmodo*, *ClassDojo*, *Socrative*. Lo cual indica que los docentes en un gran porcentaje están familiarizados con las herramientas TIC.

7. A la pregunta: ¿Qué herramientas TIC utiliza para el dictado de clase? las respuestas de los docentes fueron:

Gráfico 7. Las herramientas TIC que utiliza el docente para la enseñanza-aprendizaje

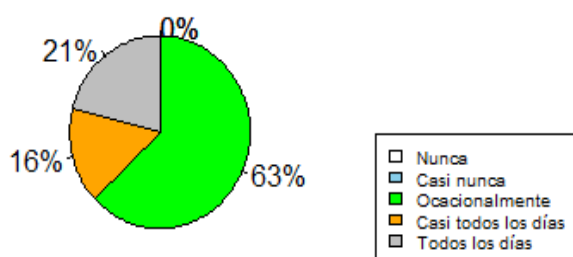


Fuente: elaboración propia

En el Gráfico 7. el 95% de los docentes mencionan que utilizan las herramientas de Google Apps para sus procesos de enseñanza-aprendizaje, de la misma forma, el 68%, emplean la herramienta YouTube, el 32% de los docentes emplean la herramienta *Dropbox*, la diferencia de docentes utilizan, también, las herramientas *WordPress* , *Office* y *Genialy*, no usan las herramientas como: *Edmodo*, *Twiter*, *Skype*, *Evernote*, *CassDojo* y *Socrative*. Lo cual indica que los docentes en un gran porcentaje emplean las herramientas TIC.

8. A la pregunta: ¿Utiliza la herramienta *Excel* para procesar datos educativos? las respuestas de los docentes fueron:

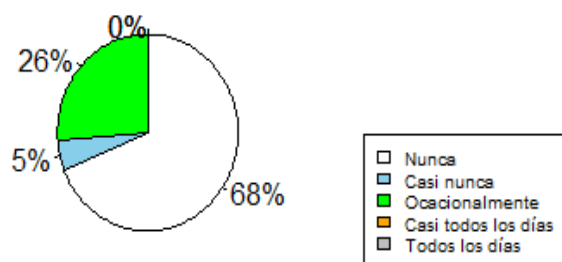
Gráfico 8. Uso docente de la herramienta *excel*



Fuente: elaboración propia

Muestra el Gráfico 8. que en las actividades diarias los docentes utilizan la herramienta *Excel*, el 63% de los docentes, ocasionalmente, y solo el 21% no las utilizan todos los días, mientras que el 16% de los docentes utilizan casi todos los días. Ésta es otra de las razones, para iniciar las simulaciones en *Excel*, siempre con la intención de aproximar a los docentes la plataforma *R*.

9. A la pregunta: ¿Ha utilizado el *software* estadístico *R*? las respuestas de los docentes fueron:

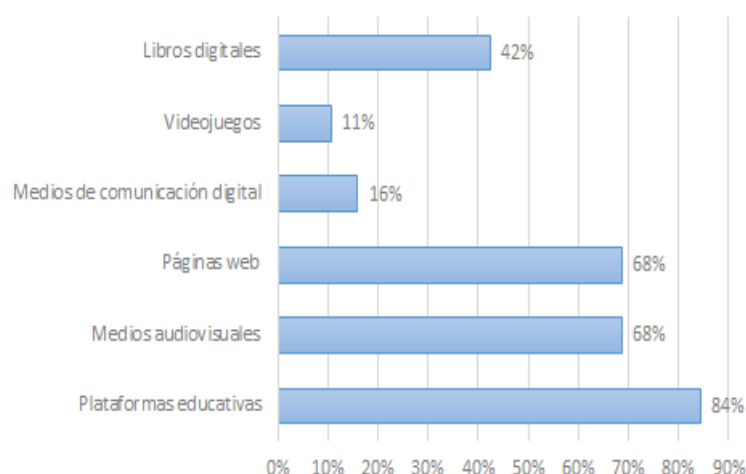
Gráfico 9. Uso docente del *software* estadístico *R*

Fuente: elaboración propia

En congruencia con las respuestas anteriores, de acuerdo al Gráfico 9. la mayoría de los docentes de la UEMIB Chibuleo (el 68%), nunca han utilizado el *Software Estadístico R*, solo el 26% de los profesores han utilizado de alguna forma, los demás (el 5%) mencionan que casi nunca han utilizado este *Software* Estadístico. Este resultado, indica que es una oportunidad de llegar a los docentes, con estrategias amigables, fomentar el uso de las técnicas del ASI y LA en el quehacer educacional.

10. Con respecto a los medios digitales que utiliza el docente en sus clases, las respuestas fueron:

Gráfico 10. Uso docente de medios digitales en clases

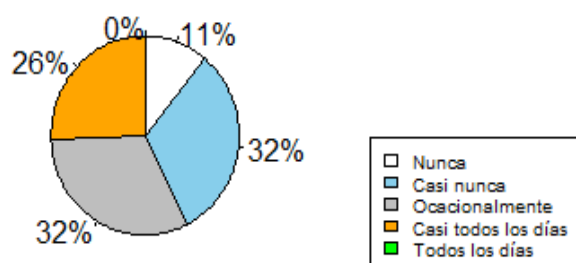


Fuente: elaboración propia

Está claro de acuerdo al Gráfico 10. que el 84% de los docentes utilizan plataformas educativas en las actividades educacionales, en el mismo porcentaje del 68% respectivamente los docentes mencionan que: utilizan medios audiovisuales o páginas web, un porcentaje bajo de docentes el 42% mencionan que: utilizan libros digitales, mientras que la diferencia de docentes el 11%, usan videojuegos y el 16% utilizan medios de comunicación digital. Al igual que en el caso anterior, este resultado, nos indica que es una oportunidad de llegar a los docentes, con las estrategias para el uso de las técnicas del ASI y LA en datos educativos generados, en el uso de medios digitales.

11. A la pregunta: ¿Usted realiza análisis estadístico de los avances de aprendizaje que tienen los diferentes grupos de estudiantes a los que dicta clases? las respuestas de los docentes fueron:

Gráfico 11. Análisis estadístico de aprendizajes

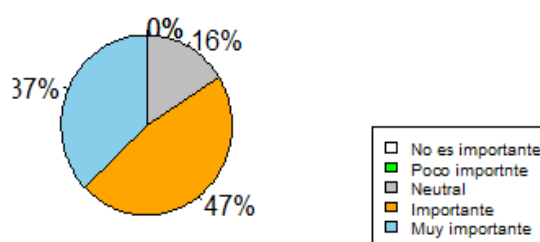


Fuente: elaboración propia

En la parte conceptual del ASI, de acuerdo al Gráfico 11. el 11% de los docentes nunca realizan el análisis estadístico de los avances de los aprendizajes de estudiantes, en el porcentaje del 32%, los docentes mencionan que casi nunca y ocasionalmente utilizan las técnicas estadísticas, mientras que la diferencia, el 26% de los docentes emplean casi todos los días, el proceso del análisis estadístico de los avances de aprendizajes de los estudiantes. Esto informa que, más de la mitad de los docentes realizan el análisis estadístico de los avances de los aprendizajes de estudiantes.

12. A la pregunta: ¿Considera importante manejar datos estadísticos de los avances de sus estudiantes en su asignatura? las respuestas de los docentes fueron:

Gráfico 12. Datos estadísticos en promociones estudiantiles

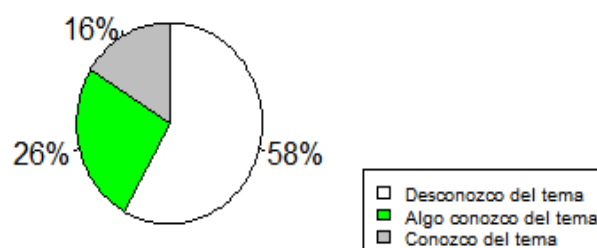


Fuente: elaboración propia

De acuerdo con el Gráfico 12. el 47% de los docentes del nivel de bachillerato, consideran que es importante utilizar datos estadísticos de los avances de sus estudiantes, mientras que el 37% de los encuestados, mencionan que es muy importante es uso de datos educativos, por otro lado, el 16 de los docentes presentan una neutralidad en esta pregunta. Estos resultados abren camino, los docentes están conscientes del uso de datos educativos para el seguimiento del avance de sus estudiantes.

13. A la pregunta: ¿Conoce para qué sirve el análisis estadístico implicativo? las respuestas de los docentes fueron:

Gráfico 13. Conocimiento de las técnicas ASI

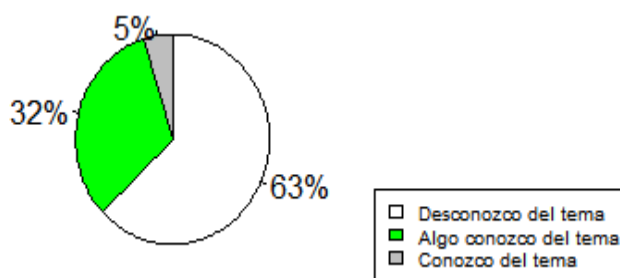


Fuente: elaboración propia

De acuerdo con el Gráfico 13. el 58% de los docentes del nivel de bachillerato, desconocen de uso de las técnicas del ASI, el 26% conocen algo del tema y la diferencia, el 16% de los docentes conocen del tema. Una oportunidad más para abrir el camino al uso de las técnicas ASI y LA en la comunidad de los docentes, en vías de la mejora de los procesos de enseñanza y aprendizajes.

14. A la pregunta: ¿Conoce las técnicas de analíticas de aprendizajes, también, conocidas por su sigla en inglés *Learning Analytics*? las respuestas de los docentes fueron:

Gráfico 14. Conocimiento de las técnicas LA

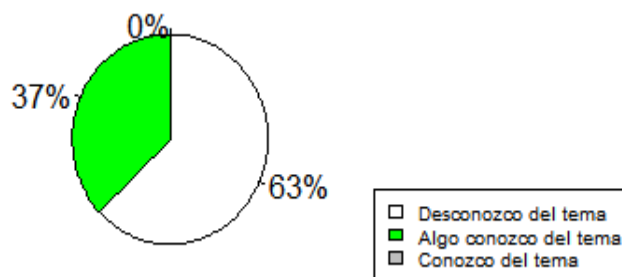


Fuente: elaboración propia

De acuerdo con el Gráfico 14. el 63% de los docentes del nivel de bachillerato, desconocen de uso de las técnicas del LA, el 32% conocen algo del tema y la diferencia, el 5% de los docentes conocen del tema. Esta información muestra la oportunidad al uso de las técnicas LA en la comunidad de los docentes, en vías de la mejora de sus procesos de enseñanza y aprendizajes.

15. A la pregunta: ¿Conoce sobre los datos educativos en el uso de las técnicas de estadística implicativa, combinadas con las técnicas de analíticas de aprendizajes? las respuestas de los docentes fueron:

Gráfico 15. Conocimiento de las técnicas del ASI y LA en datos educativos



Fuente: elaboración propia

De acuerdo con el Gráfico 15. el 63% de los docentes del nivel de bachillerato, desconocen de la combinación de las técnicas del ASI y LA sobre los datos educativos, en vías de la mejora de los procesos de enseñanza-aprendizaje, mientras que la diferencia el 37% conocen algo del tema. Esta información muestra la oportunidad para combinar las técnicas del ASI y LA sobre los datos educativos, para la mejorar de los procesos de enseñanza y aprendizajes, a través de reglas de comportamiento educativo.

3.2. Evaluación de expertos de la propuesta

La evaluación de expertos de la propuesta, se lo realiza por medio de una ficha de validación, este instrumento permite contrastar los criterios de: calidad, objetividad, actualidad, organización, suficiencia, intencionalidad, consistencia, coherencia, metodología y pertinencia de las simulaciones ASI y LA, hacia la comunidad de docentes de bachillerato.

Se cuenta con la colaboración de cuatro docentes (tres docentes universitarios, y una docente de nivel medio) (parte E del Anexo 3). El cálculo del Coeficiente V de Aiken arroja el valor de $0.95 > 0.7$, lo cual indica que es buena. Los resultados que arrojaron el proceso de revisión de expertos indican que promoverían la utilización de los simuladores con los estudiantes en las aulas, motivarían al uso de las TIC.

Figura 18. El cálculo del coeficiente V de Aiken de la propuesta

Validez de la Propuesta										
Evaluadores	INDICADORES									
	Claridad	Objetividad	Actualidad	Organización	Suficiencia	Intencionalidad	Consistencia	Coherencia	Metodología	Aplicabilidad
Evaluador 1	85	100	100	80	95	95	100	100	95	100
Evaluador 2	85	90	95	90	90	90	90	90	90	95
Evaluador 3	95	95	95	95	95	95	95	95	95	95
Evaluador 4	100	100	95	100	95	100	100	100	100	100
V de Aiken	0,91	0,96	0,96	0,91	0,94	0,95	0,96	0,96	0,95	0,98
V de Aiken Global	0,95									

Fuente: elaboración propia

3.3. Resultados del análisis cuasi experimental

En la Tabla 6. al comparar los diferentes estadísticos de la prueba pre test y post test: el promedio varía desde, 4,28 hasta 7,14 puntos, los docentes muestran una mejoría en la comprensión de los conceptos ASI y LA. De la misma forma la nota más baja, varía desde 3 puntos hasta 6 puntos, la nota más alta, también, varía desde 6 puntos hasta los 8 puntos, de acuerdo a los resultados de las pruebas, la comprensión del proceso del análisis cohesivo, aun les cuesta aproximar a los

docentes (preguntas 6, 7 y 9). El valor de la mediana, también, varía desde 4 puntos hasta los 7 puntos.

Este análisis aporta en que las simulaciones digitales causaron efecto positivo en el proceso de aproximación docente, en los referentes a los conceptos y aplicaciones del ASI y LA.

Tabla 6. Estadísticos descriptivos pruebas pre test y post test

Estadísticos	Valor Pre Test	Valor Post Test
Media	4,2857	7,1429
Mediana	4,0000	7,0000
Varianza	1,6327	0,4082
Desv. Estándar	1,3801	0,6901
Mínimo	3,0000	6,0000
Máximo	6,0000	8,0000

Fuente: elaboración propia

En el proceso de comprobación de la hipótesis de investigación, para probar la normalidad del proceso pre-test y post-test, al analizar la normalidad de los datos y continuar con pruebas paramétricas o no paramétricas, se tiene:

Hipótesis nula (H_0): Los datos que, se adquieren en pre-test y post-test, siguen una distribución normal.

Hipótesis alterna (H_1): Los datos que, se adquieren en pre-test y post-test, no siguen una distribución normal.

La regla de decisión estadística: donde p-valor es el valor de la probabilidad y α el nivel de significancia. Si $p\text{-valor} \leq 0,05$, se rechaza la hipótesis nula. En la Tabla 7. muestra este análisis.

Tabla 7. Prueba de normalidad pre-test y post-test

Shapiro – Wilk			
	Estadístico	Grados de libertad	p-valor
Notas Pre-Test	0,820	7	0,064
Notas Post-Test	0,804	7	0,045

Fuente: elaboración propia

Por lo reducido de la muestra, se aplica la prueba de Shapiro – Wilk, para la prueba pre test el estadístico es de 0,820 y el p-valor de $0,064 \geq 0,05$, no se rechaza la H_0 , es decir, que los datos siguen una distribución normal. De la misma forma en la prueba post test, el estadístico equivale a 0,804 con un p-valor de $0,045 \leq 0,05$, se rechaza la H_0 , es decir, que los datos no siguen una distribución normal. Seguidamente, se procede a aplicar la prueba estadística no paramétrica de *Wilcoxon* para muestras pareadas, bajo las hipótesis:

Hipótesis nula (H_0): La metodología de las simulaciones digitales como estrategia de aprendizaje, no facilitan la comprensión docente de las técnicas del Análisis Estadístico Ampliativo (ASI) y el *Learning Analytics* (LA).

Hipótesis alterna (H_1): La metodología de las simulaciones digitales como estrategia de aprendizaje, facilitan la comprensión docente de las técnicas del Análisis Estadístico Ampliativo (ASI) y el *Learning Analytics* (LA).

En la Tabla 8. se observa el contraste de *Wilcoxon*.

Tabla 8. Estadísticos de contraste de *Wilcoxon*

Nota Post Test – Nota Pre Test ^a	
Z	-2,379 ^b
Sig. Asintótica (bilateral)	0,017
a. Prueba de los rangos con signo de Wilcoxon b. Basado en los rangos negativos	

Fuente: elaboración propia

Al comparar, el valor de Sig. Asintótica (bilateral) 0,017 que es menor o igual que el valor de 0,05 ($p\text{-valor} \leq 0,05$). Se concluye que, se rechaza la hipótesis nula H_0 , y se acepta la hipótesis alterna H_1 , es decir que: La metodología de las simulaciones digitales como estrategia de aprendizaje, facilitan la comprensión docente de las técnicas del Análisis Estadístico Ampliativo (ASI) y el *Learning Analytics (LA)*.

CONCLUSIONES

- La fundamentación de los aspectos teóricos asociados al desarrollo de las simulaciones digitales y las técnicas del análisis estadístico implicativo y el *learning analytics*, de acuerdo al proceso investigativo, existen aportes suficientes para el empleo en el ámbito de educación de bachillerato. Es decir, existe la teoría matemática necesaria, algoritmos eficientes, tecnología desarrollada aplicable al medio, aportes de la comunidad científica en constante actualización, aspectos que fueron relevantes para la construcción del estado de arte del presente trabajo investigativo.
- El diagnóstico del nivel de conocimiento docente de las bases teóricas de las técnicas del análisis estadístico implicativo y el *learning analytics*, mediante el cuestionario desarrollado y aplicado a los profesores de bachillerato, muestra que están con la predisposición de conocer y aplicar estas técnicas estadísticas y algoritmos, para así innovar su quehacer educativo, al analizar los datos generados en procesos de enseñanza y aprendizaje que utilizan medios tecnológicos.
- La elaboración del sistema de simulaciones digitales como estrategia de aprendizaje de las técnicas del análisis estadístico implicativo y el *learning analytics* en la comunidad de docentes de bachillerato de la UEMIB Chibuleo, permite el aporte de una herramienta virtual con la información resumida de las técnicas, procesos de aprendizaje acorde a la comunidad de docentes, encaminados al análisis de datos educativos generados en el empleo de las tecnologías como medios de aprendizaje.
- La utilización de los simuladores para el proceso de aproximación docente hacia las técnicas del análisis estadístico implicativo y el *learning analytics* diseñada por medio de la metodología cuasi-experimental, a través del proceso pre intervención y post intervención, al comparar las medianas de la evaluación pre test y post test, se confirmó un mayor acercamiento de los profesores hacia las mencionadas técnicas.

- La validación de la propuesta del uso de los simuladores digitales para facilitar la comprensión docente de las técnicas del análisis estadístico implicativo y el *learning analytics*, estableció que fueron pertinentes las primeras aproximaciones con el empleo de herramientas *excel* y *R*, acorde a las capacidades de los profesores. Los expertos indican que promoverían la utilización de los simuladores con los estudiantes en el aula, aportar así al análisis de datos educativos de formato binario (1 y 0) y el uso masivo de las herramientas TIC.

RECOMENDACIONES

- Es importante que los docentes de bachillerato apropien de las técnicas comunes del análisis estadístico implicativo y el *learning analytics*, para que apliquen en sus contextos educativos, debido a que en la actualidad por el uso masivo de las TIC en procesos de enseñanza aprendizaje, se generan datos que, también, es posible analizarlos de acuerdo a estas nuevas técnicas, en otras palabras, gracias a los algoritmos presentes en las técnicas es posible descubrir reglas de desempeño del grupo estudiantil y así mejorar los procesos de enseñanza-aprendizaje.
- Se toma en consideración las normas vigentes para la redacción y presentación de textos científicos e investigaciones, que aportan a una construcción consistente del estado de arte.
- Es necesario desarrollar procesos de aproximación docente, para el análisis de datos educativos masivos presentes en formato: frecuencial, de intervalo y de razón.
- Debido al dominio docente de la herramienta *Office*, se sugiere abordar el uso de los simuladores presentados en la página web, desde las hojas *Excel* presentadas y, a continuación, el del código abierto *R*, para así impulsar la cultura de análisis de datos educativos masivos presentes en tipos: estructurados, no estructurados y semi estructurados de la tendencia *big data*.

BIBLIOGRAFÍA

- Baker, R., e Inventado, P. (2014). *Educational Data Mining and Learning Analytics*.
- Baker, R., y Siemens, G. (2014). *Educational Data Mining and Learning Analytics*. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (2 ed., pp. 253-272). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139519526.016>
- Benavides Herrera, P. V. (2014). *Propuesta de un currículo macro para educación general básica de jóvenes y adultos con escolaridad inconclusa*.
- Bender, S., y Fish, A. (2000). *The transfer of knowledge and the retention of expertise: The continuing need for global assignments*. *Journal of Knowledge Management*, 4, 125-137. doi: 10.1108/13673270010372251
- Caraballo Colmenares, R. (2007). *La andragogía en la educación superior*. *Investigación y Postgrado*, 22(2), 187-206.
- Carvajal, V. (2010). *Consideraciones para el uso de simulaciones en entornos virtuales para el aprendizaje de las estrategias de programación de computadoras*. *Innovaciones educativas*, 11, 1. doi: 10.22458/ie.v11i16.548
- Contreras Gelves, G. A., García Torres, R., y Ramírez Montoya, M. S. (2010). *Uso de simuladores como recurso digital para la transferencia de conocimiento*. [simuladores; recurso digital; transferencia de conocimiento; programación;]. *Apertura*, 2(1), 15.
- Couturier, R. (2008). CHIC: *Cohesive Hierarchical Implicative Classification* (Vol. 127).

- Couturier, R., y Pazmiño, R. (2016). *Use of Statistical Implicative Analysis in Complement of Item Analysis*. International Journal of Information and Education Technology, 6(1), 39-43. doi: 10.7763/ijiet.2016.v6.655
- Chatti, M. A., Dyckhoff, A. L., Schroeder, U., y Thüs, H. (2012). *A reference model for learning analytics*. International Journal of Technology Enhanced Learning, 4(5-6), 318-331.
- Chiappe, A., y Rodríguez, L. P. (2017). *Learning Analytics in 21st century education: a review*. Ensaio: Avaliação e Políticas Públicas em Educação, 25, 971-991. <https://doi.org/10.1590/S0104-40362017002501211>
- Daza, W. G. I. (2018). *Investigación educativa desde un enfoque cualitativo: la historia oral como método*. Voces de la educación, 3, 93-110.
- Drachsler, H., y Greller, W. (2016). *Privacy and analytics: it's a delicate issue a checklist for trusted learning analytics*. Paper presented at the Proceedings of the Sixth International Conference on Learning Analytics and Knowledge, Edinburgh, United Kingdom.
- Ferguson, R. (2012). *Learning analytics: Drivers, developments and challenges (Vol. 4)*.
- Fynn, A. (2016). *Ethical Considerations in the Practical Application of the Unisa Socio-Critical Model of Student Success (Vol. 17)*.
- Garaizar, P., y Guenaga, M. (2014). *A multimodal learning analytics view of HTML5 APIs: Technical benefits and privacy risks*. Paper presented at the 2nd International Conference on Technological Ecosystems for Enhancing Multiculturality, TEEM 2014.

García-Peñalvo, F., Pazmiño-Maji, R., Naranjo-Ordoñez, L., y Conde-González, M. (2019). *Learning Analytics in Ecuador: An Initial Analysis based in a Mapping Review*.

García-Peñalvo, F. J., Griffiths, D., Johnson, M., Sharples, P., y Sherlock, D. (2014). *Problems and opportunities in the use of technology to manage informal learning*. Paper presented at the 2nd International Conference on Technological Ecosystems for Enhancing Multiculturality, TEEM 2014.

Gil Vera, V. (2018). *Análisis del aprendizaje: una revisión sistemática de literatura (Vol. 10)*.

Gras, R., y Kuntz, P. (2009). *El Analisis Estadistico Implicativo (ASI) en respuesta a problemas que le dieron origen*.

Greller, W., y Drachsler, H. (2012). *Translating Learning into Numbers: A Generic Framework for Learning Analytics (Vol. 15)*.

Higueras, A. M. G., Campano, L. L., Olivera, V. M., Suances, A. M., y Gómez, J. L. S. (2014). *Atmospheric physics group open data (GFA Open Data): Meteorological data and tools for learning analytics*. Paper presented at the 2nd International Conference on Technological Ecosystems for Enhancing Multiculturality, TEEM 2014.

Kaufman, L., y Rousseeuw, P. J. (2009). *Finding groups in data: an introduction to cluster analysis (Vol. 344)*: John Wiley and Sons.

Knowles, M. S. (1984). *Andragogy in action*. San Francisco: Jossey-Bass.

Kuyper, M., Hoog, R., y Jong, T. (2001). *Modeling and supporting the authoring process of multimedia simulation based educational software: A knowledge*

engineering approach. *Instructional Science*, 29, 337-359. doi: 10.1023/a:1011947909796

Larrea de Granados, E., y Granados Boza, V. (2013). *EL sistema de educación superior para la sociedad del buen vivir basada en el conocimiento: el caso ecuatoriano*. El sistema de educación superior para la sociedad del buen vivir basada en el conocimiento: el caso ecuatoriano. Disponible en: <http://repositorio.ucsg.edu.ec/handle/3317/533>

Lasa, A., Sáenz de Cabezón, Á., Wilhelmi, M. R., y Abaurrea, J. (2019). *El problema del caftán: Proporcionalidad como herramienta óptima en un problema de resolución de ecuaciones* *The caftan problem: Proportionality as an optimum tool in equation solving problems*. *Educação Matemática Pesquisa : Revista do Programa de Estudos Pós-Graduados em Educação Matemática*; v. 21, n. 3 (2019): Volume 21-3 - Temático - Educação Algébrica. doi: 10.23925/1983-3156.2019vol21i3p209-230

Law, A. M., y McComas, M. G. (2001). *Building valid models: how to build valid and credible simulation models*. Paper presented at the Proceedings of the 33rd conference on Winter simulation, Arlington, Virginia.

Lerman, I.-C. (1981). *Classification et Analyse Ordinale des Données*, Dunod.

Martin, S. (2014). *Measuring cognitive load and cognition: metrics for technology-enhanced learning*. [Article]. *Educational Research and Evaluation*, 20, 592-621. doi: 10.1080/13803611.2014.997140

Mayor, E. (2015). *Learning predictive analytics with R*.

Montero, C. S., y Suhonen, J. (2014). *Emotion analysis meets learning analytics - Online learner profiling beyond numerical data*. Paper presented at the 14th

Koli Calling International Conference on Computing Education Research, Koli Calling 2014.

Naranjo, M., Maji, R. P., Conde, M., y Peñalvo, F. (2018). *Métodos de agrupamiento LA & SIA: Comparación computacional*. Paper presented at the Congreso de Ciencia y Tecnología ESPE.

Nespereira, C. G., Dai, K., Vilas, A. F., y Redondo, R. P. D. (2014). *Is the LMS access frequency a sign of students' success in face-to-face higher education?* Paper presented at the 2nd International Conference on Technological Ecosystems for Enhancing Multiculturality, TEEM 2014.

Ochoa, X., Worsley, M., Chiluzza, K., y Luz, S. (2014). *MLA'14 - Third multimodal learning analytics workshop and grand challenges*. Paper presented at the 16th ACM International Conference on Multimodal Interaction, ICMI 2014.

Olekhnik, S. N., Nesterenko, I. U. r. V., Potapov, M. K., y Bernardo, E. A. (1994). *Antiguos problemas recreativos en Rusia*: Servicio Editorial, Universidad de País Vasco.

Orús, P., Zamora, L., y Gregori, P. (2009). *Teoría y aplicaciones del Análisis Estadístico Implicativo. Primera aproximación en lengua hispana*. Universitat Jaume I, Castellón de la Plana: Innovació Digital Castelló, slu.

Oviatt, S., y Cohen, A. (2014). *Written activity, representations and fluency as predictors of domain expertise in mathematics*. Paper presented at the 16th ACM International Conference on Multimodal Interaction, ICMI 2014.

Pardo, A., y Siemens, G. (2014). *Ethical and privacy principles for learning analytics*. British Journal of Educational Technology, 45(3), 438-450. doi: 10.1111/bjet.12152

Pazmiño-Maji, R., Naranjo-Ordoñez, L., Conde-González, M., y García-Peñalvo, F. (2019). *Learning Analytics in Ecuador: An Initial Analysis based in a Mapping Review*.

Pazmiño-Maji, R. A., García-Peñalvo, F. J., y Conde-González, M. A. (2016). *Approximation of statistical implicative analysis to learning analytics: a systematic review*. Paper presented at the Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality.

Pazmiño Maji, R. A. (2021). *Aporte del Análisis Estadístico Implicativo a Learning Analytics*.

Pazmiño, B. M. B., y Maji, R. A. P. (2018). *Literatura científica sobre análisis estadístico implicativo: un mapeo sistemático de la década que transcurre*. *Ciencia Digital*, 2(4.1.), 55-69.

Pazmiño, R. (2014). *Aproximación al Análisis Estadístico Implicativo desde sus Aplicaciones Educativas*.

Peña-Ayala, A. (2017). *Learning Analytics: Fundaments, Applications, and Trends: A View of the Current State of the Art to Enhance e-Learning (Vol. 94)*.

Peñafiel, A. G., Zúñiga, B. B., y González, M. D. O. (2018). *El aprendizaje basado en simulación y el aporte de las teorías educativas*.

Pravilovic, S., Bilancia, M., Appice, A., y Malerba, D. (2016). *Using Multiple Time Series Analysis for Geosensor Data Forecasting*. *Information Sciences*, 380. doi: 10.1016/j.ins.2016.11.001

Rayón, A., Guenaga, M., y Núñez, A. (2014). *Supporting competency-assessment through a learning analytics approach using enriched rubrics*. Paper

presented at the 2nd International Conference on Technological Ecosystems for Enhancing Multiculturality, TEEM 2014.

Romero, C., y Ventura, S. (2007). *Educational data mining: A survey from 1995 to 2005 (Vol. 33)*.

Romero, C., y Ventura, S. (2010). *Educational Data Mining: A Review of the State of the Art (Vol. 40)*.

Romero, C., y Ventura, S. (2013). *Data Mining in Education (Vol. 3)*.

Ruiz, J. S., Díaz, H. J. P., Ruipérez-Valiente, J. A., Muñoz-Merino, P. J., y Kloos, C. D. (2014). *Towards the development of a learning analytics extension in open edX*. Paper presented at the 2nd International Conference on Technological Ecosystems for Enhancing Multiculturality, TEEM 2014.

Sampieri, R. H., Collado, C. F., y Lucio, P. B. (2014). *Metodología de la investigación*. McGraw-Hill Education.

Softic, S., De Vocht, L., Taraghi, B., Ebner, M., Mannens, E., y De Walle, R. V. (2014). *Leveraging learning analytics in a personal learning environment using linked data*. [Article]. Bulletin of the Technical Committee on Learning Technology, 16(4), 10-13.

Torres Barzabal, L. M. (2005). *Elementos que deben contener las páginas web educativas*. Pixel-Bit. Revista de Medios y Educación, 25, 75-83.

Valls Pla X, G. H. P. (2014). *Diseño de un paquete R para el Análisis Estadístico Implicativo Trabajo Final de Master en Matemática Computacional*. Universitat Jaume I.

Vihavainen, A., Helminen, J., y Ihantola, P. (2014). *How novices tackle their first lines of code in an IDE: Analysis of programming session traces*. Paper presented at the 14th Koli Calling International Conference on Computing Education Research, Koli Calling 2014.

Wolpers, M., Najjar, J., Verbert, K., y duval, e. (2007). *Tracking Actual Usage: The Attention Metadata Approach (Vol. 10)*.

Zamora, L., Gregori, P., y Orús, P. (2007). *Conceptos fundamentales del Análisis Estadístico Implicativo (ASI) y su soporte computacional CHIC*. ASI4, 38.

ANEXOS

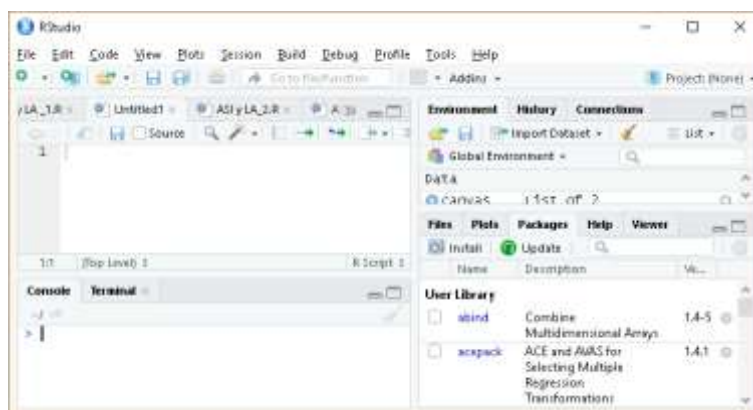
ANEXO 1

A. Instalación de los *Software: RStudio, R y Rchic*

Para instalar el *Software R Studio*:

1.- Ir a la página: <https://www.rstudio.com/>

2.- Proceder con la descarga e instalación del *Software RStudio*, de acuerdo a la versión del computador. (Guiarse en: <https://bit.ly/3mu89N9>).

