



Pontificia Universidad
Católica del Ecuador

FACULTAD DE INGENIERIA

ESCUELA DE SISTEMAS

**TEMA: Análisis de Datos en una
Empresa para Predecir Ventas
utilizando Inteligencia Artificial**

**MARÍA FERNANDA RAMOS
EDWIN VILLALOBOS**

12/02/2020

DEDICATORIA

El presente proyecto está dedicado a Dios y nuestra familia, ya que han sido la fortaleza, la sabiduría e inspiración durante todo este proceso para alcanzar este gran objetivo.

A nuestros padres porque ellos han sido el motor en nuestras vidas, por darnos siempre todo su amor, su apoyo y sacrificio incondicional para con nosotros en este largo camino, por soportar cada mal humor, amanecidas, tristezas y alegrías que conlleva este proceso, por impulsarnos de manera constante a seguir adelante y sobre todo por enseñarnos que el conocimiento, la educación, el trabajo honesto y la perseverancia es la mejor herencia que nos llevaremos en esta vida.

A nuestros familiares, amigos, compañeros y demás personas especiales que de muchas formas nos han brindado su apoyo incondicional, ya que han sido una parte importante en esta etapa universitaria, dentro de la cual hemos vivido momentos llenos de incertidumbre, pero también momentos llenos de esplendor y emotividad para cumplir nuestro objetivo.

A nuestros docentes y mentores que supieron formarnos como profesionales y seres humanos, que nos han guiado en este largo camino universitario y sobre todo supieron adoctrinarnos con su sabiduría y sapiencia dentro de las aulas y en el día a día.

A la Pontificia Universidad Católica Del Ecuador nuestra prestigiosa universidad por permitirnos instruirnos dentro de sus aulas e instalaciones.

AGRADECIMIENTO

El presente proyecto de tesis es un esfuerzo en el cual muchas personas han sido partícipes para lograr culminarlo.

Gracias a nuestros padres por ser los primeros en motivarnos y darnos aliento cada día, gracias por depositar su confianza en nosotros y en nuestros sueños, gracias por todo el amor recibido, la dedicación y la paciencia que mostraban mientras se realizaba este trabajo de tesis. Gracias a ustedes hemos logrado convertirnos en lo que ahora hemos llegado a ser y darnos cuenta de que somos capaces de cumplir nuestros objetivos.

Gracias a todos los profesores, a nuestro director de tesis Ing. Miguel Ortiz y a los revisores por siempre apoyarnos en este trabajo, corrigiendo y dándonos sugerencias de cómo mejorar y lograr culminar este proyecto.

Así mismo queremos agradecer a la Empresa DYSUR por el apoyo y la confianza depositada en nosotros para realizar este proyecto.

También agradecer a todos nuestros familiares, amigos, compañeros y demás personas especiales que apoyaron, dándonos muestra de apoyo y colaborando de diversas maneras con nosotros en la realización de este proyecto.

INDICE DE TEMAS

Contenido

DEDICATORIA	ii
AGRADECIMIENTO	iii
INDICE DE TEMAS	iv
INDICE DE ILUSTRACIONES	vi
RESUMEN	1
CAPÍTULO 1: PROBLEMÁTICA.....	2
1.1. Introducción.....	2
1.2. Planteamiento del problema.....	3
1.3. Objetivo General.....	3
1.4. Objetivos específicos	4
1.5. Justificación	4
1.6. Alcance y Limitaciones.....	4
1.6.1. Alcance	4
1.6.2. Hipótesis.....	5
1.6.3. Limitaciones	5
1.7. Unidad de análisis.....	5
CAPITULO 2 MARCO METODOLÓGICO	7
2.1. Metodología de investigación	7
2.1.1. Método	7
2.1.2. Técnica	7
2.2. Metodología informática.....	8
2.2.1. Metodología para el desarrollo de sistemas	8
2.2.2. Modelo	8
2.2.3. Procesos de ingeniería	9
CAPITULO 3 MARCO DE REFERENCIA.....	12
3.1. Marco teórico	12
3.2. Marco conceptual.....	12
3.3. Marco legal	16
CAPITULO 4 DESARROLLO DE SOFTWARE	18
4.1. Análisis	18
4.2. Diseño	20
4.2.1. Requisitos funcionales	22

4.2.2.	Diagramas de Interacción	30
4.2.3.	Diagrama de Actividades	35
4.2.4.	Modelo de datos	36
4.3.	Implementación.....	38
4.4.	Pruebas	46
4.5.	Documentación	47
4.5.1.	Requisitos.....	47
CAPITULO 5 CONCLUSIONES Y RECOMENDACIONES.....		49
5.1.	Conclusiones	49
5.2.	Recomendaciones.....	50
REFERENCIAS BIBLIOGRÁFICAS.....		51
APÉNDICE		53
MANUAL DE USUARIO		53
❖	Introducción	53
❖	Proceso.....	53

INDICE DE ILUSTRACIONES

ILUSTRACIÓN 1. ILUSTRACIÓN DEL MODELO EN CASCADA.....	9
ILUSTRACIÓN 2. ILUSTRACIÓN CICLO DE VIDA RUP.....	10
ILUSTRACIÓN 3. ILUSTRACIÓN GRÁFICA UML POR COMPOSICIÓN	11
ILUSTRACIÓN 4. EJEMPLO REGRESIÓN POLINOMIAL.....	14
ILUSTRACIÓN 5. GRÁFICO EJEMPLO DE SERIES TEMPORALES.....	14
ILUSTRACIÓN 6. GRÁFICO REPRESENTACIÓN MODELO SARIMA.....	16
ILUSTRACIÓN 7.EJEMPLOS DE FUNCIONES POLINOMIALES.....	19
ILUSTRACIÓN 8. EJEMPLOS DE FUNCIONES POLINOMIALES.....	20
ILUSTRACIÓN 9. ILUSTRACIÓN GRÁFICA DISEÑO INTERFAZ.	21
ILUSTRACIÓN 10. EJEMPLO DE PROCESO DE SOFTWARE.....	38
ILUSTRACIÓN 11. RESULTADO DE LA PROYECCIÓN CON SERIES TEMPORALES.	46

RESUMEN

Esta investigación, cuyo objetivo es diseñar y elaborar un sistema web para hacer predicción de ventas enfocado a una empresa mediante la utilización de inteligencia artificial, se presenta como un proyecto muy tentador ya que propone automatizar un proceso manual que la mayoría de las empresas lo están realizando en estos tiempos. Para el desarrollo de este sistema se abordó con una investigación aplicada la cual permite tener un resultado teórico bastante amplio en conceptos de matemáticas y programación para así poder realizar este prototipo, de la misma forma se apoyó con el método descriptivo el cual permite detallar las características del tema planteado en este proyecto. Se decidió que para realizar las predicciones los datos proporcionados tienen que ser de un histórico de hace cuatro años y deben ser de productos tangibles lo cual la empresa DYSUR es la que apoyó para que se pueda obtener esta información. Después de algunos análisis y pruebas se llegó a la conclusión de que para realizar este sistema se programarán polinomios de segundo, tercer y cuarto grado que son los que permiten visualizar cómo va la tendencia en la curva mediante la información aplicada, así mismo se utilizará series temporales ya que permite ordenar ciertas variables en periodos de tiempo pero debido a la complejidad de la información que proporcionó la empresa se tiene que hacer uso del modelo SARIMA los que permiten encontrar patrones para realizar predicciones futuras en base al aprendizaje de los datos históricos. Cuando se realiza la ejecución del programa se obtienen gráficas las cuales muestran los resultados obtenidos y permite tomar decisiones en cuanto a las ventas en los próximos periodos. Siendo este un proceso muy común que las empresas por lo general lo realizan cada año se planteó este trabajo para facilitar y apoyar mediante el uso tecnológico al crecimiento de empresas pequeñas.

CAPÍTULO 1: PROBLEMÁTICA

1.1. Introducción

"Toda actividad requiere algún sistema de estimación del volumen que se va a manejar dentro de ésta. Los estimados son el resultado de predicciones y pronósticos". (López, 2016)

Realizar un pronóstico de ventas puede resultar beneficioso para una empresa porque permite tomar decisiones con criterio basados en hechos históricos de mercadotecnia, producción, aprovisionamiento y flujo de caja. Por tanto, esto debe ser tratado y elaborado con mucha atención, ya que este análisis podría ayudar a cumplir con los objetivos de la empresa los cuales traerían consigo mejoras favorables generando así mayores utilidades y un buen posicionamiento en el mercado.

En la actualidad es importante para cualquier empresa en el mercado poder gestionar a futuro sus ventas, para satisfacer la demanda de los compradores. Para lo cual se debe poder responder las siguientes interrogantes, ¿Cuánto podría vender en los próximos 6, 9 o 12 meses?, ¿Qué temporada es la mejor para vender mis productos?, ¿Cuenta mi empresa con materia prima suficiente para cubrir mis ventas futuras?, ¿Cómo planifico las necesidades futuras de recursos?, ¿Cómo planifico el flujo de efectivo?, ¿Cómo tomo decisiones de producción y aprovisionamiento?, ¿La capacidad productiva de mi empresa es apta para afrontar el volumen de mis ventas?, etc. Al responder estas preguntas se podría tener una visión más clara de qué hacer a futuro con mi empresa. Con lo que un pronóstico de ventas solventaría estas interrogantes.

Los pronósticos de ventas son muy útiles dentro de la empresa, esto se realiza en base a datos históricos y es el fundamento para decidir sobre el adecuado manejo de los diferentes recursos y saber cuánto gastar en diversas actividades de una empresa. Además, ayuda a corroborar estimaciones financieras de las diferentes áreas y por lo tanto para la toma de decisiones más acertadas y oportunas permitiendo que todos los departamentos pueden trabajar de forma conjunta teniendo claro que el objetivo es mejorar el manejo de información sobre las ventas. Muchas empresas realizan este proceso de forma manual, es decir, basándose en resultados de ventas obtenidos en meses o años pasados, lo cual genera desventajas ante una empresa que sí realice esto con la ayuda de un sistema para predicción de ventas. Así se puede decir que el pronóstico de ventas se lo señala como un sistema de aprendizaje que permite mejorar la precisión de ventas para el futuro.

En los últimos años el avance tecnológico ha traído consigo ventajas y oportunidades para las empresas, por ejemplo, el desarrollo de la Inteligencia Artificial ha permitido establecer modelos matemáticos que ayudan a la proyección de ventas de ciertos productos o servicios.

1.2. Planteamiento del problema

"En más del 80% de las empresas los procesos de ventas no están formalizados"
(Market, 2017)

Las empresas tienen problemas con sus ventas porque este proceso no se lo realiza de una manera formalizada, es decir no se lleva un control adecuado y sistematizado sobre ventas lo que hace que no se puedan tomar las mejores decisiones al momento de gestionar y manejar las ventas.

Existen situaciones negativas que afectan a las compañías debido a la carencia de un sistema de proyección de ventas. Por ejemplo, no hay información que los gerentes de esta área puedan usar para estructurar un plan de ventas adecuado o en su defecto que exista dicha información y sea imprecisa, incompleta o incorrecta; lo cual genera que se pierdan oportunidades de negocio importantes para la empresa. También no se tiene una visión clara de qué cantidad de dinero se va a facturar aproximadamente el siguiente mes. El control de inventario también se ve afectado, debido a que no se puede examinar las tendencias o temporadas más elevadas, con el fin de determinar los picos de ventas y adquirir la mercadería necesaria que cubrirán las ventas del año o incluso para ofrecer promociones especiales en aquellos productos que tiene mayor demanda.

Otro aspecto negativo que se da dentro de las empresas es que no se tiene un modelo y proceso de ventas efectivo que esté orientado a sus giros de negocio y las empresas tienden a estar experimentando o probando diferentes técnicas que les ayuden a generar más ingresos con lo que se hace un desperdicio de recursos tanto humanos como financieros, recursos materiales y tecnológicos que hacen que no se vean las ganancias y que al final esa técnica de venta usada no sea la más adecuada en el negocio que estamos haciendo.

Las empresas también necesitan realizar proyecciones para medir sus resultados objetivamente por esta razón es que nuestro trabajo está enfocado a desarrollar un sistema que ayude a la empresa a realizar y solventar todos estos problemas para que sus ventas sean el punto más importante para poder obtener sus utilidades.

1.3. Objetivo General

- Desarrollar un sistema web basado en algoritmos de inteligencia artificial que permita predecir ventas en función de los datos de una empresa.

1.4. Objetivos específicos

- Analizar modelos matemáticos que se apliquen en la proyección y estimación de ventas.
- Determinar las variables requeridas por el modelo matemático y sintetizarlas con miras a una automatización.
- Ajustar los algoritmos en base a las variables seleccionadas.
- Identificar el lenguaje de programación y la herramienta de desarrollo más adecuada para trabajar con algoritmos de inteligencia artificial.
- Diseñar y construir un prototipo web funcional.

1.5. Justificación

“El uso de un pronóstico de ventas es necesario para la buena administración de cualquier empresa, a lo largo de la historia moderna de la administración de recursos se confirma que un buen desarrollo en los pronósticos conlleva a una mejor administración de los recursos disponibles” (Wacker & Lummus, 2002).

Con la ayuda de las nuevas tecnologías se pretende desarrollar un sistema que sea capaz de analizar datos históricos de ventas de una organización, para proyectar una situación que sea próxima a la realidad y fácil de entender; esto con el fin de que la persona encargada de estructurar los planes de ventas sea capaz en base a los datos reportados por el software tomar mejores decisiones y pueda anticiparse a la demanda de su mercado.

Los datos de ventas que una empresa almacena en su base no se aprovecha de la manera adecuada ya que si se realizase un análisis y un correcto tratamiento de estos datos se puede generar información que tenga valor.