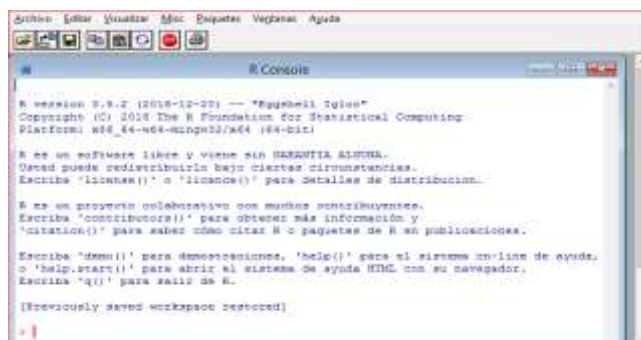
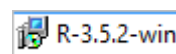


Para instalar el *Software R de acuerdo a la versión de Rchic*

1.- Por ejemplo, para la versión 0.27 de *Rchic*, es compatible la versión 3.5.2 de *R*, la cual, se descarga de la página: <https://cran.r-project.org/bin/windows/base/old/3.5.2/>



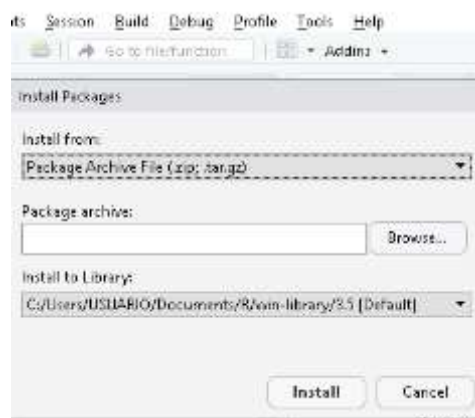
2.- Proceder con la descarga e instalación del *Software R (3.5.2-win)*, guiarse este proceso de instalación en: <https://bit.ly/3mu89N9>



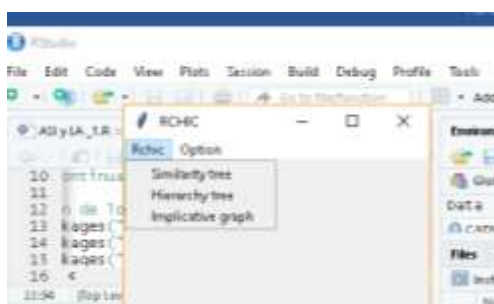
Para cargar el paquete *Rchic*

1.- Ir a la página: <https://bit.ly/3wExmJb>

- 2.- Descargar el paquete *Rchic* de acuerdo al *software* del computador y la versión de *R*, junto con los paquetes adicionales sugeridas por el desarrollador (ver la parte A del Anexo 4).
- 3.- Proceder con la carga de *Rchic* tanto en *RStudio* y en *R*, por medio del uso del menú: Tools>>Install Packages o Paquetes:



- 4.- Verificar el funcionamiento, teclear los comandos `library(rchic)` y `rchic()`, en la ventana consola tanto de *RStudio* como de *R*.



Acerca del *Software CHIC*

CHIC es una herramienta de análisis de datos basado en el Análisis Estadístico Implicativo (ASI), el objetivo es determinar las implicaciones más significativas entre las diferentes variables, existen dos formas de organizar estas implicaciones: (1) como un grafo Implicativo y (2), produce un árbol de similitud no orientado basados en la probabilidad de los enlaces entre estados (Couturier, 2008).

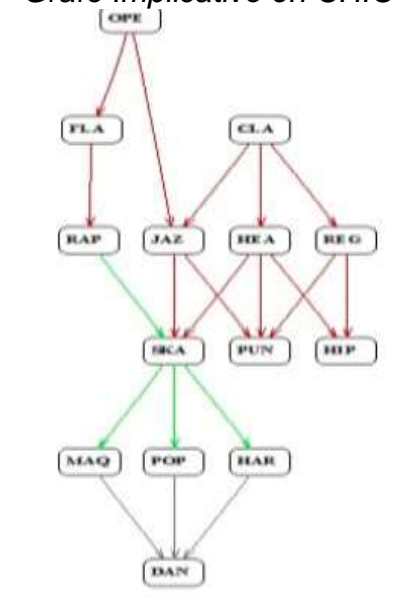
Características del Software CHIC



El entorno del Software CHIC



Grafo Implicativo en CHIC



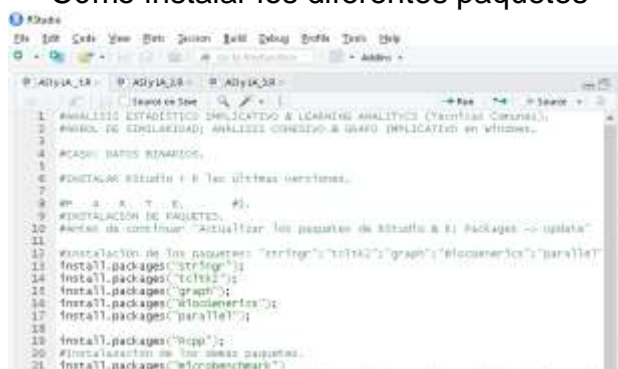
Fuente: tomado a partir de Zamora *et al.* (2007)

B. El proceso de carga de los diferentes paquetes

Proceder con la carga de los paquetes citados en la parte C del Anexo 2. En otro caso, crear el *R Script* en *RStudio* y proceder con la verificación, teclear en:

 **Run**, tener conexión a Internet.

Cómo instalar los diferentes paquetes



```

1 #ANÁLISIS ESTADÍSTICO IMPLICATIVO & LEARNING ANALYTICS (Técnicas Comunes)
2 #ARBOOL DE SIMILARIDAD | ANÁLISIS CONEJIVO & GRAFO IMPLICATIVO EN WINDOWS
3
4 #CASO: DATOS BINARIOS.
5
6 #INSTALAR RStudio e R las últimas versiones.
7
8 #R -> A, T, E. #
9 #INSTALACION DE PAQUETES.
10 #Antes de continuar "Actualizar" los paquetes de RStudio e R Packages -> update
11
12 #Instalación de los paquetes "string", "tcltk2", "graph", "wlocamerics", "parallel"
13 install.packages("string")
14 install.packages("tcltk2")
15 install.packages("graph")
16 install.packages("wlocamerics")
17 install.packages("parallel")
18
19 install.packages("Rcpp")
20 #Instalación de los otros paquetes.
21 install.packages("microbenchmark")

```

Verificación de los diferentes paquetes



```

1 #ANÁLISIS ESTADÍSTICO IMPLICATIVO & LEARNING ANALYTICS (Técnicas Comunes)
2 #ARBOOL DE SIMILARIDAD | ANÁLISIS CONEJIVO & GRAFO IMPLICATIVO EN WINDOWS
3
4 #CASO: DATOS BINARIOS.
5
6 #R -> A, T, E. #
7 #VERIFICACIÓN DE LAS LIBRERÍAS.
8 library("microbenchmark")
9 library("tcltk2")
10 library("ggplot2")
11 library("factextra")
12 library("cluster")
13 library("fastcluster")
14 library("clustfx")

```

ANEXO 2

A. Herramientas informáticas del *Learning Analytics* (LA)

Herramientas de extracción, transformación y carga	
Nombre de la herramienta informática.	Referencia
<i>Pentaho Data Integration</i>	http://www.pentaho.com/product/dataintegration
<i>Talend Data Integration</i>	https://www.talend.com/products/dataintegration
<i>Yahoo! Pipes</i>	https://pipes.yahoo.com/
Herramienta del análisis web	
<i>Google Analytics</i>	http://www.google.com/analytics/
<i>Open Web Analytics (OWA)</i>	http://www.openwebanalytics.com/
<i>Piwik</i>	http://piwik.org/
Herramienta de Inteligencia de Negocios	
<i>Metrics That Matter</i>	http://www.executiveboard.com/echd/human-resources/metrics/all-matters/index.page
<i>Microsoft BI suite</i>	https://powerbi.microsoft.com/es-es/
<i>Pentaho</i>	http://www.pentaho.com/product/business-visualization-analytics
Herramientas de la visualización de la información	
<i>Cytoscape</i>	http://www.cytoscape.org/
<i>Gephi</i>	https://gephi.github.io/
<i>Tableau</i>	https://www.tableausoftware.com/
Herramienta de análisis de las redes sociales	
<i>NodeXL</i>	http://nodexl.codeplex.com/
<i>NetMiner</i>	http://www.netminer.com/
<i>Pajek</i>	http://pajek.usfm.si/doku.php
<i>SNAPP</i>	http://www.snappvis.org/
Herramientas del análisis de textos	
<i>Leximancer</i>	http://info.leximancer.com/
<i>TagHelper</i>	http://www.cs.cmu.edu/~cpenso/TagHelper.html
<i>Wordle</i>	http://www.wordle.net/
Herramientas de análisis de propósito general	
<i>R</i>	https://www.r-project.org/
<i>SAS (Statistical Analysis System)</i>	http://www.sas.com/
<i>SPSS (Statistical Package for the Social Sciences)</i>	http://www-01.ibm.com/software/analytics/spss/
Herramienta de minería de datos.	
<i>Coron</i>	http://coron.loria.fr/
<i>KEEL (Knowledge Extraction based on</i>	http://keel.es/

<i>Evolutionary Learning</i>	
<i>Rapid Miner</i>	http://rapidminer.com/
<i>Weka</i>	http://www.cs.waikato.ac.nz/ml/weka/index.html
Herramientas de analíticas de aprendizajes de propósito especial	
<i>BNT-SM (Bayes Net Toolbox for Student Modelling)</i>	http://www.cs.cmu.edu/~listen/BNT-SM/
<i>Degree Compass</i>	http://www.brightspace.com/products/degreecompass/
<i>myClass</i>	http://nextell.eu/portfolios/myclass/
<i>ProNFA (Probabilistic Non-Invasive Formative Assessment)</i>	http://nextell.eu/portfolios/pronifa/
<i>PSLC DataShop</i>	https://pslcdatahop.web.cmu.edu/index.jsp
<i>Student Success System</i>	http://www.brightspace.com/products/insights/
Análisis en sistemas de aprendizaje electrónico	
<i>Blackboard Analytics for Learn</i>	https://bit.ly/3uxypwf
<i>Student Activity Monitor</i>	http://www.role-widgestore.eu/content/student-activity-monitor
<i>LOCO Analysis</i>	http://jelenajevanovic.net/LOCOAnalyst/index.html
<i>Engagement Analysis for Moodle</i>	https://moodle.org/plugins/view.php?plugin=report_engagement https://docs.moodle.org/22/en/Engagement_Analytics_Plugin
<i>Compod</i>	http://esskmi.tugraz.at/8080/compod/web/

Nota. Tabla que muestra las diferentes herramientas para las Analíticas de Aprendizajes.

Fuente. (<http://leas-box.cognitive-science.at/downloads/D3.1.pdf>)

B. Simulaciones digitales en Excel para las técnicas del ASI

Técnica: Árbol de similaridad

Simulación en Excel para el Índice de Similitud

Definiciones: Según Orús *et al.* (2009), "la similitud es una medida de correspondencia o semejanza entre los objetos que van a ser agrupados".

Menciona Orús *et al.* (2009), "sea un conjunto I formado por n individuos y un conjunto A formado por p características, $A = \{a_1, a_2, \dots, a_p\}$, donde: $A_i = \{x \in I : a(x) = 1\}$, $\text{Card}(I) = n$ y $\text{Card}(A_i) = n_i$;

Al número dado por:

$$s(a_i, a_j) = \text{Prob} \left[\frac{\text{Card}(X_i \cap X_j) - \frac{n_i n_j}{n}}{\sqrt{\frac{n_i n_j}{n}}} \leq K_c \right] = \frac{1}{\sigma} \int_{-\infty}^{K_c} e^{-t^2/2} dt$$

Se le llama Índice de similitud de la pareja $\{a_i, a_j\}$.

Proceso de Aproximación a la definición: Ejemplo cotidiano: "A un grupo de n estudiantes se pueden verificar a características".

"La palabra 'Card' quiere decir 'Cantidad de'; en este caso, $\text{Card}(X_i \cap X_j)$ quiere decir cantidad de estudiantes que coinciden en características X_i y a la vez con X_j ".

"La expresión 'Prob', calcula la probabilidad de aquellas coincidencias llamadas copresencias".

Uso de la herramienta digital en vis: Proceso del cálculo en Excel del elemento denominado "Índice de Similitud".

Pasos:
 1.- "Se inicia con una tabla que refleje los resultados de una encuesta" (pueden ser resultados de una evaluación/observación, con datos de tipo binario, es decir datos de la forma: sí/no, verdadero/falso, admite/no admite, con propiedad/no propiedad, aprueba/desaprueba, admitido/no admitido, promovido/no promovido, motivado/desmotivado), estos datos convertidos a 1 y 0 respectivamente, tal como se muestra en la figura (acompañar con el documento Excel en línea descargado).

1	Indice de similitud	0,1	0,1	0,1	0,1	0,1
2	Indice de similitud	1	1	1	1	1
3	Indice de similitud	0,1	0,1	1	0,1	0,1
4	Indice de similitud	0,1	0,1	0,1	1	0,1
5	Indice de similitud	0,1	0,1	0,1	0,1	1
6	Indice de similitud	0,1	0,1	0,1	0,1	0,1
7	Indice de similitud	0,1	0,1	0,1	0,1	0,1

- 2.- Calcular los totales de cada columna (los n_{aj}).
- 3.- Se forma una nueva tabla.

H	I	J	K	L	M	N	O
Un primer acercamiento al cálculo del índice de similitud							
Indice de similitud		Card(A _i ∩ A _j)		n _i * n _j		K _c	
HIP	JAZ	11	11	10	10	0,584148	0,584148
HIP	HEA	7	7	5,5	5,5	0,856789	0,856789
HIP	REG	10	10	7	7	0,9458524	0,9458524
HIP	PUN	11	11	8	8	0,856789	0,856789
JAZ	HEA	10	10	8	8	0,856789	0,856789
JAZ	REG	10	10	8	8	0,856789	0,856789
JAZ	PUN	11	11	8	8	0,856789	0,856789
HEA	REG	10	10	8	8	0,856789	0,856789
HEA	PUN	11	11	8	8	0,856789	0,856789
REG	PUN	11	11	8	8	0,856789	0,856789

4.- Se determina las coincidencias (llamadas copresencias) entre A_i y A_j , es decir el:

$$\text{Card}(A_i \cap A_j)$$

Ejemplos: "hay 10 estudiantes que coinciden en sus opiniones entre A_i =HIP y A_j = PUN, hay 7 estudiantes que coinciden en sus opiniones entre PUN y JAZ, entre otros".

5.- "Se obtiene el valor de $(n_i * n_j) + n$, al multiplicar las filas y dividir el resultado entre el total".

Ejemplo: " $(11 * 11) + 20 = 141$; $(11 * 10) + 20 = 130$, entre otros".

6.- Se calculan las copresencias estandarizadas K_c : **Ejemplo:** $(10 / 141) = 0,071$; $(8 / 130) = 0,062$, entre otros.

7.- "Para la columna $s(a_i, a_j)$ se calcula el área bajo la curva normal estándar hasta K_c . **Ejemplo:** El área bajo la curva normal estándar hasta 1,61 es 0,9458524; El área bajo la curva normal estándar hasta 1,07 es 0,856789; entre otros".

8.- "Este número hallado, se llama el índice de similitud".

"Recordando que, si este número se acerca a 1, hay más similitud entre este grupo de datos". **Ejemplo:** "Hay más similitud entre HIP y PUN, pues $s(\text{HIP}, \text{PUN}) = 0,9458524$ ".

Actividades propuestas: En el documento Excel descargado, se solicita al lector, revisar la conformación de las siguientes tablas:

3.- Matriz cruzada de todas las Copresencias (K)

Card(A _i ∩ A _j)	HIP	JAZ	HEA	REG	PUN
HIP	11	6	7	8	11
JAZ	6	10	8	8	7
HEA	7	8	10	8	8
REG	8	8	8	10	8
PUN	11	7	8	8	11

4.- Determinar el valor $n_i * n_j + n$ de todas las variables.

n _i * n _j + n	HIP	JAZ	HEA	REG	PUN
HIP	141	130	130	130	141
JAZ	130	141	130	130	130
HEA	130	130	141	130	130
REG	130	130	130	141	130
PUN	141	130	130	130	141

5.- Determinar el valor K_c :

K _c	HIP	JAZ	HEA	REG	PUN
HIP	0,071	0,062	0,062	0,062	0,071
JAZ	0,062	0,071	0,062	0,062	0,062
HEA	0,062	0,062	0,071	0,062	0,062
REG	0,062	0,062	0,062	0,071	0,062
PUN	0,071	0,062	0,062	0,062	0,071

6.- Determinar la Matriz de similitud.

S(a _i , a _j)	HIP	JAZ	HEA	REG	PUN
HIP	0,9458524	0,584148	0,7387844	0,856789	0,9458524
JAZ	0,584148	0,9458524	0,9101438	0,9101438	0,7387844
HEA	0,7387844	0,9101438	0,9458524	0,9101438	0,856789
REG	0,856789	0,9101438	0,9101438	0,9458524	0,856789
PUN	0,9458524	0,7387844	0,856789	0,856789	0,9458524

Matriz de similitud al nivel cero

Se solicita al lector a completar los siguientes ejercicios de aplicación.

- Pasos:**
- 1.- "Registrar (cambiar) valores (o 1 o 0, a su gusto), en la tabla 1".
 - 2.- "Observar el cambio de los valores totales en cada columna de la tabla 1 y su registro en la tabla 2".

3.- "Ahora con estos cambios, podemos responder: ¿A cuántos estudiantes les gusta tanto PUN y REG?, entre otras".

4.- "Comprobar el cálculo del valor de $(n_i * n_j) + n$, para PUN y JAZ", y otros.

5.- "Comprobar el cálculo del valor de K_c , para PUN y JAZ", y otros.

6.- "Calcule el área bajo la curva normal para los nuevos valores de K_c ".

7.- Se solicita "observar los nuevos valores índices de similitud, y como ejemplo, identificar el valor de $s(\text{PUN}, \text{HEA})$ ".

Evaluación: "Para verificar la aproximación a la comprensión docente de la definición de ésta técnica, se solicita al usuario-docente, contestar una encuesta de los procesos explicados versus las actividades planteadas, la misma que se encuentra en el enlace: <https://bit.ly/2FRx2B0> o haciendo clic en el archivo Excel en el botón: 'Encuesta 1'".

RECOMENDACIÓN. - Para continuar con las simulaciones, se solicita al lector no guardar los cambios efectuados en el archivo Excel, es decir, trabajar con los datos originales, hasta la revisión de toda la documentación, después puede hacer uso de la documentación, como a bien tuviere.

Parte 2

Simulación en Excel para el Árbol de Similitud

Definiciones. “Niveles de jerarquía”: Según Lerman (1970), citado por Zamora et al. (2007), “se buscan los nuevos Índices de Similitud al combinar la clase (a_i, a_j), con mayor índice en el paso anterior, así: Nivel cero de jerarquía es la matriz de similitud generada en pasos anteriores”.

S(a _i , a _j)	HIP	JAZ	HEA	REG	PUN
HIP	0,5844148	0,5844148	0,73878436	0,85678899	0,84654843
JAZ	0,5844148	0,91014375	0,91014375	0,73878436	0,73878436
HEA	0,73878436	0,91014375	0,91014375	0,91014375	0,85678899
REG	0,85678899	0,73878436	0,91014375	0,91014375	0,85678899
PUN	0,84654843	0,73878436	0,85678899	0,85678899	0,85678899

Matriz de similitud al nivel cero

“Nivel uno de jerarquía: Se obtiene por la fórmula expuesto en”: Zamora et al. (2007) que menciona: “cada clase previamente aislada: $s((a_i, a_j), a_k) = \{ \text{Max}[s(a_i, a_k); s(a_j, a_k)] \}^2$, con las clases formadas: $s((a_i, a_j), (a_{k1}, a_{k2}, a_{k3})) = \{ \text{Max}[s(a_i, a_{k1}); s(a_i, a_{k2}); s(a_i, a_{k3}); s(a_j, a_{k1}); s(a_j, a_{k2}); s(a_j, a_{k3})] \}^{2^3}$.”

Que sería el nivel tres de jerarquía. En general: $s(C_1, C_2) = \{ \text{Max}[s(a_i, a_j); \forall a_j \in C_1, a_k \in C_2] \}^{C_1 \times C_2 \times C_2}$, donde C1 y C2 son dos clase previamente formadas”.

Uso de la herramienta digital en vías de aproximación a la comprensión de la técnica Proceso de asimilación: “El nivel cero de jerarquía es la matriz de similitud”. Para lo cual, se solicita al lector estar en la Hoja correspondiente al “Árbol de Similitud” y observar la Tabla J0.

S(a _i , a _j)	HIP	JAZ	HEA	REG	PUN
HIP	0,84654843	0,73878436	0,85678899	0,84654843	0,84654843
JAZ	0,5844148	0,91014375	0,91014375	0,73878436	0,73878436
HEA	0,73878436	0,91014375	0,91014375	0,91014375	0,85678899
REG	0,85678899	0,73878436	0,91014375	0,91014375	0,85678899
PUN	0,84654843	0,73878436	0,85678899	0,85678899	0,85678899

Matriz de similitud al nivel uno

“Para el nivel uno de jerarquía: 1.- Hallamos el valor máximo de la tabla J0. Localizamos las variables que cumplen con esta condición. Estas variables (en ejemplo. HIP y PUN) se unen, para las demás celdas de la Tabla J1, se aplica la fórmula”:

$$s((a_i, a_j), a_k) = \{ \text{Max}[s(a_i, a_k); s(a_j, a_k)] \}^2$$

S(a _i , a _j)	HIP, PUN	JAZ	HEA	REG
HIP, PUN	0,91014375	0,73878436	0,85678899	0,84654843
JAZ	0,5844148	0,91014375	0,91014375	0,73878436
HEA	0,73878436	0,91014375	0,91014375	0,91014375
REG	0,85678899	0,73878436	0,91014375	0,91014375

Matriz de similitud al nivel uno

“Para el ejemplo se aplica la forma distributiva entre (HIP, PUN) y JAZ, obteniendo el máximo de la unión de las celdas (HIP, JAZ) y (PUN, JAZ), este resultado se eleva al cuadrado, sucesivamente con el resto de esta fila (por simetría, también se procede con la columna), las demás celdas quedan tal cual la tabla J0”.

“Para el nivel dos de jerarquía: 1.- Hallamos el valor máximo de la tabla J1 (para el ejemplo: 0,91014375). Localizamos las variables que cumplen con esta condición. Estas variables (en ejemplo. JAZ y HEA) se unen.

Para las demás celdas de la tabla J2, se aplica la fórmula”:

$$s((a_i, a_j), (a_{k1}, a_{k2})) = \{ \text{Max}[s(a_i, a_{k1}); s(a_i, a_{k2}); s(a_j, a_{k1}); s(a_j, a_{k2})] \}^4$$

S(a _i , a _j)	HIP	JAZ	HEA	REG	PUN
HIP	0,84654843	0,73878436	0,85678899	0,84654843	0,84654843
JAZ	0,5844148	0,91014375	0,91014375	0,73878436	0,73878436
HEA	0,73878436	0,91014375	0,91014375	0,91014375	0,85678899
REG	0,85678899	0,73878436	0,91014375	0,91014375	0,85678899
PUN	0,84654843	0,73878436	0,85678899	0,85678899	0,85678899

Matriz de similitud al nivel cero

S(a _i , a _j)	HIP, PUN	HEA, JAZ	REG
HIP, PUN	0,91014375	0,85678899	0,84654843
JAZ, HEA	0,5844148	0,91014375	0,91014375
REG	0,73878436	0,91014375	0,91014375

Matriz de similitud al nivel dos

“Para el ejemplo, se aplica la combinación entre (HIP, PUN) y (HEA, JAZ) obteniendo el máximo de la unión de las celdas (HIP, HEA), (HIP, JAZ), (PUN, HEA) y (PUN, JAZ), este resultado se eleva al

exponente 4, las demás celdas como tienen el nivel de jerarquía uno y se aplica esta regla (también por simetría, las celdas columnas)”.

“Para el nivel tres de jerarquía: 1.- Hallamos el valor máximo de la tabla J2 (para el ejemplo 0,82836165). Localizamos las variables que cumplen con esta condición. Estas variables (en el ejemplo. [JAZ, HEA] y REG) se unen, para las demás celdas de la Tabla J3, se aplica la fórmula”:

$$s((a_i, a_j), (a_{k1}, a_{k2}, a_{k3})) = \{ \text{Max}[s(a_i, a_{k1}); s(a_i, a_{k2}); s(a_i, a_{k3}); s(a_j, a_{k1}); s(a_j, a_{k2}); s(a_j, a_{k3})] \}^{2^3}$$

S(a _i , a _j)	HIP	JAZ	HEA	REG	PUN
HIP	0,91014375	0,85678899	0,84654843	0,84654843	0,84654843
JAZ	0,5844148	0,91014375	0,91014375	0,73878436	0,73878436
HEA	0,73878436	0,91014375	0,91014375	0,91014375	0,85678899
REG	0,85678899	0,73878436	0,91014375	0,91014375	0,85678899
PUN	0,84654843	0,73878436	0,85678899	0,85678899	0,85678899

Matriz de similitud al nivel cero

“Para el ejemplo, se aplica la distributiva entre (HIP, PUN), ((HEA, JAZ), REG), obteniendo el máximo de las celdas (HIP, HEA), (HIP, JAZ), (HIP, REG), (PUN, HEA), (PUN, JAZ) y (PUN, REG), este resultado se eleva al exponente 6”.

“Este proceso continúa en el caso de más variables”.

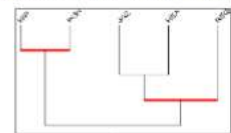
Actividades de refuerzo Se solicita al lector, revisar la hoja de Excel, la hoja correspondiente a los pasos citados en la parte anterior y mejorar la comprensión.

Actividades propuestas “Para el nivel de jerarquía cero - Se solicita al lector: Revisar la tabla J0 - Hacer doble clic en las Intersecciones (HIP, JAZ) y verificar (comprender) la fórmula del Excel - Hacer doble clic en la celda H12 y verificar la fórmula del Excel (del máximo de un grupo de celdas). “Para el nivel de jerarquía uno - Se solicita al lector: Revisar la tabla J1 - Hacer doble clic en las intersecciones ((HIP, PUN), JAZ) y verificar (comprender) la fórmula del Excel

- Hacer doble clic en la celda G20 y verificar la fórmula del Excel (del máximo de un grupo de celdas)”.

“Para el nivel de jerarquía dos - Se solicita al lector: Revisar la tabla J2 - Hacer doble clic en las intersecciones ((HIP, PUN), (HEA, JAZ)) y verificar (comprender) la fórmula del Excel - Hacer doble clic en la celda F28 y verificar la fórmula del Excel (del máximo de un grupo de celdas)”.

“Para el nivel de jerarquía tres - Se solicita al lector: Revisar la tabla J3 - Hacer doble clic en las intersecciones ((HIP, PUN), ((HEA, JAZ), REG)) y verificar (comprender) la fórmula del Excel”. “Comprender la conformación del árbol de similitud”:



Fuente: Zamora et al. (2007).

Evaluación Se solicita al lector que compare los resultados de las tablas J1, J2 y J3, con el árbol de similitud.

“Para verificar la aproximación a la comprensión docente de la definición de ésta técnica, se solicita al usuario-docente, contestar una encuesta de los procesos explicados versus las actividades planteadas, la misma que se encuentra en el enlace: <https://bit.ly/2RH2iKW> o haciendo clic en el archivo Excel en el botón: ‘Encuesta 2’”.

Parte 3

Simulación en Excel para los Nodos Significativos en un Arbol de Similaridad

Definiciones. Según (Orús et al., 2009), "los nodos significativos de un árbol de

similitud (jerárquico) son los nodos correspondientes a una clasificación compatible lo mejor posible con los valores y la calidad de los valores de similitud".

Para ello se debe:

1.- Según (Orús et al., 2009), "Determinar el preorden inicial y global Ω sobre $A \times A$, es decir:

$$G_1(\Omega) = \{(a, b) | (c, d) : s(a, b) < s(c, d)\}''$$

"Determinar el conjunto $s\pi_1$ pares separados al nivel k".

"Determinar el conjunto $r\pi_1$ pares que ya se han reunido hasta este nivel k".

"El conjunto $G_k(\Omega) \cap [s\pi_1, r\pi_1]$ conformado por pares de parejas que la nivel k respetan el preorden inicial", según (Orús et al., 2009), "por ejemplo, si se tiene: $s(e, f) < s(a, b)$, entonces $((e, f) | (a, b)) \in G_k(\Omega)$, y si al nivel k, e y f están aún separados, mientras que a y b se reúnen en la clase formada, la pareja $((e, f) | (a, b)) \in G_k(\Omega) \cap [s\pi_1, r\pi_1]$ ".

Siguen (Orús et al., 2009) quienes mencionan: "el cardinal de este último conjunto es función de este nivel k, y es un indicador del acuerdo entre el preorden inicial Ω y el preorden π_k inducido, al cardinal de $G_k(\Omega) \cap [s\pi_1, r\pi_1]$ se le asocia el índice aleatorio $G_k(\Omega^*) \cap [s\pi_1, r\pi_1]$, donde Ω^* es un preorden aleatorio en general, provisto de una probabilidad uniforme, de todos los preordenes del mismo tipo cardinal que Ω , este índice tiene, por esperanza: $0.5s_k r_k$, por varianza: $\frac{s_k r_k (s_k + r_k + 1)}{12}$; siendo: $s_k = \text{Card}[s\pi_k]$ y $r_k = \text{Card}[r\pi_k]$. El índice centrado se define como:

$$S(\Omega, k) = \frac{\text{Card}(G(\Omega) \cap [s\pi_k, r\pi_k]) - 0.5s_k r_k}{\sqrt{\frac{s_k r_k (s_k + r_k + 1)}{12}}}$$

2.- De la misma manera, según (Orús et al., 2009) "determinar el nivel significativo, que es un nivel que corresponde a un máximo local de $S(\Omega, k)$ durante la construcción de la jerarquía".

Un nodo significativo según (Orús et al., 2009) "es cualquier nodo formado a un nivel que corresponde a un máximo local de $V(\Omega, k)$, donde: $V(\Omega, k) = S(\Omega, k) - S(\Omega, k-1)$ ".

Para este respecto (Orús et al., 2009), sugieren el siguiente proceso: "Denotar por:

$$t = \text{Card}(A), A_i = \{a_j | j \in \mathbb{N}, j \in \mathbb{N}, j \in \mathbb{N}\},$$

a_i los elementos del preorden inicial Ω , con $i = 1, \dots, 2 * C_t$,

$p_i = \{p_j | j \in \mathbb{N}, j \in \mathbb{N}, j \in \mathbb{N}\}$ donde s_i es un valor cualesquiera que puede tomar el índice de similitud, y $i = 1, \dots, d$, siendo d, la cantidad de valores diferentes que toma el índice de similitud".

El ejercicio termina, al completar una tabla con la información del índice centrado $S(\Omega, k)$ y el nodo significativo respectivo $V(\Omega, k)$.

Uso de la herramienta digital en vías de aproximación a la comprensión de la técnica

Con este respecto, se ha preparado una hoja de Excel en línea <https://bit.ly/32GhyZ8>

Se va a proceder la sugerencia de Zamora et al. (2007) un ejemplo desarrollado, que vamos a acompañar con la hoja de Excel en mención, para ello:

1.- Organizando los datos bajo el siguiente análisis:

1.1. Identificamos la matriz de similitud, con todos sus elementos.

S(a _i , a _j)	HIP	JAZ	HEA	REG	PUN
HIP	1	0.5844	0.7388	0.8568	0.9459
JAZ	0.5844	1	0.9101	0.9101	0.7388
HEA	0.7388	0.9101	1	0.9101	0.8568
REG	0.8568	0.9101	0.9101	1	0.8568
PUN	0.9459	0.7388	0.8568	0.8568	1

Matriz de similitud al nivel cero.

Identificamos en esta matriz, los diferentes valores similares y las variables a las que corresponden.

1.2. Clasificamos estos valores, desde el menor hasta el mayor, estos son los valores del preorden inicial.

$$\text{Valores de } G_1(\Omega) = \{ (0.5844, 0.7388, 0.8568, 0.9101, 0.9459) \}$$

Zamora et al. (2007), lo clasificaría como:



Fuente: Zamora et al. (2007).