Así el propósito fundamental al llevar a cabo este trabajo es la evaluación y aplicación de métodos de pronósticos, con el fin de generar información importante que sea útil para una empresa al momento de realizar una planeación estratégica de los recursos.

1.6. Alcance y Limitaciones

1.6.1. Alcance

Este trabajo se centra en poder realizar una investigación y un análisis sobre modelos matemáticos para predicción de ventas y usar los mismos para desarrollar un sistema web identificando las variables necesarias para así poder predecir las ventas de la organización.

El software que se va a elaborar proyectará los resultados de las ventas basándose en datos históricos para ayudar fundamentalmente a la toma de decisiones

con el fin de que se puedan estructurar planes de ventas correctamente y ver las posibles situaciones que la empresa tendrá que enfrentar a futuro en el tema de ventas.

1.6.2. Hipótesis

¿La aplicación de la inteligencia artificial a través de modelos matemáticos facilita el análisis de datos?

1.6.3. Limitaciones

El software se va a desarrollar hasta tener un prototipo web funcional, así de esta forma la empresa puede tener un acceso más rápido y directo al sitio web ya que se lo aloja en un servidor y solamente se tiene que conectar a una dirección web lo que hace que no se tenga que instalar otros programas y así optimizar espacio en máquinas locales.

Las proyecciones y el análisis se van a elaborar en función de datos vigentes proporcionados por la empresa los cuales permitirán tener resultados más reales y certeros que faciliten a la misma poder trabajar con esa información obtenida.

En el sistema web se podrá visualizar los resultados obtenidos presentados de manera gráfica que es en donde se va a apreciar de forma más clara y precisa la información resultante, no hace falta ver como es el proceso de ejecución del algoritmo en el sistema web ya que con la base teórica se comprende cómo funciona y como se obtienen los resultados presentados en las gráficas.

Para la realización de la predicción de las ventas se desarrollan los algoritmos matemáticos los cuales son polinomios de segundo, tercer y cuarto grado, así mismo se usa las series temporales y el modelo ARIMA que permiten tener una predicción más acertada a los datos y con un margen de error mínimo lo que hace que se pueda confiar en los resultados que arroja el programa.

1.7. Unidad de análisis

Examinando los diferentes tipos de productos que comercializan las empresas, se ha llegado a la conclusión de que, para la presente investigación se va a trabajar con empresas que vendan productos tangibles, ya que así se facilita el tratamiento y análisis de los datos de las ventas.

De esta forma la empresa con la que se está trabajando es DYSUR, la cual es la distribuidora líder en el sur de Quito, que pone a disposición licores, productos de consumo masivo, confites y productos para el hogar.

Para el proyecto web la empresa proporcionó los datos de dos tipos de licores que ellos ofertan los cuales son Norteño y Ron Pon Pon ya que se ha observado que estos productos son de mucho consumo por parte del cliente y así con los datos que generan en ventas con estos productos permite tener un mayor registro histórico para que el sistema web funcione de forma más precisa.

CAPITULO 2 MARCO METODOLÓGICO

2.1. Metodología de investigación

Se realizó una investigación descriptiva la cual permite visualizar la realidad de las empresas para entender la naturaleza de su mercado, con uno de los métodos como la entrevista tanto con la investigación subjetiva que da características cualitativas y la investigación objetiva que da propiedades cuantitativas.

Se realizó esta investigación antes de realizar toda la parte técnica para comprender y puntualizar las características de la población en estudio.

Así se leyó, investigó e infirió lo importante del problema planteado luego de haber realizado la entrevista con fuentes primarias para entender el contexto del tema.

2.1.1. Método

El método que se va a usar es el descriptivo, ya que, es parte de la metodología de la investigación, este método tiene como finalidad definir, clasificar y caracterizar el caso de estudio, también permite describir y evaluar ciertas características de la información en distintos puntos del tiempo.

Aquí se analiza el conjunto de datos para definir cuales van a ser las variables que están relacionadas entre sí para tener los mejores resultados, así el método supone una observación sistemática lo que permite que la información pueda ser utilizada y aplicada en otros ámbitos.

Este método tiene pocas limitaciones para su uso, es muy útil para explicar sistemas naturales o eventos pasados mediante el análisis de una información histórica, así mismo se lo utiliza en la descripción de hechos y fenómenos actuales para luego analizarlos e interpretarlos de una manera imparcial.

2.1.2. Técnica

“Es el conjunto de procedimientos o recopilación de datos para verificar los métodos empleados en lo investigado, para llegar a la verdad del suceso estudiado, teniendo las pruebas y una serie de pasos que se llevan a cabo para comprobar la hipótesis planteada.” (gestiopolis, 2019)

Por eso las técnicas de investigación que se van a usar en este trabajo son la observación y la entrevista, todos estos procesos ayudan a recolectar de la mejor manera los datos que se necesitan.

Entonces, para obtener información a partir de los datos proporcionados la técnica de la observación es muy acertada para este trabajo. Con esta técnica se pudo determinar cómo se registrarán los datos, para luego analizarlos e interpretarlos de manera que sean de gran ayuda para el propósito de este proyecto.

Otra técnica que se utilizó es la entrevista por ser un método que permite recolectar cualitativamente la información a través de preguntas abiertas, con el fin de hacer un análisis más amplio de la información de una empresa estándar del tipo de empresas del contexto analizado. En el caso de este trabajo se tuvo una conversación con la empresa, lo que hizo que se pueda profundizar más en el tema de ventas y también conocer como esta maneja sus datos, para dar un mejor tratamiento de estos y usarlos en los algoritmos matemáticos. Esto da gran valor para que los modelos se puedan entrenar de mejor manera y así los datos sean más precisos para tener una buena predicción de ventas.

2.2. Metodología informática

2.2.1. Metodología para el desarrollo de sistemas

En su mayoría cada método contendrá una operación matemática que ayudará a la elaboración del modelo matemático que se utilizará en el sistema para la proyección de las ventas. No obstante, con el uso de la metodología no solo se facilitará el diseño y la implementación del sistema, sino, también ayudará en los futuros mantenimientos que se den al software. Todo esto para conseguir un producto de calidad y eficiente.

2.2.2. Modelo

Un modelo de ciclo de vida de software ayuda a que se lleve a cabo un proceso ordenado con el cual se pueda administrar el progreso del desarrollo del programa y que se pueda definir cuáles van a ser las etapas que se tengan que ejecutar para culminar con el desarrollo del software.

“Metodología en Cascada, también llamado Lineal secuencial, es el enfoque metodológico que ordena rigurosamente las etapas del proceso para el desarrollo de software, de tal forma que el inicio de cada etapa debe esperar a la finalización de la etapa anterior.” (Ingeniería de Software, 2019)

Por esta razón el modelo que se va a utilizar para este prototipo es el modelo en cascada.

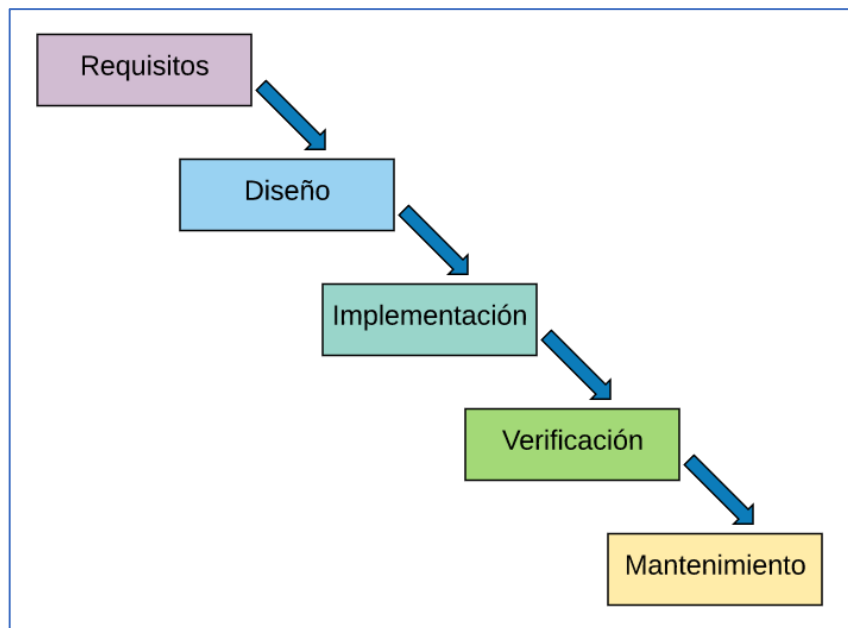


Ilustración 1. Ilustración de la metodología en cascada.

(2019). Metodología en cascada. [Imagen]. Recuperado de <https://openclassrooms.com/en/courses/4309151-gestiona-tu-proyecto-de-desarrollo/4538221-en-que-consiste-el-modelo-en-cascada>

Esta metodología es precisa para proyectos con determinados requerimientos como éste, el método en cascada se lo va a usar con más detenimiento en el capítulo 4 en donde se detallan todos los pasos que tiene el ciclo de vida de la metodología para poder culminar y que se pueda completar todo el desarrollo del software en donde cada paso se tiene que completar totalmente para poder seguir con el siguiente y así sucesivamente, este proceso hace que se pueda llevar un orden al momento de realizar el desarrollo del software lo que permite que se cumpla con todos los requerimientos y todas las fases para que se obtengan los resultados propuestos.

2.2.3. Procesos de ingeniería

Los procesos de ingeniería permiten tener un mejor manejo sobre todo lo que se tiene que hacer en el transcurso de un proyecto para que se administre a la perfección todos sus procesos.

El proceso de ingeniería que se va a usar es el RUP

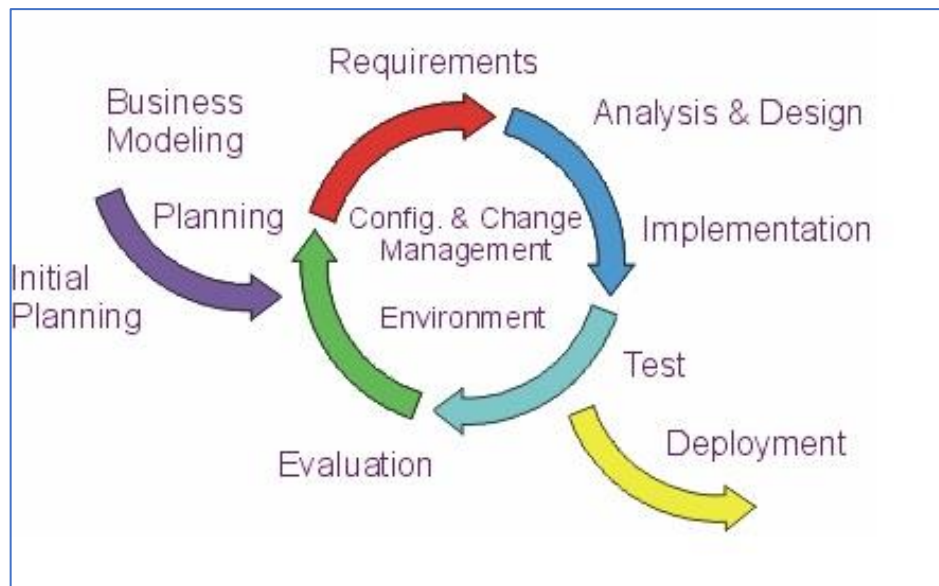


Ilustración 2. Ilustración ciclo de vida RUP.

(2019). Ciclo de Vida RUP. [Imagen]. Recuperado de <https://softwarerecopilation.wordpress.com/modelo-rup/>

En este proyecto se pretende utilizar el proceso de desarrollo de software RUP, debido a que no es un proceso en el cual sus metodologías estén firmemente formalizadas, sino que, su grupo de metodologías pueden adaptarse a las necesidades de alguna persona u organización que requiera utilizarlo.

Esto es favorable en este proyecto ya que, el sistema que se elaborará no tendrá requerimientos por el lado del cliente, si no, estos requerimientos son definidos por las personas que desarrollarán el sistema. También cabe mencionar que el sistema será diseñado, implementado, verificado y puesto en despliegue como prototipo funcional y no como producto final.

Para el desarrollo del proceso RUP se necesita emplear el UML el cual ayuda a que se pueda especificar, construir y documentar sistemas. UML apoya al momento de estructurar y desarrollar el prototipo ya que requiere ir desde el análisis al diseño y así completar todo el proceso de desarrollo.

El UML permite representar gráficamente todos los modelos o subsistemas que son necesarios para el desarrollo del sistema.

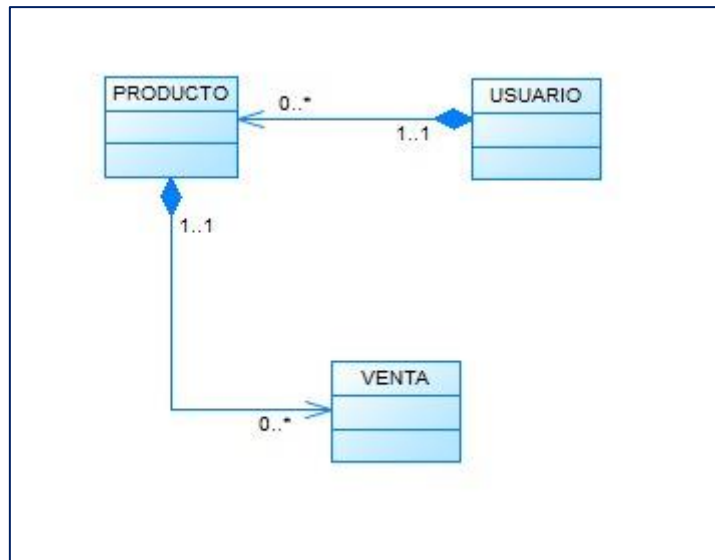


Ilustración 3. Ilustración gráfica UML por composición

(2019). UML por composición. [Recorte]. Recuperado de <https://softwarerecopilation.wordpress.com/modelo-rup/>

CAPITULO 3 MARCO DE REFERENCIA

3.1. Marco teórico

La predicción de ventas hoy en día es más común de lo que parece ya que así las empresas pueden tomar diferentes decisiones y monitorear diversos movimientos que se vayan a realizar para el siguiente periodo de tiempo, esto se lo realiza la mayoría de las veces manualmente, así las empresas implementan métodos como el Delphi, juicios de expertos en el área, encuestas o realizan un feedback de lo que se ha hecho en el periodo de tiempo calculado, sin embargo, con el avance tecnológico estos procesos manuales se pueden optimizar y las empresas pueden disminuir el uso de recursos.

Con el apoyo de las tecnologías existentes este proyecto se va a desarrollar utilizando herramientas que facilitan el desarrollo de sistemas web, permitiendo al usuario tener una interacción amigable con el sistema. De esta manera se ha planteado usar el framework de PHP LARAVEL que trabaja en MVC (modelo vista controlador), así se lo combina con el motor de base de datos relacional SQL SERVER, esta herramienta de Microsoft tiene dos versiones gratuitas: Developer y Express; para este proyecto se va a usar la versión Developer 18 la cual proporciona un modelo de licencia por usuario, aquí se podrá crear, probar y demostrar aplicaciones de forma eficiente ya que permite tener mayor flexibilidad, escalabilidad, confiabilidad, y seguridad al momento de usarlo en proyectos, lo que lo hace apto para soportar y trabajar con grandes volúmenes de información.

Para la proyección de las ventas se va a usar la regresión polinomial y series temporales con los modelos ARIMA y SARIMA, con la ayuda del lenguaje Python los cálculos matemáticos que estos métodos de proyección utilizan se facilitan de gran manera, así se podrá integrar dos lenguajes que se complementan al momento de ser ejecutados, y esta integración se la realizará mediante el editor de código fuente Visual Studio Code esta herramienta de Microsoft es de software libre que se distribuye bajo la licencia MIT gratuita, es una licencia simple y permisiva que solo requiere que se conserve los derechos de autor y condiciones de uso, no tiene limitaciones y su uso es comercial y privado permitiendo modificar y distribuir este software. Se trabajará a la par con el framework LARAVEL, los lenguajes de programación PHP y Python y el motor de base de datos SQL SERVER.

3.2. Marco conceptual

Los diferentes temas que se pueden manejar para este trabajo aportan de diversas maneras para tener un marco conceptual amplio que apoye el trabajo investigativo para dar la solución al problema.

Se necesita primero analizar la información que la empresa proporciona ya que es un archivo histórico en Excel el cual posee todas sus variables que como empresa manejan para sus registros, por esta razón se debe realizar diferentes pruebas con una parte de los datos que se tienen para filtrar información que no sea de utilidad al momento de realizar la proyección de ventas. Las pruebas se las realizaron en Excel, se tomó una muestra de datos del mes de agosto del año 2016, se clasificaron los mismos teniendo en cuenta algunas variables manejadas por la empresa que son: fecha, salida, cantidad y costo total, para luego juntar la información en una tabla la cual presenta los datos agrupados por fecha, unidades y precio esto permite tener los datos para realizar las gráficas y visualizar que los polinomios de tercer y cuarto grado se ajustan y tienen una mejor tendencia al tipo de datos que la empresa maneja.

El sistema web está basado en el framework PHP LARAVEL creado en 2011, posee un sistema de plantillas conocido como Blade el cual tolera usar código PHP simple dentro de las vistas teniendo un desarrollo más elegante y simple, permitiendo que su implementación sea amigable para el programador ya que evita el “código espagueti”, en sí es muy potente para el desarrollo web.

El lenguaje principal para desarrollar el proyecto es PHP (preprocesador de hipertexto), es de código abierto lo cual indica que puede ser usado de forma libre y gratuita, además es adecuado para realizar desarrollo web de forma dinámica y se complementa con HTML para la parte gráfica. Usado en la mayoría de los servidores web y se adapta con diferentes sistemas operativos.

Por otro lado, para realizar los cálculos matemáticos se realizarán en el lenguaje de scripting PYTHON, es una plataforma independiente y orientado a objetos, fue creado en 1991 por Guido Van Rossum, es adaptable para realizar cualquier aplicación o incluso páginas web, una de las grandes características que posee es que es un lenguaje interpretado lo que significa que no necesita que el código se compile para que pueda ser ejecutado lo que permite trabajar de forma más rápida.

El objetivo principal del proyecto es predecir ventas con la ayuda de algoritmos matemáticos. Para lo cual, se usará la Regresión Polinomial, que es una variante de la regresión lineal ya que se necesita agregar predictores adicionales al elevar cada uno de los predictores originales a una potencia, esta regresión se usa para proporcionar un ajuste no lineal a los datos que es lo que se necesitará para el tipo de datos que la empresa proporcionó para realizar el sistema web.

Para ejecutar la regresión polinomial se trabajará con un histórico de datos de cinco años así el grafico resultante se puede proyectar mejor con un error reducido en sus proyecciones

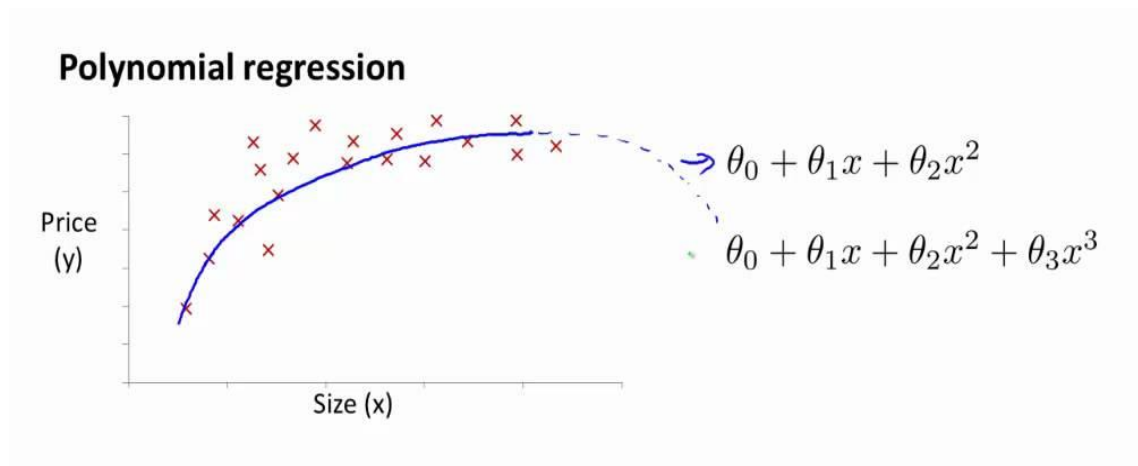


Ilustración 4. Ejemplo regresión polinomial

(2019). Regresión polinomial [Imagen]. Recuperado de <https://www.feedingthemachine.cl/regresion-polinomial-machine-learning/>

Se usará series temporales las cuales ayudan a que se pueda visualizar una gráfica de manera ordenada, ya que se hace una serie de observaciones de las variables tomadas en el transcurso del tiempo, los datos se agrupan en diferentes tiempos o intervalos lo que permite tener mejor precisión y acomodación en las gráficas resultantes, esto se hace porque las series temporales son las que se basan bastante en la evolución de los datos a través del tiempo ya que para este proyecto se necesita utilizar el histórico de datos de las ventas de la empresa Dysur.

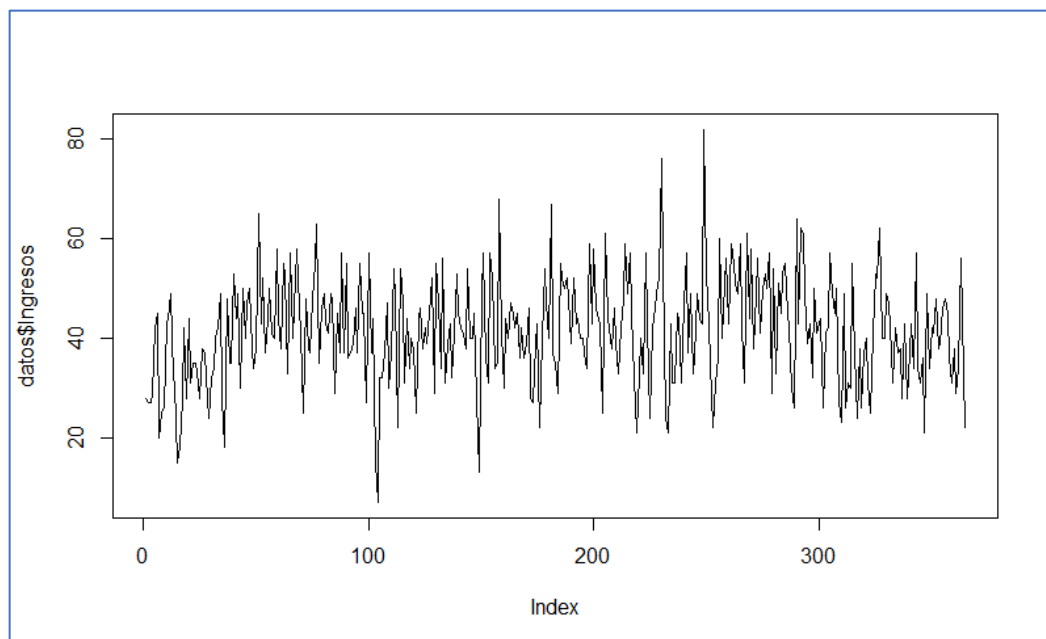


Ilustración 5. Gráfico ejemplo de series temporales

(2018). Ejemplo series temporales [Imagen]. Recuperado de <https://www.ciencias.com/pregunta/106323/contar-series-temporales-de-datos-para-llegadas-de-emergencia-a-hospitales>

Otro método que se implementará para realizar las predicciones son las Series Temporales con el modelo ARIMA ya que analizando los datos proporcionados se puede notar que estos tienen gran fluctuación en cuanto a sus ventas y mediante esta técnica se tiene una proyección apropiada para este tipo de datos.

“Cuando hablamos de series temporales nos referimos a conjunto de datos que están agrupados en intervalos de tiempo secuencialmente ordenados. Estos intervalos pueden ser diarios, semanales, quincenales, mensuales, etc.” (Bagnato, 2019).

Se tiene un conjunto de datos agrupados en intervalos de tiempo lo que se puede usar este método ya que ayuda a pronosticar datos usando valores históricos es decir toma registros anteriores dentro de los intervalos de tiempo y mediante el modelado predice los valores a futuro este proceso se lo conoce como pronóstico de serie de tiempo univariante que se complementa con el modelo ARIMA (promedio móvil integrado autorregresivo) este es un algoritmo de pronóstico que se basa en la misma idea de las series temporales porque solo puede predecir los datos teniendo una información anterior como es el caso.

“El modelo ARIMA tiene 3 términos: p , d y q donde:

- *p es el orden del término AR.*
- *q es el orden del término MA.*
- *d es la cantidad de diferenciación requerida para hacer estacionarias las series de tiempo.”*

(Prabhakaran, 2019)

Para ejecutar el modelo ARIMA se tiene que transformar a las series temporales en estacionarias ya que se basa en un modelo de regresión lineal que usa sus retrasos como predictores los cuales tienen que ser independientes y no correlacionales, el término d es el número mínimo necesario para que la serie sea estacionaria si este ya es el caso el término toma el valor de cero, en cambio p es el número de rezagos de “ y ” que son utilizados como predictores y q es la cantidad de errores de pronóstico rezagados que se incluyen en el modelo ARIMA.

“Un modelo puro de regresión automática (solo AR) es aquel en el que Y_t depende solo de sus propios retrasos. Es decir, Y_t es una función de los “retrasos de Y_t ”.

[OBJ]

donde, Y_{t-1} es el rezago1 de la serie, β_1 es el coeficiente de lag1 que estima el modelo y α es el término de intercepción, también estimado por el modelo.

Del mismo modo, un modelo de media móvil pura (solo MA) es aquel en el que Y_t depende solo de los errores de pronóstico rezagados.

[OBJ]

donde los términos de error son los errores de los modelos autorregresivos de los respectivos retrasos.” (Prabhakaran, 2019)

Para complementar la investigación se usará el modelo SARIMA, el cual es muy usado en series temporales ya que poseen componentes estacionales que se repiten en periodos u observaciones. Para trabajar la estacionalidad de los datos a los modelos ARIMA se los generaliza y aparece el modelo SARIMA (Seasonal Autoregressive Integrated Moving Average Model). Es un modelo estacional puro (P, D, Q) es decir que capta únicamente el comportamiento puro de una serie; se lo representa así:

$$\phi_p(B^s) \nabla_s^D (Z_{t-u}) = \theta_Q(B^s) a_t$$

donde la constante μ es el nivel del proceso original Z_t .

$$\phi_p(B^s) = 1 - \phi_s B^s - \phi_{2s} B^{2s} - \dots - \phi_{ps} B^{ps}$$

Es un polinomio autorregresivo estacional de orden P

$$\theta_Q(B^s) = 1 - \theta_s B^s - \theta_{2s} B^{2s} - \dots - \theta_{Qs} B^{Qs}$$

Es un polinomio de promedios móviles estacional de orden Q .

a_t : Es un proceso de ruido blanco.

Ilustración 6. Gráfico representación modelo SARIMA

(2019). Gráfico modelo SARIMA [Imagen]. Recuperado de
Obtenido de: http://www.scielo.org.bo/scielo.php?script=sci_arttext&pid=S2077-33231997000100005

Todos estos conceptos matemáticos se usaron en el desarrollo del sistema web, ya que de esta manera se puede hacer las diferentes operaciones para dar un tratamiento adecuado a los datos y así poder automatizar este proceso necesario en la empresa para llegar a tener los resultados y predecir las ventas de una manera eficiente y real.