1.3. Realizamos el conteo de cada clase, conformando según Zamora et al. (2007), "el conjunto de pares separados al nivel k".

$$s_k = \text{Card}[S\pi_k] = \{ 2, 4, 6, 6, 2 \}$$

1.4. Identificamos las variables en análisis:

$$A^* = \{ \text{HIP}, \text{JAZ}, \text{HEA}, \text{REG}, \text{PUN} \}$$

1.5. Identificamos el número de variables en análisis.

$$t = \text{Card}[A^*] = 5$$

1.6. Identificamos el número de elementos del preorden inicial.

$$m = \text{Card}[\Omega] = 20$$

1.7. Identificamos el número de elementos del preorden inducido por S.

m = Card[Ω]	=	20
Card(G ₁ (Ω))	=	152

Es decir, por medio de las celdas de Excel, se desarrolla la fórmula:

$$\text{Card}(G_k(\Omega)) = \frac{m(m-1)}{2} - \frac{\text{Card}(P_k) \times (\text{Card}(P_k) + 1)}{2}$$

2.- Comprendiendo el cálculo de los valores del nivel 1 de jerarquía:

2.1. Identificamos en la matriz de similitud, la primera pareja de variables con el mismo índice de similitud.

S(a _i , a _j)	HIP	JAZ	HEA	REG	PUN
HIP	1	0.5844	0.7388	0.8568	0.9459
JAZ	0.5844	1	0.9101	0.9101	0.7388
HEA	0.7388	0.9101	1	0.9101	0.8568
REG	0.8568	0.9101	0.9101	1	0.8568
PUN	0.9459	0.7388	0.8568	0.8568	1

2.2. Proceder a efectuar y comprender los siguientes cálculos:

```

Nivel 1 de la jerarquía
Card[Ω] = 20 // La primera pareja: P1 = {HIP, PUN}
Card[P1] = 2 // La que falta para C=20
A = {JAZ} // A = Ω - P1
s1 = 2 // s1 = Card(ai | aj ∈ A, i < j, i(ai) = s(ai))
r1 = 18 // r1 = Card(ai | aj ∈ A, i < j, i(ai) = s(ai))
Card[G1(Ω) ∩ [P1, A]] = 152 // Card(G1(Ω) ∩ [P1, A]) = Σ (si * (si + 1) / 2)
    
```

Obsérvese que es la aplicación de la fórmula propuesta, a la forma de Excel.

```

Nivel 2 de la jerarquía
s2 = Card[P1] = 2 // Las dos parejas: P2 = {HIP, JAZ, HEA, REG}
r2 = Card[Ω] - 2 // La que falta para C=20
A = {PUN} // A = Ω - P2
s2 = 4 // s2 = Card(ai | aj ∈ A, i < j, i(ai) = s(ai))
r2 = 16 // r2 = Card(ai | aj ∈ A, i < j, i(ai) = s(ai))
Card[G2(Ω) ∩ [P1, P2]] = 20 // Card(G2(Ω) ∩ [P1, P2]) = Σ (si * (si + 1) / 2)
    
```

3.- Comprendiendo el cálculo de los valores del nivel 2 de jerarquía:

3.1. Identificamos en la matriz de similitud, dos o más parejas de variables con el mismo índice de similitud.

S(a _i , a _j)	HIP	JAZ	HEA	REG	PUN
HIP	1				0.9459
JAZ		1	0.9101		
HEA			1		
REG				1	
PUN	0.9459				1

3.2. Proceder a efectuar y comprender los siguientes cálculos:

```

Nivel 2 de la jerarquía
s3 = Card[P2] = 4 // Las dos parejas: P3 = {HIP, JAZ, HEA, REG}
r3 = Card[Ω] - 4 // La que falta para C=20
A = {PUN} // A = Ω - P3
s3 = 4 // s3 = Card(ai | aj ∈ A, i < j, i(ai) = s(ai))
r3 = 16 // r3 = Card(ai | aj ∈ A, i < j, i(ai) = s(ai))
Card[G3(Ω) ∩ [P2, P3]] = 20 // Card(G3(Ω) ∩ [P2, P3]) = Σ (si * (si + 1) / 2)
    
```

Obsérvese que es la aplicación de la fórmula propuesta, a la forma de Excel.

```

Nivel 3 de la jerarquía
s4 = Card[P3] = 4 // Las dos parejas: P4 = {HIP, PUN}
r4 = Card[Ω] - 4 // La que falta para C=20
A = {JAZ, HEA, REG} // A = Ω - P4
s4 = 3 // s4 = Card(ai | aj ∈ A, i < j, i(ai) = s(ai))
r4 = 13 // r4 = Card(ai | aj ∈ A, i < j, i(ai) = s(ai))
Card[G4(Ω) ∩ [P3, A]] = 20 // Card(G4(Ω) ∩ [P3, A]) = Σ (si * (si + 1) / 2)
    
```

4.- Comprendiendo el cálculo de los valores del nivel 3 de jerarquía:

4.1. Identificamos en la matriz de similitud, dos o más parejas de variables con el mismo índice de similitud.

S(a _i , a _j)	HIP	JAZ	HEA	REG	PUN
HIP	1				
JAZ		1			
HEA			1		
REG				1	
PUN					1

4.2. Proceder a efectuar y comprender los siguientes cálculos:

Técnica: Grafo Implicativo

Simulación en Excel para el proceso: Análisis Estadístico Implicativo

Definiciones Según (Orús et al., 2009), "de forma similar que se trató en el Análisis Clasificadorio, en el AEI se parte de un conjunto I formado por n individuos y un conjunto A formado por p características,

$A = \{a_1, a_2, \dots, a_p\}$, donde: $A_i = \{x \in I : a_i(x) = 1\}$, $Card(I) = n$, $Card(A_i) = n_{a_i}$ y $Card(\bar{A}_i) = n - n_{a_i}$. Se debe calcular el índice de implicación y de cohesión,

en el análisis cohesivo se pueden determinar la tipicalidad y contribución de sujetos y variables, así como los nodos significativos, en la formación de las clases Zamora et al. (2007). En este contexto, la variable aleatoria de estudio es $Card(X_i \cap \bar{X}_j)$, es decir, la cantidad de individuos que poseen la característica a_i , pero no poseen la característica a_j . Lo explicado anteriormente para la variable aleatoria $Card(X_i \cap X_j)$ es también válida para $Card(X_i \cap \bar{X}_j)$."

Al concepto de la *cuasi-implicación*, Zamora et al. (2007), nos introduce con un ejemplo: "En la lógica matemática, la regla ' $a_i \rightarrow a_j$ ' es verdadera si para todo x , $a_j(x)$ sólo es nulo cuando $a_i(x)$ lo sea también; (...), en el caso de un cuestionario de conocimientos se podría observar a algunos raros alumnos que vencen el ítem a_i y no consiguen vencer el a_j , sin que se contradiga la tendencia a tener a_j cuando se tiene a_i , (...), la cuasi-implicación $a_i \rightarrow a_j$ significa que 'cuando a_i está presente entonces generalmente a_j están también presente'", para un mayor acercamiento a este concepto y detalles adicionales ver (Gras, 1999) y Zamora et al. (2007).

Bajo las observaciones tratadas en Zamora et al. (2007), consideramos:

1.- Según (Orús et al., 2009), "la Intensidad Implicativa de la regla $a_i \rightarrow a_j$, se define como:

$$\phi(a_i, a_j) = 1 - P(K \leq n_{a_i a_j}) \quad \text{si } n_{a_i} \neq n_{a_j} \text{ en caso contrario } \phi(a_i, a_j) = 0."$$

2.- De la misma forma, complementa (Orús et al., 2009), "el Índice de Implicación de la regla $a_i \rightarrow a_j$, se define como:

$$q(a_i, \bar{a}_j) = \frac{n_{a_i \bar{a}_j} - \frac{n_{a_i} n_{\bar{a}_j}}{n}}{\sqrt{\frac{n_{a_i} n_{\bar{a}_j}}{n}}}$$

Se puede demostrar que $\phi(a_i, a_j) = -q(a_i, \bar{a}_j) \sqrt{\frac{n_{a_i} n_{a_j}}{n}}$

3.- Según, (Orús et al., 2009), "la condición necesaria y suficiente para que la implicación $a_i \rightarrow a_j$ sea admisible al nivel de confianza $1 - \alpha$, es que $\phi(a_i, a_j) > 1 - \alpha$ ".

Grafo de implicación, según R. Gras y P. Kuntz

Es un grafo orientado, ponderado (hacia las intensidades de implicación) (Gras & Kuntz, 2009).

Uso de la herramienta digital en vías de aproximación a la comprensión de la técnica

Proceso del cálculo en Excel del elemento denominado "Índice de Implicación"

Pasos:

1.- Partimos de la tabla 1 (descargado) que refleja los resultados de una encuesta (pueden ser resultados de una evaluación/observación, con datos de tipo binario, es decir datos de la forma: si/no, verdadero/falso, admite/no admite, con propiedad/sin propiedad, aprueba/desaprueba, admitido/no admitido, promovido/no promovido, motivado/desmotivado), estos datos convertidos a 1 y 0 respectivamente, en la tabla 2 se refleja el conjunto complemento (en sentido opuesto a la tabla 1) tal como se muestra en la figura:

Jose1	1	1	0	0	0	1	1	Jose1	0	0	1	1	1	0	0
Jose2	0	0	1	1	0	1	1	Jose2	0	0	1	0	0	1	0
Jose3	0	0	0	0	1	1	1	Jose3	0	0	0	0	1	1	0
Jose4	0	1	1	1	1	1	0	Jose4	1	0	0	0	0	0	1
Jose5	0	0	1	1	1	0	1	Jose5	0	1	1	0	0	1	0
Jose6	0	0	1	0	0	1	0	Jose6	1	1	0	1	1	0	1
Jose7	0	1	1	1	1	0	1	Jose7	0	1	0	0	0	1	0
Jose8	0	1	1	0	0	1	0	Jose8	1	0	0	1	1	0	1
Jose9	0	1	0	1	1	1	1	Jose9	0	1	0	1	0	0	0
Jose10	0	0	0	0	0	0	0	Jose10	1	1	1	1	1	1	1
Jose11	0	1	1	1	1	1	1	Jose11	1	1	0	0	0	0	0
Jose12	0	0	1	1	0	1	1	Jose12	0	1	0	0	1	0	1
Jose13	0	0	1	0	0	1	1	Jose13	0	1	1	0	1	1	0
Jose14	0	1	0	0	1	1	0	Jose14	1	1	0	1	1	1	0
Jose15	0	1	1	0	0	1	0	Jose15	1	0	0	1	1	0	1
Jose16	0	0	1	1	0	1	0	Jose16	0	1	1	0	0	1	0
Jose17	0	1	0	1	1	1	0	Jose17	0	1	0	1	0	0	1
Jose18	0	0	0	0	0	0	0	Jose18	1	1	1	1	1	1	1
Jose19	1	1	1	1	1	1	1	Jose19	0	0	0	0	0	0	0
Jose20	0	0	1	0	1	0	1	Jose20	1	1	1	0	1	0	1
na1 o na2	12	5	10	10	10	12	12	na1 o na2	9	15	10	10	10	8	9
n=	20							n=	20						

Los cálculos de los valores totales, sigue la misma explicación dada para el caso del índice de similitud.

2.- Con los totales de cada variable (los n_{a_i}) y sus complementos (los $n_{\bar{a}_i}$), también notados como ($n_{\bar{a}_i}$). Se forma una nueva tabla.

Tabla 3. Un primer acercamiento al cálculo
 $K_{ij} = Card(A_i \cap \bar{A}_j)$

VARIABLES	HIP	OPE	FLA	RAP	OPE	FLA	RAP	HIP	HIA
HIP	12	5	10	10	10	12	12	9	15
HIP	12	5	10	10	10	12	12	9	15
FLA	10	10	8	1	0	0	0	10	10
OPE	5	10	8	0	0	0	0	5	10
HIP	10	10	8	0	0	0	0	10	10
HIA	10	10	8	0	0	0	0	10	10

Interpretaciones:
Hay 9 coincidencias entre HIP y OPE.
Hay 1 coincidencia entre HIP y PUN.
etc.

Complementando con el cálculo respectivo del índice de implicación.

Tabla 4. Un primer acercamiento al cálculo del índice de implicación.

VARIABLES	HIP	OPE	FLA	RAP	HIP	OPE	FLA	RAP	HIP	HIA
HIP	12	5	10	10	10	12	12	9	15	0.82553
HIP	12	5	10	10	10	12	12	9	15	0.82553
FLA	10	10	8	1	0	0	0	10	10	0.84423
OPE	5	10	8	0	0	0	0	5	10	0.87442
HIP	10	10	8	0	0	0	0	10	10	0.84423
HIA	10	10	8	0	0	0	0	10	10	0.84423

3.- "Para ello, se determina las coincidencias (llamadas copresencias) entre A_i y \bar{A}_j ".

Ejemplo: "Hay 9 alumnos que coinciden en sus preferencias entre HIP y OPE", hay 1 alumno que prefiere en su opinión entre FLA y RAP", entre otros".

4.- Según Zamora et al. (2007), "se debe determinar el valor de $p = (n_{a_i} n_{\bar{a}_j}) / (n^2)$, al multiplicar las filas y dividir el resultado entre el producto de los totales".

Ejemplo: $(11 \times 15) / (20^2) = 0.4125$; entre otros.

5.- Con este valor de p, se calculan para los diferentes valores de K_{ij} , bajo el modelo binomial $P(K = K_{ij}) = C_{K_{ij}}^n p^{K_{ij}} (1-p)^{n-K_{ij}}$.

Ejemplo:

$$P(K = 0) = C_0^{20} p^0 (1-p)^{20} = 1 \cdot 1 \cdot (1 - 0.4125)^{20} = 0.3675^{20} = 0.00025997 = 2^{-4}$$

entre otros.

6.- Para el término $\phi(A_i, A_j)$, se resta de la unidad, la suma de los valores calculados en el paso anterior, desde $K_{ij}=0$ hasta $K_{ij}=9$, (en otros casos, hasta el límite de K_{ij}).

Ejemplo: Se ha calculado que $\phi(HIP, OPE) = 1 - P(K \leq 9) = 0.282553$; entre otros.

7.- Este número hallado, se llama la Intensidad Implicativa.

Actividades de refuerzo: Se solicita al lector, recorrer por la hoja Excel descargado y verificar los pasos anteriores en la generación de las nuevas tablas 4, 5 y 6.

Excel screenshot showing the calculation of the intensity of implication for various variable pairs.

n. Valores intensidad implicativa *	
(A_i, A_j)	
HIP	0.80000
OPE	0.16106
FLA	0.14702
HIP	0.80000
OPE	0.16106
FLA	0.14702
HIP	0.80000
OPE	0.16106
FLA	0.14702

De acuerdo a Zamora *et al.* (2007) y los cálculos en *Excel* facilitados, observe que:

Las reglas HIP → PUN y PUN → HIP son admisibles al 95% de confiabilidad (en orden de centésimas, redondear al menor entero).

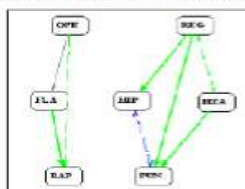
Las reglas FLA → RAP; HEA → REG y REG → HEA son admisibles al 90% de confiabilidad.

Las reglas

OPE → RAP; HEA → PUN; REG → HIP; REG → PUN; RAP → FLA son

admisibles al 85% de confiabilidad, el resto de reglas no son admisibles.

Este análisis se complementa con el siguiente grafo implicativo:



Fuente: Zamora *et al.* (2007).

Actividades propuestas. Después de recorrer la hoja de *Excel* descargado, se solicita al lector, completar los siguientes ejercicios de aplicación:



Evaluación “Para verificar la aproximación a la comprensión docente de la definición de ésta técnica, se solicita al usuario-docente, contestar una encuesta de los procesos explicados versus las actividades planteadas, la misma que se encuentra en el enlace: <https://bit.ly/2EfcMxB> o haciendo clic en el archivo *Excel* en el botón: Encuesta 4”.

Técnica: Cohesión

Simulación en *Excel* para el proceso: Análisis Cohesivo

Definiciones Según (Orús *et al.*, 2009), “la cohesión surge como una medida de la calidad implicativa de la R-regla, el objetivo es descubrir R-reglas (R → R’) con una fuerte relación implicativa entre las componentes de R’ y las de R”, (...), se debe contrastar con el desorden de una experiencia aleatoria, la entropía es bastante conveniente para medir este desorden”.

Consideremos una regla $a_i \rightarrow a_j$ de orden 1, definimos una variable

aleatoria Y como indicadora del evento: $X_{a_i/a_j} \geq n_{a_i/a_j} - z$, con:

$$P(Y=1) = P(X_{a_i/a_j} \geq n_{a_i/a_j} - z) = \varphi(a_i, a_j) \text{ y } P(Y=0) = 1 - \varphi(a_i, a_j)$$

Según (Orús *et al.*, 2009), “la entropía de este experimento es:

$$E = -p \log_2(p) - (1-p) \log_2(1-p) \text{ con } p = \varphi(a_i, a_j)$$

También (Orús *et al.*, 2009) sostiene, “la Cohesión de una R-regla

$a_i \rightarrow a_j$ de grado 1 es:

$$\text{Coh}(a_i, a_j) = \sqrt{1 - E^2}$$
, cuando $p \geq 0.5$; 0 caso contrario.

si $\varphi(a_i, a_j) < 0.5$, la regla $a_i \rightarrow a_j$ es neutral en el sentido de que da lugar al número de contraejemplos esperado bajo independencia”.

De la misma forma (Orús *et al.*, 2009) menciona: “La Cohesión de la clase de variables $R = (a_1, \dots, a_n)$ es la media geométrica de las cohesiones de los pares de variables que la conforman:

$$\text{Coh}(R) = \sqrt[n]{\prod_{i < j} \text{Coh}(a_i, a_j)}$$

Según Zamora *et al.* (2007), “la intensidad de implicación de una clase A sobre una clase B se define como:

$$\Psi(A, B) = \sqrt[n]{\prod_{i=1}^n \varphi(a_i, a_j)}$$

Uso de la herramienta digital en vías Proceso del cálculo en *Excel* del elemento denominado “índice de cohesión”, acompañar este proceso con el documento en *Excel* descargado.

Pasos:
1.- Partimos de la tabla de valores de la intensidad implicativa, calculada anteriormente:

p=FLA \ q=	HIP	OPE	REG	HEA	PUN	RAP	HEA
HIP	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
OPE	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
FLA	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
REG	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
HEA	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
PUN	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000

2.- Hallamos una nueva matriz que refleja los valores de la entropía, asociada a cada valor de $p = \varphi(a_i, a_j)$.

p=FLA \ q=	HIP	OPE	REG	HEA	PUN	RAP	HEA
HIP	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
OPE	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
FLA	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
REG	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
HEA	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
PUN	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000

3.- Complementando con el cálculo respectivo del índice de cohesión.

Clase(A) \ Clase(B)	HIP	OPE	FLA	REG	HEA	RAP	HEA
HIP	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
OPE	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
FLA	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
REG	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
HEA	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
RAP	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
HEA	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000

Ejemplo: Para la pareja (HIP, FLA), el valor de la Intensidad Implicativa es $\varphi(\text{HIP}, \text{FLA}) = 0.15781 < 0.5$, por lo que $\text{Coh}(\text{HIP}, \text{FLA}) = 0$; de la misma forma, para la pareja (OPE, RAP), el valor de la Intensidad Implicativa es $\varphi(\text{OPE}, \text{RAP}) = 0.878423 > 0.5$, por lo que

$$\text{Coh}(\text{OPE}, \text{RAP}) = \sqrt{1 - 0.53387599^2} = 0.845563$$
, entre otros.

Observar que la matriz no es simétrica.

Proceso de cálculo de la Intensidad de Implicación entre clases de variables.

- 1.- Partimos de la tabla 3, de los valores del índice de cohesión al nivel cero de jerarquía, calculada anteriormente.
- 2.- De esta matriz, determinamos el valor máximo de cohesión de las clases.

Tabla 17: Intensidad de implicación entre clases de variables al nivel 1 de la jerarquía.

Nivel 1	(HIP, PUN)	OPE	FLA	HEA	REG	RAP
(HIP, PUN)	0,000000	0,079237	0,024717	0,772094	0,009378	0,000000
OPE	0,000000	0,000000	0,547125	0,000000	0,000000	0,543303
FLA	0,000000	0,024717	0,000000	0,000000	0,029290	0,511852
HEA	0,772094	0,000000	0,000000	0,000000	0,000000	0,000000
REG	0,009378	0,000000	0,000000	0,000000	0,000000	0,000000
RAP	0,000000	0,543303	0,511852	0,000000	0,000000	0,000000
Máx.	0,772094	0,543303	0,547125	0,772094	0,009378	0,543303

3.- Estas clases se unen para conformar el nivel 1. Para hallar los nuevos valores (intensidad de implicación de una regla A sobre una variable B), observe:

Proceso del cálculo de $\Psi(A, B)$: Intensidad de implicación de la regla A sobre la variable B.

A=(HIP, PUN)	Con(A)=	A=(OPE)	Con(A)=
B=(OPE)	0,000000	B=(HIP, PUN)	0,000000
$\Psi(a,b)$	HEA	REG	HEA
HEA	0,247125	REG	0,000000
REG	0,000000	HEA	0,247125

Las demás celdas se procede de forma similar, sus resultados se muestran en la siguiente tabla:

Tabla 18: Intensidad de implicación $\Psi(A, B)$ entre clases de variables al nivel 1 de la jerarquía.

Nivel 1	(HIP, PUN)	OPE	FLA	HEA	REG	RAP
(HIP, PUN)	0,000000	0,079237	0,024717	0,772094	0,009378	0,000000
OPE	0,000000	0,000000	0,547125	0,000000	0,000000	0,543303
FLA	0,024717	0,000000	0,000000	0,000000	0,029290	0,511852
HEA	0,772094	0,000000	0,000000	0,000000	0,000000	0,000000
REG	0,009378	0,000000	0,000000	0,000000	0,000000	0,000000
RAP	0,000000	0,543303	0,511852	0,000000	0,000000	0,000000

4.- De la matriz conformada J1, hallamos el valor máximo.

Tabla 11: Intensidad de implicación $\Psi(A, B)$ entre clases de variables al nivel 2 de la jerarquía.

Nivel 2	(HIP, PUN)	OPE	FLA	HEA	REG	RAP
(HIP, PUN)	0,000000	0,079237	0,024717	0,772094	0,009378	0,000000
OPE	0,000000	0,000000	0,547125	0,000000	0,000000	0,543303
FLA	0,024717	0,000000	0,000000	0,000000	0,029290	0,511852
HEA	0,772094	0,000000	0,000000	0,000000	0,000000	0,000000
REG	0,009378	0,000000	0,000000	0,000000	0,000000	0,000000
RAP	0,000000	0,543303	0,511852	0,000000	0,000000	0,000000
Máx.	0,772094	0,543303	0,547125	0,772094	0,009378	0,543303

5.- Estas clases se unen para conformar el nivel 2. Para hallar los nuevos valores (intensidad de implicación de una regla A sobre una variable B), observe:

Proceso del cálculo de $\Psi(A, B)$: Intensidad de implicación de la regla A sobre la variable B.

A=(FLA, RAP)	Con(A)=	A=(HIP, PUN)	Con(A)=
B=(HIP, PUN)	0,000000	B=(OPE)	0,000000
$\Psi(a,b)$	HEA	REG	HEA
HEA	0,024717	REG	0,000000
REG	0,000000	HEA	0,024717

Las demás celdas se procede de forma similar, sus resultados se muestran en la siguiente tabla:

Tabla 12: Intensidad de implicación $\Psi(A, B)$ entre clases de variables al nivel 2 de la jerarquía.

Nivel 2	(HIP, PUN)	OPE	FLA	HEA	REG
(HIP, PUN)	0,000000	0,079237	0,024717	0,772094	0,009378
OPE	0,000000	0,000000	0,547125	0,000000	0,000000
FLA	0,024717	0,000000	0,000000	0,000000	0,029290
HEA	0,772094	0,000000	0,000000	0,000000	0,000000
REG	0,009378	0,000000	0,000000	0,000000	0,000000
RAP	0,000000	0,543303	0,511852	0,000000	0,000000

6.- De la matriz conformada J2, hallamos el valor máximo.

Tabla 13: Intensidad de implicación $\Psi(A, B)$ entre clases de variables al nivel 3 de la jerarquía.

Nivel 3	(HIP, PUN)	OPE	FLA	(HEA, REG)
(HIP, PUN)	0,000000	0,079237	0,024717	0,772094
OPE	0,000000	0,000000	0,547125	0,000000
FLA	0,024717	0,000000	0,000000	0,029290
(HEA, REG)	0,772094	0,000000	0,000000	0,000000
RAP	0,000000	0,543303	0,511852	0,000000
Máx.	0,772094	0,543303	0,547125	0,772094

2.- Llenar los valores faltantes en la sección del análisis:

Proceso del cálculo de $\Psi(A, B)$: Intensidad de implicación de la regla A sobre la variable B.

A=(RAP, FLA)	Con(A)=	A=(RAP, FLA)	Con(A)=
B=(OPE)	ejercicio	B=(OPE)	ejercicio
$\Psi(a,b)$	ejercicio	ejercicio	ejercicio
RAP	ejercicio	ejercicio	ejercicio
FLA	ejercicio	ejercicio	ejercicio

3.- Observar la conformación de la tabla J4.

Las demás celdas se procede de forma similar, sus resultados se muestran en la siguiente tabla:

Tabla 14: Intensidad de implicación $\Psi(A, B)$ entre clases de variables al nivel 3 de la jerarquía.

Nivel 3	(HIP, PUN)	OPE	(HEA, REG)
(HIP, PUN)	0,000000	0,079237	0,772094
OPE	0,000000	0,000000	0,000000
FLA	0,024717	0,000000	0,029290
(HEA, REG)	0,772094	0,000000	0,000000
(RAP, FLA)	0,000000	0,543303	0,511852

4.- Comparar los valores calculados y la conformación del árbol jerárquico cohesivo indicado en la parte anterior.

Evaluación

“Para verificar la aproximación a la comprensión docente de la definición de esta técnica, se solicita al usuario-docente, contestar una encuesta de los procesos explicados versus las actividades planteadas, la misma que se encuentra en el enlace: <https://bit.ly/33JJxa3> o haciendo clic en el archivo Excel en el botón: Encuesta 5”.

Nivel 3	(HIP, PUN)	OPE	FLA	(HEA, REG)
(HIP, PUN)	0,000000	0,079237	0,024717	0,772094
OPE	0,000000	0,000000	0,547125	0,000000
FLA	0,024717	0,000000	0,000000	0,029290
(HEA, REG)	0,772094	0,000000	0,000000	0,000000
(RAP, FLA)	0,000000	0,543303	0,511852	0,000000
Máx.	0,772094	0,543303	0,547125	0,772094

7.- Estas clases se unen para conformar el nivel 3. Para hallar los nuevos valores (intensidad de implicación de una regla A sobre la variable B), observe:

Proceso del cálculo de $\Psi(A, B)$: Intensidad de implicación de la regla A sobre la variable B.

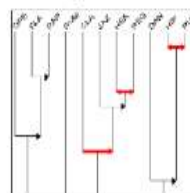
A=(HEA, REG)	Con(A)=	A=(HEA, REG)	Con(A)=
B=(HIP, PUN)	0,000000	B=(OPE)	0,000000
$\Psi(a,b)$	HEA	REG	HEA
HEA	0,000000	REG	0,000000
REG	0,000000	HEA	0,000000

Las demás celdas se procede de forma similar, sus resultados se muestran en la siguiente tabla:

Tabla 25: Intensidad de implicación $\Psi(A, B)$ entre clases de variables al nivel 3 de la jerarquía.

Nivel 3	(HIP, PUN)	OPE	FLA	(HEA, REG)
(HIP, PUN)	0,000000	0,079237	0,024717	0,772094
OPE	0,000000	0,000000	0,547125	0,000000
FLA	0,024717	0,000000	0,000000	0,029290
(HEA, REG)	0,772094	0,000000	0,000000	0,000000
(RAP, FLA)	0,000000	0,543303	0,511852	0,000000
Máx.	0,772094	0,543303	0,547125	0,772094

De acuerdo a Zamora *et al.* (2007), “el proceso continúa hasta que no existan más clases que no puedan ser unidas, el árbol jerárquico cohesivo que se obtiene es”:



Fuente: Zamora *et al.* (2007).

Actividades de Se solicita al lector, recorrer por la hoja Excel descargado y verificar refuerzo. los pasos anteriores en la generación de las tablas: J0, J1, J2, J3 y J4.

Actividades Después de recorrer la hoja de Excel descargado, se solicita al lector, propuestas. completar los siguientes ejercicios de aplicación:

- 1.- Observar el cálculo del valor máximo en la tabla J3.

Parte 2

Simulación para el cálculo de los Nodos Significativos en un Árbol Cohesivo Jerárquico

Definiciones: Según (Orús et al., 2009), "la determinación de los índices de intensidad implicativa entre clases de variables nos permite construir un árbol jerárquico, al igual que se hizo en el caso del árbol de similitud, los nodos significativos de este árbol jerárquico son los nodos correspondientes a una clasificación compatible lo mejor posible con los valores y la calidad de los valores de implicación y de cohesión". Para el cálculo de los niveles y nodos significativos se sigue el proceso descrito en la tabla 5.

Uso de la herramienta: Para este respecto, se solicita al lector que nos acompañe con el documento de Excel descargado, se va a trabajar con las variables digital en vías de aproximación a la comprensión de la técnica

A* = {JAZ, HEA, CLA, REG}

1.- Cálculo de los valores del nivel 1 de jerarquía:

1.1. Identificamos la Matriz de similitud, con todos sus elementos.

Tabla 10.- Intensidad implicativa entre clases de var

Nivel 0	JAZ	HEA	CLA	REG
JAZ		0,908740	0,671648	0,908740
HEA	0,908740		0,671648	0,908740
CLA	0,733052	0,733052		0,733052
REG	0,908740	0,908740	0,671648	

Identificamos en esta matriz de Intensidad Implicativa, los diferentes valores similares y las variables a las que corresponden.

1.2. Clasificamos estos valores, desde el menor hasta el mayor, estos son los valores del *preorden inicial*.

Valores de $G_5(\Omega) = \{ 0,671648 \quad 0,733052 \quad 0,90874 \}$

Zamora et al. (2007), lo clasificaría como:

$$\begin{bmatrix} Coh(JAZ, CLA) \\ Coh(HEA, CLA) \\ Coh(REG, CLA) \end{bmatrix} < \begin{bmatrix} Coh(CLA, HEA) \\ Coh(CLA, JAZ) \\ Coh(CLA, REG) \end{bmatrix} < \begin{bmatrix} Coh(JAZ, REG) \\ Coh(HEA, REG) \\ Coh(REG, JAZ) \\ Coh(HEA, JAZ) \\ Coh(JAZ, HEA) \end{bmatrix}$$

Fuente: Zamora et al. (2007).

1.3. Realizamos el conteo de la cantidad de cada clase, *conjunto de pares separados al nivel k*.

$s_k = Card[S\Omega_k] = \{ 3 \quad 3 \quad 6 \}$

1.4. Identificamos las variables en análisis:

$A^* = \{ JAZ, HEA, CLA, REG \}$

1.5. Identificamos el número de variables en análisis.

$t = Card(A^*) = 4$

1.6. Identificamos el número de elementos del *preorden inicial*.

$m = Card[\Omega] = 12$

1.7. Identificamos el número de elementos del *preorden inducido por S*.

$m = Card[\Omega] = 12$				suma
	3	3	15	21

$Card(G_2(\Omega)) = \frac{-(12^2 + 12 \cdot 2) / 2}{21} + 24$ entos en $Card(G_2(\Omega)) = 45$

Es decir, por medio de las celdas de Excel, se desarrolla la fórmula:

$$Card(G_2(\Omega)) = \frac{m(m-1)}{2} - \frac{4 \cdot Card(P_2) \cdot (Card(P_2) + 1)}{2}$$

1.8. Identificamos en la matriz de intensidad implicativa, las primeras parejas de variables con valores similares.

Nivel 0	JAZ	HEA	CLA	REG	2
JAZ		0,908740	0,671648	0,908740	
HEA	0,908740		0,671648	0,908740	
CLA	0,733052	0,733052		0,733052	
REG	0,908740	0,908740	0,671648		10

1.9. Proceder a efectuar y comprender los siguientes cálculos:

Nivel 1 de la jerarquía:

$r_1 = Card(R\Omega_1) = 2$ // La primera pareja $R\Omega_1 = \{(JAZ, HEA); (HEA, JAZ)\}$

$s_1 = Card(S\Omega_1) = 10$ // Lo que falta para $\Omega = 12$.

$t_1 = \{12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1\}$ // $t_k = U / n_k \cdot R\Omega_k$

$f_1^* = 4$ // $f_k^* = Card\{a_j \in I_{k-1} \mid s(a_j) < t_k \cdot s(a_j)\}$

$f_1' = 4$ // $f_k' = Card\{a_j \in I_{k-1} \mid s(a_j) = t_k \cdot s(a_j)\}$

$Card(G_{1,1}(\Omega) \cap [S\Omega_1, R\Omega_1]) = 12$ // $Card(G_{k,1}(\Omega) \cap [S\Omega_k, R\Omega_k]) = \frac{s_k}{m} \cdot \frac{t_k(t_k+1)}{2} - \frac{f_k' \cdot f_k}{2}$

Obsérvese que es la aplicación de la fórmula propuesta, a la forma de Excel.

Nivel 1 de la jerarquía:

$r_1 = Card(R\Omega_1) = 0$ // La primera pareja $R\Omega_1 = \{(JAZ, HEA); (HEA, JAZ)\}$

$s_1 = Card(S\Omega_1) = 10$ // Lo que falta para $\Omega = 12$.