3.3. Marco legal

Para el presente trabajo se toma en cuenta diferentes bases legales que apoyan a la investigación con las cuales se llega a tener un fundamento justo y adaptado sobre las diversas reglas y leyes que permite tener alcance con este trabajo las cuales son leyes que van ligadas a esta investigación como la Ley de Protección de Datos

Personales, las de derecho de autor en sus diferentes modalidades y la de software protegido por derechos de autor las cuales siendo analizadas permiten dar paso para que el trabajo se genere conociendo y aplicando estas normas para no caer en algún delito que ponga en riesgo los datos de la empresa DYSUR la cual proporcionó su información de ventas y la fidelidad de este proyecto web.

“Desarrollo tecnológico e inversión. - Se debe promover el desarrollo y la utilización de nuevos servicios, redes y tecnologías de la información y las comunicaciones y su acceso universal a toda la población y fomentar la inversión pública y privada.”
(Ley Organica de Telecomunicaciones. Art 94 Capitulo I)

Los datos proporcionados son información netamente de la empresa DYSUR y por esta razón en este proyecto solo se usan los datos con fines académicos, el uso y manipulación de los datos son como ejemplo para que se pueda tener resultados reales y verídicos cuando se ponga en funcionamiento el sitio web realizado en el proyecto.

“Art.- 35.- Derecho al acceso universal a las tecnologías de la información y comunicación. - Todas las personas tienen derecho a acceder, capacitarse y usar las tecnologías de información y comunicación para potenciar el disfrute de sus derechos y oportunidades de desarrollo.” (Telégrafo, 2013)

Además, como está escrito en el Artículo 35 Sección II de la Ley Orgánica de Comunicación todas las personas tienen derecho a acceder a las tecnologías de la información y comunicación para su beneficio y mejorar oportunidades de desarrollo como es en este caso ya que el sistema web permite automatizar el proceso de predicción de ventas de la empresa DYSUR mediante el uso de la tecnología lo que ayuda a que se economice tiempo y recursos humanos para realizar este procedimiento.

CAPITULO 4 DESARROLLO DE SOFTWARE

4.1. Análisis

En este capítulo se va a realizar una descripción de los diferentes componentes que se va a usar en este proyecto, se necesita generar diversos modelos de pruebas y de datos para obtener respuestas distintas mediante el uso de algunas herramientas que apoyan el desarrollo del prototipo web, también se va a especificar cómo será el funcionamiento de este proyecto y como se va a lograr cumplir con los objetivos propuestos.

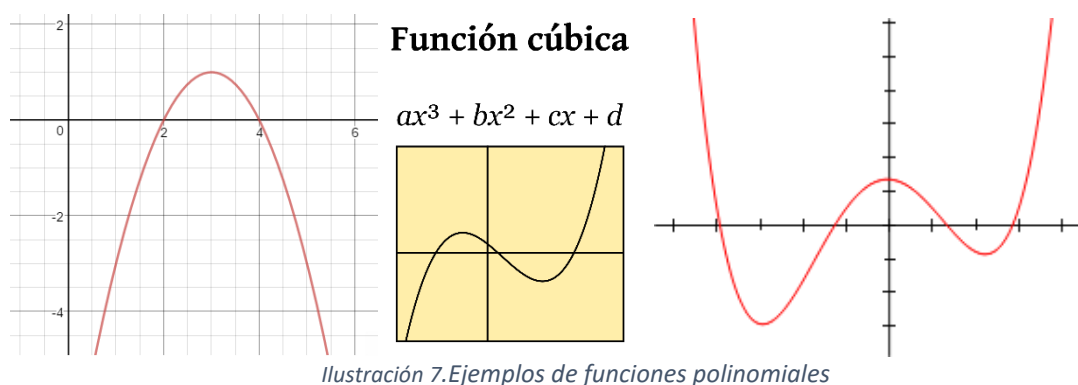
Mediante los diversos análisis y pruebas que se hicieron en base a los datos que se tienen de la empresa se pudo determinar cuál va a ser el lenguaje y la herramienta más optima a utilizar para generar el código fuente del sitio web, así como también se llegó a decidir el algoritmo matemático que más se adaptó a la estructura de los datos que dio la empresa para generar una predicción de ventas correcta.

Para empezar, se necesita obtener la información la cual se la consigue mediante algunas técnicas como la entrevista, se tuvo una conversación con personas de la empresa (gerente y la supervisora) los cuales manifestaron que tipos de productos son los que más se venden en su línea de negocio, así mismo comentaron como su empresa ha ido ganando mercado para poder tener una buena posición dentro del sur de Quito, esta charla ayuda a que se pueda generar un análisis rápido de los datos proporcionados y a focalizar que tipos de algoritmos matemáticos se pueden usar, de inmediato se piensa que los datos son tan variables que se los puede representar mediante una curva que por medio de un polinomio de cuarto grado se los podría representar para realizar la proyección.

Después de la entrevista y de haber hecho un análisis rápido de lo que son los datos, la información es entregada en un archivo digital en un formato que las empresas manejan para sus reportes de ventas de los cinco últimos años de dos de los productos estrella que la empresa ha notado que tienen un mayor consumo por parte del cliente lo que ayuda que se tenga mucha información y de esta manera se pueda entrenar a los algoritmos para realizar la predicción más precisa.

Una vez obtenida la información se necesita tomar decisiones para saber que algoritmos se tienen que trabajar para realizar la predicción, primero se puede observar que los datos en los años 2014 y 2015 son relativamente bajos pero desde ahí la empresa empieza a tener un crecimiento progresivo y se nota una diferencia en la curva esto se realizó dentro de Excel simulando el cálculo de polinomios de tercer y cuarto grado lo que permite tomar una primera decisión que es empezar a modelar estos algoritmos ya en PHP porque se nota que la curva que arroja el cálculo del polinomio se va adaptando a la forma en cómo se van presentando los datos.

Mientras se va realizando los cálculos matemáticos con las funciones polinomiales de segundo, tercer y cuarto grado podemos llegar a un punto importante en el análisis ya que este tipo de algoritmos tienden a la figura de su grado es decir si se tiene un polinomio de segundo grado se grafica una parábola, un polinomio de tercer grado una función cúbica y uno de cuarto grado una gráfica polinomial, así cada función polinómica se adapta a la figura que presenta su grado como se puede observar en la imagen a continuación:



(2019). Ejemplos de funciones polinomiales. [Imagen]. Recuperado de <http://cimanet.uoc.edu/cursMates0/IniciacionMatematicas/pdf/C%2023Las%20Funciones%20Polinomicas.pdf>

con estos argumentos se decide que las proyecciones no serían óptimas mientras se realicen estas funciones y no sirve para tener los mejores resultados al momento de proyectar el tipo de datos que la empresa tiene ya que la gráfica resultante tiene otra forma que ya no se adapta a estas funciones.

Por esta razón se decide buscar y analizar otro tipo de modelo matemático que se pueda ajustar a los datos permitiendo entrenarlos para sacar las gráficas resultantes con una buena proyección así se empieza a investigar las series temporales con los modelos ARIMA y SARIMA ya que estos están hechos para predecir series como lo dice su nombre, estos modelos son más confiables debido a que proporciona predicciones óptimas en corto plazo ya que trabaja con diferentes rangos mediante sea el caso de los datos.

Se dispone a ejecutar el modelo ARIMA en donde se debe tener presente que es un proceso estocástico y no estacional es decir que el modelo usa variables aleatorias juntamente con un parámetro como el tiempo que hace que sea estacional, esta razón hace que se pruebe a la par el modelo SARIMA que tiene una diferencia peculiar este algoritmo trabaja con datos estacionales lo que hace que sea el que mejor se adapte a la tendencia de estos datos, después de esta larga investigación se llega a la conclusión de que la mejor proyección para esta empresa es con el modelo SARIMA ya que da un margen de confiabilidad del noventa y cinco por ciento lo que hace que su error sea mínimo y sus resultados sean reales.

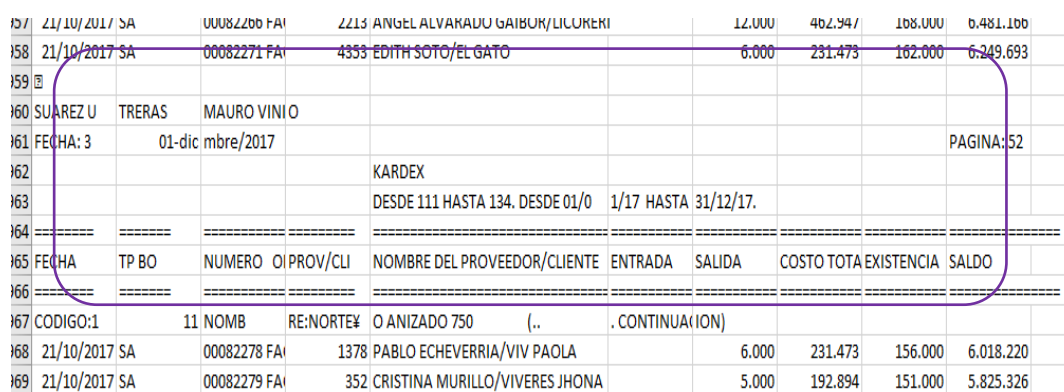
Así se va a desarrollar en el sistema web los dos tipos de algoritmos matemáticos la función polinomial y las series temporales para poder realizar una comparación de como los modelos actúan con el tipo de datos que se tiene en la base y poder observar los diferentes resultados para la toma de decisiones.

4.2. Diseño

Este trabajo consiste en realizar un prototipo web que permita predecir ventas en base a un histórico de datos de una empresa, para esto se va a analizar modelos matemáticos que se aplican para realizar proyecciones, y se determinarán las variables requeridas para los modelos, se identificará el lenguaje de programación y la herramienta más adecuada para trabajar con Inteligencia Artificial, y de esta manera se diseñará y construirá el prototipo web funcional.

El prototipo se lo desarrollará en PHP en conjunto con PYTHON, en donde se programarán los algoritmos matemáticos en base a los modelos de predicción de ventas, luego se cargarán los datos históricos para que los algoritmos puedan hacer el proceso de aprendizaje, al final después de varias pruebas generar la respectiva gráfica y así el usuario pueda dar un buen uso a la información generada por este prototipo.

La parte del diseño empieza cuando se tiene que hacer una limpieza de los datos, ya que el archivo que la empresa entrega no siempre está adecuado a la forma de la base con la que se va a trabajar, el archivo está en formato .xlsx con toda la información pero observando el mismo se nota que cada ciertos periodos de tiempo existe una separación de celdas en las cuales están escritos los encabezados que tiene cada columna como se puede observar en la siguiente imagen:



21/10/2017	SA	00082278	FAI	1378	PABLO ECHEVERRIA/VIV PAOLA	6.000	231.473	156.000	6.018.220
21/10/2017	SA	00082279	FAI	352	CRISTINA MURILLO/VIVERES JHONA	5.000	192.894	151.000	5.825.326

Ilustración 8. Ejemplos de funciones polinomiales

(2019). Ejemplo de archivo Excel. [Recorte].

esto hace que se tenga que borrar estos enunciados para que la información continúe uniformemente y no tenga espacios demás ya que así no se podrá leer el archivo en la base de datos porque serían campos que no forman parte de la estructura de datos que se necesita y el motor de base de datos generaría errores.

Es momento de cargar los datos al motor de base SQL Server con el cual se va a trabajar, el archivo una vez depurado se lo tiene que convertir a un .txt. El archivo consta solo de tres columnas (PRODUCTO_ID, VENTA_TOTAL_UNIDADES, VENTA_FECHA) las cuales poseen toda la información necesaria para realizar la proyección, este archivo puede estar localizado en cualquier ubicación dentro del computador.

Se deberá tener una visión general de cómo va a estar la interfaz para el sistema web, debe contener una pantalla principal en donde se muestre un menú desplegable con las diferentes opciones que se necesitan para ejecutar el sistema, cada que se seleccione una opción desplegará nuevas ventanas en donde se podrá visualizar las diversas interfaces con cajas de texto, labels, botones y combo box que hacen que se tenga un sistema completo para poder ejecutarlo y que presente los resultados en gráficas y tablas para que se las pueda apreciar mejor. Cada ventana del sistema web estará dividida en la estructura que permite el HTML para que así todos sus componentes se puedan ver de una forma amigable y ordenada.

La interfaz del sistema web se la hace con BLADE el motor de plantillas de LARAVEL esta parte se aloja dentro de la carpeta de vistas del proyecto de Laravel, aprovechando los beneficios de este motor se crea la página principal con el uso de secciones y herencia de plantillas para que así se pueda organizar y estructurar de una forma más eficiente y amigable la página, se tiene la parte del header que permite presentar el encabezado del sitio web además se encuentra la barra de navegación, el login del usuario y la selección del dashboard, en la sección del nav se tiene el sidebar o la barra lateral que permite presentar el listado de opciones del dashboard, en el div se tiene el espacio para mostrar todo el contenido principal del sitio web y para concluir se tiene el footer en donde se encuentra los datos informativos del desarrollo del sistema web.