$t_1 = \{12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1\}$ // $t_k = U / n_k \cdot R\Omega_k$

$f_1^* = 4$ // $f_k^* = Card\{a_j \in I_{k-1} \mid s(a_j) < t_k \cdot s(a_j)\}$

$f_1' = 0$ // $f_k' = Card\{a_j \in I_{k-1} \mid s(a_j) = t_k \cdot s(a_j)\}$

$Card(G_{1,1}(\Omega) \cap [S\Omega_1, R\Omega_1]) = 10$ // $Card(G_{k,1}(\Omega) \cap [S\Omega_k, R\Omega_k]) = \frac{s_k}{m} \cdot \frac{t_k(t_k+1)}{2} - \frac{f_k' \cdot f_k}{2}$

2.- Comprendiendo el cálculo de los valores del nivel 2 de jerarquía:

2.1. Identificamos en la matriz de intensidad implicativa, las primeras parejas de variables con valores similares.

Nivel 0	JAZ	HEA	CLA	REG	4
JAZ		0,908740	0,671648	0,908740	
HEA	0,908740		0,671648	0,908740	
CLA	0,733052	0,733052		0,733052	
REG	0,908740	0,908740	0,671648		8

2.2. Proceder a efectuar y comprender los siguientes cálculos:

Nivel 2 de la jerarquía:

$r_2 = Card(R\Omega_2) = 4$ // La primera pareja $R\Omega_2 = \{(JAZ, HEA); (HEA, JAZ); (REG, JAZ); (REG, HEA)\}$

$s_2 = Card(S\Omega_2) = 8$ // Lo que falta para $\Omega = 12$.

$t_2 = \{12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1\}$ // $t_k = U / n_k \cdot R\Omega_k$

$f_2^* = 4$ // $f_k^* = Card\{a_j \in I_{k-1} \mid s(a_j) < t_k \cdot s(a_j)\}$

$f_2' = 0$ // $f_k' = Card\{a_j \in I_{k-1} \mid s(a_j) = t_k \cdot s(a_j)\}$

$Card(G_{2,1}(\Omega) \cap [S\Omega_2, R\Omega_2]) = 24$ // $Card(G_{k,1}(\Omega) \cap [S\Omega_k, R\Omega_k]) = \frac{s_k}{m} \cdot \frac{t_k(t_k+1)}{2} - \frac{f_k' \cdot f_k}{2}$

Obsérvese que es la aplicación de la fórmula propuesta, a la forma de Excel.

Nivel 2 de la jerarquía:

$r_2 = Card(R\Omega_2) = 0$ // La primera pareja $R\Omega_2 = \{(JAZ, HEA); (HEA, JAZ); (REG, JAZ); (REG, HEA)\}$

$s_2 = Card(S\Omega_2) = 8$ // Lo que falta para $\Omega = 12$.

$t_2 = \{12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1\}$ // $t_k = U / n_k \cdot R\Omega_k$

$f_2^* = 4$ // $f_k^* = Card\{a_j \in I_{k-1} \mid s(a_j) < t_k \cdot s(a_j)\}$

$f_2' = 0$ // $f_k' = Card\{a_j \in I_{k-1} \mid s(a_j) = t_k \cdot s(a_j)\}$

$Card(G_{2,1}(\Omega) \cap [S\Omega_2, R\Omega_2]) = 24$ // $Card(G_{k,1}(\Omega) \cap [S\Omega_k, R\Omega_k]) = \frac{s_k}{m} \cdot \frac{t_k(t_k+1)}{2} - \frac{f_k' \cdot f_k}{2}$

3.- Comprendiendo el cálculo de los valores del nivel 3 de jerarquía:

3.1. Identificamos en la matriz de intensidad implicativa, las primeras parejas de variables con valores similares.

Nivel 0	JAZ	HEA	CLA	REG	7
JAZ		0,908740	0,671648	0,908740	
HEA	0,908740		0,671648	0,908740	
CLA	0,733052	0,733052		0,733052	
REG	0,908740	0,908740	0,671648		5

3.2. Proceder a efectuar y comprender los siguientes cálculos:

Nivel 3 de la jerarquía:

$r_3 = Card(R\Omega_3) = 7$ // La primera pareja $R\Omega_3 = \{(JAZ, HEA); (HEA, JAZ); (REG, JAZ); (REG, HEA); (CLA, JAZ); (CLA, HEA)\}$

$s_3 = Card(S\Omega_3) = 5$ // Lo que falta para $\Omega = 12$.

$t_3 = \{12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1\}$ // $t_k = U / n_k \cdot R\Omega_k$

$f_3^* = 3$ // $f_k^* = Card\{a_j \in I_{k-1} \mid s(a_j) < t_k \cdot s(a_j)\}$

$f_3' = 0$ // $f_k' = Card\{a_j \in I_{k-1} \mid s(a_j) = t_k \cdot s(a_j)\}$

$Card(G_{3,1}(\Omega) \cap [S\Omega_3, R\Omega_3]) = 21$ // $Card(G_{k,1}(\Omega) \cap [S\Omega_k, R\Omega_k]) = \frac{s_k}{m} \cdot \frac{t_k(t_k+1)}{2} - \frac{f_k' \cdot f_k}{2}$

4.- Calculando el nivel significativo.

4.1. Los valores anteriores, se resume en la siguiente tabla:

Nivel	s_k	r_k	$Card(G_{k,1}(\Omega) \cap [S\Omega_k, R\Omega_k])$	$S(\Omega, k)$	$V(\Omega, k)$
1	10	2	12	0,42967	0,429669
2	8	4	24	1,35873	0,920064
3	5	7	21	0,56884	-0,790334

Valores del índice centrado y de la función $V(\Omega, k)$.

4.2. Determinamos los valores de S y V:

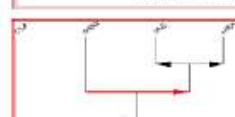
$S(\Omega, k) = \frac{Card(G_{k,1}(\Omega) \cap [S\Omega_k, R\Omega_k]) - 0,5s_k r_k}{\sqrt{s_k r_k (s_k + r_k + 1)}}$ y $V(\Omega, k)$ definido:

Comprender por medio de la hoja, las respectivas fórmulas.

Nivel	s_k	r_k	$Card(G_{k,1}(\Omega) \cap [S\Omega_k, R\Omega_k])$	$S(\Omega, k)$	$V(\Omega, k)$
1	10	2	$= (12 - 0,5 * 10 * 2) / \sqrt{(10 * 2 * (10 + 2 + 1)) / (12)}$	0,42967	0,429669
2	8	4	24	1,35873	0,920064
3	5	7	21	0,56884	-0,790334

Valores del índice centrado y de la función $V(\Omega, k)$.

4.3. Graficamos los valores y comparamos con el Árbol Cohesivo Jerárquico:



	Según los cálculos anteriores, el nodo 1 y el nodo 2 poseen valores positivos (en este caso se observa el máximo local, es decir, el nodo 2), la misma que en el Árbol Cohesivo Jerárquico aparece con el color rojo.
Actividades propuestas.	Se solicita al lector revisar nuevamente la conformación de los niveles: 1, 2 y 3 de jerarquía.
Evaluación	“Para verificar la aproximación a la comprensión docente de la definición de ésta técnica, se solicita al usuario-docente, contestar una encuesta de los procesos explicados versus las actividades planteadas, la misma que se encuentra en el enlace: https://bit.ly/3kvvN9J o haciendo clic en el archivo <i>Excel</i> en el botón: Encuesta 6”.

C. Paquetes a ser instaladas para el funcionamiento de las simulaciones en R

<u>NOMBRE DEL PAQUETE.</u>	<u>REFERENCIA</u>	<u>VERSION</u>	<u>INSTRUCCIONES DE CARGA</u> “Ejecutar la instrucción, en la consola de RStudio, R y RStudio Cloud”	<u>VERIFICACIÓN Y ACTIVACIÓN.</u> “Ejecutar el comando, en la consola de RStudio R y RStudio Cloud”
Stringr	https://cran.r-project.org/web/packages/stringr/stringr.pdf	1.4.0	install.packages("stringr");	library("stringr");
tcltk2	https://cran.r-project.org/web/packages/tcltk2/tcltk2.pdf	1.2-11	install.packages("tcltk2");	library("tcltk2");
Graph	https://stackoverflow.com/questions/47073836/r-how-to-install-package-graph	1.58.0	install.packages("graph");	library("graph");
BiocGenerics	https://bioconductor.org/packages/release/bioc/html/BiocGenerics.html	0.26.0	install.packages("BiocGenerics");	library("BiocGenerics");
Parallel	https://stat.ethz.ch/R-manual/R-devel/library/parallel/doc/parallel.pdf	3.5.1	install.packages("parallel");	library("parallel");
Rcpp	https://cran.r-project.org/web/packages/Rcpp/Rcpp.pdf	1.0.4.6	install.packages("Rcpp");	library("Rcpp");
microbenchmark	(Mersmann, Beileites, HUrlin & Friedman, 2014)	1.4-6	install.packages("microbenchmark")	library("microbenchmark");
rchic	https://members.femto-st.fr/raphael-couturier/en/rchic	0.27	Descargar el CRAN desde: http://members.femto-st.fr/raphael-couturier/en/rchic Realizar la carga manual. (Ver Apéndice A.)	library("rchic"); rchic()
ggplot2	(Wichham & Chang, 2016)	3.1.0	install.packages("ggplot2");	library("ggplot2");
factoextra	(Kassambara & Mundt, 2016)	1.0.5	install.packages("factoextra");	library("factoextra");
cluster	(Maechler et al., 2017)	2.0.7-1	install.packages("cluster");	library("cluster");
fastcluster	(Müller, 2017)	1.1.25	install.packages("fastcluster");	library("fastcluster");
caTools	https://cran.r-project.org/web/packages/caTools/caTools.pdf	1.18.0	install.packages("caTools");	library("caTools");
CluMix	https://github.com/cran/CluMix/blob/master/DESCRIPTION (Hommel, Edelmann, & Kopp-Schneider, 2017)	2.3.1	install.packages("CluMix");	library("CluMix");
dplyr	https://cran.r-project.org/web/packages/dplyr/dplyr.pdf	0.8.5	install.packages("dplyr");	library("dplyr");
replayr	https://mran.microsoft.com/snapshot/2017-06-09/web/packages/replayr/replayr.pdf	0.3.01	install.packages("replayr");	library("replayr");
Arules	https://cran.r-project.org/web/packages/arules/arules.pdf (Hahsler et al., 2018)	1.6-5	install.packages("arules");	library("arules");
arulesViz	https://cran.r-project.org/web/packages/arulesViz/arulesViz.pdf	1.3-3	install.packages("arulesViz");	library("arulesViz");
Vcd	https://cran.r-project.org/web/packages/vcd/vcd.pdf	1.4-7	install.packages("vcd");	library("vcd");
Readr	https://cran.r-project.org/web/packages/readr/readr.pdf	1.3.1	install.packages("readr");	library("readr");
Rgraphviz	https://www.bioconductor.org/packages/release/bioc/html/Rgraphviz.html	2.52.0	Pegar en la consola: ## try http:// if https:// URLs are not supported source("https://bioconductor.org/biocLite.R") biocLite("Rgraphviz")	library("Rgraphviz");

Nota. A manera de ejemplo, estos archivos, se encuentran disponibles en: <https://bit.ly/3g7DzYD>

D. Simulaciones digitales en R para las técnicas del ASI

Técnica: Árbol de Similaridad en R

Simulación en R: Técnica → Árbol de Similaridad del ASI (clustering)

Definiciones	<p>La parte conceptual: Ver las definiciones en las tablas 3, 4 y 5.</p> <p>¿Cuál es el nombre de la técnica en R?: <i>CallSimilarityTree</i></p> <p>¿Cómo buscar la técnica en R?: <i>help(callSimilarityTree)</i></p>																																				
Uso de la técnica a través del Software R	<p>Un ejemplo de la forma del uso educativo de la técnica: En Matemática, con la finalidad de comprender: "Si la proporcionalidad es una herramienta óptima en la resolución de un problema de ecuaciones a un grupo de estudiantes de entre 13 y 14 años que estén asistiendo de forma regular a una Unidad Educativa, se puede plantear un ejercicio como: El amo contrató a un trabajador para un año y prometió pagarle 12 rublos y un caftán. Pero éste, después de haber trabajado 7 meses quiso marcharse. Una vez hechas las cuentas recibió el caftán y 5 rublos. ¿Cuánto cuesta el caftán?", es un problema de tipo aritmético-algebraico denominado "del caftán", de la tradición rusa (Olekhnik, Nesterenko, Potapov, & Bernardo, 1994), citado en (Lasa, Sáenz de Cabezón, Wilhelmi, & Abaurrea, 2019), las soluciones entregadas por los estudiantes, lo clasifica como:</p> <p>"V1.- Resuelve por un método de procedimiento de proporcionalidad simple</p> <p>V2.- Resuelve por un sistema de ecuaciones, método de reducción</p> <p>V3.- Resuelve por una ecuación de una incógnita, método de igualación</p> <p>V4.- Resuelve por un procedimiento de proporcionalidad compuesta</p> <p>V5.- Resuelve por falsa proporción: el caftán vale 2 rublos</p> <p>V6.- Revisa el resultado y lo comprueba por más de un método</p> <p>V7.- La solución es correcta (precio del caftán es de 4,8 rublos)</p> <p>V8.- Muestra un nivel algebraico consolidado (nivel 3)"</p> <p>Este análisis contempla la siguiente matriz de similaridad.</p> <table border="1"> <thead> <tr> <th></th> <th>V1</th> <th>V3</th> <th>V6</th> <th>V7</th> <th>V8</th> </tr> </thead> <tbody> <tr> <th>V1</th> <td>1</td> <td>0,98</td> <td>0,97</td> <td>1</td> <td>0,9</td> </tr> <tr> <th>V3</th> <td>0,98</td> <td>1</td> <td>0,92</td> <td>0,93</td> <td>0,99</td> </tr> <tr> <th>V6</th> <td>0,97</td> <td>0,92</td> <td>1</td> <td>0,95</td> <td>0,99</td> </tr> <tr> <th>V7</th> <td>1,00</td> <td>0,93</td> <td>0,95</td> <td>1</td> <td>0,97</td> </tr> <tr> <th>V8</th> <td>0,99</td> <td>0,99</td> <td>0,97</td> <td>0,97</td> <td>1</td> </tr> </tbody> </table> <p>Fuente: (Lasa et al., 2019).</p>		V1	V3	V6	V7	V8	V1	1	0,98	0,97	1	0,9	V3	0,98	1	0,92	0,93	0,99	V6	0,97	0,92	1	0,95	0,99	V7	1,00	0,93	0,95	1	0,97	V8	0,99	0,99	0,97	0,97	1
	V1	V3	V6	V7	V8																																
V1	1	0,98	0,97	1	0,9																																
V3	0,98	1	0,92	0,93	0,99																																
V6	0,97	0,92	1	0,95	0,99																																
V7	1,00	0,93	0,95	1	0,97																																
V8	0,99	0,99	0,97	0,97	1																																

(Lasa et al., 2019) extrae las siguientes observaciones:

"La variable 1 guarda relación con las variables 3, 6 y 7. Es decir, el uso de procedimiento de proporcionalidad simple aparece ligado al uso del método de igualación (ambos procedimientos se emplean de manera complementaria) y a la obtención de la solución correcta. La variable 3, por su parte, guarda relación, además de con la variable 1, con las variables 6, 7 y 8. En este sentido, los estudiantes que emplean el método de igualación emplean técnicas de proporcionalidad para plantear la ecuación a resolver, y obtienen la solución correcta, mediante una propuesta que emplea un nivel de lenguaje algebraico consolidado (nivel 3)".

Para este tipo de análisis se puede proceder como:

1.- Debemos disponer de una base de datos (en un sitio localizable) con el registro de las variables en formato Binario y en la extensión *.csv, así:

	v1	v2	v3	v4	v5	v6
Grup 1	0	0	0	0	1	0
Grup 2	0	0	0	0	1	0
Grup 3	0	0	0	1	0	0
Grup 4	1	0	0	0	0	0
Grup 5	0	1	1	1	0	0
Grup 6	0	0	0	0	1	1
Grup 7	1	0	1	0	0	0

Nota: V1 hasta V8, son las variables en análisis; Grup 1 hasta Grup n, pueden ser los estudiantes; los datos son de tipo binario, es decir datos de la forma (si/no, verdadero/falso, admite/no admite, con propiedad/sin propiedad, aprueba/desaprueba, admitido/no admitido, promovido/no promovido, motivado/desmotivado) convertidas a 1 y 0 respectivamente.

- En RStudio, activamos Rchic, tecleando en la "consola" los comandos: `library(rchic)` y `rchic()`.
- En la ventana Rchic, seleccionamos la opción: "Similarity tree"



4.- Ubicamos la dirección de la base de datos creada y observamos los resultados:

4.1.- En la ventana "consola" de RStudio:

Ejemplo 1

```
CallSimilarityTree(1, 2, 3, 4, 5, 6, 7, 8)
Classification 1: (1,2) (3,4) (5,6,7,8)
Classification 2: (1,2) (3,4) (5,6,7,8)
Classification 3: (1,2) (3,4) (5,6,7,8)
Classification 4: (1,2) (3,4) (5,6,7,8)
Classification 5: (1,2) (3,4) (5,6,7,8)
Classification 6: (1,2) (3,4) (5,6,7,8)
Classification 7: (1,2) (3,4) (5,6,7,8)
Classification 8: (1,2) (3,4) (5,6,7,8)
Classification 9: (1,2) (3,4) (5,6,7,8)
Classification 10: (1,2) (3,4) (5,6,7,8)
Classification 11: (1,2) (3,4) (5,6,7,8)
Classification 12: (1,2) (3,4) (5,6,7,8)
Classification 13: (1,2) (3,4) (5,6,7,8)
Classification 14: (1,2) (3,4) (5,6,7,8)
Classification 15: (1,2) (3,4) (5,6,7,8)
Classification 16: (1,2) (3,4) (5,6,7,8)
Classification 17: (1,2) (3,4) (5,6,7,8)
Classification 18: (1,2) (3,4) (5,6,7,8)
Classification 19: (1,2) (3,4) (5,6,7,8)
Classification 20: (1,2) (3,4) (5,6,7,8)
Classification 21: (1,2) (3,4) (5,6,7,8)
Classification 22: (1,2) (3,4) (5,6,7,8)
Classification 23: (1,2) (3,4) (5,6,7,8)
Classification 24: (1,2) (3,4) (5,6,7,8)
Classification 25: (1,2) (3,4) (5,6,7,8)
Classification 26: (1,2) (3,4) (5,6,7,8)
Classification 27: (1,2) (3,4) (5,6,7,8)
Classification 28: (1,2) (3,4) (5,6,7,8)
Classification 29: (1,2) (3,4) (5,6,7,8)
Classification 30: (1,2) (3,4) (5,6,7,8)
Classification 31: (1,2) (3,4) (5,6,7,8)
Classification 32: (1,2) (3,4) (5,6,7,8)
Classification 33: (1,2) (3,4) (5,6,7,8)
Classification 34: (1,2) (3,4) (5,6,7,8)
Classification 35: (1,2) (3,4) (5,6,7,8)
Classification 36: (1,2) (3,4) (5,6,7,8)
Classification 37: (1,2) (3,4) (5,6,7,8)
Classification 38: (1,2) (3,4) (5,6,7,8)
Classification 39: (1,2) (3,4) (5,6,7,8)
Classification 40: (1,2) (3,4) (5,6,7,8)
Classification 41: (1,2) (3,4) (5,6,7,8)
Classification 42: (1,2) (3,4) (5,6,7,8)
Classification 43: (1,2) (3,4) (5,6,7,8)
Classification 44: (1,2) (3,4) (5,6,7,8)
Classification 45: (1,2) (3,4) (5,6,7,8)
Classification 46: (1,2) (3,4) (5,6,7,8)
Classification 47: (1,2) (3,4) (5,6,7,8)
Classification 48: (1,2) (3,4) (5,6,7,8)
Classification 49: (1,2) (3,4) (5,6,7,8)
Classification 50: (1,2) (3,4) (5,6,7,8)
Classification 51: (1,2) (3,4) (5,6,7,8)
Classification 52: (1,2) (3,4) (5,6,7,8)
Classification 53: (1,2) (3,4) (5,6,7,8)
Classification 54: (1,2) (3,4) (5,6,7,8)
Classification 55: (1,2) (3,4) (5,6,7,8)
Classification 56: (1,2) (3,4) (5,6,7,8)
Classification 57: (1,2) (3,4) (5,6,7,8)
Classification 58: (1,2) (3,4) (5,6,7,8)
Classification 59: (1,2) (3,4) (5,6,7,8)
Classification 60: (1,2) (3,4) (5,6,7,8)
Classification 61: (1,2) (3,4) (5,6,7,8)
Classification 62: (1,2) (3,4) (5,6,7,8)
Classification 63: (1,2) (3,4) (5,6,7,8)
Classification 64: (1,2) (3,4) (5,6,7,8)
Classification 65: (1,2) (3,4) (5,6,7,8)
Classification 66: (1,2) (3,4) (5,6,7,8)
Classification 67: (1,2) (3,4) (5,6,7,8)
Classification 68: (1,2) (3,4) (5,6,7,8)
Classification 69: (1,2) (3,4) (5,6,7,8)
Classification 70: (1,2) (3,4) (5,6,7,8)
Classification 71: (1,2) (3,4) (5,6,7,8)
Classification 72: (1,2) (3,4) (5,6,7,8)
Classification 73: (1,2) (3,4) (5,6,7,8)
Classification 74: (1,2) (3,4) (5,6,7,8)
Classification 75: (1,2) (3,4) (5,6,7,8)
Classification 76: (1,2) (3,4) (5,6,7,8)
Classification 77: (1,2) (3,4) (5,6,7,8)
Classification 78: (1,2) (3,4) (5,6,7,8)
Classification 79: (1,2) (3,4) (5,6,7,8)
Classification 80: (1,2) (3,4) (5,6,7,8)
Classification 81: (1,2) (3,4) (5,6,7,8)
Classification 82: (1,2) (3,4) (5,6,7,8)
Classification 83: (1,2) (3,4) (5,6,7,8)
Classification 84: (1,2) (3,4) (5,6,7,8)
Classification 85: (1,2) (3,4) (5,6,7,8)
Classification 86: (1,2) (3,4) (5,6,7,8)
Classification 87: (1,2) (3,4) (5,6,7,8)
Classification 88: (1,2) (3,4) (5,6,7,8)
Classification 89: (1,2) (3,4) (5,6,7,8)
Classification 90: (1,2) (3,4) (5,6,7,8)
Classification 91: (1,2) (3,4) (5,6,7,8)
Classification 92: (1,2) (3,4) (5,6,7,8)
Classification 93: (1,2) (3,4) (5,6,7,8)
Classification 94: (1,2) (3,4) (5,6,7,8)
Classification 95: (1,2) (3,4) (5,6,7,8)
Classification 96: (1,2) (3,4) (5,6,7,8)
Classification 97: (1,2) (3,4) (5,6,7,8)
Classification 98: (1,2) (3,4) (5,6,7,8)
Classification 99: (1,2) (3,4) (5,6,7,8)
Classification 100: (1,2) (3,4) (5,6,7,8)
Classification 101: (1,2) (3,4) (5,6,7,8)
Classification 102: (1,2) (3,4) (5,6,7,8)
Classification 103: (1,2) (3,4) (5,6,7,8)
Classification 104: (1,2) (3,4) (5,6,7,8)
Classification 105: (1,2) (3,4) (5,6,7,8)
Classification 106: (1,2) (3,4) (5,6,7,8)
Classification 107: (1,2) (3,4) (5,6,7,8)
Classification 108: (1,2) (3,4) (5,6,7,8)
Classification 109: (1,2) (3,4) (5,6,7,8)
Classification 110: (1,2) (3,4) (5,6,7,8)
Classification 111: (1,2) (3,4) (5,6,7,8)
Classification 112: (1,2) (3,4) (5,6,7,8)
Classification 113: (1,2) (3,4) (5,6,7,8)
Classification 114: (1,2) (3,4) (5,6,7,8)
Classification 115: (1,2) (3,4) (5,6,7,8)
Classification 116: (1,2) (3,4) (5,6,7,8)
Classification 117: (1,2) (3,4) (5,6,7,8)
Classification 118: (1,2) (3,4) (5,6,7,8)
Classification 119: (1,2) (3,4) (5,6,7,8)
Classification 120: (1,2) (3,4) (5,6,7,8)
Classification 121: (1,2) (3,4) (5,6,7,8)
Classification 122: (1,2) (3,4) (5,6,7,8)
Classification 123: (1,2) (3,4) (5,6,7,8)
Classification 124: (1,2) (3,4) (5,6,7,8)
Classification 125: (1,2) (3,4) (5,6,7,8)
Classification 126: (1,2) (3,4) (5,6,7,8)
Classification 127: (1,2) (3,4) (5,6,7,8)
Classification 128: (1,2) (3,4) (5,6,7,8)
Classification 129: (1,2) (3,4) (5,6,7,8)
Classification 130: (1,2) (3,4) (5,6,7,8)
Classification 131: (1,2) (3,4) (5,6,7,8)
Classification 132: (1,2) (3,4) (5,6,7,8)
Classification 133: (1,2) (3,4) (5,6,7,8)
Classification 134: (1,2) (3,4) (5,6,7,8)
Classification 135: (1,2) (3,4) (5,6,7,8)
Classification 136: (1,2) (3,4) (5,6,7,8)
Classification 137: (1,2) (3,4) (5,6,7,8)
Classification 138: (1,2) (3,4) (5,6,7,8)
Classification 139: (1,2) (3,4) (5,6,7,8)
Classification 140: (1,2) (3,4) (5,6,7,8)
Classification 141: (1,2) (3,4) (5,6,7,8)
Classification 142: (1,2) (3,4) (5,6,7,8)
Classification 143: (1,2) (3,4) (5,6,7,8)
Classification 144: (1,2) (3,4) (5,6,7,8)
Classification 145: (1,2) (3,4) (5,6,7,8)
Classification 146: (1,2) (3,4) (5,6,7,8)
Classification 147: (1,2) (3,4) (5,6,7,8)
Classification 148: (1,2) (3,4) (5,6,7,8)
Classification 149: (1,2) (3,4) (5,6,7,8)
Classification 150: (1,2) (3,4) (5,6,7,8)
Classification 151: (1,2) (3,4) (5,6,7,8)
Classification 152: (1,2) (3,4) (5,6,7,8)
Classification 153: (1,2) (3,4) (5,6,7,8)
Classification 154: (1,2) (3,4) (5,6,7,8)
Classification 155: (1,2) (3,4) (5,6,7,8)
Classification 156: (1,2) (3,4) (5,6,7,8)
Classification 157: (1,2) (3,4) (5,6,7,8)
Classification 158: (1,2) (3,4) (5,6,7,8)
Classification 159: (1,2) (3,4) (5,6,7,8)
Classification 160: (1,2) (3,4) (5,6,7,8)
Classification 161: (1,2) (3,4) (5,6,7,8)
Classification 162: (1,2) (3,4) (5,6,7,8)
Classification 163: (1,2) (3,4) (5,6,7,8)
Classification 164: (1,2) (3,4) (5,6,7,8)
Classification 165: (1,2) (3,4) (5,6,7,8)
Classification 166: (1,2) (3,4) (5,6,7,8)
Classification 167: (1,2) (3,4) (5,6,7,8)
Classification 168: (1,2) (3,4) (5,6,7,8)
Classification 169: (1,2) (3,4) (5,6,7,8)
Classification 170: (1,2) (3,4) (5,6,7,8)
Classification 171: (1,2) (3,4) (5,6,7,8)
Classification 172: (1,2) (3,4) (5,6,7,8)
Classification 173: (1,2) (3,4) (5,6,7,8)
Classification 174: (1,2) (3,4) (5,6,7,8)
Classification 175: (1,2) (3,4) (5,6,7,8)
Classification 176: (1,2) (3,4) (5,6,7,8)
Classification 177: (1,2) (3,4) (5,6,7,8)
Classification 178: (1,2) (3,4) (5,6,7,8)
Classification 179: (1,2) (3,4) (5,6,7,8)
Classification 180: (1,2) (3,4) (5,6,7,8)
Classification 181: (1,2) (3,4) (5,6,7,8)
Classification 182: (1,2) (3,4) (5,6,7,8)
Classification 183: (1,2) (3,4) (5,6,7,8)
Classification 184: (1,2) (3,4) (5,6,7,8)
Classification 185: (1,2) (3,4) (5,6,7,8)
Classification 186: (1,2) (3,4) (5,6,7,8)
Classification 187: (1,2) (3,4) (5,6,7,8)
Classification 188: (1,2) (3,4) (5,6,7,8)
Classification 189: (1,2) (3,4) (5,6,7,8)
Classification 190: (1,2) (3,4) (5,6,7,8)
Classification 191: (1,2) (3,4) (5,6,7,8)
Classification 192: (1,2) (3,4) (5,6,7,8)
Classification 193: (1,2) (3,4) (5,6,7,8)
Classification 194: (1,2) (3,4) (5,6,7,8)
Classification 195: (1,2) (3,4) (5,6,7,8)
Classification 196: (1,2) (3,4) (5,6,7,8)
Classification 197: (1,2) (3,4) (5,6,7,8)
Classification 198: (1,2) (3,4) (5,6,7,8)
Classification 199: (1,2) (3,4) (5,6,7,8)
Classification 200: (1,2) (3,4) (5,6,7,8)
Classification 201: (1,2) (3,4) (5,6,7,8)
Classification 202: (1,2) (3,4) (5,6,7,8)
Classification 203: (1,2) (3,4) (5,6,7,8)
Classification 204: (1,2) (3,4) (5,6,7,8)
Classification 205: (1,2) (3,4) (5,6,7,8)
Classification 206: (1,2) (3,4) (5,6,7,8)
Classification 207: (1,2) (3,4) (5,6,7,8)
Classification 208: (1,2) (3,4) (5,6,7,8)
Classification 209: (1,2) (3,4) (5,6,7,8)
Classification 210: (1,2) (3,4) (5,6,7,8)
Classification 211: (1,2) (3,4) (5,6,7,8)
Classification 212: (1,2) (3,4) (5,6,7,8)
Classification 213: (1,2) (3,4) (5,6,7,8)
Classification 214: (1,2) (3,4) (5,6,7,8)
Classification 215: (1,2) (3,4) (5,6,7,8)
Classification 216: (1,2) (3,4) (5,6,7,8)
Classification 217: (1,2) (3,4) (5,6,7,8)
Classification 218: (1,2) (3,4) (5,6,7,8)
Classification 219: (1,2) (3,4) (5,6,7,8)
Classification 220: (1,2) (3,4) (5,6,7,8)
Classification 221: (1,2) (3,4) (5,6,7,8)
Classification 222: (1,2) (3,4) (5,6,7,8)
Classification 223: (1,2) (3,4) (5,6,7,8)
Classification 224: (1,2) (3,4) (5,6,7,8)
Classification 225: (1,2) (3,4) (5,6,7,8)
Classification 226: (1,2) (3,4) (5,6,7,8)
Classification 227: (1,2) (3,4) (5,6,7,8)
Classification 228: (1,2) (3,4) (5,6,7,8)
Classification 229: (1,2) (3,4) (5,6,7,8)
Classification 230: (1,2) (3,4) (5,6,7,8)
Classification 231: (1,2) (3,4) (5,6,7,8)
Classification 232: (1,2) (3,4) (5,6,7,8)
Classification 233: (1,2) (3,4) (5,6,7,8)
Classification 234: (1,2) (3,4) (5,6,7,8)
Classification 235: (1,2) (3,4) (5,6,7,8)
Classification 236: (1,2) (3,4) (5,6,7,8)
Classification 237: (1,2) (3,4) (5,6,7,8)
Classification 238: (1,2) (3,4) (5,6,7,8)
Classification 239: (1,2) (3,4) (5,6,7,8)
Classification 240: (1,2) (3,4) (5,6,7,8)
Classification 241: (1,2) (3,4) (5,6,7,8)
Classification 242: (1,2) (3,4) (5,6,7,8)
Classification 243: (1,2) (3,4) (5,6,7,8)
Classification 244: (1,2) (3,4) (5,6,7,8)
Classification 245: (1,2) (3,4) (5,6,7,8)
Classification 246: (1,2) (3,4) (5,6,7,8)
Classification 247: (1,2) (3,4) (5,6,7,8)
Classification 248: (1,2) (3,4) (5,6,7,8)
Classification 249: (1,2) (3,4) (5,6,7,8)
Classification 250: (1,2) (3,4) (5,6,7,8)
Classification 251: (1,2) (3,4) (5,6,7,8)
Classification 252: (1,2) (3,4) (5,6,7,8)
Classification 253: (1,2) (3,4) (5,6,7,8)
Classification 254: (1,2) (3,4) (5,6,7,8)
Classification 255: (1,2) (3,4) (5,6,7,8)
Classification 256: (1,2) (3,4) (5,6,7,8)
Classification 257: (1,2) (3,4) (5,6,7,8)
Classification 258: (1,2) (3,4) (5,6,7,8)
Classification 259: (1,2) (3,4) (5,6,7,8)
Classification 260: (1,2) (3,4) (5,6,7,8)
Classification 261: (1,2) (3,4) (5,6,7,8)
Classification 262: (1,2) (3,4) (5,6,7,8)
Classification 263: (1,2) (3,4) (5,6,7,8)
Classification 264: (1,2) (3,4) (5,6,7,8)
Classification 265: (1,2) (3,4) (5,6,7,8)
Classification 266: (1,2) (3,4) (5,6,7,8)
Classification 267: (1,2) (3,4) (5,6,7,8)
Classification 268: (1,2) (3,4) (5,6,7,8)
Classification 269: (1,2) (3,4) (5,6,7,8)
Classification 270: (1,2) (3,4) (5,6,7,8)
Classification 271: (1,2) (3,4) (5,6,7,8)
Classification 272: (1,2) (3,4) (5,6,7,8)
Classification 273: (1,2) (3,4) (5,6,7,8)
Classification 274: (1,2) (3,4) (5,6,7,8)
Classification 275: (1,2) (3,4) (5,6,7,8)
Classification 276: (1,2) (3,4) (5,6,7,8)
Classification 277: (1,2) (3,4) (5,6,7,8)
Classification 278: (1,2) (3,4) (5,6,7,8)
Classification 279: (1,2) (3,4) (5,6,7,8)
Classification 280: (1,2) (3,4) (5,6,7,8)
Classification 281: (1,2) (3,4) (5,6,7,8)
Classification 282: (1,2) (3,4) (5,6,7,8)
Classification 283: (1,2) (3,4) (5,6,7,8)
Classification 284: (1,2) (3,4) (5,6,7,8)
Classification 285: (1,2) (3,4) (5,6,7,8)
Classification 286: (1,2) (3,4) (5,6,7,8)
Classification 287: (1,2) (3,4) (5,6,7,8)
Classification 288: (1,2) (3,4) (5,6,7,8)
Classification 289: (1,2) (3,4) (5,6,7,8)
Classification 290: (1,2) (3,4) (5,6,7,8)
Classification 291: (1,2) (3,4) (5,6,7,8)
Classification 292: (1,2) (3,4) (5,6,7,8)
Classification 293: (1,2) (3,4) (5,6,7,8)
Classification 294: (1,2) (3,4) (5,6,7,8)
Classification 295: (1,2) (3,4) (5,6,7,8)
Classification 296: (1,2) (3,4) (5,6,7,8)
Classification 297: (1,2) (3,4) (5,6,7,8)
Classification 298: (1,2) (3,4) (5,6,7,8)
Classification 299: (1,2) (3,4) (5,6,7,8)
Classification 300: (1,2) (3,4) (5,6,7,8)
Classification 301: (1,2) (3,4) (5,6,7,8)
Classification 302: (1,2) (3,4) (5,6,7,8)
Classification 303: (1,2) (3,4) (5,6,7,8)
Classification 304: (1,2) (3,4) (5,6,7,8)
Classification 305: (1,2) (3,4) (5,6,7,8)
Classification 306: (1,2) (3,4) (5,6,7,8)
Classification 307: (1,2) (3,4) (5,6,7,8)
Classification 308: (1,2) (3,4) (5,6,7,8)
Classification 309: (1,2) (3,4) (5,6,7,8)
Classification 310: (1,2) (3,4) (5,6,7,8)
Classification 311: (1,2) (3,4) (5,6,7,8)
Classification 312: (1,2) (3,4) (5,6,7,8)
Classification 313: (1,2) (3,4) (5,6,7,8)
Classification 314: (1,2) (3,4) (5,6,7,8)
Classification 315: (1,2) (3,4) (5,6,7,8)
Classification 316: (1,2) (3,4) (5,6,7,8)
Classification 317: (1,2) (3,4) (5,6,7,8)
Classification 318: (1,2) (3,4) (5,6,7,8)
Classification 319: (1,2) (3,4) (5,6,7,8)
Classification 320: (1,2) (3,4) (5,6,7,8)
Classification 321: (1,2) (3,4) (5,6,7,8)
Classification 322: (1,2) (3,4) (5,6,7,8)
Classification 323: (1,2) (3,4) (5,6,7,8)
Classification 324: (1,2) (3,4) (5,6,7,8)
Classification 325: (1,2) (3,4) (5,6,7,8)
Classification 326: (1,2) (3,4) (5,6,7,8)
Classification 327: (1,2) (3,4) (5,6,7,8)
Classification 328: (1,2) (3,4) (5,6,7,8)
Classification 329: (1,2) (3,4) (5,6,7,8)
Classification 330: (1,2) (3,4) (5,6,7,8)
Classification 331: (1,2) (3,4) (5,6,7,8)
Classification 332: (1,2) (3,4) (5,6,7,8)
Classification 333: (1,2) (3,4) (5,6,7,8)
Classification 334: (1,2) (3,4) (5,6,7,8)
Classification 335: (1,2) (3,4) (5,6,7,8)
Classification 336: (1,2) (3,4) (5,6,7,8)
Classification 337: (1,2) (3,4) (5,6,7,8)
Classification 338: (1,2) (3,4) (5,6,7,8)
Classification 339: (1,2) (3,4) (5,6,7,8)
Classification 340: (1,2) (3,4) (5,6,7,8)
Classification 341: (1,2) (3,4) (5,6,7,8)
Classification 342: (1,2) (3,4) (5,6,7,8)
Classification 343: (1,2) (3,4) (5,6,7,8)
Classification 344: (1,2) (3,4) (5,6,7,8)
Classification 345: (1,2) (3,4) (5,6,7,8)
Classification 346: (1,2) (3,4) (5,6,7,8)
Classification 347: (1,2) (3,4) (5,6,7,8)
Classification 348: (1,2) (3,4) (5,6,7,8)
Classification 349: (1,2) (3,4) (5,6,7,8)
Classification 350: (1,2) (3,4) (5,6,7,8)
Classification 351: (1,2) (3,4) (5,6,7,8)
Classification 352: (1,2) (3,4) (5,6,7,8)
Classification 353: (1,2) (3,4) (5,6,7,8)
Classification 354: (1,2) (3,4) (5,6,7,8)
Classification 355: (1,2) (3,4) (5,6,7,8)
Classification 356: (1,2) (3,4) (5,6,7,8)
Classification 357: (1,2) (3,4) (5,6,7,8)
Classification 358: (1,2) (3,4) (5,6,7,8)
Classification 359: (1,2) (3,4) (5,6,7,8)
Classification 360: (1,2) (3,4) (5,6,7,8)
Classification 361: (1,2) (3,4) (5,6,7,8)
Classification 362: (1,2) (3,4) (5,6,7,8)
Classification 363: (1,2) (3,4) (5,6,7,8)
Classification 364: (1,2) (3,4) (5,6,7,8)
Classification 365: (1,2) (3,4) (5,6,7,8)
Classification 366: (1,2) (3,4) (5,6,7,8)
Classification 367: (1,2) (3,4) (5,6,7,8)
Classification 368: (1,2) (3,4) (5,6,7,8)
Classification 369: (1,2) (3,4) (5,6,7,8)
Classification 370: (1,2) (3,4) (5,6,7,8)
Classification 371: (1,2) (3,4) (5,6,7,8)
Classification 372: (1,2) (3,4) (5,6,7,8)
Classification 373: (1,2) (3,4) (5,6,7,8)
Classification 374: (1,2) (3,4) (5,6,7,8)
Classification 375: (1,2) (3,4) (5,6,7,8)
Classification 376: (1,2) (3,4) (5,6,7,8)
Classification 377: (1,2) (3,4) (5,6,7,8)
Classification 378: (1,2) (3,4) (5,6,7,8)
Classification 379: (1,2) (3,4) (5,6,7,8)
Classification 380: (1,2) (3,4) (5,6,7,8)
Classification 381: (1,2) (3,4) (5,6,7,8)
Classification 382: (1,2) (3,4) (5,6,7,8)
Classification 383: (1,2) (3,4) (5,6,7,8)
```