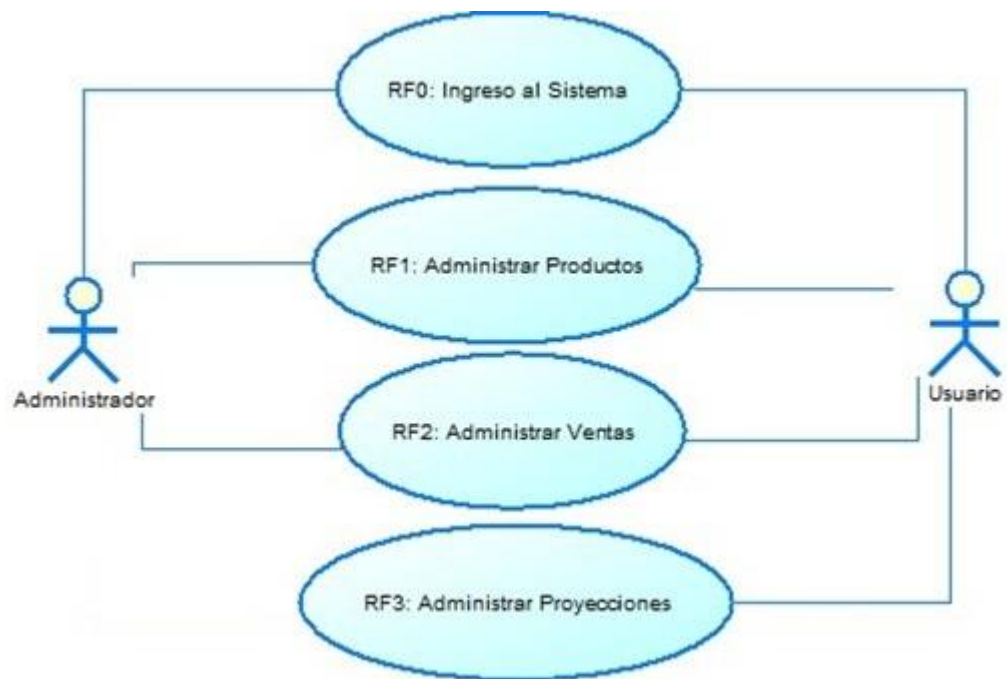


Ilustración 9. Ilustración gráfica Diseño Interfaz.

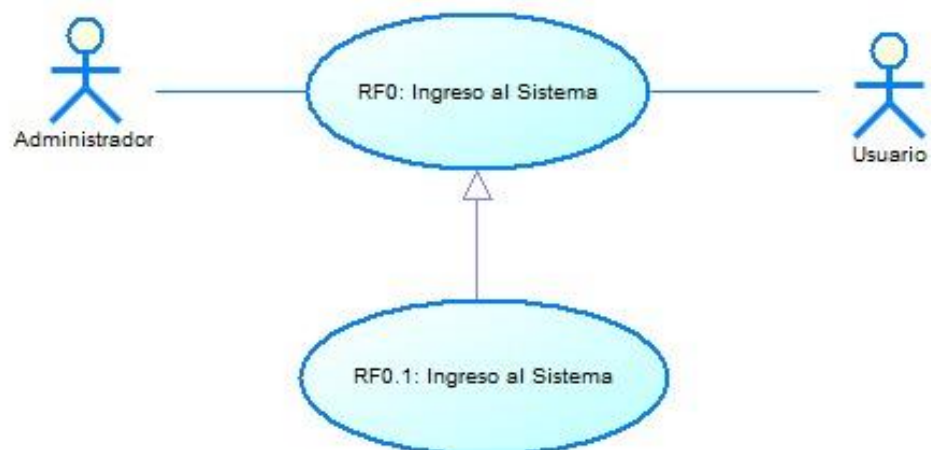
(2019). Interfaz principal [Captura de imagen].

4.2.1. Requisitos funcionales

4.2.1.1. Diagrama General



- Segundo Nivel RF0: Ingreso al Sistema



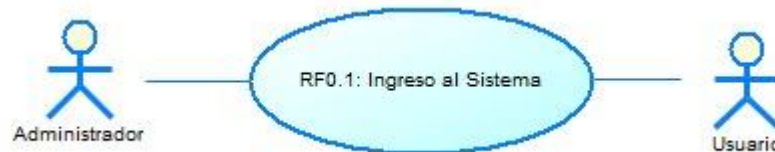
- A Detalle RF0

1) **Nombre:** RF0.1 Ingreso al Sistema.

Descripción: El sistema actúa conforme a los siguientes requerimientos para poder ingresar al sistema.

Actores: Administrador, usuario.

Gráfico:



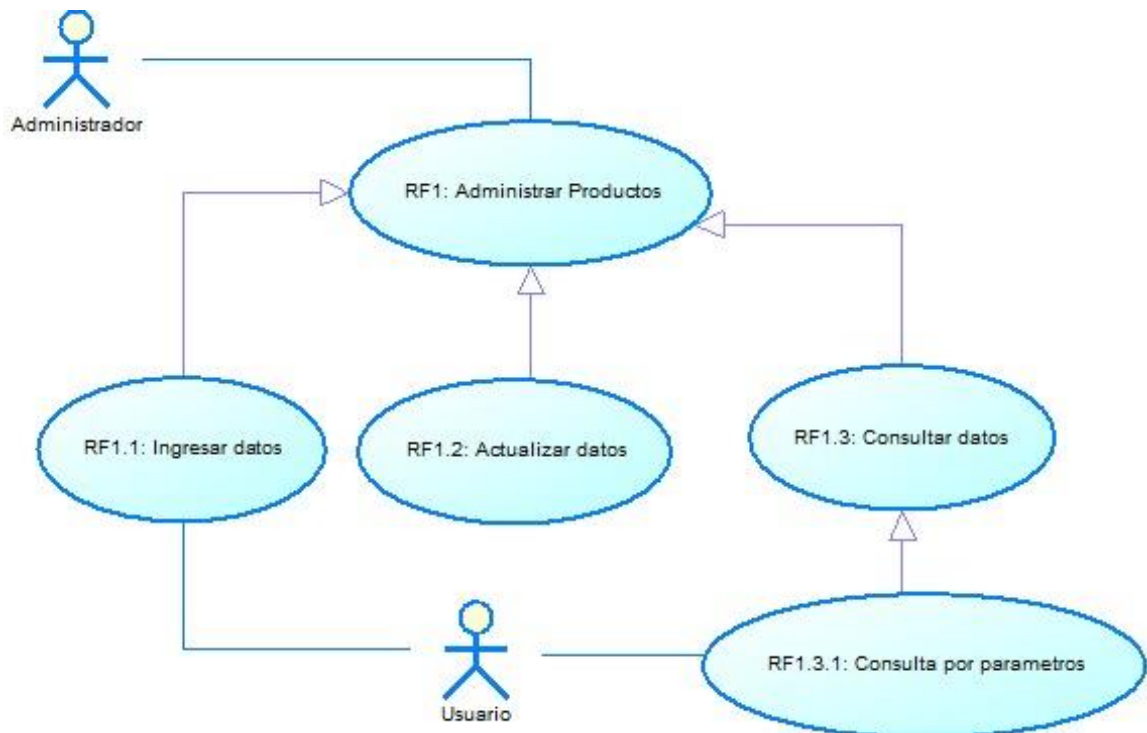
Flujo Principal:

1. El actor va a llenar los campos “username y password” correspondientes.
2. El actor es el que va a presionar la opción de “Iniciar sesión”. E1
3. El sistema muestra la ventana con las opciones disponibles que puede seleccionar.

Excepciones:

Problema	Detalle	Acción
E1.	Ingreso al sistema incorrecto.	Solución: Ingresar correctamente los datos.

- Segundo Nivel RF1: Administrar Productos



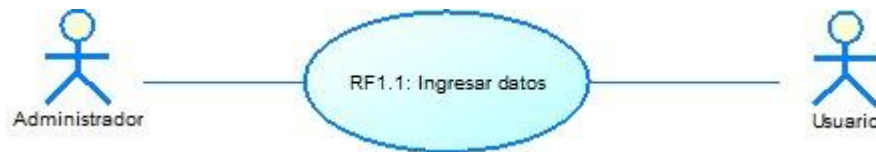
- A Detalle RF1

1) Nombre: RF1.1 Ingresar datos.

Descripción: El sistema actúa conforme a los siguientes requerimientos para poder ingresar datos.

Actores: Administrador, usuario.

Gráfico:



Flujo Principal:

1. El actor es el que va a seleccionar “Cargar Datos” dentro del menú de la pantalla principal del sistema.
2. El sistema presenta la nueva ventana que es Cargar Datos.
3. El actor debe pulsar el botón “Seleccionar archivo”.
4. El sistema muestra la ventana “Abrir” en la cual se debe ubicar el archivo en donde está la información para cargar y se da clic en “aceptar”.
5. El actor selecciona del combo box la tabla “Productos” en la cual se van a ingresar los datos.
6. El actor da clic en el botón “Cargar datos”.
7. El sistema muestra un mensaje informativo del estado de la carga de los datos. E1
8. El sistema almacena la información en la base de datos.

Excepciones

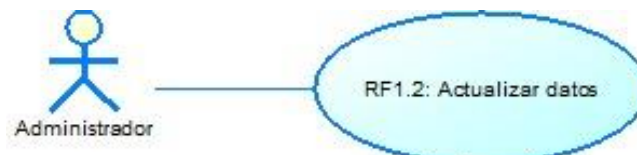
Problema	Detalle	Acción
E1.	Posible error en la conexión con la base de datos o información incorrecta	Solución: Volver a intentar desde paso 1.

2) Nombre: RF1.2 Actualizar datos.

Descripción: El sistema actúa conforme a los siguientes requerimientos para poder actualizar datos.

Actores: Administrador.

Gráfico:



Flujo Principal:

1. El actor es el que va a seleccionar “Productos” dentro del menú de la pantalla principal del sistema.
2. El sistema presenta la nueva ventana que es Listado de Productos.
3. El actor debe pulsar el botón “Editar”.

4. El sistema permite la modificación del nombre del producto.
5. El actor da clic en el botón “Actualizar”.
6. El sistema muestra un mensaje de actualizado correctamente. E1

Excepciones

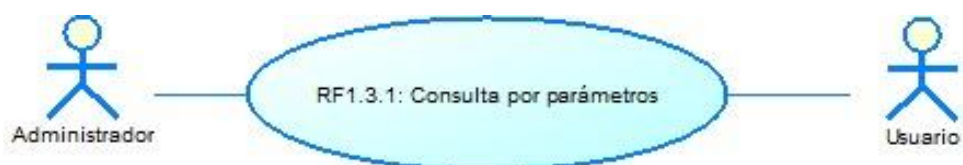
Problema	Detalle	Acción
E1.	Posible error en la conexión con la base de datos o error de sintaxis en el script de actualizar	Solución: Regresar al paso 3.

3) Nombre: RF1.3.1 Consultar por parámetros.

Descripción: El sistema actúa conforme a los siguientes requerimientos para cuando el actor necesita realizar una consulta de manera parametrizada de la información que ya se encuentra en la base de datos.

Actores: Administrador, usuario.

Gráfico:



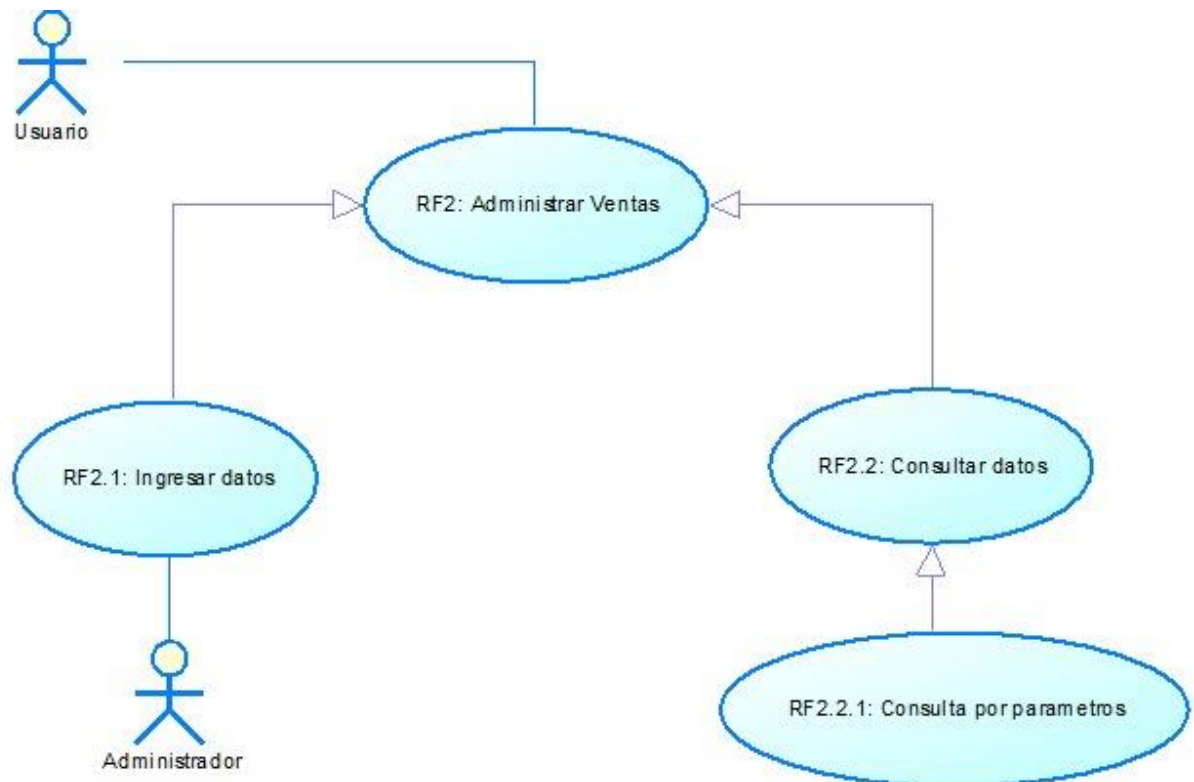
Flujo Principal:

1. El actor es el que va a seleccionar “Productos” dentro del menú de la pantalla principal del sistema.
2. El sistema presenta la nueva ventana que es Listado de Productos.
3. El actor debe pulsa el textbox “Ingrese el código de producto”.
4. El actor ingresa el código del producto que quiere consultar y presiona el botón “Buscar”.
5. El sistema realiza la búsqueda en base al código ingresado por el actor.
6. El sistema despliega un listado con los productos que contengan el código que fue ingresado por el actor. E1

Excepciones

Problema	Detalle	Acción
E1.	Posible error en la conexión con la base de datos o código ingresado no existe en la base de datos.	Solución: Volver al paso 3.