Técnica: Árbol Jerárquico Cohesivo en R

Simulación en R: Técnica →Árbol Jerárquico Cohesivo del ASD (clustering)

Definiciones	La parte conceptual. Ver las definiciones en las tablas 6 y 7. ¿Cuál es el nombre de la técnica en R?: <code>callHierarchyTree</code> ¿Cómo buscar la técnica en R?: <code>help(callHierarchyTree)</code>
Uso de la técnica a través del Software R	Continuando con el problema propuesto en la tabla 9. Después de ingresar los datos en un archivo de extensión *.csv, de activar el paquete <code>Rshiny</code> en RStudio. Para este tipo de análisis se puede proceder como: 1.- En la ventana <code>Rshiny</code> , seleccionamos la opción: "Hierarchy tree"  2.- Ubicamos la dirección de la base de datos creada y observamos los resultados. 2.1.- En la ventana "consola" de RStudio: Ejemplo 1 <pre>## Ejemplo 1 ## Ejemplo 2 ## Ejemplo 3 ## Ejemplo 4 ## Ejemplo 5 ## Ejemplo 6 ## Ejemplo 7 ## Ejemplo 8 ## Ejemplo 9 ## Ejemplo 10 ## Ejemplo 11 ## Ejemplo 12 ## Ejemplo 13 ## Ejemplo 14 ## Ejemplo 15 ## Ejemplo 16 ## Ejemplo 17 ## Ejemplo 18 ## Ejemplo 19 ## Ejemplo 20 ## Ejemplo 21 ## Ejemplo 22 ## Ejemplo 23 ## Ejemplo 24 ## Ejemplo 25 ## Ejemplo 26 ## Ejemplo 27 ## Ejemplo 28 ## Ejemplo 29 ## Ejemplo 30 ## Ejemplo 31 ## Ejemplo 32 ## Ejemplo 33 ## Ejemplo 34 ## Ejemplo 35 ## Ejemplo 36 ## Ejemplo 37 ## Ejemplo 38 ## Ejemplo 39 ## Ejemplo 40 ## Ejemplo 41 ## Ejemplo 42 ## Ejemplo 43 ## Ejemplo 44 ## Ejemplo 45 ## Ejemplo 46 ## Ejemplo 47 ## Ejemplo 48 ## Ejemplo 49 ## Ejemplo 50 ## Ejemplo 51 ## Ejemplo 52 ## Ejemplo 53 ## Ejemplo 54 ## Ejemplo 55 ## Ejemplo 56 ## Ejemplo 57 ## Ejemplo 58 ## Ejemplo 59 ## Ejemplo 60 ## Ejemplo 61 ## Ejemplo 62 ## Ejemplo 63 ## Ejemplo 64 ## Ejemplo 65 ## Ejemplo 66 ## Ejemplo 67 ## Ejemplo 68 ## Ejemplo 69 ## Ejemplo 70 ## Ejemplo 71 ## Ejemplo 72 ## Ejemplo 73 ## Ejemplo 74 ## Ejemplo 75 ## Ejemplo 76 ## Ejemplo 77 ## Ejemplo 78 ## Ejemplo 79 ## Ejemplo 80 ## Ejemplo 81 ## Ejemplo 82 ## Ejemplo 83 ## Ejemplo 84 ## Ejemplo 85 ## Ejemplo 86 ## Ejemplo 87 ## Ejemplo 88 ## Ejemplo 89 ## Ejemplo 90 ## Ejemplo 91 ## Ejemplo 92 ## Ejemplo 93 ## Ejemplo 94 ## Ejemplo 95 ## Ejemplo 96 ## Ejemplo 97 ## Ejemplo 98 ## Ejemplo 99 ## Ejemplo 100</pre> Ejemplo 2 <pre>## Ejemplo 1 ## Ejemplo 2 ## Ejemplo 3 ## Ejemplo 4 ## Ejemplo 5 ## Ejemplo 6 ## Ejemplo 7 ## Ejemplo 8 ## Ejemplo 9 ## Ejemplo 10 ## Ejemplo 11 ## Ejemplo 12 ## Ejemplo 13 ## Ejemplo 14 ## Ejemplo 15 ## Ejemplo 16 ## Ejemplo 17 ## Ejemplo 18 ## Ejemplo 19 ## Ejemplo 20 ## Ejemplo 21 ## Ejemplo 22 ## Ejemplo 23 ## Ejemplo 24 ## Ejemplo 25 ## Ejemplo 26 ## Ejemplo 27 ## Ejemplo 28 ## Ejemplo 29 ## Ejemplo 30 ## Ejemplo 31 ## Ejemplo 32 ## Ejemplo 33 ## Ejemplo 34 ## Ejemplo 35 ## Ejemplo 36 ## Ejemplo 37 ## Ejemplo 38 ## Ejemplo 39 ## Ejemplo 40 ## Ejemplo 41 ## Ejemplo 42 ## Ejemplo 43 ## Ejemplo 44 ## Ejemplo 45 ## Ejemplo 46 ## Ejemplo 47 ## Ejemplo 48 ## Ejemplo 49 ## Ejemplo 50 ## Ejemplo 51 ## Ejemplo 52 ## Ejemplo 53 ## Ejemplo 54 ## Ejemplo 55 ## Ejemplo 56 ## Ejemplo 57 ## Ejemplo 58 ## Ejemplo 59 ## Ejemplo 60 ## Ejemplo 61 ## Ejemplo 62 ## Ejemplo 63 ## Ejemplo 64 ## Ejemplo 65 ## Ejemplo 66 ## Ejemplo 67 ## Ejemplo 68 ## Ejemplo 69 ## Ejemplo 70 ## Ejemplo 71 ## Ejemplo 72 ## Ejemplo 73 ## Ejemplo 74 ## Ejemplo 75 ## Ejemplo 76 ## Ejemplo 77 ## Ejemplo 78 ## Ejemplo 79 ## Ejemplo 80 ## Ejemplo 81 ## Ejemplo 82 ## Ejemplo 83 ## Ejemplo 84 ## Ejemplo 85 ## Ejemplo 86 ## Ejemplo 87 ## Ejemplo 88 ## Ejemplo 89 ## Ejemplo 90 ## Ejemplo 91 ## Ejemplo 92 ## Ejemplo 93 ## Ejemplo 94 ## Ejemplo 95 ## Ejemplo 96 ## Ejemplo 97 ## Ejemplo 98 ## Ejemplo 99 ## Ejemplo 100</pre> 2.2.- Nuevas ventanas de <code>Rshiny</code> : Ejemplo 1  Ejemplo 2




3.- Se emiten las respectivas conclusiones.

Nota.- Para más ejemplos, ver el Apéndice F.

Actividades propuestas	Se solicita al lector: 1.- Crear una base de datos binarios de su contexto educativo, es decir datos de la forma (sí/no, verdadero/falso, admite/no admite, con propiedad/sin propiedad, aprueba/desaprueba, admitido/no admitido, promovido/no promovido, motivado/desmotivado) convertidas a 1 y 0 respectivamente. 2.- Guardar estos datos en formato *.csv, en la carpeta "Documentos" o en el sitio donde podamos localizarlo. 3.- Activar <code>Rshiny</code> y seleccionar la opción: "Hierarchy tree". 4.- Ubicar la dirección de la base de datos creada y observar los resultados.
Evaluación	"Para verificar la comprensión docente, del uso de esta técnica en contextos educativos, se solicita al usuario que conteste la encuesta del siguiente enlace: https://bit.ly/2TIRPKR ".

Técnica: `hclust.vector` en R

Simulación en R: Técnica →`hclust.vector` del LA (clustering)

Definiciones	Según (Naranjo et al., 2018): <code>hclust.vector</code> –Agrupación jerárquica y aglomerativa rápida de datos vectoriales, es una función que proporciona agrupamiento cuando la entrada es de datos vectoriales. Utiliza algoritmos de ahorro de memoria que permiten el procesamiento de conjunto de datos más grandes que <code>hclust</code> . Los métodos "ward", "centroid" y "median" requieren <code>metric="euclidean"</code> y agrupan el conjunto de datos con respecto a las distancias euclidianas. Para el agrupamiento de ligamiento "único", se puede elegir cualquier medida de semejanza. Actualmente, las mismas métricas se implementan como proporciona la función <code>dist</code> . ¿Cuál es el nombre de la técnica en R?: <code>hclust.vector</code> ¿Cómo buscar la técnica en R?: <code>help(hclust.vector)</code>
Uso de la técnica a través del Software R	Se puede proceder como: 1.- Disponer de una base de datos binarios del contexto educativo, es decir datos de la forma (sí/no, verdadero/falso, admite/no admite, con propiedad/sin propiedad, aprueba/desaprueba, admitido/no admitido, promovido/no promovido, motivado/desmotivado) convertidas a 1 y 0 respectivamente. 2.- Guardar los datos en un archivo de extensión *.csv (ejemplo: "Café.csv"). 3.- En la ventana "consola" de RStudio, ejecutar el algoritmo (ver Apéndice D.3). 4.- Observar los resultados: 

5.- Se emiten las respectivas conclusiones.

Actividades propuestas	Se solicita al lector: 1.- Crear una base de datos binarios de su contexto educativo, es decir datos de la forma (sí/no, verdadero/falso, admite/no admite, con propiedad/sin propiedad, aprueba/desaprueba, admitido/no admitido, promovido/no promovido, motivado/desmotivado) convertidas a 1 y 0 respectivamente. 2.- Guardar estos datos en formato *.csv, en la carpeta "Documentos" o en el sitio donde podamos localizarlo. 3.- En la ventana "consola" de RStudio, ejecutar el algoritmo (ver Apéndice D.3) asociada a la base de datos creada. 4.- Ubicar la dirección de la base de datos creada y observar los resultados.
Evaluación	"Para verificar la comprensión docente, del uso de esta técnica en contextos educativos, se solicita al usuario que conteste la encuesta del siguiente enlace: https://bit.ly/3J5OVy ".

Técnica: *dendro.variables* en R

Simulación en R: *Técnica dendro.variables del LA (clustering)*

Definiciones Según (Naranjo et al., 2018): *dendro.variables*-Variables dendrogram, es una función que permite la agrupación de variables (1) basada en la similitud al usar medidas de asociación, (2) basada en la similitud al usar la correlación de distancias, o (3) mediante el enfoque *ClustOfVar*, que utiliza el análisis de componentes principales para datos mixtos.

¿Cuál es el nombre de la técnica en R?: *dendro.variables*

¿Cómo buscar la técnica en R?: *help(dendro.variables)*

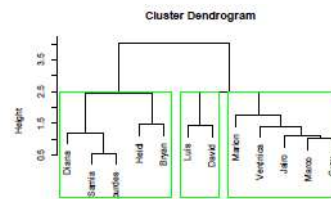
Uso de la técnica a través del Software R Se puede proceder como:

1.- Disponer de una base de datos numéricas del contexto educativo, es decir datos que pertenezcan a un intervalo de número reales.

2.- Guardar los datos en un archivo de extensión *.csv (ejemplo: "Notas_Estudiantes.csv").

3.- En la ventana "consola" de RStudio, ejecutar el algoritmo *dendro.variables* (ver Apéndice D.3).

4.- Observar los resultados:



dato(Datos)
hclust("average")

Una breve interpretación: Basándose en los datos proporcionados, se observan la conformación de grupos de estudiantes, de acuerdo a sus diferentes capacidades.

Actividades propuestas: Se solicita al lector:

1.- Crear una base de datos continuos de su contexto educativo, pueden ser notas de sus estudiantes, de la forma:

Estudiante	Matemática	Lenguaje Castellano	Lenguaje Extranjero	Arte	Ed. Física	Historia	Geografía	Biología
1. Sara	7.00	8.25	6.70	8.00	6.70	6.20	6.00	6.00
2. Luis	6.24	7.20	5.00	7.00	7.00	7.24	7.24	7.20
3. Lucía	5.57	6.40	6.70	6.50	7.94	6.00	6.00	6.00
4. Verónica	5.00	7.00	5.70	7.50	7.94	7.94	7.94	7.94
5. Marcos	7.10	7.00	6.50	7.50	7.94	7.94	7.94	7.94
6. Heidi	6.50	6.20	6.00	6.24	6.47	6.70	6.00	6.00
7. Cory	6.00	7.00	6.00	6.24	6.47	6.70	6.00	6.00
8. Bryan	6.00	6.00	6.00	6.24	6.47	6.70	6.00	6.00
9. David	6.00	6.00	6.00	6.24	6.47	6.70	6.00	6.00
10. Diana	7.00	7.00	6.74	6.50	6.00	7.24	7.24	7.24
11. Sara	6.00	7.00	6.24	7.00	7.00	7.00	7.00	7.00
12. Heidi	6.00	6.24	6.00	7.00	7.00	6.00	6.00	6.00

2.- Guardar estos datos en formato *.csv, en la carpeta "Documentos" o en el sitio donde se pueda localizar.

3.- En la ventana "consola" de RStudio, ejecutar el algoritmo *dendro.variables* (ver Apéndice D.3.) asociada a la base de datos creada.

4.- Ubicar la dirección de la base de datos creada y observar los resultados.

Evaluación

"Para verificar la comprensión docente, del uso de esta técnica en contextos educativos, se solicita al usuario que conteste la encuesta del siguiente enlace: <https://bit.ly/2Tly45Y>".

Técnica: *dendro.diana* en R

Simulación en R: *Técnica dendro.diana del LA (clustering)*

Definiciones Según (Naranjo et al., 2018): *diana*-Divisive Analysis Clustering, es una función que calcula una agrupación jerárquica divisiva del conjunto de datos que devuelve un objeto de la clase *diana*, probablemente sea único en el cálculo de una jerarquía divisiva, mientras que la mayoría de los otros (software), para la agrupación jerárquica son aglomerativos, el *diana*-algoritmo construye una jerarquía de agrupamientos, al comenzar con un grupo grande que contiene todas las *n* observaciones, los grupos se dividen hasta que cada grupo contiene solo una observación, en cada etapa, se selecciona el clúster con el diámetro más grande, el diámetro de un grupo es la mayor diferencia entre dos de sus observaciones, para dividir el clúster seleccionado, el algoritmo primero busca su observación más dispersa (es decir, que tiene la mayor diferencia de promedio con las otras observaciones del clúster seleccionado)".

¿Cuál es el nombre de la técnica en R?: *dendro.diana*

¿Cómo buscar la técnica en R?: *help(dendro.diana)*

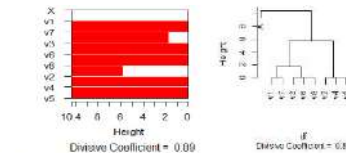
Uso de la técnica a través del Software R Se puede proceder como:

1.- Disponer de una base de datos binarios del contexto educativo, es decir datos de la forma (si/no, verdadero/falso, admite/no admite, con propiedad/sin propiedad, aprueba/desaprueba, admitido/no admitido, promovido/no promovido, motivado/desmotivado) convertidas a 1 y 0 respectivamente.

2.- Guardar los datos en un archivo de extensión *.csv (ejemplo: "Caféin.csv").

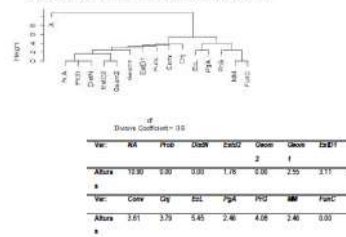
3.- En la ventana "consola" de RStudio, ejecutar el algoritmo (ver Apéndice D.3.)

4.- Observar los resultados:



Más ejemplos:

Dendrogram of diana = d1.metric = "euclidean", stand = TRUE



5.- Emitir las respectivas conclusiones.

Actividades propuestas: Se solicita al lector:

1.- Crear una base de datos binarios de su contexto educativo, es decir datos de la forma (si/no, verdadero/falso, admite/no admite, con propiedad/sin propiedad, aprueba/desaprueba, admitido/no admitido, promovido/no promovido, motivado/desmotivado) convertidas a 1 y 0 respectivamente.

2.- Guardar estos datos en formato *.csv, en la carpeta "Documentos" o en el sitio donde se pueda localizar.

3.- En la ventana "consola" de RStudio, ejecutar el algoritmo *diana* (ver Apéndice D.3.) asociada a la base de datos creada.

4.- Ubicar la dirección de la base de datos creada y observar los resultados.

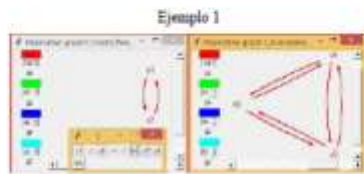
Técnica: Grafo Implicativo en R

Simulación en R: Técnica → Grafo Implicativo del ASI (Reglas de Asociación de Minería)

Definiciones	La parte conceptual. Ver las definiciones de la tabla 6. ¿Cuál es el nombre de la técnica en R? <code>implicativeGraph</code> ¿Cómo buscar la técnica en R? <code>help(implicativeGraph)</code>
Uso de la técnica a través del Software R	Continuando con el problema propuesto en las tablas 9 y 10. Después de ingresar los datos en un archivo de extensión *.csv, de activar el paquete <code>Rclicr</code> en <code>RStudio</code> , para este tipo de análisis se puede proceder como: 1.- En la ventana <code>Rclicr</code> , seleccionamos la opción: “ <code>Implicative graph</code> ”



- 4.- Observamos los resultados:
- 4.1.- Nuevas ventanas de `Rclicr`:



- 5.- Se emiten las respectivas conclusiones.
Una breve interpretación del ejemplo 2. El tema de `EstD` implica al tema de `EstD2` y `GeoM2`, con una intensidad de implicación del 50% y 57%.

El tema de `EstD2` implica al tema de `GeoM2` y viceversa, con una intensidad de implicación del 57% y 75%. El tema de `Fgd` implica al tema de `EoL` y el tema de `EoL` implica al tema de `Com`, con una intensidad de implicación del 50% al 57% y 57% al 75% respectivamente.
Nota.- Para más ejemplos, ver el Apéndice F.

Actividades propuestas	Se solicita al lector: 1.- Crear una base de datos binarios de su contexto educativo, es decir datos de la forma (s/no, verdadero/falso, admite/no admite, con propiedad/sin propiedad, aprueba/desaprueba, admitido/no admitido, promovido/no promovido, motivado/desmotivado) convertidas a 1 y 0 respectivamente. 2.- Guardar estos datos en formato *.csv, en la carpeta “Documentos” o en el tino donde podamos localizarlo. 3.- Activar <code>Rclicr</code> y seleccionar la opción: “ <code>Implicative graph</code> ”.
Evaluación	“Para verificar la comprensión docente, del uso de esta técnica en contextos educativos, se solicita al usuario que conteste la encuesta del siguiente enlace: https://bit.ly/3oGHbT4 ”.

- 4.- Ubicar la dirección de la base de datos creada y observar los resultados.

Técnica: Apriori en R

Simulación en R: Técnica → Apriori del LA (Minería de Asociación)

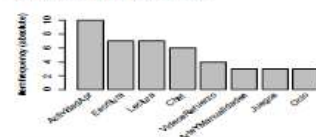
Definiciones	Según (Narjain et al., 2018): Apriori, es un conjunto de elementos frecuentes, reglas de asociación o hiper bordes de asociación que utilizan el algoritmo Apriori. El algoritmo Apriori emplea la búsqueda a nivel de conjuntos de elementos frecuentes. Incluye la implementación de un árbol de prefijos y clasificación de elementos. Método: <ul style="list-style-type: none"> • Deje $k=1$ • Genera conjuntos de elementos frecuentes de longitud 1 • Repite hasta que no se identifiquen nuevos conjuntos de elementos frecuentes • Genera conjuntos de elementos de longitud $(k+1)$ a partir de la longitud k conjuntos de elementos frecuentes • Poda conjuntos de elementos candidatos que contienen subconjuntos de longitud k que son infrecuentes • Cuenta el apoyo de cada candidato al escanear la DB • Elimina candidatos que son poco frecuentes, al dejar solo aquellos que son frecuentes La implementación de Apriori, es a través de uso de Hash Tree. ¿Cuál es el nombre de la técnica en R? <code>apriori</code> ¿Cómo buscar la técnica en R? <code>help(apriori)</code>
---------------------	--

Uso de la técnica a través del Software R
Se puede proceder como:
1.- Al tratar de técnicas de reglas de asociación, disponer de una base de datos del contexto educativo, agrupadas en transacción y actividad.

transacción	transacción
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10

Nota.- Es posible que el estudiante1, tienda a realizar: Actividades de aprendizaje, lectura, arte y manualidades, entre otras actividades, que podemos agruparlos por “transacciones”.

- 2.- Guardar los datos en un archivo de extensión *.csv (ejemplo: “Diario_Estudantil.csv”).
- 3.- En la ventana “consola” de `RStudio`, ejecutar el algoritmo `apriori` (ver Apéndice D.4.).
- 4.- Observar los resultados:



```
Algorithmic control:
*filter tree heap remove load sort verbose
0.1 TRUE TRUE FALSE TRUE 2 TRUE

absolute minimum support count: 1

set item appearances ... [9 items] done [0.002]
set transactions ... [8 items], 10 transactions done [0.002]
sorting and recoding items ... [8 items] done [0.002]
creating transaction tree ... done [0.002]
checking subsets of size 1 2 3 4 5 6 done [0.002]
pruning ... [177 rules] done [0.002]
creating R object ... done [0.002]

... 8 transactions data frame: 0 obs. of 0 variables
... quality: data frame: 140 obs. of 4 variables
... $ support: num [1:140] 0.3 0.3 0.3 0.3 0.3 0.3 0.3 0.3 0.3 0.3 ...
... $ confidence: num [1:140] 1 1 1 1 1 1 1 1 1 1 ...
... $ lift: num [1:140] 1.67 1.33 1.33 1.33 1.33 ...
... $ conv: num [1:140] 2 2 2 2 2 2 2 2 2 2 ...
... $ info: list of 4
... $ data: symbol transactions
... $ transactions: int 20
... $ support: num 0.1
... $ confidence: num 0.9
```

5.- Existen reglas, reflexionamos.

arte/manualidades, escritura, videosrefuerzo	=>	(lectura)	0.3	1.1428571	3
arte/manualidades, lectura, videosrefuerzo	=>	(escritura)	0.3	1.1428571	3

Más ejemplos:

Transacción	(A, B, C, D)
(A, B, D)	(A, D)
(B, C, D)	(B, C, D)
(D, C)	(D, C)

Regla	Confianza	Contorno
[B] => [B]	0.431241 + 0	
[B] => [B]	0.431241 + 0.3	
[B] => [B]	0.431241 + 0.5	
[B] => [B]	0.431241 + 0.75	
[B] => [B]	0.431241 + 0.8	
[B] => [B]	0.431241 + 0.8	
[B] => [B]	0.431241 + 0.75	
[B] => [B]	0.431241 + 0.8	

Fuente: https://rpubs.com/Joaquin_AR/397172

Una breve Interpretación. – Según Amat (2018), “supóngase que se desean únicamente reglas con una confianza igual o superior a 0.7, es decir que la regla se cumpla en un 70%, (...), de todas las reglas posibles, únicamente (C) \Rightarrow (D) y (C) \Rightarrow (B) superan el límite de confianza”.

Actividades propuestas	Se solicita al lector: 1.- Crear una base de datos del contexto educativo, agrupadas en transacción y actividad. 2.- Guardar los datos en un archivo de extensión *.csv (ejemplo: "Diario_Estudiantill.csv"). 3.- En la ventana "consola" de RStudio, ejecutar el algoritmo (ver Apéndice D.4. <i>apriori</i>), relacionada a la base de datos con la que esté trabajando. 4.- Observar los resultados (las reglas generadas) y reflexionar.
Evaluación	"Para verificar la comprensión docente, del uso de esta técnica en contextos educativos, se solicita al usuario que conteste la encuesta del siguiente enlace: https://bit.ly/2JpVhYU ".

Técnica: weclat en R

Basado en R. Técnica basada del LA (Análisis de Asociaciones)

Definición: Según (Narayo et al., 2018) el algoritmo weclat, realiza combinaciones optimizadas de reglas y ponderaciones de transacción para implementar la extracción de reglas de asociación ponderadas (WARM).
¿Cuál es el nombre de la técnica en R? weclat
¿Cómo llama la técnica en R? weclat

Uso de la técnica a través del software R
Se puede proceder como:
1.- Al tener de técnicas de reglas de asociación, disponer de una base de datos del contexto educativo, agrupadas en transacción y actividad.



Nota: Es posible que el estudiante, pueda realizar Actividades de generalización, lecturas, etc y manipulaciones, entre otras actividades, que pueden agruparse por "transacciones"

2.- Guardar los datos en un archivo de extensión *.csv (ejemplo: "Diario_Estudiantill.csv")

3.- En la ventana "consola" de RStudio, ejecutar el algoritmo weclat (ver Apéndice D.4.)

4.- Observar los resultados, analizar reglas, reflexionar:

Regla	Confianza	Contorno
[B] => [B]	0.431241 + 0	
[B] => [B]	0.431241 + 0.3	
[B] => [B]	0.431241 + 0.5	
[B] => [B]	0.431241 + 0.75	
[B] => [B]	0.431241 + 0.8	
[B] => [B]	0.431241 + 0.8	
[B] => [B]	0.431241 + 0.75	
[B] => [B]	0.431241 + 0.8	

Actividades propuestas:
Se solicita al lector:
1.- Crear una base de datos del contexto educativo, agrupadas en transacción y actividad.
2.- Guardar los datos en un archivo de extensión *.csv (ejemplo: "Diario_Estudiantill.csv").

3.- En la ventana "consola" de RStudio, ejecutar el algoritmo (ver Apéndice D.4. *weclat*), relacionada a la base de datos con la que esté trabajando.

4.- Observar los resultados (las reglas generadas) y reflexionar.

Evaluación
"Para verificar la comprensión docente, del uso de esta técnica en contextos educativos, se solicita al usuario que conteste la encuesta del siguiente enlace: <https://bit.ly/2H56ND8>".

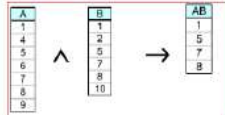
Técnica: *eclat* en R

Simulación en R: Técnica \rightarrow *eclat* del LA (Minería de Asociación)

Definiciones Según Hurairah, como se citó en (Naranjo et al., 2018): el algoritmo *Eclat* “*Equivalence Class Transformation*”, usa operaciones simples de intersección para el agrupamiento de clase de equivalencia junto con el recorrido de retícula ascendente.

Generación

- Determinar el soporte de cualquier k-conjunto de elementos mediante la intersección de listas de dos de sus subconjuntos (k-1).



Fuente: Hurairah como se citó en (Naranjo et al., 2018)).

- Tres enfoques transversales: de arriba hacia abajo, de abajo hacia arriba y híbridos
- Ventaja: recuento de soporte muy rápido
- Desventajas: las listas intermedias pueden convertirse también grande para la memoria.

La generación frecuente de conjuntos de elementos computacionalmente costosa.

Algoritmo ECLAT

- Agrupación de clase de equivalencia y de abajo hacia arriba Lattice Transversal - ECLAT
- Método para la generación frecuente de conjuntos de elementos
- Busca de manera DFS
- Representa los datos en forma vertical

¿Cuál es el nombre de la técnica en R?: *eclat*

¿Cómo buscar la técnica en R?: *help(eclat)*

Uso de la Se puede proceder como:

técnica a través del Software R 1.- Al tratar de técnicas de reglas de asociación, disponer de una base de datos de tipo “caracteres” del contexto educativo, agrupadas en filas.

PropSimple	Mé Igualación	Revis Resultad	Solución Correcta
PropComp	Falsa Prop	PropSimple	
PropSimple	Solución Cor Sistec	PropComp	
Revis Resultad	PropSimple	Solución Correcta	
Solución Cor Sistec	Revis Resultad	Falsa Prop	
Falsa Prop	Revis Resultad	Solución Cor PropSimple	
PropSimple	Falsa Prop		

Nota. - Es posible que el estudiante, tienda a utilizar estrategias:

Proporcionalidad simple, método de igualación, revisión de resultados y así llegar a la solución correcta, entre otras, que podemos agrupar por “filas”.

2.- Guardar los datos en un archivo de extensión *.csv (ejemplo:

“ReglasAso.csv”).

3.- En la ventana “consola” de *RStudio*, ejecutar el algoritmo (ver Apéndice D.4. *eclat*).

4.- Observamos los resultados, reflexionamos:

```

library(MASS)
support_count = 0

create_transactions = function(x) {
  for (i in 1:nrow(x)) {
    trans = x[i,]
    support_count = support_count + 1
  }
}

create_transactions(x)
support_count

# [1] 10

```

Más Ejemplos:

Transacciones	Soporte
T1	4/10
T2	1/10
T3	3/10
T4	4/10
T5	1/10
T6	4/10
T7	2/10
T8	2/10
T9	3/10
T10	3/10

Itemset (k = 2)	Transacciones	Soporte
{1,2}	{2,3}	0.20
{1,4}	{3}	0.10
{1,6}	{2,3,8}	0.20
{2,3}	{3}	0.10
{2,4}	{2,3}	0.20
{3,4}	{2,3,7,8}	0.40
{3,6}	{2,3,7,8}	0.40
{4,6}	{3,7,8}	0.30

Itemset (k = 3)	Transacciones	Soporte
{1,4,6}	{2,3}	0.20
{2,4,6}	{2,3}	0.20

- Nota.** - Según Amat (2018), “El soporte mínimo para considerar un *itemset* frecuente es del 20%. Este algoritmo finaliza cuando no hay más *itemsets* frecuentes, este algoritmo solo identifica los *itemsets* frecuentes no genera las reglas de asociación, para este último se puede aplicar la segunda parte del algoritmo *a priori*”.
- Fuente: https://rpubs.com/Joaquin_AR/397172

Actividades propuestas	Se solicita al lector:
	1.- Crear una base de datos de tipo “caracteres” del contexto educativo, agrupadas en filas.
	2.- Guardar los datos en un archivo de extensión *.csv (ejemplo: “ReglasAso1.csv”).
	3.- En la ventana “consola” de <i>RStudio</i> , ejecutar el algoritmo (ver Apéndice D.4. <i>eclat</i>), relacionada a la base de datos con la que esté trabajando.
	4.- Observar los resultados y reflexionar.
Evaluación	“Para verificar la comprensión docente, del uso de esta técnica en contextos educativos, se solicita al usuario que conteste la encuesta del siguiente enlace: https://bit.ly/37X2CJw ”.

ANEXO 3

A. Modelo de encuesta aplicada a los docentes



Departamento de Postgrado e Investigación

Tema: Simulaciones digitales para facilitar la comprensión docente en: Análisis Estadístico Implicativo y el *Learning Analytics*

ENCUESTA PARA APLICAR A LOS DOCENTES

Estimado (a) docente. El cuestionario que, se entrega es parte de una investigación de Maestría en Ciencias de la Educación. El objetivo es diagnosticar tu conocimiento en aspectos generales de análisis estadístico, herramientas TIC y el software estadístico R. Por lo cual, solicito contestar las siguientes preguntas con absoluta sinceridad.

Los resultados ayudarán en la elaboración de una propuesta, para el uso de las técnicas de estadística implicativa y el *Learning Analytics*, en el quehacer educativo.

Instrucción. Según tu autoevaluación, seleccione la opción de respuesta que consideres adecuado en cada enunciado.

Datos generales del docente

Fecha: _____ Sexo: Femenino Masculino Edad: ____ años

Años de servicio docente: _____ N° de estudiantes a quienes imparte asignatura (s): _____

Asignatura con mayor carga horaria: _____

CUESTIONARIO

1) ¿Con cuáles herramientas TIC, se encuentra familiarizado? *Puede escoger 1 o varias.*

Google Apps (Derive, Gmail, Office Online, etc)	<input type="checkbox"/>	Dropbox	<input type="checkbox"/>
Twitter	<input type="checkbox"/>	Edmodo	<input type="checkbox"/>
Skype	<input type="checkbox"/>	ClassDojo	<input type="checkbox"/>
YouTube	<input type="checkbox"/>	WordPress	<input type="checkbox"/>
Evernote	<input type="checkbox"/>	Socrative	<input type="checkbox"/>
		Otro: _____	

2) ¿Qué herramientas TIC utiliza para el dictado de clase? *Puede escoger 1 o varias.*

Google Apps (Derive, Gmail, Office Online, etc)	<input type="checkbox"/>	YouTube	<input type="checkbox"/>
Twitter	<input type="checkbox"/>	Evernote	<input type="checkbox"/>
Skype	<input type="checkbox"/>	Dropbox	<input type="checkbox"/>
		Edmodo	<input type="checkbox"/>

ClassDojo Socrative
 WordPress Otro: _____

3) ¿Utiliza la herramienta *Excel* para procesar datos educativos?

Nunca	Casi nunca	Ocasionalmente	Casi todos los días	Todos los días
1	2	3	4	5

4) ¿Ha utilizado el *software* estadístico *R*?

Nunca	Casi nunca	Ocasionalmente	Casi todos los días	Todos los días
1	2	3	4	5

5) Seleccione los *medios digitales* que utiliza en sus clases. *Puede escoger 1 o varias.*

Plataformas educativas Videojuegos
 Medios audiovisuales Libros digitales
 Páginas web Otro (detalle): _____
 Medios de comunicación digital

6) ¿Usted realiza análisis estadístico de los avances de aprendizaje que tienen los diferentes grupos de estudiantes a los que dicta clases?

Nunca	Casi nunca	Ocasionalmente	Casi todos los días	Todos los días
1	2	3	4	5

7) ¿Considera importante manejar datos estadísticos de los avances de sus estudiantes en su asignatura?

No es importante	Poco importante	Neutral	Importante	Muy importante
1	2	3	4	5

8) ¿Conoce para qué sirve el análisis estadístico implicativo?

Desconozco del tema	Algo conozco del tema	Conozco del tema
1	2	3

9) ¿Conoce las técnicas de analíticas de aprendizajes, también, conocidas por su sigla en inglés *learning analytics*?

Desconozco del tema	Algo conozco del tema	Conozco del tema
1	2	3

10) ¿Conoce sobre los datos educativos que pueden observar usando técnicas de estadística implicativa combinadas con las técnicas de analíticas de aprendizajes?

Desconozco del tema	Algo conozco del tema	Conozco del tema
1	2	3

Muchas gracias por sus respuestas

B. Validación del Instrumento de diagnóstico por juicio de expertos

Evaluador 1

Ambato, 24 de mayo del 2022

Magister.

Diana Guamboa

Presente. -

De mi consideración:

Deseando éxitos en sus labores, por medio del presente y conociendo su alto espíritu de colaboración en el área de Investigación Científica, con la finalidad de seguir fortaleciendo la propuesta: SIMULACIONES DIGITALES PARA FACILITAR LA COMPRENSIÓN DOCENTE EN ANALISIS ESTADISTICO IMPLICATIVO Y EL *LEARNING ANALYTICS*, a darse en la PUCE-Sede Ambato. Solicito de la manera más comedida, emitir su criterio y más información en calidad de **Revisor Experto** del **Instrumento de Recolección de Datos (ENCUESTA)**, para la cual, adjunto a la presente, la documentación pertinente (Doc1 y 2.- **Leer**, Doc3.- **Llenar** y **enviar** a mi WhatsApp o al correo y jose.mullo@educación.gob.ec).

Por la atención favorable que digne dar a la presente, anticipo mis más sinceros agradecimientos.

Atentamente,



Dr. José Luis Mullo G.

0602986887

jose.mullo@educación.gob.ec

0985093883

Evaluador 2

Ambato, 24 de mayo del 2022

Magister.

Efrén Santana


Presente. -

De mi consideración:

Deseando éxitos en sus labores, por medio del presente y conociendo su alto espíritu de colaboración en el área de Investigación Científica, con la finalidad de seguir fortaleciendo la propuesta: SIMULACIONES DIGITALES PARA FACILITAR LA COMPRENSIÓN DOCENTE EN ANALISIS ESTADISTICO IMPLICATIVO Y EL *LEARNING ANALYTICS*, a darse en la PUCE-Sede Ambato. Solicito de la manera más comedida, emitir su criterio y más información en calidad de **Revisor Experto** del **Instrumento de Recolección de Datos (ENCUESTA)**, para la cual, adjunto a la presente, la documentación pertinente (Doc1 y 2.- **Leer**, Doc3.- **Llenar** y **enviar** a mi WhatsApp o al correo y jose.mullo@educación.gob.ec).

Por la atención favorable que digno dar a la presente, anticipo mis más sinceros agradecimientos.

Atentamente,



Dr. José Luis Mullo G.

0602986887

jose.mullo@educación.gob.ec

0985093883

Evaluador 3

Ambato, 24 de mayo del 2022

Magister.

Paola Villalón

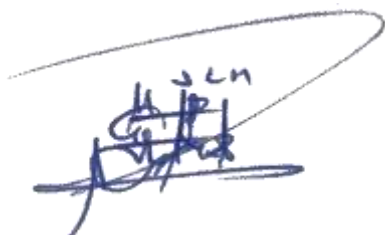
Presente. -

De mi consideración:

Deseando éxitos en sus labores, por medio del presente y conociendo su alto espíritu de colaboración en el área de Investigación Científica, con la finalidad de seguir fortaleciendo la propuesta: SIMULACIONES DIGITALES PARA FACILITAR LA COMPRENSIÓN DOCENTE EN ANALISIS ESTADISTICO IMPLICATIVO Y EL *LEARNING ANALYTICS*, a darse en la PUCE-Sede Ambato. Solicito de la manera más comedida, emitir su criterio y más información en calidad de **Revisor Experto** del **Instrumento de Recolección de Datos (ENCUESTA)**, para la cual, adjunto a la presente, la documentación pertinente (Doc1 y 2.- **Leer**, Doc3.- **Llenar** y **enviar** a mi WhatsApp o al correo y jose.mullo@educación.gob.ec).

Por la atención favorable que digno dar a la presente, anticipo mis más sinceros agradecimientos.

Atentamente,



Dr. José Luis Mullo G.

0602986887

jose.mullo@educación.gob.ec

0985093883

FICHA PARA VALIDACIÓN DE INSTRUMENTOS DE INVESTIGACIÓN
Para revisor

INDICADORES	CRITERIOS	DEFICIENTE				BAJA				REGULAR				BUENO				MUY BUENO					
		5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100		
1.- Claridad	Se ha formulado con lenguaje apropiado.																				x		
2.- Objetividad	Está expresado en conductas observables.																					x	
3.- Actualidad	Está acorde a los aportes recientes en la disciplina de estudio.																					x	
4.- Organización	Hay una organización lógica.																					x	
5.- Suficiencia	Comprende las dimensiones de la investigación en cantidad y calidad.																					x	
6.- Intencionalidad	Está adecuado para valorar la variable seleccionada.																					x	
7.- Consistencia	Está basado en aspectos teóricos y científicos.																					x	
8.- Coherencia	Hay relación entre indicadores dimensiones e índice.																					x	
9.- Metodología	El instrumento se relaciona con el método planteado en el proyecto.																					x	
10.- Aplicabilidad	El instrumento es de fácil aplicación.																						x

REVISOR

Nombres y Apellidos: Villalón Muñoz Paola Alejandra
 Título de Tercer Nivel: Doctor en Matemáticas
 Título de Cuarto Nivel: Magister en Matemática Básica
 Cédula: 0601908510

Evaluador 4

Ambato, 24 de mayo del 2022

Magister.

Susana Zurita

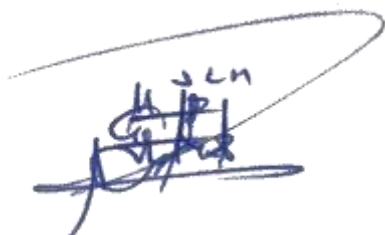
Presente. -

De mi consideración:

Deseando éxitos en sus labores, por medio del presente y conociendo su alto espíritu de colaboración en el área de Investigación Científica, con la finalidad de seguir fortaleciendo la propuesta: SIMULACIONES DIGITALES PARA FACILITAR LA COMPRENSIÓN DOCENTE EN ANALISIS ESTADISTICO IMPLICATIVO Y EL *LEARNING ANALYTICS*, a darse en la PUCE-Sede Ambato. Solicito de la manera más comedida, emitir su criterio y más información en calidad de **Revisor Experto** del **Instrumento de Recolección de Datos (ENCUESTA)**, para la cual, adjunto a la presente, la documentación pertinente (Doc1 y 2.- **Leer**, Doc3.- **Llenar** y **enviar** a mi WhatsApp o al correo y jose.mullo@educación.gob.ec).

Por la atención favorable que digno dar a la presente, anticipo mis más sinceros agradecimientos.

Atentamente,



Dr. José Luis Mullo G.

0602986887

jose.mullo@educación.gob.ec

0985093883

FICHA PARA VALIDACIÓN DE INSTRUMENTOS DE INVESTIGACIÓN
Para revisor

INDICADORES	CRITERIOS	DEFICIENTE		BAJA				REGULAR				BUENO				MUY BUENO					
		5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100
1.- Claridad	Se ha formulado con lenguaje apropiado.																				X
2.- Objetividad	Está expresado en conductas observables.																				X
3.- Actualidad	Está acorde a los aportes recientes en la disciplina de estudio.																				X
4.- Organización	Hay una organización lógica.																				X
5.- Suficiencia	Comprende las dimensiones de la investigación en cantidad y calidad.																				X
6.- Intencionalidad	Está adecuado para valorar la variable seleccionada.																				X
7.- Consistencia	Está basado en aspectos teóricos y científicos.																				X
8.- Coherencia	Hay relación entre indicadores dimensiones e índice.																				X
9.- Metodología	El instrumento se relaciona con el método planteado en el proyecto.																				X
10.- Aplicabilidad	El instrumento es de fácil aplicación.																				X

REVISOR

Nombres y Apellidos: SUSANA DEL ROCIO ZURITA LOPEZ

Título de Tercer Nivel: LICENCIADA EN CIENCIAS DE LA EDUCACION ESPECIALIZACION FISICA Y MATEMATICA

Título de Cuarto Nivel: MAGISTER EN HERRAMIENTAS TECNOLOGICAS PARA LA PRACTICA DOCENTE

Cédula: 1802144319

Evaluador 5

Ambato, 24 de mayo del 2022

Magíster.

Isabel Escudero

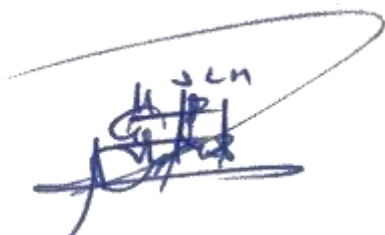
Presente. -

De mi consideración:

Deseando éxitos en sus labores, por medio del presente y conociendo su alto espíritu de colaboración en el área de Investigación Científica, con la finalidad de seguir fortaleciendo la propuesta: SIMULACIONES DIGITALES PARA FACILITAR LA COMPRENSIÓN DOCENTE EN ANALISIS ESTADISTICO IMPLICATIVO Y EL *LEARNING ANALYTICS*, a darse en la PUCE-Sede Ambato. Solicito de la manera más comedida, emitir su criterio y más información en calidad de **Revisor Experto** del **Instrumento de Recolección de Datos (ENCUESTA)**, para la cual, adjunto a la presente, la documentación pertinente (Doc1 y 2.- **Leer**, Doc3.- **Llenar** y **enviar** a mi WhatsApp o al correo y jose.mullo@educación.gob.ec).

Por la atención favorable que digno dar a la presente, anticipo mis más sinceros agradecimientos.

Atentamente,



Dr. José Luis Mullo G.

0602986887

jose.mullo@educación.gob.ec

0985093883

TIPO DE INSTRUMENTO DE EVALUACIÓN: CUESTIONARIO
Para revisor

INDICADORES	CRITERIOS	DEFICIENTE 0-20%		BAJA 21-40%				REGULAR 41-60%				BUENO 61-90%				MUY BUENO 81-100%					
		5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100
1.- Claridad	Se ha formulado con lenguaje apropiado.																				X
2.- Objetividad	Está expresado en conductas observables.																				X
3.- Actualidad	Está acorde a los aportes recientes en la disciplina de estudio.																				X
4.- Organización	Hay una organización lógica.																				X
5.- Suficiencia	Comprende las dimensiones de la investigación en cantidad y calidad.																				X
6.- Intencionalidad	Está adecuado para valorar la variable seleccionada.																				X
7.- Consistencia	Está basado en aspectos teóricos y científicos.																				X
8.- Coherencia	Hay relación entre indicadores dimensiones e índice.																				X
9.- Metodología	El instrumento se relaciona con el método planteado en el proyecto.																				X
10.- Aplicabilidad	El instrumento es de fácil aplicación.																				X

DATOS GENERALES DEL REVISOR

AMALIA ISABEL ESCUDERO VILLA

Nombres y Apellidos: Isabel Escudero V.

Título de Tercer Nivel: Ing. Estadística Informática

Título de Cuarto Nivel: Estadística Matemática y Aplicada

Cédula: 0603980467

Evaluador 6

Ambato, 20 de agosto del 2022

Doctor.

Fernando Procel

Presente. -

De mi consideración:

Deseando éxitos en sus labores, por medio del presente y conociendo su alto espíritu de colaboración en el área de Investigación Científica, con la finalidad de seguir fortaleciendo la propuesta: SIMULACIONES DIGITALES PARA FACILITAR LA COMPRENSIÓN DOCENTE EN ANALISIS ESTADISTICO IMPLICATIVO Y EL *LEARNING ANALYTICS*, a darse en la PUCE-Sede Ambato. Solicito de la manera más comedida, emitir su criterio y más información en calidad de **Revisor Experto** del **Instrumento de Recolección de Datos (ENCUESTA)**, para la cual, adjunto a la presente, la documentación pertinente (Doc1 y 2.- **Leer**, Doc3.- **Llenar** y **enviar** a mi WhatsApp o al correo y jose.mullo@educación.gob.ec).

Por la atención favorable que digno dar a la presente, anticipo mis más sinceros agradecimientos.

Atentamente,



Dr. José Luis Mullo G.

0602986887

jose.mullo@educación.gob.ec

0985093883

TIPO DE INSTRUMENTO DE EVALUACIÓN: CUESTIONARIO
Para revisor

INDICADORES	CRITERIOS	DEFICIENTE 0-20%				BAJA 21-40%				REGULAR 41-60%				BUENO 61-90%				MUY BUENO 91-100%			
		5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100
1.- Claridad	Se ha formulado con lenguaje apropiado.																				
2.- Objetividad	Está expresado en conductas observables.																				
3.- Actualidad	Está acorde a los aportes recientes en la disciplina de estudio.																				
4.- Organización	Hay una organización lógica.																				
5.- Suficiencia	Comprende las dimensiones de la investigación en cantidad y calidad.																				
6.- Intencionalidad	Está adecuado para valorar la variable seleccionada.																				
7.- Consistencia	Está basado en aspectos teóricos y científicos.																				
8.- Coherencia	Hay relación entre indicadores dimensiones e índice.																				
9.- Metodología	El instrumento se relaciona con el método planteado en el proyecto.																				
10.- Aplicabilidad	El instrumento es de fácil aplicación.																				

DATOS GENERALES DEL REVISOR

Nombres y Apellidos: Fernando Efraín Procel Orozco
 Título de Tercer Nivel: Doctor en Matemática
 Título de Cuarto Nivel: Magister en Ciencias de la Educación, mención Aprendizaje de la Matemática
 Cédula: 0601878291

Validez del Instrumento de Recolección de Datos (ENCUESTA)										
Evaluadores	INDICADORES									
	Claridad	Objetividad	Actualidad	Organización	Suficiencia	Intencionalidad	Consistencia	Coherencia	Metodología	Aplicabilidad
Evaluador 1	90	90	90	90	90	95	100	95	100	95
Evaluador 2	100	100	100	100	100	100	100	100	100	100
Evaluador 3	90	95	95	95	90	90	90	90	90	100
Evaluador 4	100	100	100	100	100	100	100	100	100	100
Evaluador 5	100	100	95	100	100	95	100	95	100	100
Evaluador 6	100	95	95	100	95	95	95	95	95	95
V de Aiken	0,97	0,97	0,96	0,98	0,96	0,96	0,98	0,96	0,98	0,98
V de Aiken Global	0,97									

C. Modelo prueba pre y post test para aplicar a los docentes



PONTIFICIA UNIVERSIDAD CATÓLICA DE ECUADOR
Sede AMBATO

Departamento de Postgrado e Investigación

Tema: Simulaciones digitales para facilitar la comprensión docente en: Análisis Estadístico Implicativo y el *Learning Analytics*

INSTRUMENTO DE EVALUACIÓN DE ESTADÍSTICA IMPLICATIVA PARA LA COMUNIDAD DE DOCENTES

Estimado docente, con la finalidad de conocer tu nivel de conocimiento en temas de Estadística Implicativa (ASI) y el Learning Analytics (LA). Solicito contestar las siguientes preguntas con absoluta sinceridad.

Instrucción. Favor de colocar un en la opción de respuesta que usted considere correcto. Cada pregunta tiene una valoración de UN PUNTO (1p).

Datos generales del docente

Fecha de Aplicación: _____

Apellidos y Nombres: _____

A. Preguntas relacionadas con las Técnicas del ASI y LA

- 1) El *Índice de similaridad*:
 - a) Es la similaridad de conjuntos (a_i, a_j).
 - b) Es la similaridad de puntos (a_i, a_j).
 - c) Representa las similaridades de una pareja (a_i, a_j) al nivel cero de la jerarquía.
- 2) La *cohesión* es:
 - a) Es la adherencia equilibrada de $\neg a \vee b$ sobre $a \wedge \neg b$.
 - b) Una medida de desequilibrio de las frecuencias de los eventos $\neg a \vee b$ y $a \wedge \neg b$ a favor del primero.
 - c) Es la congruencia relacionada de $\neg a \vee b$ sobre $a \wedge \neg b$.
- 3) El *grafo de implicación* es:
 - a) Un grafo orientado, ponderado (hacia las intensidades de implicación).
 - b) Un grafo no orientado, no ponderado (hacia las intensidades de implicación).
 - c) Un grafo orientado, no ponderado (hacia las intensidades de implicación).
- 4) A un grupo de 20 estudiantes, se analiza el uso de las diferentes plataformas, la tabla de análisis de similaridad de eventos, se asemeja el siguiente cuadro:

Un primer acercamiento al cálculo del Índice de Similaridad.					
VARIABLES		Card($A_i \cap A_j$)	$\frac{n_i \cdot n_j}{n}$	Kc	s(a_i, a_j)
KanA	11 Ytube 11	10	6,05	1,61	0,9458524
KanA	11 Face 10	6	5,5	0,21	0,5844148
KanA	11 Whp 10	7	5,5	0,64	0,7387844

El valor de *Índice de Similitud* entre el uso de las plataformas: *Khan Academy* y *YouTube* es:

- a) 0,9458.
- b) 0,5844.
- c) 0,7387.

5) Para transformar la tabla de similitudes, desde un nivel cero hacia un nivel uno, se procede así:

Tabla J0.- Nivel cero de jerarquía.						Tabla J1.- Nivel uno de jerarquía.					
S(a _i ,a _j)	HIP	JAZ	HEA	REG	PUN	S(a _i ,a _j)	HIP, PUN	JAZ	HEA	REG	
HIP		0,5844148	0,73878436	0,85678899	0,94585243	HIP, PUN		0,54580233	0,73408737	0,73408737	
JAZ	0,5844148		0,91014375	0,91014375	0,73878436	JAZ	0,54580233		0,91054375	0,91014375	
HEA	0,73878436	0,91014375		0,91014375	0,85678899	HEA	0,73408737	0,91054375		0,91014375	
REG	0,85678899	0,91014375	0,91014375		0,85678899	REG	0,73408737	0,91014375	0,91014375		
PUN	0,94585243	0,73878436	0,85678899	0,85678899							

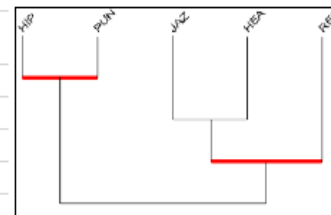
Matriz de similitud al nivel cero Matriz de similitud al nivel uno

Max= 0,94585243 Max= 0,91014375

- a) Se halla el valor **mínimo** de la **tabla J0**, se eliminan la fila y columna correspondiente a este valor, los demás valores de la nueva tabla, son el **cuadrado** del: valor máximo entre el valor de la celda y el valor a eliminarse.
- b) Se halla el valor **máximo** de la **tabla J0**, se eliminan la fila y columna correspondiente a este valor, los demás valores de la nueva tabla, son el **cuadrado** del: valor máximo entre el valor de la celda y el valor a eliminarse.
- c) Se halla el valor **máximo** de la **tabla J0**, se eliminan la fila y columna correspondiente a este valor, los demás valores de la nueva tabla, vienen de la **raíz cuadrada** del: valor máximo entre el valor de la celda y el valor a eliminarse.

6) Cómo, se relacionan los valores V de la siguiente tabla, con el diagrama del árbol:

Nivel	s _k	r _k	Card[G(Ω) ∩ [ΣΠ _k x RΠ _k]]	S(Ω, k)	V(Ω, k)
1	19	1	18	1,47408695	1,47408695
2	18	2	29	1,38586973	-0,08821721
3	16	4	51	1,79533125	0,40946151
4	10	10	38	-0,90711474	-2,70244598

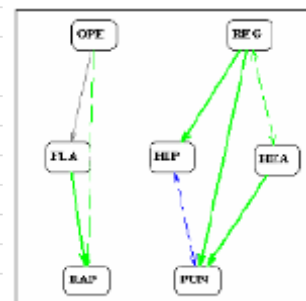


Valores del índice centrado y de la función V(Ω, k).

- a) Los valores **V positivos**, se visualizan en el diagrama de árbol, con el **color negro**.
- b) Los valores **V positivos**, se visualizan en el diagrama de árbol, con el **color rojo**.
- c) Los valores **V negativos**, se visualizan en el diagrama de árbol, con el **color rojo**.

7) Para la siguiente tabla de valores de la *Intensidad Implicativa* y su correspondiente *Grafo Implicativo*.

6.- Valores Intensidad Implicativa: φ							
φ(A _i ,A _j)	HIP'	OPE'	FLA'	HEA'	REG'	RAP'	PUN'
HIP	0,00000	0,282553	0,15781	0,680576	0,842024	0,130137	0,974314
OPE	0,18106	0,00000	0,733052	0,234668	0,234668	0,878423	0,18106
FLA	0,142627	0,671648	0,00000	0,214218	0,585158	0,930825	0,142627
HEA	0,690122	0,317129	0,214218	0,00000	0,90874	0,086693	0,860566
REG	0,860566	0,317129	0,585158	0,90874	0,00000	0,195792	0,860566
RAP	0,145454	0,747994	0,892913	0,113331	0,227728	0,00000	0,145454
PUN	0,974314	0,282553	0,15781	0,842024	0,842024	0,130137	0,00000



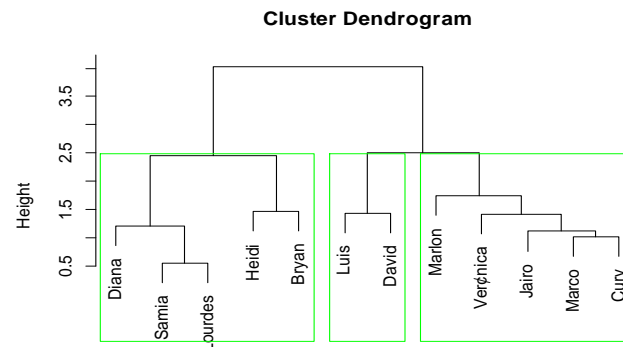
La interpretación acertada es:

- a) φ(HIP, PUN) = 0,9743; φ(FLA, RAP) = 0,1426; φ(REG, HIP) = 0,8605 .
- b) φ(HIP, PUN) = 0,9743; φ(FLA, RAP) = 0,9308; φ(REG, HIP) = 0,9087 .
- c) φ(HIP, PUN) = 0,9743; φ(FLA, RAP) = 0,9308; φ(REG, HIP) = 0,8605 .

- 8) Al aplicar el algoritmo *dendro.variables* en *RStudio* a una tabla de notas,

Estudiantes	Matemáticas	Lengua_Castellana	Lengua_Kichwa	Ingls	F.sica
1 Samia	7.65	8.37	8.15	8.65	8.15
2 Luis	6.34	7.27	5.98	7.06	7.07
3 Lourdes	7.67	8.45	8.16	8.52	7.94
4 Verónica	7.08	7.89	5.77	7.92	7.35
5 Marco	7.12	7.97	6.69	7.52	7.38
6 Heidi	9.50	9.27	8.99	9.34	9.67
7 Cury	6.98	7.64	6.39	8.15	7.73

Se refleja la imagen:



Se puede observar:

- a) Agrupación de personas, de acuerdo a sus gustos.
 b) Grupos de estudiantes, de acuerdo a sus procedencias.
 c) La conformación de grupos de estudiantes, de acuerdo a sus notas.
- 9) Al aplicar el algoritmo *implicative.graph* en *RCHIC* a una tabla de notas de una asignatura, se refleja la imagen:



Cuál de las conclusiones es la más acertada:

- a) Las temáticas de *EstD2* no implica a las temáticas de *Geom2* y viceversa, con una intensidad de implicación del 57% y 75%.
 b) Las temáticas de *EstD2* implica a las temáticas de *Geom2* y viceversa, con una intensidad de implicación del 57% y 75%.
 c) Las temáticas de *EstD2* implica a las temáticas de *Geom2* y viceversa, con una intensidad de implicación del 50% y 57%.
- 10) Al aplicar el algoritmo *apriori* en *RStudio* a una tabla de actividades de estudiantes, se refleja el reporte:

Transaccion	Actividad
1	ActividadApr
1	Lectura
1	ArteyManualidades
1	Escritura
1	VideosRefuerzo
1	Chat
2	ActividadApr
2	Lectura
2	Escritura
3	ActividadApr
3	Lectura
3	Escritura
3	VideosRefuerzo
4	ActividadApr

{ArteyManualidades, Escritura, VideosRefuerzo}	=>	{Lectura}	0.3	1 1.428571	3
{ArteyManualidades, Lectura, VideosRefuerzo}	=>	{Escritura}	0.3	1 1.428571	3

Cuál de las conclusiones es la más acertada:

- a) Con una confianza del 0.7: Si un estudiante realiza las actividades de arte y manualidades, lectura y observa videos de refuerzo, podría lograr las destrezas de la escritura.
- b) Con una confianza del 0.3: Si un estudiante realiza las actividades de arte y manualidades, lectura y observa videos de refuerzo, podría lograr las destrezas de la escritura.
- c) Con una confianza del 0.5: Si un estudiante realiza las actividades de arte y manualidades, lectura y observa videos de refuerzo, podría lograr las destrezas de la escritura.

B. Comentarios y sugerencias

Muchas gracias por su colaboración.

D. Validación del Instrumento Pre Test y Post Test: Juicio del experto

Evaluador 1

Ambato, 31 de agosto del 2022

Docente evaluador:

Deseando éxitos en sus labores, por medio del presente se le pide su aporte, que luego de un análisis riguroso de los Items del *Instrumento de Evaluación de Estadística Implicativa para Docentes de Bachillerato* que le mostramos. Marque con una X el casillero en la tabla de validación de contenido, conforme a su criterio y experiencia profesional, denotando si cuenta o no con los requisitos mínimos de formulación para su aplicación posterior.

INFORMACION GENERAL DEL INVESTIGADOR

Investigador	Dr. José Luis Múlo Gualán
Tema del Proyecto de Investigación	"Simulaciones digitales para facilitar la comprensión docente de las técnicas del análisis estadístico implicativo y el <i>learning analytics</i> "
Programa de Estudio	Maestría en Ciencias de la Educación
Institución	Pontificia Universidad Católica del Ecuador Sede Ambato
Objetivo General de la Investigación	Diseñar un sistema de simulaciones digitales para que el profesor de nivel de bachillerato, familiarice con las técnicas comunes entre el Análisis Estadístico Implicativo (AEI) y el <i>Learning Analytics (LA)</i>
Herramienta a evaluar	Instrumento de evaluación de estadística implicativa para docentes de bachillerato
Objetivo del Instrumento	<i>Medir la aproximación docente a las técnicas comunes entre la estadística implicativa y el learning analytics.</i>

INFORMACION GENERAL DEL EVALUADOR

Evaluador	Mg. Efrén Bonifacio Santana
Universidad o Institución Educativa, a la que pertenece	Unidad Educativa ATAHUALPA
Cargo	Docente
Años de experiencia en el cargo	25
Grado Académico	Tercer Nivel () Cuarto Nivel (X) PhD ()
Nivel o área a la que pertenece en la Universidad o Institución Educativa	Área de Matemáticas Nivel: Bachillerato

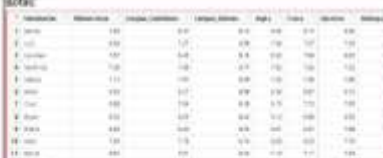
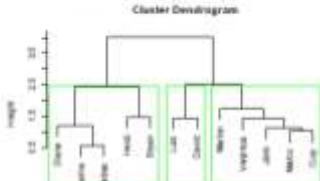

TABLA DE VALIDACIÓN

NOTA.- Para cada pregunta se considera la escala de 1 al 5, donde:

1.- Muy poco. 2.- Poco. 3.- Regular. 4.- Aceptable. 5.- Muy aceptable

Nº	ITEMS	Puntajes				
		1	2	3	4	5
1	El Índice de similitud: a) <input type="checkbox"/> Es la similitud de conjuntos (A, B) . b) <input type="checkbox"/> Es la similitud de partes (x, y) . c) <input type="checkbox"/> Representa las similitudes de cada pareja (x, y) al nivel cero de la jerarquía.					X
2	La relación es: a) <input type="checkbox"/> Es la adiferencia equilibrada de $-a+b$ sobre $a+b$. b) <input type="checkbox"/> Una medida de desequilibrio de las frecuencias de los eventos $\bar{x}+b$ y $a+\bar{x}$ a favor del primer. c) <input type="checkbox"/> Es la congruencia relacionada de $-a+b$ sobre $a+b$.					X
3	El gráfico de implicación es: a) <input type="checkbox"/> Un gráfico orientado, poseído (hacia las atomidades de implicación). b) <input type="checkbox"/> Es un gráfico en coordenadas cartesianas y las atomidades de implicación. c) <input type="checkbox"/> Es un gráfico en coordenadas, asociado a las atomidades de implicación.					X
4	A un grupo de 20 estudiantes se analiza el uso de los diferentes plataformas, la tabla de análisis de similitud de eventos se muestra el siguiente cuadro:  El valor de Índice de Similitud entre el uso de las plataformas: <i>Edx Academy</i> y <i>EdxTube</i> es: a) <input type="checkbox"/> 0,9438 b) <input type="checkbox"/> 0,3844 c) <input type="checkbox"/> 0,7387					X
5	Para transformar la tabla de similitudes, desde un nivel cero hacia un nivel uno, se procede así:					X

<p>a) <input type="checkbox"/> Se halla el valor máximo de la tabla B), se eliminan la fila y columna correspondiente a este valor, los demás valores de la nueva tabla, con el cuadrado del: valor máximo entre el valor de la celda y el valor a eliminarse.</p> <p>b) <input type="checkbox"/> Se halla el valor máximo de la tabla B), se eliminan la fila y columna correspondiente a este valor, los demás valores de la nueva tabla, con el cuadrado del: valor máximo entre el valor de la celda y el valor a eliminarse.</p> <p>c) <input type="checkbox"/> Se halla el valor máximo de la tabla, se eliminan la fila y columna correspondiente a este valor, los demás valores de la nueva tabla, con el cuadrado del: valor máximo entre el valor de la celda y el valor a eliminarse.</p>					
<p>6) Cómo se relacionan los valores V de la siguiente tabla, con el diagrama del árbol:</p>  <p>a) <input type="checkbox"/> Los valores V positivos, se visualizan en el diagrama de árbol, con el color rojo.</p> <p>b) <input type="checkbox"/> Los valores V positivos, se visualizan en el diagrama de árbol, con el color negro.</p> <p>c) <input type="checkbox"/> Los valores V negativos, se visualizan en el diagrama de árbol, con el color rojo.</p>					X
<p>7) Para la siguiente tabla de valores de <i>Intensidad Implicativa</i> y su correspondiente Gráfico Implicativo:</p>  <p>La interpretación acertada es:</p> <p>a) <input type="checkbox"/> $\#(NP, PCN) = 0,743$, $\#(LA, RAP) = 0,428$, $\#(EDU, SEP) = 0,803$</p> <p>b) <input type="checkbox"/> $\#(NP, PCN) = 0,743$, $\#(LA, RAP) = 0,538$, $\#(EDU, SEP) = 0,807$</p> <p>c) <input type="checkbox"/> $\#(NP, PCN) = 0,743$, $\#(LA, RAP) = 0,538$, $\#(EDU, SEP) = 0,803$</p>					X
<p>8) Al aplicar el algoritmo descrito conjetura en <i>EdxTube</i> a una tabla de</p>					X

<p>NOTA:</p>  <p>Se refleja la imagen:</p>  <p>Se puede observar a:</p> <p>a) <input type="checkbox"/> Agrupación de personas, de acuerdo a sus gustos.</p> <p>b) <input type="checkbox"/> Grupos de estudiantes, de acuerdo a sus procedencias.</p> <p>c) <input type="checkbox"/> La conformación del grupo de estudiantes, de acuerdo a sus notas.</p>					
<p>9) Al aplicar el algoritmo implicativo graph en RCHDC a una tabla de notas: algunas temáticas en sus asignaturas, se refleja la imagen:</p>  <p>¿Cuál de las conclusiones es la más acertada?</p> <p>a) <input type="checkbox"/> Las temáticas de <i>EdxDU</i> no implica a las temáticas de <i>Geo2</i> y <i>viernes</i>, con una intensidad de implicación del 37% y 73%.</p> <p>b) <input type="checkbox"/> Las temáticas de <i>EdxDU</i> implica a las temáticas de <i>Geo2</i> y <i>viernes</i>, con una intensidad de implicación del 37% y 73%.</p> <p>c) <input type="checkbox"/> Las temáticas de <i>EdxDU</i> implica a las temáticas de <i>Geo2</i> y <i>viernes</i>, con una intensidad de implicación del 30% y 57%.</p>					X

10 Al aplicar el algoritmo *apriori* en *Fitchit* a una tabla de actividades de estudiantes, se refleja el reporte:

¿Cuál de las conclusiones es la más acertada?

a) Con una confianza del 0.7: Si un estudiante realiza las actividades de arte y manualidades, lectura y observa videos de refuerzo, podrá lograr las destrezas de la escritura.

b) Con una confianza del 0.3: Si un estudiante realiza las actividades de arte y manualidades, lectura y observa videos de refuerzo, podrá lograr las destrezas de la escritura.

c) Con una confianza del 3: Si un estudiante realiza las actividades de arte y manualidades, lectura y observa videos de refuerzo, podrá lograr las destrezas de la escritura.

OBSERVACIONES

El instrumento se ha elaborado correctamente y es pertinente.

Este documento, certifica la revisión y análisis del contenido: "Instrumento de evaluación de estadística implícita para docentes de bachillerato" de recolección de información.

Para la constancia de lo expuesto, firma:

Mg. Elvira Bonifacio Santos
C.I: 1801976828
Grado Académico: Magister en Docencia Matemática

Evaluador 2

Ambato, 31 de agosto del 2022

Docente evaluador:

Desearo éxitos en sus labores, por medio del presente se le pide su aporte, que luego de un análisis riguroso de los ítems del Instrumento de Evaluación de Estadística Implícita para Docentes de Bachillerato que le mostramos. Marque con una X el casillero en la tabla de validación de contenido, conforme a su criterio y experiencia profesional, denotando si cuenta o no con los requisitos mínimos de fundamentación para su aplicación posterior.

INFORMACION GENERAL DEL INVESTIGADOR

Investigador	Dr. José Luis Mulla Guzmán
Tema del Proyecto de Investigación	"Simulaciones digitales para facilitar la comprensión docente de las técnicas del análisis estadístico implícito y el <i>learning analytics</i> "
Programa de Estudio	Maestría en Ciencias de la Educación
Institución	Pontificia Universidad Católica del Ecuador Sede Ambato
Objetivo General de la Investigación	Diseñar un sistema de simulaciones digitales para que el profesor de nivel de bachillerato, familiarice con las técnicas comunes entre el Análisis Estadístico Implícito (AEI) y el <i>Learning Analytics</i> (LA).
Herramienta a evaluar	Instrumento de evaluación de estadística implícita para docentes de bachillerato
Objetivo del Instrumento	Medir la aptitud docente a las técnicas comunes entre la estadística implícita y el <i>learning analytics</i> .

INFORMACION GENERAL DEL EVALUADOR

Evaluador	Isabel Escudero
Universidad o Institución Educativa, a la que pertenece	ESPOCH
Cargo	Docente
Años de experiencia en el cargo	11 años
Grado Académico	Tercer Nivel (), Cuarto Nivel (x), PhD ()
Nivel o área a la que pertenece en la Universidad o Institución Educativa	Estadística

TABLA DE VALIDACIÓN

NOTA.- Para cada pregunta se considera la escala de 1 al 5, donde:

1.- Muy poca. 2.- Poca. 3.- Regular. 4.- Aceptable. 5.- Muy aceptable

N°	ITEMS	Puntajes				
		1	2	3	4	5
1	El Índice de similitud. <ul style="list-style-type: none"> a) <input type="checkbox"/> Es la similitud de conjuntos (x, y). b) <input type="checkbox"/> Es la similitud de puntos (x, y). c) <input type="checkbox"/> Representa las similitudes de cada pareja (x, y) al nivel cero de la jerarquía. 					X
2	La cobertura es: <ul style="list-style-type: none"> a) <input type="checkbox"/> Es la adherencia equilibrada de $-a+b$ sobre $a+b$. b) <input type="checkbox"/> Una medida de desequilibrio de las frecuencias de los eventos $x=b$ y $x=b$ a favor del primer. c) <input type="checkbox"/> Es la congruencia relacionada de $-a+b$ sobre $a+b$. 					X
3	El gráfico de implicación es: <ul style="list-style-type: none"> a) <input type="checkbox"/> Un gráfico orientado, ponderado (hacia las instancias de implicación). b) <input type="checkbox"/> Es un gráfico en coordenadas cartesianas y las instancias de implicación. c) <input type="checkbox"/> Es un gráfico en coordenadas, asociados a las instancias de implicación. 					X
4	A un grupo de 20 estudiantes se analiza el uso de las diferentes plataformas, la tabla de análisis de similitud de eventos se muestra el siguiente cuadro: <p>El valor de Índice de Similitud entre el uso de las plataformas: <i>Don Academy</i> y <i>TeaTiba</i> es:</p> <ul style="list-style-type: none"> a) <input type="checkbox"/> 0.9438 b) <input type="checkbox"/> 0.7844 c) <input type="checkbox"/> 0.7387 					X
5	Para transformar la tabla de similitudes, desde un nivel cero hacia un nivel uno, se procede así:					x

Evaluador 3

Ambato, 31 de agosto del 2022

Docente evaluador:

Deseario éxitos en sus labores, por medio del presente se le pide su aporte, que luego de un análisis riguroso de los ítems del Instrumento de Evaluación de Estadística Implicativa para Docentes de Bachillerato que le mostramos. Marque con una X el casillero en la tabla de validación de contenido, conforme a su criterio y experiencia profesional, denotando si cuenta o no con los requisitos mínimos de formulación para su aplicación posterior.

INFORMACION GENERAL DEL INVESTIGADOR

Investigador	Dr. José Luis Muñoz Guzmán
Tema del Proyecto de Investigación	"Simulaciones digitales para facilitar la comprensión docente de las técnicas del análisis estadístico implicativo y el learning analytics"
Programa de Estudio	Maestría en Ciencias de la Educación
Institución	Pontificia Universidad Católica del Ecuador Sede Ambato
Objetivo General de la Investigación	Diseñar un sistema de simulaciones digitales para que el profesor de nivel de bachillerato, familiarice con las técnicas comunes entre el Análisis Estadístico Implicativo (AEI) y el Learning Analytics (LA)
Herramienta a evaluar	Instrumento de evaluación de estadística implicativa para docentes de bachillerato
Objetivo del Instrumento	Medir la apropiación docente a las técnicas comunes entre la estadística implicativa y el learning analytics.

INFORMACION GENERAL DEL EVALUADOR

Evaluador	Msc. Diana Alexandra Gamboa Castillo
Universidad o Institución Educativa, a la que pertenece	Unidad Educativa "Píthagos"
Cargo	Docente
Años de experiencia en el cargo	12 años
Grado Académico	Tercer Nivel () Cuarto Nivel (x) PhD ()
Nivel o área a la que pertenece en la Universidad o Institución Educativa	FISCAL

TABLA DE VALIDACION


NOTA: Para cada pregunta se considera la escala de 1 al 5, donde:

1- Muy poco; 2- Poco; 3- Regular; 4- Aceptable; 5- Muy aceptable

Nº	ITEMS	Puntaje				
		1	2	3	4	5
1	El Índice de similitud es: a) <input type="checkbox"/> Es la similitud de conjuntos (A, B) . b) <input type="checkbox"/> Es la similitud de puntos (a, b) . c) <input type="checkbox"/> Representa las similitudes de cada grupo (a, b) al nivel uno de la jerarquía.				X	
2	La relación es: a) <input type="checkbox"/> Es la similitud equilibrada de $-a+b$ tales $a>=b$. b) <input type="checkbox"/> Una medida de desequilibrio de las frecuencias de los eventos $b < a$ y $a < b$ a favor del primero. c) <input type="checkbox"/> Es la congruencia relacionada de $-a+b$ tales $a>=b$.				X	
3	El gráfico de implicación es: a) <input type="checkbox"/> Un gráfico orientado, ponderado (hacia las actividades de aplicación). b) <input type="checkbox"/> Un gráfico no orientado, no ponderado (hacia las actividades de aplicación). c) <input type="checkbox"/> Un gráfico orientado, no ponderado (hacia las actividades de aplicación).				X	X
4	A un grupo de 20 estudiantes se analiza el uso de los diferentes plataformas, la tabla de análisis de similitud de eventos se muestra el siguiente cuadro: El valor de Índice de similitud entre el uso de las plataformas Khan Academy y YouTube es: a) <input type="checkbox"/> 0,3438 b) <input type="checkbox"/> 0,3844 c) <input type="checkbox"/> 0,7387					X
5	Para transformar la tabla de similitudes, desde un nivel cero hacia un nivel uno, se procede así:					

<p>a) <input type="checkbox"/> Se halla el valor máximo de la tabla M, se eliminan la fila y columna correspondiente a este valor, los demás valores de la nueva tabla, son el cuadrado del valor máximo entre el valor de la celda y el valor a eliminarse.</p> <p>b) <input type="checkbox"/> Se halla el valor máximo de la tabla M, se eliminan la fila y columna correspondiente a este valor, los demás valores de la nueva tabla, son el cuadrado del valor máximo entre el valor de la celda y el valor a eliminarse.</p> <p>c) <input type="checkbox"/> Se halla el valor máximo de la tabla, se eliminan la fila y columna correspondiente a este valor, los demás valores de la nueva tabla, vienen de la raíz cuadrada del valor máximo entre el valor de la celda y el valor a eliminarse.</p>					X	
<p>6. Como se relacionan los valores V de la siguiente tabla, con el diagrama del árbol:</p> <p>a) <input type="checkbox"/> Los valores V positivos, se visualizan en el diagrama de árbol, con el color rojo.</p> <p>b) <input type="checkbox"/> Los valores V positivos, se visualizan en el diagrama de árbol, con el color azul.</p> <p>c) <input type="checkbox"/> Los valores V negativos, se visualizan en el diagrama de árbol, con el color rojo.</p>					X	X
<p>7. Para la siguiente tabla de valores de Intensidad Implicativa y su correspondiente Gráfico Implicativo:</p> <p>La interpretación correcta es:</p> <p>a) <input type="checkbox"/> $A(HP, PCN)=0,7345$, $A(FL, RAP)=0,1426$, $A(EO, JEP)=0,2801$</p> <p>b) <input type="checkbox"/> $A(HP, PCN)=0,7345$, $A(FL, RAP)=0,1426$, $A(EO, JEP)=0,6807$</p> <p>c) <input type="checkbox"/> $A(HP, PCN)=0,7345$, $A(FL, RAP)=0,1426$, $A(EO, JEP)=0,2807$</p>					X	X
8. Al aplicar el algoritmo de árbol correlado en Etluco a una tabla de						

<p>8. Al aplicar el algoritmo de árbol correlado en Etluco a una tabla de datos, algunos temáticas en sus asignaturas, se refleja la imagen:</p> <p>Se puede observar a:</p> <p>a) <input type="checkbox"/> Agrupación de personas, de acuerdo a sus gustos.</p> <p>b) <input type="checkbox"/> Grupos de estudiantes, de acuerdo a sus procedencias.</p> <p>c) <input type="checkbox"/> La conformación del grupo de estudiantes, de acuerdo a sus notas.</p>						X
<p>9. Al aplicar el algoritmo implicativo graph en RC3DC a una tabla de datos, algunas temáticas en sus asignaturas, se refleja la imagen:</p> <p>¿Cuál de las conclusiones es la más acertada?</p> <p>a) <input type="checkbox"/> Las temáticas de EtD2 no implica a las temáticas de Gues2 y viceversa, con una intensidad de implicación del 37% y 75%.</p> <p>b) <input type="checkbox"/> Las temáticas de EtD2 implica a las temáticas de Gues2 y viceversa, con una intensidad de implicación del 37% y 75%.</p> <p>c) <input type="checkbox"/> Las temáticas de EtD2 implica a las temáticas de Gues2 y viceversa, con una intensidad de implicación del 30% y 57%.</p>						X

10	Al aplicar el algoritmo <i>a priori</i> en RStudio a una tabla de actividades de estudiantes, se refleja el reporte:						
	 <p>¿Cuál de las conclusiones es la más acertada?</p> <p>a) <input type="checkbox"/> Con una confianza del 0.7: Si un estudiante realiza las actividades de autoevaluación, lectura y observa videos de referencias, podrá lograr las destrezas de la escritura.</p> <p>b) <input type="checkbox"/> Con una confianza del 0.3: Si un estudiante realiza las actividades de autoevaluación, lectura y observa videos de referencias, podrá lograr las destrezas de la escritura.</p> <p>c) <input type="checkbox"/> Con una confianza del 0.5: Si un estudiante realiza las actividades de autoevaluación, lectura y observa videos de referencias, podrá lograr las destrezas de la escritura.</p>						X
							X
							X

OBSERVACIONES

Fomentar el uso permanente de los instrumentos presentados en pro de la educación de los estudiantes

Este documento, certifica la revisión y análisis del contenido: "Instrumento de evaluación de analíticas implícitas para docentes de bachillerato" de recolección de información.

Para la constancia de lo expuesto, firma:



C.I. 1903903234

Grado Académico: Cuarto Nivel

Evaluador 4

Ambato, 31 de agosto del 2022

Documento evaluado:

Deseario exitos en sus labores, por medio del presente se le pide su aporte, que luego de un analisis riguroso de las items del *Instrumento de Evaluación de Estadística Implícita para Docentes de Bachillerato* que le mandamos. Marque con una X el casillero en la tabla de validación de contenido, conforme a su criterio y experiencia profesional, desotando si cuenta o no con los requisitos mínimos de formulación para su aplicación posterior.

INFORMACION GENERAL DEL INVESTIGADOR

Investigador	Dr. José Luis Méndez Guillón
Tema del Proyecto de Investigación	"Simulaciones digitales para facilitar la comprensión docente de las técnicas del análisis estadístico implícito y el <i>learning analytics</i> "
Programa de Estudio	Maestría en Ciencias de la Educación
Institución	Pontificia Universidad Católica del Ecuador Sede Ambato
Objetivo General de la Investigación	Diseñar un sistema de simulaciones digitales para que el profesor de nivel de bachillerato, familiarice con las técnicas comunes entre el Análisis Estadístico Implícito (AEI) y el <i>Learning Analytics</i> (LA).
Instrumentos a evaluar	Instrumento de evaluación de estadística implícita para docentes de bachillerato
Objetivo del Instrumento	Medir la aproximación docente a las técnicas comunes entre la estadística implícita y el <i>learning analytics</i> .


INFORMACION GENERAL DEL EVALUADOR

Evaluador	Manuel Antonio Méndez Freire
Universidad o Institución Educativa, a la que pertenece	Universidad Nacional de Chimborazo
Cargo	Docente de la Facultad de Ingeniería
Años de experiencia en el cargo	23
Grado Académico	Tercer Nivel () Cuarto Nivel () PhD (x)
Nivel o área a la que pertenece en la Universidad o Institución Educativa	Área de Estadística y Matemática


TABLA DE VALIDACIÓN

NOTA.- Para cada pregunta se considera la escala de 1 al 5, donde:

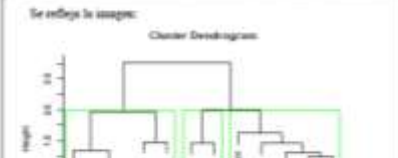
1.- Muy poco. 2.- Poco. 3.- Regular. 4.- Aceptable. 5.- Muy aceptable

Nº	ITEMS	Puntajes				
		1	2	3	4	5
1	El Índice de similitud:					x
	a) <input type="checkbox"/> Es la similitud de conjuntos (a, b) b) <input type="checkbox"/> Es la similitud de puntos (a, b) c) <input type="checkbox"/> Representa las similitudes de cada pareja (a, b) al nivel cero de la premisa					
2	La covarianza es:					x
	a) <input type="checkbox"/> Es la adherencia equilibrada de $-a < b$ sobre $a < -b$. b) <input type="checkbox"/> Una medida de desequilibrio de las frecuencias de los eventos $b < a$ y $a < b$ a favor del primero. c) <input type="checkbox"/> Es la congruencia relacionada de $-a < b$ sobre $a < -b$					
3	El gráfico de implicación es:					x
	a) <input type="checkbox"/> Un gráfico orientado, ponderado (hacia las atomidades de implicación). b) <input type="checkbox"/> Un gráfico no orientado, no ponderado (hacia las atomidades de implicación). c) <input type="checkbox"/> Un gráfico orientado, no ponderado (hacia las atomidades de implicación).					
4	A un grupo de 20 estudiantes se analizó el uso de los diferentes plataformas, la tabla de análisis de consistencia de eventos se muestra el siguiente cuadro:					x
	 <p>El valor de Índice de Similitud entre el uso de las plataformas: Khan Academy y FutureLearn es:</p> <p>a) <input type="checkbox"/> 0.0458 b) <input type="checkbox"/> 0.5844 c) <input type="checkbox"/> 0.7387</p>					
5	Para transformar la tabla de similitudes, desde un nivel cero hacia un nivel uno, se procede así:					x

6) Al aplicar el algoritmo de selección en RStudio a una tabla de datos:




Se refleja la imagen:



Se puede observar a:

- Agrupación de personas, de acuerdo a sus gustos.
- Grupos de estudiantes, de acuerdo a sus precedencias.
- La conformación del grupo de estudiantes, de acuerdo a sus actas.


7) Para la siguiente tabla de valores de densidad de probabilidad y su correspondiente Gráfico de Probabilidad:



La interpretación acertada es:

- $f(NDP, PCN) = 0.370$, $f(FLA, RLP) = 0.028$, $f(NDG, SDP) = 0.880$
- $f(NDP, PCN) = 0.370$, $f(FLA, RLP) = 0.028$, $f(NDG, SDP) = 0.880$
- $f(NDP, PCN) = 0.370$, $f(FLA, RLP) = 0.028$, $f(NDG, SDP) = 0.880$


8) Como se seleccionan los valores V de la siguiente tabla, con el diagrama del árbol:



Se puede observar a:

- Los valores V positivos, se visualizan en el diagrama de árbol, con el color negro.
- Los valores V positivos, se visualizan en el diagrama de árbol, con el color rojo.
- Los valores V negativos, se visualizan en el diagrama de árbol, con el color rojo.


9) Al aplicar el algoritmo de selección en RStudio a una tabla de datos, se refleja la imagen:



¿Cuál de las conclusiones es la más acertada?

- Los resultados de E(D) se aplican a los resultados de Geom2 y viceversa, con una intensidad de implicación del 77% y 73%.
- Los resultados de E(D) se aplican a los resultados de Geom2 y viceversa, con una intensidad de implicación del 77% y 73%.
- Los resultados de E(D) se aplican a los resultados de Geom2 y viceversa, con una intensidad de implicación del 70% y 57%.

10) Al aplicar el algoritmo de selección en RStudio a una tabla de actividades de estudiantes, se refleja el reporte:



¿Cuál de las conclusiones es la más acertada?

- Con una confianza del 0.7: Si un estudiante realiza las actividades de: arte/manualidades, lecturas y observa videos de refuerzo, podrá lograr las destrezas de la escritura.
- Con una confianza del 0.3: Si un estudiante realiza las actividades de: arte/manualidades, lecturas y observa videos de refuerzo, podrá lograr las destrezas de la escritura.
- Con una confianza del 0.5: Si un estudiante realiza las actividades de: arte/manualidades, lecturas y observa videos de refuerzo, podrá lograr las destrezas de la escritura.

OBSERVACIONES
 El instrumento presentado es excelente para la investigación planteada.
 Este documento, certifica la revisión y análisis del contenido: "Instrumento de evaluación de estadística inferencial para docentes de bachillerato" de recolección de información.
 Para la constancia de lo expuesto, firma:



C.I. 1802518349
 Grado Académico: PhD en Estadística E Investigación Operativa

Evaluador 5

Análisis 31 de agosto del 2022

Docente evaluador:

Desarrollando éxitos en sus labores, por medio del presente se le pide su apoyo, que luego de un análisis riguroso de los ítems del Instrumento de Evaluación de Estadística Implícita para Docentes de Bachillerato que le mostramos. Marque con una X el casillero en la tabla de validación de contenido, conforme a su criterio y experiencia profesional, desotando si cuenta o no con los requisitos mínimos de formulación para su aplicación posterior.

INFORMACION GENERAL DEL INVESTIGADOR

Investigador	Dr. José Luis Mollo Guzmán
Tema del Proyecto de Investigación	"Simulaciones digitales para facilitar la comprensión docente de las técnicas del análisis estadístico implícito y el <i>learning analytics</i> "
Programa de Estudio	Maestría en Ciencias de la Educación
Institución	Postulada Universidad Católica del Ecuador Sede Ambato
Objetivo General de la Investigación	Desarrollar un sistema de simulaciones digitales para que el profesor de nivel de bachillerato, familiarice con las técnicas comunes entre el Análisis Estadístico Implícito (AEI) y el <i>Learning Analytics</i> (LA).
Recursos a evaluar	Instrumentos de evaluación de estadística implícita para docentes de bachillerato
Objetivo del Instrumento	Medir la <i>aprovechación docente</i> a las técnicas comunes entre el <i>análisis estadístico implícito</i> y el <i>learning analytics</i> .

INFORMACION GENERAL DEL EVALUADOR

Evaluador	Dr. Fernando Procel
Universidad e Institución Educativa, a la que pertenece	Unidad Educativa Isabel de Goda de Kumbura
Cargo	Docente de Matemática en Bachillerato
Años de experiencia en el cargo	16
Grado Académico	Tercer Nivel () Cuarto Nivel (X) PhD ()
Nivel a área a la que pertenece en la Universidad e Institución Educativa	Área de matemática y Física


TABLA DE VALIDACIÓN

NOTA: Para cada pregunta se considera la escala de 1 al 5, donde:


1- Muy poca; 2- Poca; 3- Regular; 4- Aceptable; 5- Muy aceptable

Nº	ITEMS	Puntajes				
		1	2	3	4	5
1	El Índice de similitud es: a) <input type="checkbox"/> Es la similitud de conjuntos (a, b). b) <input type="checkbox"/> Es la similitud de pares (a, b). c) <input type="checkbox"/> Representa las similitudes de cada pareja (a, b) al nivel uno de la jerarquía.					X
2	La cubiende es: a) <input type="checkbox"/> Es la adiferencia equilibrada de $-a+b$ sobre $a+b$. b) <input type="checkbox"/> Una medida de desequilibrio de las frecuencias de los eventos $X=y$ y $a+b$ a fines del primer. c) <input type="checkbox"/> Es la congruencia relacionada de $-a+b$ sobre $a+b$.					X
3	El gráfico de implicación es: a) <input type="checkbox"/> Un gráfico orientado, ponderado (hacia las atomidades de implicación). b) <input type="checkbox"/> Un gráfico no orientado, no ponderado (hacia las atomidades de implicación). c) <input type="checkbox"/> Un gráfico no orientado, no ponderado (hacia las atomidades de implicación).					X
4	A un grupo de 20 estudiantes se analiza el uso de los diferentes plataformas, la tabla de análisis de similitud de eventos se muestra el siguiente cuadro:  El valor de Índice de Similitud entre el uso de las plataformas: Chess Academy y TinTubo es: a) <input type="checkbox"/> 0,9438 b) <input type="checkbox"/> 0,5844 c) <input type="checkbox"/> 0,7387					X
5	Para transformar la tabla de similitudes, desde un nivel uno, se procede así:					X

<p>a) <input type="checkbox"/> Se halla el valor mínimo de la tabla B, se eliminan la fila y columna correspondiente a este valor, los demás valores de la nueva tabla, son el cuadrado del valor máximo entre el valor de la celda y el valor a eliminarse.</p> <p>b) <input type="checkbox"/> Se halla el valor máximo de la tabla B, se eliminan la fila y columna correspondiente a este valor, los demás valores de la nueva tabla, son el cuadrado del valor máximo entre el valor de la celda y el valor a eliminarse.</p> <p>c) <input type="checkbox"/> Se halla el valor máximo de la tabla, se eliminan la fila y columna correspondiente a este valor, los demás valores de la nueva tabla, vienen de la raíz cuadrada del valor mínimo entre el valor de la celda y el valor a eliminarse.</p>																																																																																						
<p>6</p> <p>Cómo se relacionan los valores V de la siguiente tabla, con el diagrama del árbol:</p> <table border="1"> <thead> <tr> <th>Nivel</th> <th>V1</th> <th>V2</th> <th>V3</th> <th>V4</th> <th>V5</th> <th>V6</th> <th>V7</th> <th>V8</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>10</td> <td>10</td> <td>10</td> <td>10</td> <td>10</td> <td>10</td> <td>10</td> <td>10</td> </tr> <tr> <td>2</td> <td>10</td> <td>10</td> <td>10</td> <td>10</td> <td>10</td> <td>10</td> <td>10</td> <td>10</td> </tr> <tr> <td>3</td> <td>10</td> <td>10</td> <td>10</td> <td>10</td> <td>10</td> <td>10</td> <td>10</td> <td>10</td> </tr> <tr> <td>4</td> <td>10</td> <td>10</td> <td>10</td> <td>10</td> <td>10</td> <td>10</td> <td>10</td> <td>10</td> </tr> </tbody> </table> <p>a) <input type="checkbox"/> Los valores V positivos, se visualizan en el diagrama de árbol, con el color negro.</p> <p>b) <input type="checkbox"/> Los valores V positivos, se visualizan en el diagrama de árbol, con el color rojo.</p> <p>c) <input type="checkbox"/> Los valores V negativos, se visualizan en el diagrama de árbol, con el color rojo.</p>	Nivel	V1	V2	V3	V4	V5	V6	V7	V8	1	10	10	10	10	10	10	10	10	2	10	10	10	10	10	10	10	10	3	10	10	10	10	10	10	10	10	4	10	10	10	10	10	10	10	10					X																																				
Nivel	V1	V2	V3	V4	V5	V6	V7	V8																																																																														
1	10	10	10	10	10	10	10	10																																																																														
2	10	10	10	10	10	10	10	10																																																																														
3	10	10	10	10	10	10	10	10																																																																														
4	10	10	10	10	10	10	10	10																																																																														
<p>7</p> <p>Para la siguiente tabla de valores de Intersidad Implícita y su correspondiente Gráfico Implícito:</p> <table border="1"> <thead> <tr> <th>Evento</th> <th>1</th> <th>2</th> <th>3</th> <th>4</th> <th>5</th> <th>6</th> <th>7</th> <th>8</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> </tr> <tr> <td>2</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> </tr> <tr> <td>3</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> </tr> <tr> <td>4</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> </tr> <tr> <td>5</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> </tr> <tr> <td>6</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> </tr> <tr> <td>7</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> </tr> <tr> <td>8</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> <td>0,10</td> </tr> </tbody> </table> <p>La interpretación correcta es:</p> <p>a) <input type="checkbox"/> $I(ASP, PCN) = 0,9743$, $I(PLA, RAP) = 0,5426$, $I(RES, SEP) = 0,8803$</p> <p>b) <input type="checkbox"/> $I(ASP, PCN) = 0,9743$, $I(PLA, RAP) = 0,5938$, $I(RES, SEP) = 0,9087$</p> <p>c) <input type="checkbox"/> $I(ASP, PCN) = 0,9743$, $I(PLA, RAP) = 0,5938$, $I(RES, SEP) = 0,8803$</p>	Evento	1	2	3	4	5	6	7	8	1	0,10	0,10	0,10	0,10	0,10	0,10	0,10	0,10	2	0,10	0,10	0,10	0,10	0,10	0,10	0,10	0,10	3	0,10	0,10	0,10	0,10	0,10	0,10	0,10	0,10	4	0,10	0,10	0,10	0,10	0,10	0,10	0,10	0,10	5	0,10	0,10	0,10	0,10	0,10	0,10	0,10	0,10	6	0,10	0,10	0,10	0,10	0,10	0,10	0,10	0,10	7	0,10	0,10	0,10	0,10	0,10	0,10	0,10	0,10	8	0,10	0,10	0,10	0,10	0,10	0,10	0,10	0,10					X
Evento	1	2	3	4	5	6	7	8																																																																														
1	0,10	0,10	0,10	0,10	0,10	0,10	0,10	0,10																																																																														
2	0,10	0,10	0,10	0,10	0,10	0,10	0,10	0,10																																																																														
3	0,10	0,10	0,10	0,10	0,10	0,10	0,10	0,10																																																																														
4	0,10	0,10	0,10	0,10	0,10	0,10	0,10	0,10																																																																														
5	0,10	0,10	0,10	0,10	0,10	0,10	0,10	0,10																																																																														
6	0,10	0,10	0,10	0,10	0,10	0,10	0,10	0,10																																																																														
7	0,10	0,10	0,10	0,10	0,10	0,10	0,10	0,10																																																																														
8	0,10	0,10	0,10	0,10	0,10	0,10	0,10	0,10																																																																														
<p>8</p> <p>Al aplicar el algoritmo de árbol de variables en RStudio a una tabla de datos, algunos resultados en sus siguientes, se refleja la imagen:</p>  <p>a) <input type="checkbox"/> Agrupación de personas, de acuerdo a sus gustos.</p> <p>b) <input type="checkbox"/> Grupos de estudiantes, de acuerdo a sus procedencias.</p> <p>c) <input type="checkbox"/> La conformación del grupo de estudiantes, de acuerdo a sus gustos.</p>					X																																																																																	

<p>9</p> <p>Al aplicar el algoritmo implícito gráfico en RStudio a una tabla de datos, algunos resultados en sus siguientes, se refleja la imagen:</p>  <p>¿Cuál de las conclusiones es la más acertada?</p> <p>a) <input type="checkbox"/> Las técnicas de EISD no aplica a las técnicas de Grupos 2 y viceversa, con una intensidad de aplicación del 37% y 75%.</p> <p>b) <input type="checkbox"/> Las técnicas de EISD aplica a las técnicas de Grupos 2 y viceversa, con una intensidad de aplicación del 37% y 75%.</p> <p>c) <input type="checkbox"/> Las técnicas de EISD aplica a las técnicas de Grupos 2 y viceversa, con una intensidad de aplicación del 30% y 37%.</p>					X
--	--	--	--	--	---

10 Al aplicar el algoritmo a priori en EStudio a una tabla de actividades de estudiantes, se refleja el reporte:



¿Cuál de las conclusiones es la más acertada?

a) Con una confianza del 0.7: Si un estudiante realiza las actividades de: artes/manualidades, lecturas y observa videos de refuerzo, podría lograr las destrezas de la escritura.

b) Con una confianza del 0.3: Si un estudiante realiza las actividades de: artes/manualidades, lecturas y observa videos de refuerzo, podría lograr las destrezas de la escritura.

c) Con una confianza del 0.5: Si un estudiante realiza las actividades de: artes/manualidades, lecturas y observa videos de refuerzo, podría lograr las destrezas de la escritura.

OBSERVACIONES

Este documento, certifica la revisión y análisis del contenido: "Instrumento de evaluación de estadística inferencial para docentes de bachillerato" de recolección de información.

Para la constancia de lo expuesto, firma:

Mgs. Fernando Procel

C.I. 0601878291

Grado Académico: Magister en Ciencias de la Educación, mención Aprendizaje de la Matemática.

Validez del Instrumento de Evaluación (Pre Test / Post Test)

Evaluadores	ITEMS										Total Fila
	1	2	3	4	5	6	7	8	9	10	
Evaluador 1	5	5	5	5	5	5	5	5	5	5	50
Evaluador 2	5	5	5	5	5	5	5	5	5	5	50
Evaluador 3	4	4	5	4	4	5	5	5	5	5	46
Evaluador 4	5	5	5	5	5	5	5	5	5	5	50
Evaluador 5	5	5	5	5	5	5	5	5	5	5	50
Total Columna	24	24	25	24	24	25	25	25	25	25	246
Promedio	4,80	4,80	5,00	4,80	4,80	5,00	5,00	5,00	5,00	5,00	49,20
Desv. Estándar	0,43	0,43	0,00	0,43	0,43	0,00	0,00	0,00	0,00	0,00	1,79
Varianzas	0,16	0,16	0,00	0,16	0,16	0,00	0,00	0,00	0,00	0,00	2,56
Alfa de Cronbach:											
$\alpha = \frac{K}{K-1} \left[1 - \frac{\sum_{i=1}^K S_i^2}{S^2} \right]$											$\alpha = \frac{10}{10-1} \left[1 - \frac{0,64}{2,56} \right] = 0,8333333$

Evaluador 2

FICHA PARA VALIDACION DE LA PROPUESTA

Ambato, 1 de octubre del 2022

Docente evaluador:

Deseando éxitos en sus labores, por medio del presente, conociendo su trayectoria profesional y estrecha vinculación en el campo de Investigación Científica y Tecnológica, con la finalidad de seguir fortaleciendo la propuesta: "SIMULACIONES DIGITALES PARA FACILITAR LA COMPRESIÓN DOCENTE DE LAS TÉCNICAS DEL ANÁLISIS ESTADÍSTICO IMPLICATIVO Y EL LEARNING ANALYTICS", a darse en la PUCE-Sede Ambato. Solicito de la manera más comedida, emitir su criterio y más información en calidad de "Revisor experto de la propuesta" (disponible en la página: <https://sites.google.com/view/pazmino-r-mullo-jt%C3%A9cnicas-comunes-del-asi-y-el-la>), para lo cual adjunto a la presente un documento con la información adicional para la respectiva validación.

INFORMACION GENERAL DEL INVESTIGADOR

Investigador	Dr. José Luis Mullo Gualán
Tema del Proyecto de Investigación	"Simulaciones digitales para facilitar la comprensión docente de las técnicas del análisis estadístico implicativo y el learning analytics"
Programa de Estudio	Maestría en Ciencias de la Educación
Institución	Pontificia Universidad Católica del Ecuador Sede Ambato
Objetivo General de la Investigación	Diseñar un sistema de simulaciones digitales para que el profesor de nivel de bachillerato, familiarice con las técnicas comunes entre el Análisis Estadístico Implicativo (ASI) y el Learning Analytics (LA)
Herramienta a evaluar	Sistema de simulaciones para el uso docente de las técnicas comunes al Análisis Estadístico Implicativo y el Learning Analytics, disponible en: https://sites.google.com/view/pazmino-r-mullo-jt%C3%A9cnicas-comunes-del-asi-y-el-la
Objetivo de la validación	Emitir a Juicio de Expertos, el sistema de simulaciones digitales para el uso docente de las técnicas ASI y LA.

INFORMACION GENERAL DEL EVALUADOR

Evaluador	Diana Alexandra Gamboa Castillo
Universidad o Institución Educativa, a la que pertenece	Unidad Educativa "Pilalun"
Cargo	Docente
Años de experiencia en el cargo	12
Grado Académico	Tercer Nivel () Cuarto Nivel (x) PhD ()
Nivel o área a la que pertenece en la Universidad o Institución Educativa	Bachillerato

Indicación. – Estimado evaluador, marque con una X el casillero, en la tabla de validación de contenido, conforme a su criterio y experiencia profesional, denotando si cuenta o no con los requisitos mínimos de formulación para su aplicación posterior.

FICHA DE VALIDACIÓN


INDICADORES	CRITERIOS	NIVEL DE CALIFICACIÓN																				
		DEFICIENTE 0-20%				BAJA 21-40%				REGULAR 41-60%				BUENO 61-90%				MUY BUENO 91-100%				
		3	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100	
1.- Claridad	Las simulaciones digitales se han formulado con lenguaje apropiado																			X		
2.- Objetividad	Las simulaciones digitales, están expresado en conductas observables.																				X	
3.- Actualidad	Las simulaciones digitales está acorde a los aportes recientes en la disciplina de estudio																					X
4.- Organización	En las simulaciones digitales presentadas hay una organización lógica																					X
5.- Suficiencia	Las simulaciones digitales presentadas comprende los aspectos en cantidad y calidad																					X
6.- Intencionalidad	Las simulaciones digitales presentadas está adecuado para valorar la variable seleccionada																					X
7.- Consistencia	Las simulaciones digitales presentadas, está basado en aspectos teóricos y científicos																					X
8.- Coherencia	Con las simulaciones digitales presentadas, hay relación entre las dimensiones y los indicadores																					X
9.- Metodología	La estrategia responde al propósito del diagnóstico																					X
10.- Pertinencia	La estrategia de las simulaciones presentadas, es útil y adecuado para la aplicación																					X

OBSERVACIONES

Se deberá promover la utilización de los simuladores para que sean aplicados con los estudiantes en las aulas motivando el uso de las Tics.

Este documento, certifica la revisión y análisis del contenido de la propuesta: "Simulaciones digitales para el uso docente de las técnicas comunes al Análisis Estadístico Implicativo (ASI) y el Learning Analytics (LA)".

Para la constancia de lo expuesto, firma:



C.I: 1803903234

Grado Académico: 4to nivel

Evaluador 3

FICHA PARA VALIDACION DE LA PROPUESTA

Ambato, 1 de octubre del 2022

Docente evaluador:

Desearo éxitos en sus labores, por medio del presente, conociendo su trayectoria profesional y estrecha vinculación en el campo de Investigación Científica y Tecnológica, con la finalidad de seguir fortaleciendo la propuesta: "SIMULACIONES DIGITALES PARA FACILITAR LA COMPRENSIÓN DOCENTE DE LAS TÉCNICAS DEL ANÁLISIS ESTADÍSTICO IMPLICATIVO Y EL LEARNING ANALYTICS", a darse en la PUCE-Sede Ambato. Solicito de la manera más comedida, emitir su criterio y más información en calidad de "Revisor experto de la propuesta" (disponible en la página: <https://sites.google.com/view/pazmino-r-mullo-i/%C3%A9caicas-comunes-del-asi-y-el-la>), para lo cual adjunto a la presente un documento con la información adicional para la respectiva validación.

INFORMACION GENERAL DEL INVESTIGADOR

Investigador	Dr. José Luis Mullo Gualán
Tema del Proyecto de Investigación	"Simulaciones digitales para facilitar la comprensión docente de las técnicas del análisis estadístico implicativo y el <i>learning analytics</i> "
Programa de Estudio	Maestría en Ciencias de la Educación
Institución	Pontificia Universidad Católica del Ecuador Sede Ambato
Objetivo General de la Investigación	Diseñar un sistema de simulaciones digitales para que el profesor de nivel de bachillerato, familiarice con las técnicas comunes entre el Análisis Estadístico Implicativo (ASI) y el <i>Learning Analytics (LA)</i>
Herramienta a evaluar	Sistema de simulaciones para el uso docente de las técnicas comunes al Análisis Estadístico Implicativo y el <i>Learning Analytics</i> , disponible en: https://sites.google.com/view/pazmino-r-mullo-i/%C3%A9caicas-comunes-del-asi-y-el-la .
Objetivo de la validación	Emitir a Juicio de Expertos, el sistema de simulaciones digitales para el uso docente de las técnicas ASI y LA.

INFORMACION GENERAL DEL EVALUADOR

Evaluador	Paola Alejandra Villalón Muñoz
Universidad o Institución Educativa, a la que pertenece	Universidad Estatal Amazónica
Cargo	Profesor Titular Auxiliar I
Años de experiencia en el cargo	11 años
Grado Académico	Tercer Nivel () Cuarto Nivel () PhD ()
Nivel o área a la que pertenece en la Universidad o Institución Educativa	Ciencias de la Vida, Área de física y matemática

Indicación. – Estimado evaluador, marque con una X el casillero, en la tabla de validación de contenido, conforme a su criterio y experiencia profesional, denotando si cuenta o no con los requisitos mínimos de formulación para su aplicación posterior.

FICHA DE VALIDACIÓN


INDICADORES	CRITERIOS	DEFICIENTE 0-20%				BAJA 21-40%				REGULAR 41-60%				BUENO 61-90%				MUY BUENO 81-100%			
		5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100
1.- Claridad	Las simulaciones digitales se han formulado con lenguaje apropiado																				X
2.- Objetividad	Las simulaciones digitales, están expresado en conductas observables																				X
3.- Actualidad	Las simulaciones digitales está acorde a los aportes recientes en la disciplina de estudio																				X
4.- Organización	En las simulaciones digitales presentadas hay una organización lógica																				X
5.- Suficiencia	Las simulaciones digitales presentadas comprende los aspectos en cantidad y calidad																				X
6.- Intencionalidad	Las simulaciones digitales presentadas está adecuado para valorar la variable seleccionada																				X
7.- Consistencia	Las simulaciones digitales presentadas, está basado en aspectos teóricos y científicos																				X
8.- Coherencia	Con las simulaciones digitales presentadas, hay relación entre las dimensiones y los indicadores																				X
9.- Metodología	La estrategia responde al propósito del diagnóstico																				X
10.- Pertinencia	La estrategia de las simulaciones presentadas, es útil y adecuado para la aplicación																				X

OBSERVACIONES

.....

Este documento, certifica la revisión y análisis del contenido de la propuesta: "Simulaciones digitales para el uso docente de las técnicas comunes al Análisis Estadístico Implicativo (ASI) y el *Learning Analytics (LA)*".

Para la constancia de lo expuesto, firma:

 PAOLA ALEJANDRA VILLALÓN MUÑOZ

C.I: 0601908510

Grado Académico: Magister

Validez de la Propuesta										
Evaluadores	INDICADORES									
	Claridad	Objetividad	Actualidad	Organización	Suficiencia	Intencionalidad	Consistencia	Coherencia	Metodología	Aplicabilidad
Evaluador 1	85	100	100	80	95	95	100	100	95	100
Evaluador 2	85	90	95	90	90	90	90	90	90	95
Evaluador 3	95	95	95	95	95	95	95	95	95	95
Evaluador 4	100	100	95	100	95	100	100	100	100	100
V de Aiken	0,91	0,96	0,96	0,91	0,94	0,95	0,96	0,96	0,95	0,98
V de Aiken Global	0,95									

ANEXO 4

Algoritmos para la ejecución de los diferentes componentes

A. Algoritmo para la instalación de los paquetes

```

##INSTALAR RStudio y R; con versiones
aconsejadas por los autores.
##INSTALACIÓN DE PAQUETES.
##Antes de continuar "Actualizar los paquetes de
RStudio & R: Packages -> Update"
##Instalación de los paquetes:
"string";"tcltk2";"graph";"BiocGenerics";"parallel";"R
cpp".
install.packages("string");
install.packages("tcltk2");
##install.package("graph");
##install.package("BiocGenerics");
##install.package("parallel");
install.packages("Rcpp");
##Instalación de los demás paquetes.
install.packages("microbenchmark")
##el paquete rchic; descargar del sitio:
http://members.femto-st.fr/raphael-couturier/en/rchic
##instalar el paquete rchic, de forma manual.
install.packages("ggplot2");
install.packages("factoextra");
install.packages("cluster");
install.packages("fastcluster");
install.packages("caTools");
install.packages("CluMix");
install.packages("dplyr");
install.packages("replyr");
install.packages("arules");
install.packages("arulesViz");
install.packages("vcd");
install.packages("readr");
##Instalación del paquete: "Rgraphviz"
##PEGAR EN COSOLA (Tanto en RStudio como
en R):
## try http:// if https:// URLs are not supported
source("https://bioconductor.org/biocLite.R")
biocLite("Rgraphviz")

```

B. Algoritmo para la instalación del paquete: *Rchic*
 ##Descargar del sitio: <http://members.femto-st.fr/raphael-couturier/en/rchic>
 ##Instalar de forma manual el paquete *Rchic*, (tanto en *RStudio* como en *R*).

```

##Instalación de Rgraphviz
##PEGAR EN COSOLA (tanto en RStudio como en R):
##=====
## try http:// if https:// URLs are not supported
source("https://bioconductor.org/biocLite.R")
biocLite("Rgraphviz")
##=====

```

```

##Instalación de Rgraphviz en la versión 4.2.0 de R
##PEGAR EN COSOLA (tanto en RStudio como en R):
##=====
if (!requireNamespace("BiocManager", quietly = TRUE))
  install.packages("BiocManager")

```

```
BiocManager::install("Rgraphviz")
```

```
##=====
```

```

##Instalación de paquetes complementarios
install.packages("string");
install.packages("tcltk2");
install.packages("Rcpp");

```

```

##Prueba de funcionamiento de Rchic
library(rchic)
rchic()

```

C. Algoritmo para las técnicas: *hclust.vector*;
dendro.variables; *dendro.diana*

```

##ANÁLISIS ESTADÍSTICO IMPLICATIVO &
LEARNING ANALITYCS (Técnicas Comunes).
##CASO: DATOS BINARIOS.
##VERIFICACIÓN DE LAS LIBRERIAS.

```

```

##library("microbenchmark");
library("rchic");
library("ggplot2");
library("factoextra");
library("cluster");
library("fastcluster");
library("CluMix");
library("dplyr");
library("replyr");
library("arules");
library("arulesViz");
library("vcd");
library("Rgraphviz");
library("readr");

```

```
##FUNCIONES.
```

```

##Método CLUSTER
hclust_vector<-function(x){
  df<-read.csv(x, sep = ";")
  hc <- hclust.vector(df, "cen")
  hc<-hclust.vector(df, method="single", members=NULL,
metric='euclidean', p=NULL)
  plot(hc)
  return(hc)
}

```

```

##Método DENDRO.VARIABLES
dendro_variables<-function(x){
  df<-read.csv(x, sep = ";")
  dend1 <- dendro.variables(df, method="distcor")
  plot(dend1)#opcional
  return(dend1)
}

```

```

##Método DENDRO.DIANA
dendro_diana<-function(x){
  df<-read.csv(x, sep = ";")
  dend <- diana(df, metric = "euclidean", stand = TRUE)
  plot(dend)#opcional
  return(dend)
}

```

```
##EJECUCIÓN.
```

```

##
1=====
##Leer en la matriz DataBase, el archivo Caftán.csv
(Datos Cotidianos 1 & 0) ubicados en la carpeta del
Usuario (Documentos)

```

```

DataBase<-read.csv("Caftán.csv", sep = ";")
#Proceso de preparación del Archivo DataBase
f<-paste('_DCaftán',toString(1),'.csv',sep="_")
T<-paste('_TCaftán',toString(1),'.csv',sep="_")
write.table(DataBase, file=f, sep=";", quote = FALSE)
DataBaseT<-t(DataBase)
write.table(DataBaseT, file = T ,sep=";")
nVar<- ncol(DataBase)
nFilas<- nrow(DataBase)
fc<-paste(toString(nVar),'X',toString(nFilas))

##EJECUCIÓN
hclust_vector(T)
dendro_variables(T)
dendro_diana(T)

##ALGORITMO PARALELO PARA LA TÉCNICA
dendro.variables
install.packages("ClustOfVar", dependencies = TRUE);
install.packages("Hmisc", dependencies = TRUE);
install.packages("DescTools", dependencies = TRUE);
install.packages("extracat", dependencies = TRUE);
install.packages("marray", dependencies = TRUE);
if (!requireNamespace("BiocManager", quietly = TRUE))
  install.packages("BiocManager")
  BiocManager::install("marray")

if (!requireNamespace("BiocManager", quietly = TRUE))
  install.packages("BiocManager")
BiocManager::install("Biobase")

install.packages("FD", dependencies = TRUE);
install.packages("gplots", dependencies = TRUE);
install.packages("Matrix", dependencies = TRUE);
install.packages("CluMix", dependencies = TRUE);
install.packages("arules", dependencies = TRUE);
install.packages("CluMix")

library(ClustOfVar);
library(fastcluster);
library(arules);
#library(ClustOfVar)
#library(Hmisc);
#library(DescTools);
#library(extracat);
#library(marray);
#library(FD);
library(gplots);
#library(CluMix);
library(microbenchmark);
library(ggplot2);
library(factoextra);
library(cluster);
library(dplyr);
#library(replyr);
library(arulesViz);
library(vcd);
library(Rgraphviz);
library(readr);

##Indica la carpeta donde se va a trabajar
setwd("C:/Users/WinUser/Documents")
##Leer el archivo a analizar
Datos <- read.csv(file = "Notas_Estudiantes.csv", header
= TRUE, sep = ";", dec = ".", row.names = 1)
str(Datos)
require(fastcluster)
require(graphics)
hc <- hclust(dist(Datos), "ave")
plot(hc)
plot(hc, hang = 0.1)
rect.hclust(hc, k=3, border = "green")
require(CluMix)
require(graphics)

##Indica la carpeta donde se va a trabajar
setwd("C:/Users/WinUser/Documents")
##Leer el archivo a analizar
df <- read.csv(file = "Notas_Estudiantes.csv", header =
TRUE, sep = ";", dec = ".", row.names = 1)
str(df)
dend1 <- dendro.variables(df, method =
c("associationMeasures", "distcor", "ClustOfVar"),
linkage="ward.D2", associationFun =
association, check.psd = TRUE)
plot(dend1)#opcional
return(dend1)

D. Algoritmo para las técnicas: Apriori; Weclat y Eclat
##ALGORITMO APRIORI
install.packages("arules", dependencies = TRUE)
install.packages("plyr")

library(arules)

##Indica la carpeta donde se va a trabajar
setwd("C:/Users/WinUser/Documents")

##Leer el archivo a analizar
Datos_a_priori <- read.csv("Diario_Estudiantil.csv", sep
= ";", dec = ".", header = TRUE)

##Las transacciones deben ser numéricas y y los
productos factores
str(Datos_a_priori)

##Dar formato de transacción a Datos_a_priori
library(plyr) #instalar desde la consola
install.packages("plyr", dependencies = TRUE)

##Por cada transacción se crea una fila conteniendo
todos los productos separados con un coma
Lista_actividades <- ddply(Datos_a_priori,
c("Transaccion"), function(df1)paste(df1$Actividad,
collapse = ";"))

Lista_actividades$Transaccion <- NULL

##Guarda las actividades como csv
##Es muy importante incluir row.names para que no
agregue una columna con el número de
write.csv(Lista_actividades, "Lista_actividades.csv",
quote=FALSE, row.names = FALSE)

##Lee el archivo con las transacciones
transacciones <-
read.transactions("Lista_actividades.csv",
format = "basket", sep = ";", header
= TRUE)

##Desplega las transacciones en la consola
inspect(transacciones)

##Grafica las actividades con sus frecuencias
itemFrequencyPlot(transacciones, topN=10, type=
"absolute")

##Se obtiene las reglas de asociación por medio del
algoritmo apriori
reglas <- apriori(transacciones, parameter =
list(supp=0.3, conf=0.9, minlen=4))

##Se ordena las reglas de acuerdo a la "confianza"
reglas <- sort(reglas, by="confidence", decreasing =
TRUE)

##Se observa la estructura que tiene el objeto reglas
str(reglas)

```

```

##Se despliega las reglas en la pantalla
inspect(reglas)

##Se revisa si hay reglas duplicadas
duplicated(reglas)

##Se verifica si hay reglas redundantes
redundantes<- is.redundant(reglas)

##Para saber cuáles son las reglas redundantes
which(is.redundant(reglas))

##Se quita las reglas redundantes
reglas_podadas <- reglas[!redundantes]

##Se despliega las reglas ya sin las redundantes
inspect(reglas_podadas)

##se crea un grafo con las reglas
library(arulesViz) ## install.packages("arulesViz",
dependencies = TRUE)
plot(reglas_podadas, method= "graph",
engine="interactive", shading="confidence")

##ALGORITMO WECLAT
install.packages("arules", dependencies = TRUE)
library(arules)

##Se indica la carpeta donde se va a trabajar
setwd("C:/Users/WinUser/Documents")

##Se lee el archivo a analizar
Datos_weclat <- read.csv("Diario_Estudiantil.csv", sep =
";", dec = ".", header = TRUE)

##Las transacciones deben ser numéricas y las
actividades factores
str(Datos_weclat)

##Se da el formato de transacción a Datos_weclat
library(plyr) #instalar desde la consola
install.packages("plyr", dependencies = TRUE)

##Por cada transacción se crea una fila conteniendo
todos los productos separados con un coma
Lista_actividades <- ddply(Datos_weclat,
c("Transaccion"), function(df1)paste(df1$Actividad,
collapse = ";"))

Lista_actividades$Transaccion <- NULL

##Se guarda las actividades como csv

```

```

##Es muy importante incluir row.names para que no
agregue una columna con el número de
write.csv(Lista_actividades, "Lista_actividades.csv",
quote=FALSE, row.names = FALSE)

##Se leen el archivo con las transacciones
transacciones <-
read.transactions("Lista_actividades.csv",
format = "basket", sep = ";", header
= TRUE)

##Se despliega las transacciones en la consola
inspect(transacciones)

##Se obtiene las reglas de asociación por medio del
algoritmo Weclat
s <- weclat(data = transacciones, parameter =
list(supp=0.3, conf=0.9, minlen=4), control = list(verbose
= TRUE))

##Se despliega las transacciones en la consola
inspect(sort(s, by = "supp"))

##Crea reglas de asociación
reglas <- ruleInduction(s, confidence = .9)
inspect(reglas)
##Se revisa si hay reglas duplicadas
duplicated(reglas)

##Se verifica si hay reglas redundantes
redundantes<- is.redundant(reglas)

##Para saber cuáles son las reglas redundantes
which(is.redundant(reglas))

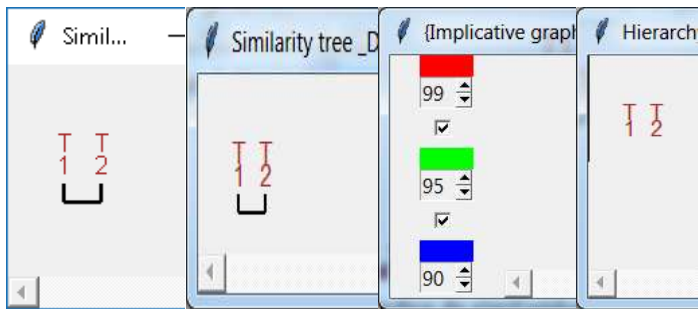
##Se quita las reglas redundantes
reglas_podadas <- reglas[!redundantes]

##Se despliega las reglas ya sin las redundantes
inspect(reglas_podadas)

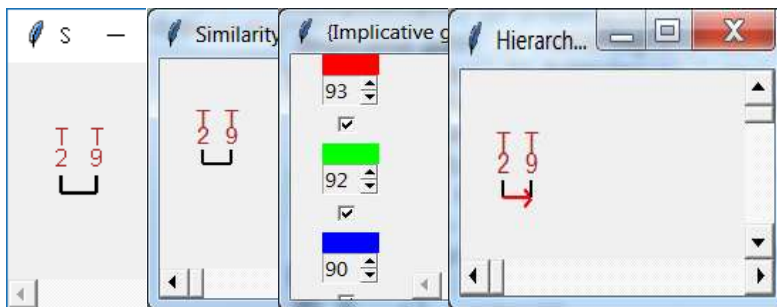
##ALGORITMO ECLAT
install.packages("arules", dependencies = TRUE)
library(arules)
##Se lee el archivo a analizar
base <- read.transactions("ReglasAso.csv", sep = ";",
rm.duplicates = TRUE)
inspect(base)
##se obtiene las reglas de asociación por medio del
algoritmo Eclat
itemsets <- eclat(data = base, parameter = list(supp =
0.5, minlen = 3))
##Se despliega las transacciones en la consola
inspect(sort(itemsets, by = "supp"))

```

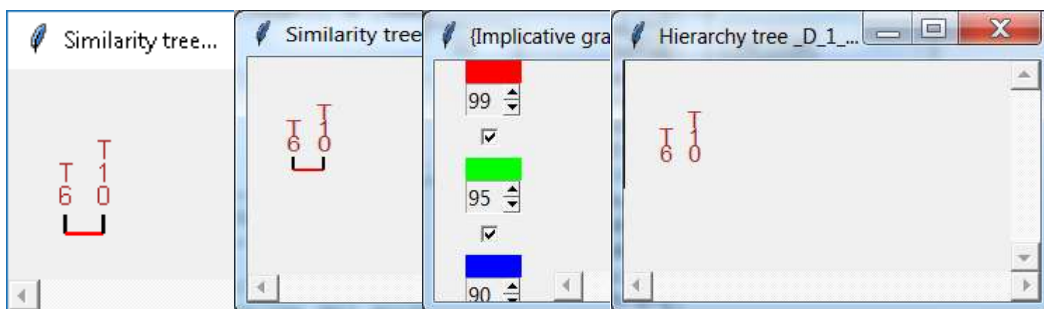

Ejemplo 2. Cuando los datos poseen diferente índice de similitud



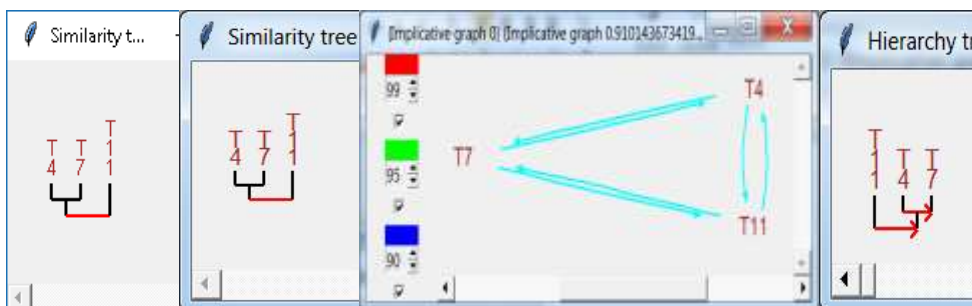
Ejemplo 3. Cuando los datos poseen índice de similitud alta y baja



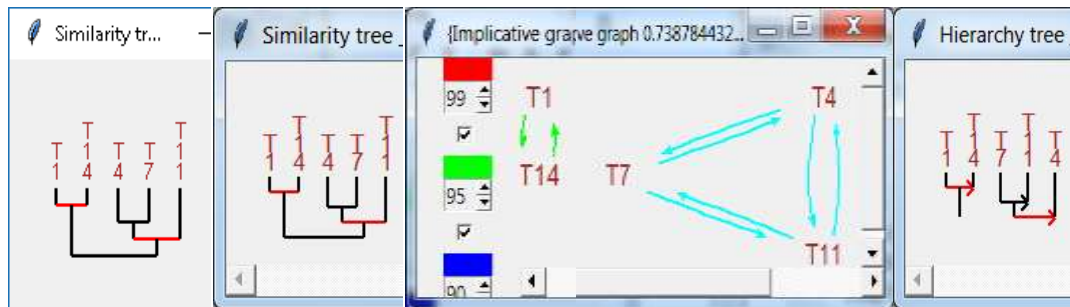
Ejemplo 4. Cuando los datos poseen índices de similitud alta



Ejemplo 5. Análisis de tres o más variables



Ejemplo 6. Análisis de más variables

Ejemplo 7.- Análisis de datos facilitados en Zamora *et al.* (2007).