- Segundo Nivel RF2: Administrar Ventas



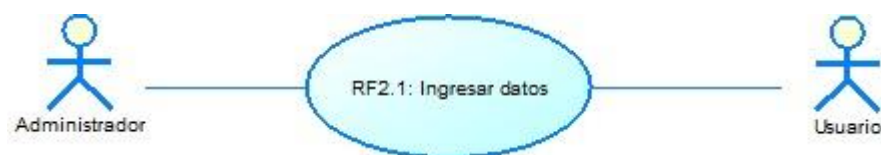
- A Detalle RF2

1) Nombre: RF2.1 Ingresar datos.

Descripción: El sistema actúa conforme a los siguientes requerimientos para poder ingresar datos.

Actores: Administrador, usuario.

Gráfico:



Flujo Principal:

1. El actor es el que va a seleccionar “Cargar Datos” dentro del menú de la pantalla principal del sistema.
2. El sistema presenta la nueva ventana que es Cargar Datos.
3. El actor debe pulsar el botón “Seleccionar archivo”.
4. El sistema muestra la ventana “Abrir” en la cual se debe ubicar el archivo en donde está la información para cargar y se da clic en “aceptar”.
5. El actor selecciona del combo box la tabla “Ventas” en la cual se van a ingresar los datos.
6. El actor da clic en el botón “Cargar datos”.

7. El sistema muestra un mensaje informativo del estado de la carga de los datos. E1
8. El sistema almacena la información en la base de datos.

Excepciones

Problema	Detalle	Acción
E1.	Posible error en la conexión con la base de datos o información incorrecta	Solución: Volver a intentar desde paso 1.

2) Nombre: RF2.2.1 Consulta por parámetros.

Descripción: El sistema actúa conforme a los siguientes requerimientos para cuando el actor necesita realizar una consulta de manera parametrizada de la información que ya se encuentra en la base de datos.

Actores: Administrador, usuario.

Gráfico:



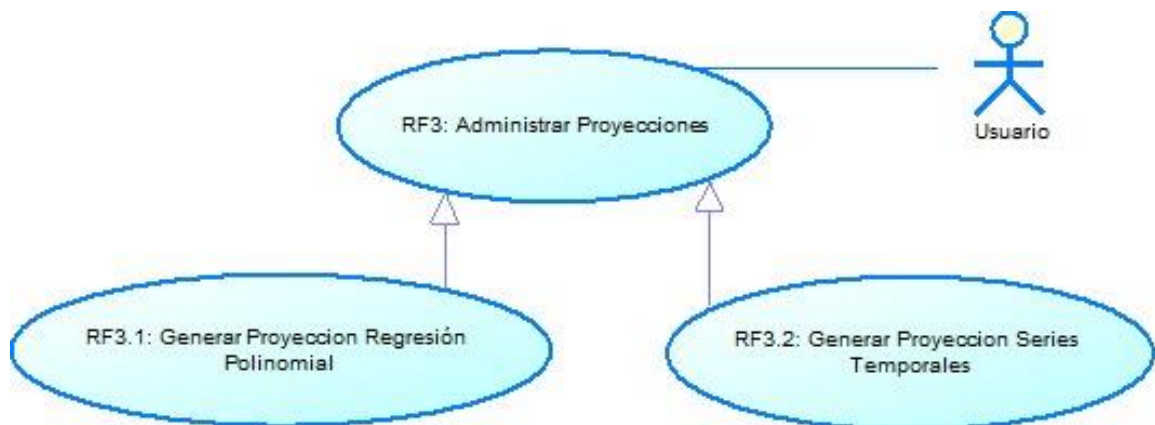
Flujo Principal:

1. El actor es el que va a seleccionar “Ventas” dentro del menú de la pantalla principal del sistema.
2. El sistema presenta la nueva ventana que es Listado de Ventas.
3. El actor debe pulsa el textbox “Ingrese el nombre de producto”.
4. El actor ingresa el nombre del producto que quiere consultar y presiona el botón “Buscar”.
5. El sistema realiza la búsqueda en base al nombre ingresado por el actor.
6. El sistema despliega un listado con las ventas del producto ingresado por el actor. E1

Excepciones

Problema	Detalle	Acción
E1.	Posible error en la conexión con la base de datos o código ingresado no existe en la base de datos.	Solución: Volver al paso 3.

- Segundo Nivel RF3: Administrar Proyecciones



- A Detalle RF3

1) **Nombre:** RF3.1 Generar Proyección Regresión Polinomial.

Descripción: El sistema actúa conforme a los siguientes requerimientos para poder generar la proyección.

Actores: Usuario.

Gráfico:



Flujo Principal:

1. El actor es el que va a seleccionar “Regresión Polinomial” dentro del menú de la pantalla principal del sistema.
2. El sistema presenta la nueva ventana que es Proyección de Ventas Regresión Polinomial.
3. El actor debe seleccionar del combo box el producto que desea realizar la proyección.
4. El actor selecciona el textbox donde se ingresa el grado del polinomio.
5. El actor ingresa el valor requerido y presiona el botón “Proyectar”. E1
6. El sistema muestra un gráfico donde se visualiza el polinomio generado con los datos del producto elegido.
7. El sistema muestra un cuadro en el cual se visualiza el producto elegido, el grado y el polinomio generado y el residuo entre los datos del producto con los datos proyectados.
8. El sistema muestra un gráfico con la proyección generada versus los datos reales.

Excepciones

Problema	Detalle	Acción
E1.	Posible error en la conexión con la base de datos o en el cálculo de la proyección.	Solución: Volver al paso 1.

2) **Nombre:** RF3.2 Generar Proyección Series Temporales.

Descripción: El sistema actúa conforme a los siguientes requerimientos para poder generar la proyección.

Actores: Usuario.

Gráfico:



Flujo Principal:

1. El actor es el que va a seleccionar “Series Temporales” dentro del menú de la pantalla principal del sistema.
2. El sistema presenta la nueva ventana que es Proyección de Ventas Series Temporales.
3. El actor debe seleccionar del combo box el producto que desea realizar la proyección.
4. El actor selecciona del combo box el número de periodos a proyectar.
5. El actor presiona el botón “Proyectar”. E1
6. El sistema muestra una gráfica con los datos reales y la proyección generada.
7. El sistema muestra una tabla en donde se puede observar los datos proyectados en el periodo seleccionado.

Excepciones

Problema	Detalle	Acción
E1.	Posible error en la conexión con la base de datos o en el cálculo de la proyección.	Solución: Volver al paso 1.

4.2.2. Diagramas de Interacción

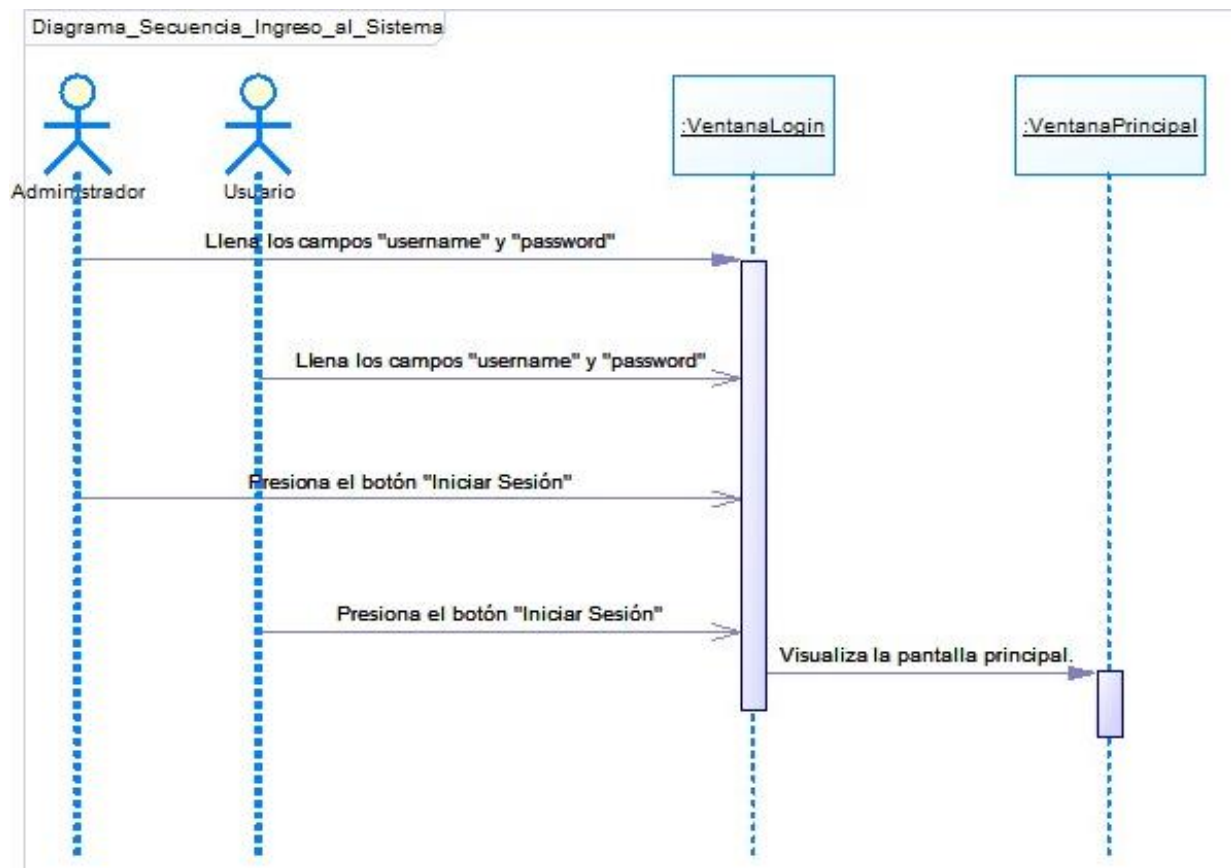
4.2.2.1. Diagramas de Secuencia

- RF0 Ingreso al Sistema

Caso de Uso: Ingreso al Sistema

Identificador: RF0.1

Gráfico:

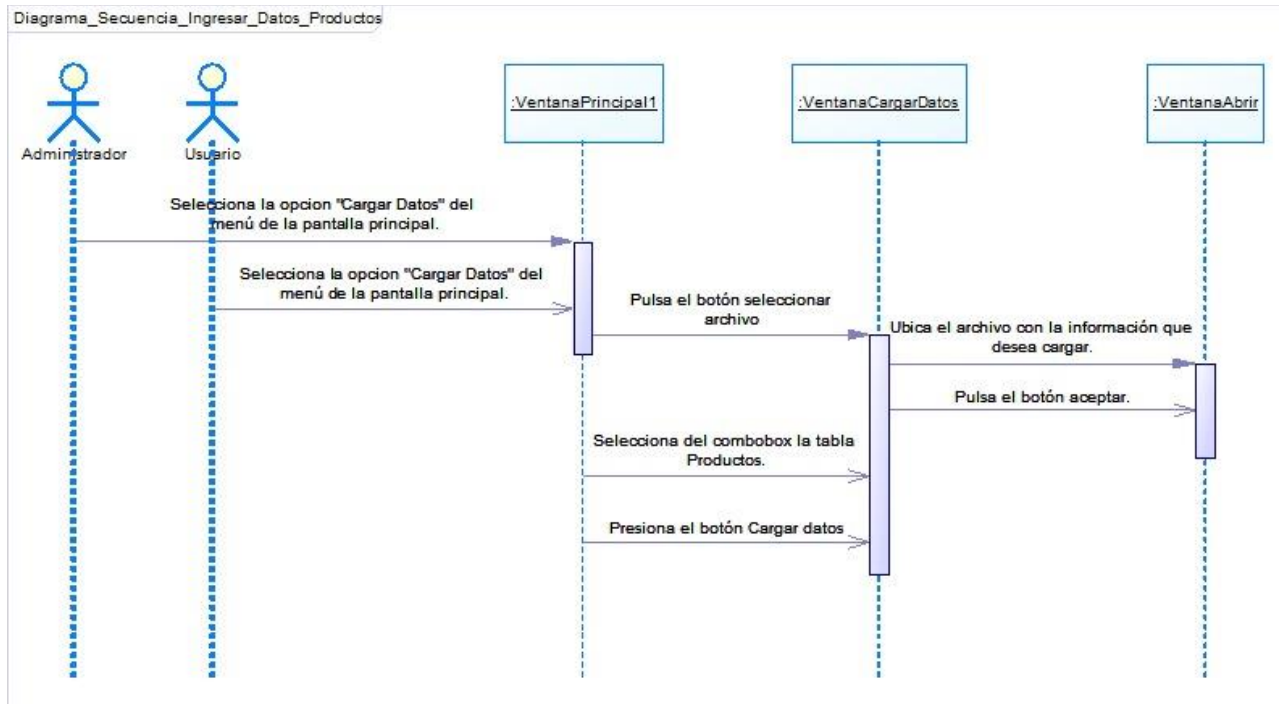


- RF1 Administrar Productos

Caso de Uso: Ingresar Datos

Identificador: RF1.1

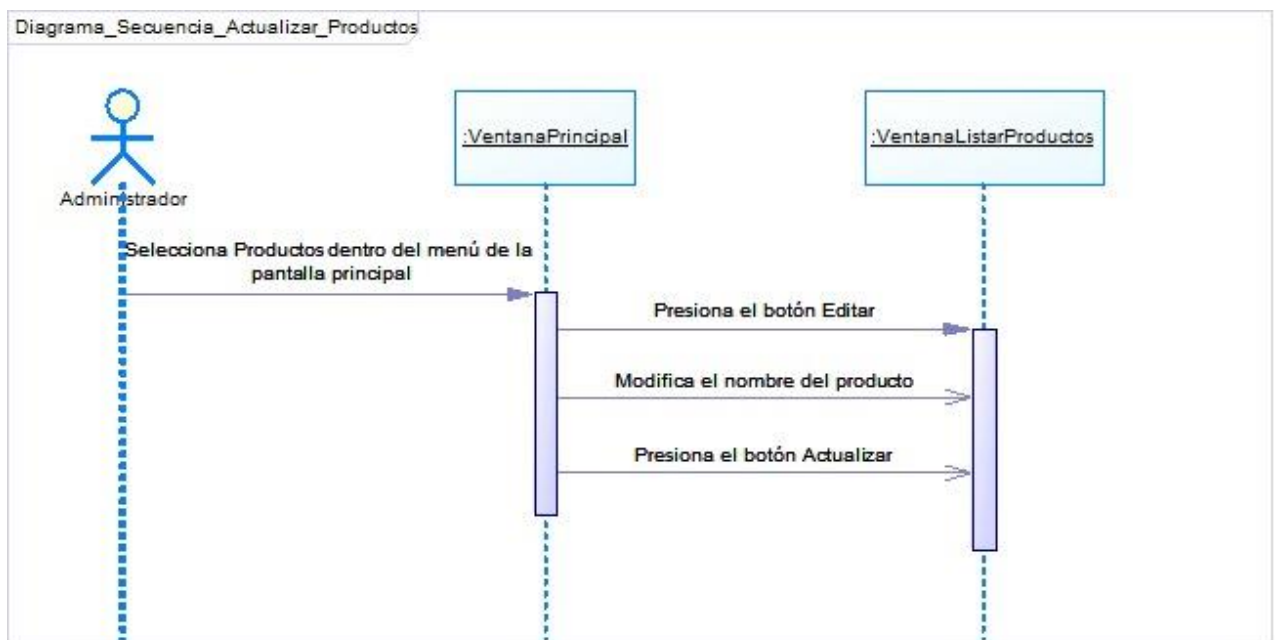
Gráfico:



Caso de Uso: Actualizar Datos

Identificador: RF1.2

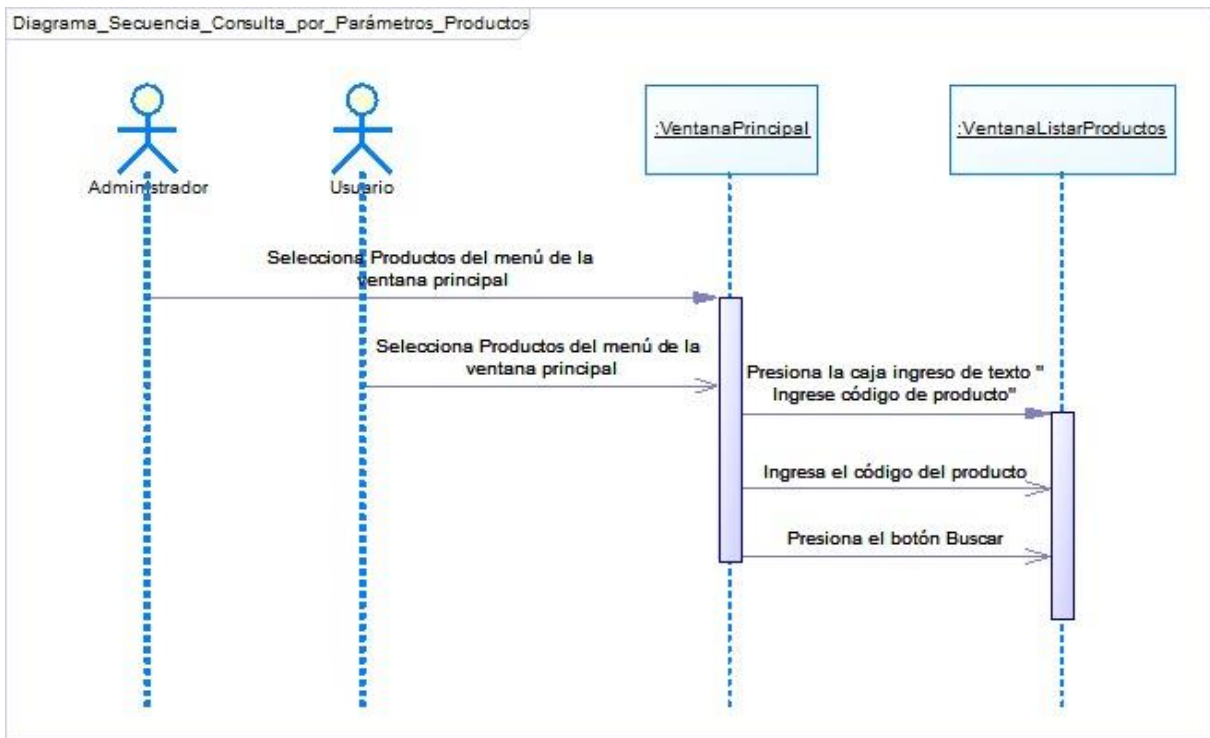
Gráfico:



Caso de Uso: Consulta por Parámetros

Identificador: RF1.3.1

Gráfico:

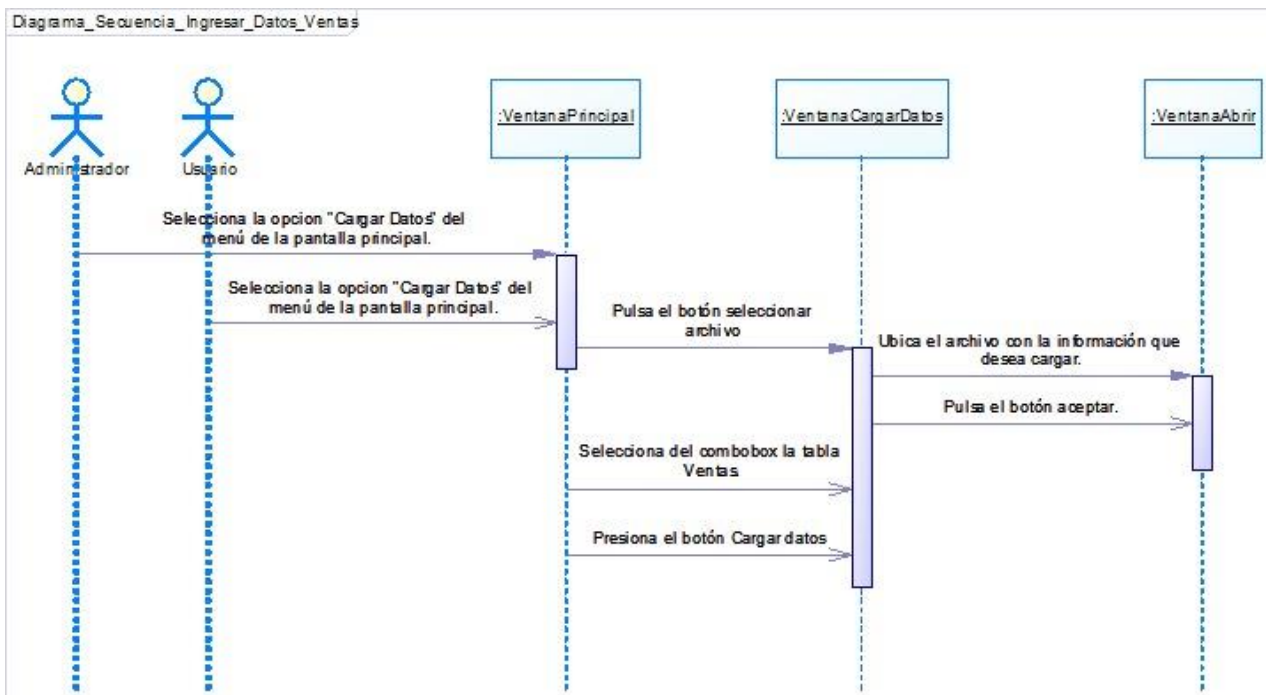


- RF2 Administrar Ventas

Caso de Uso: Ingresar Datos

Identificador: RF2.1

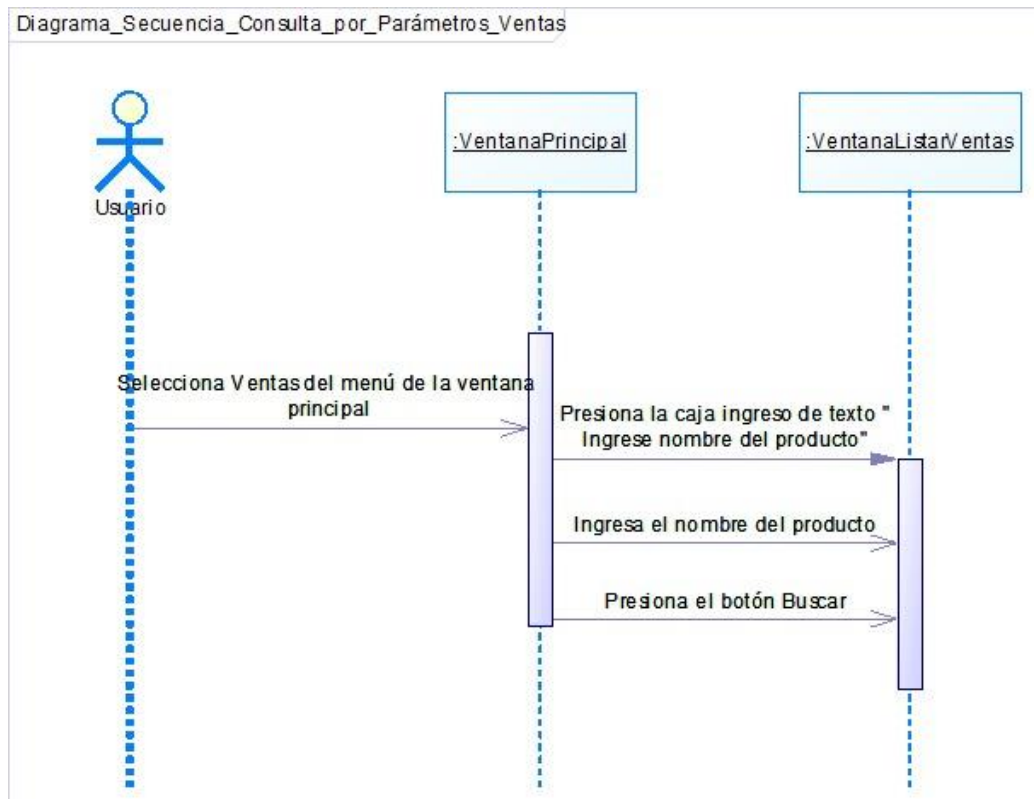
Gráfico:



Caso de Uso: Consulta por Parámetros

Identificador: RF2.2.1

Gráfico:

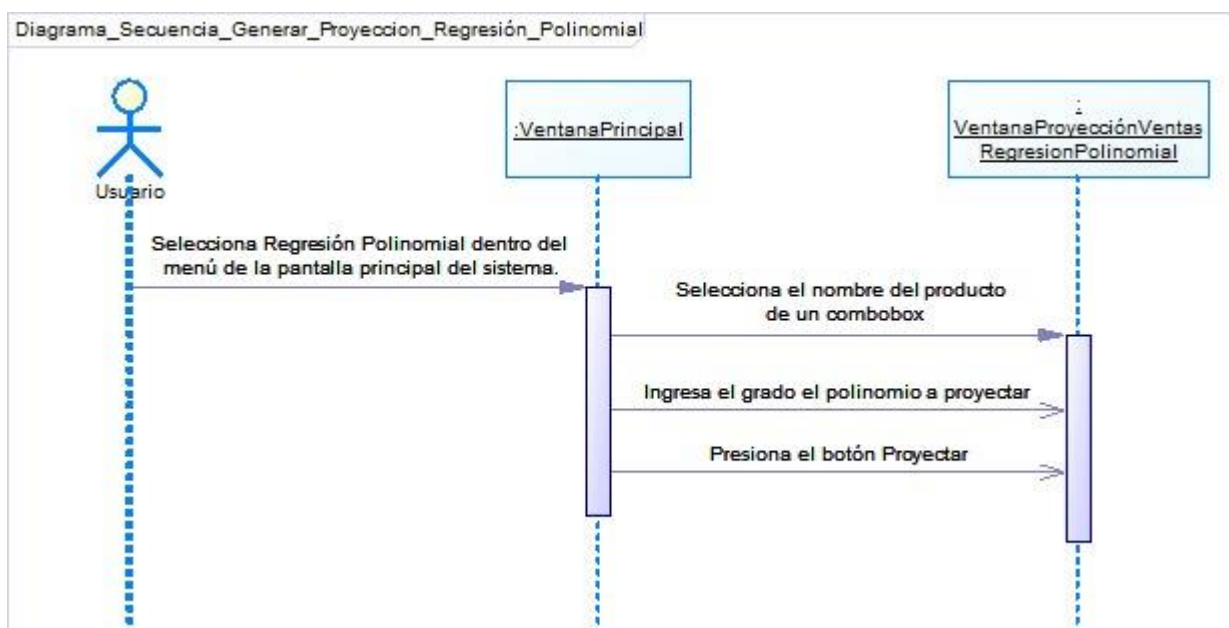


- RF3 Administrar Proyecciones

Caso de Uso: Generar proyección regresión polinomial

Identificador: RF3.1

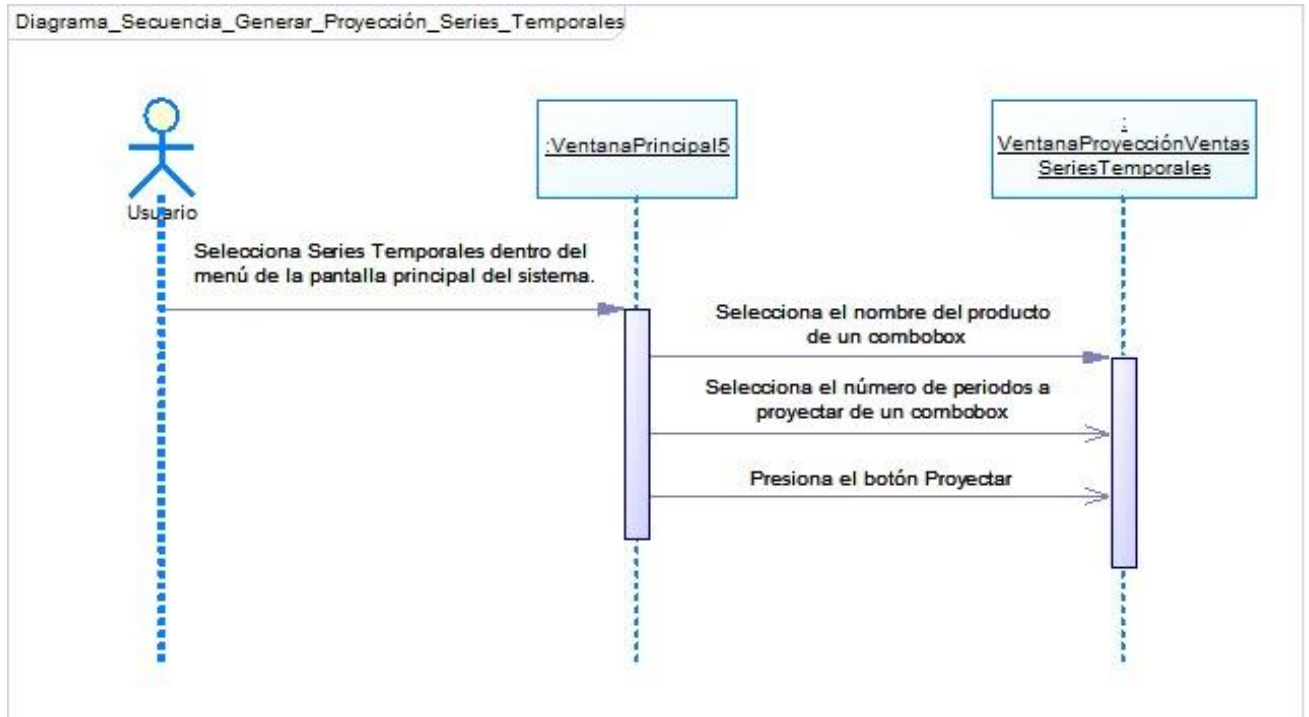
Gráfico:



Caso de Uso: Generar proyección series temporales

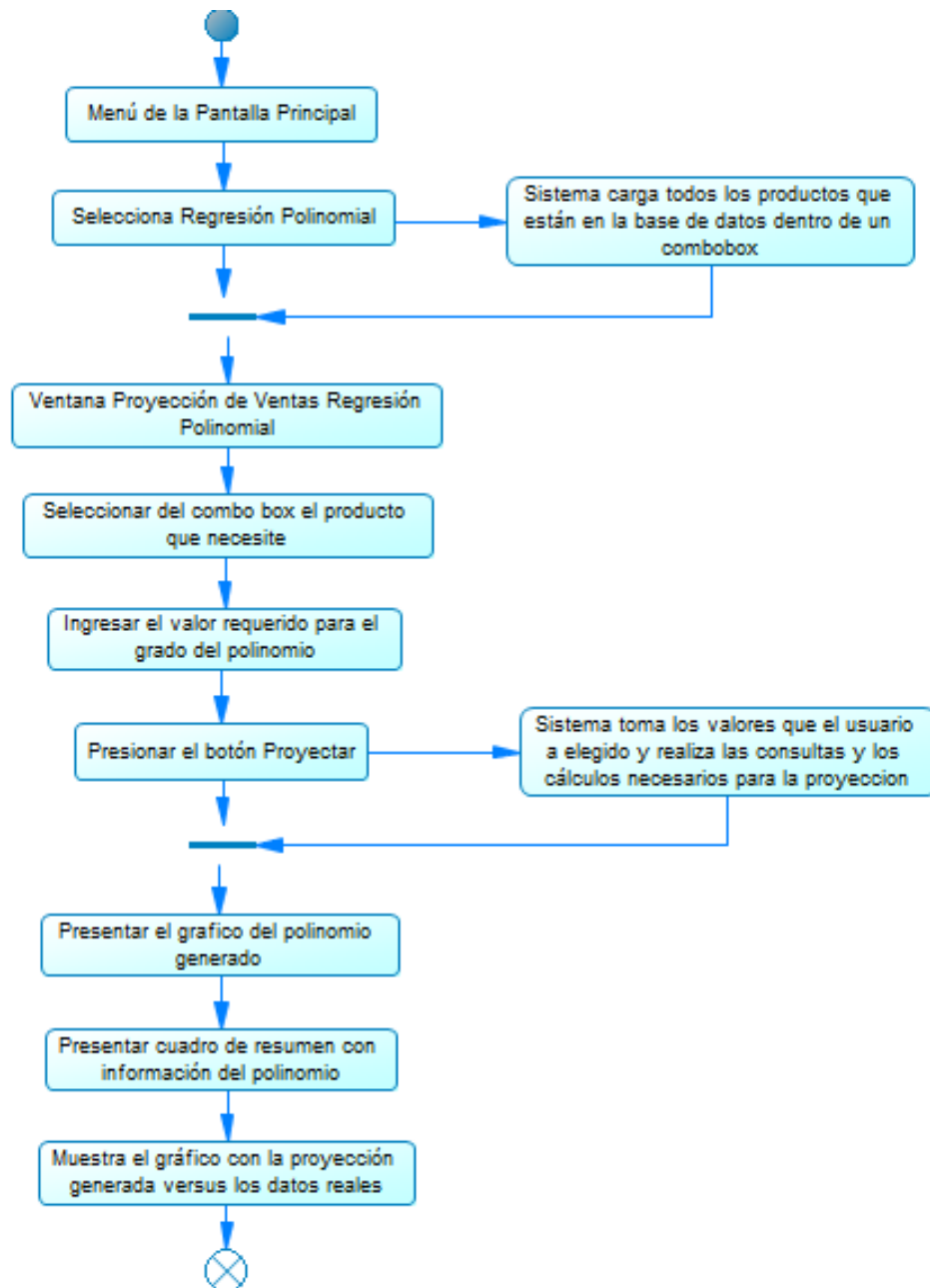
Identificador: RF3.2

Gráfico:

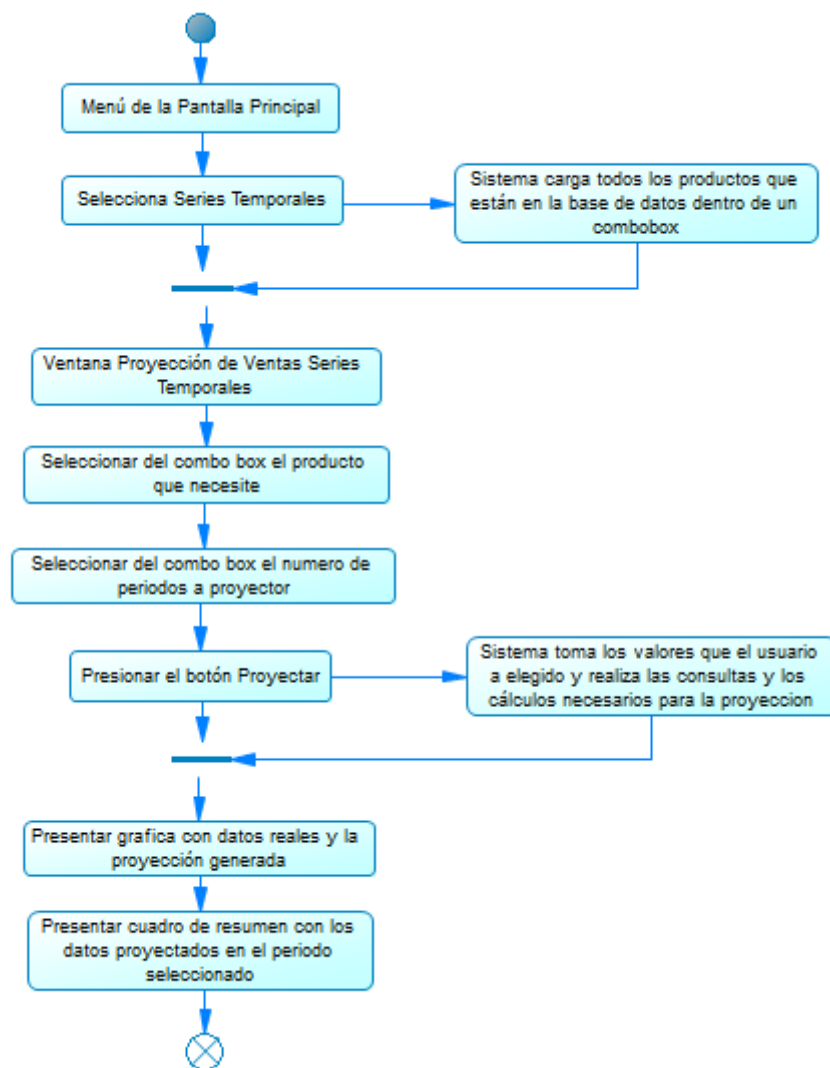


4.2.3. Diagrama de Actividades

- Generar Proyección Regresión Polinomial

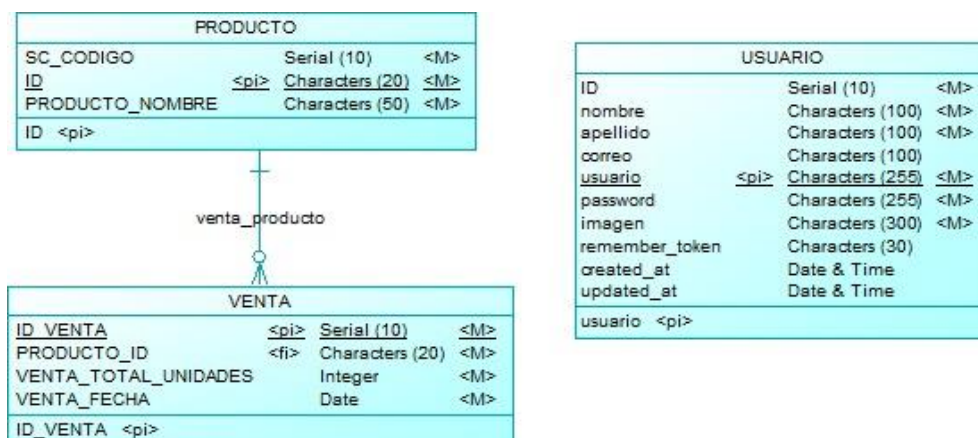


- Generar Proyección Series Temporales

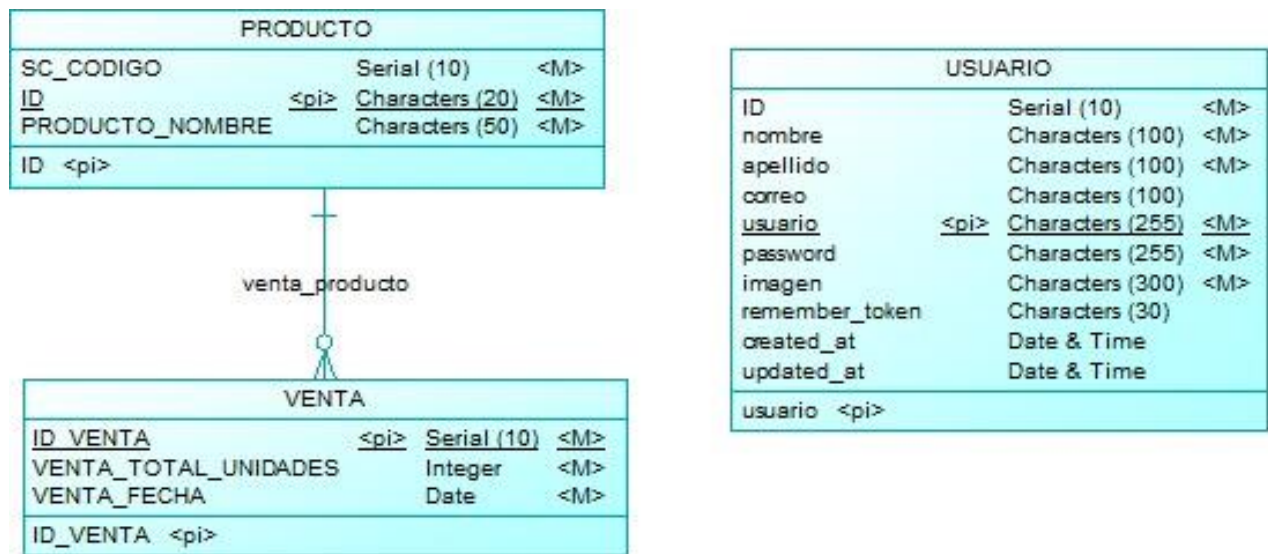


4.2.4. Modelo de datos

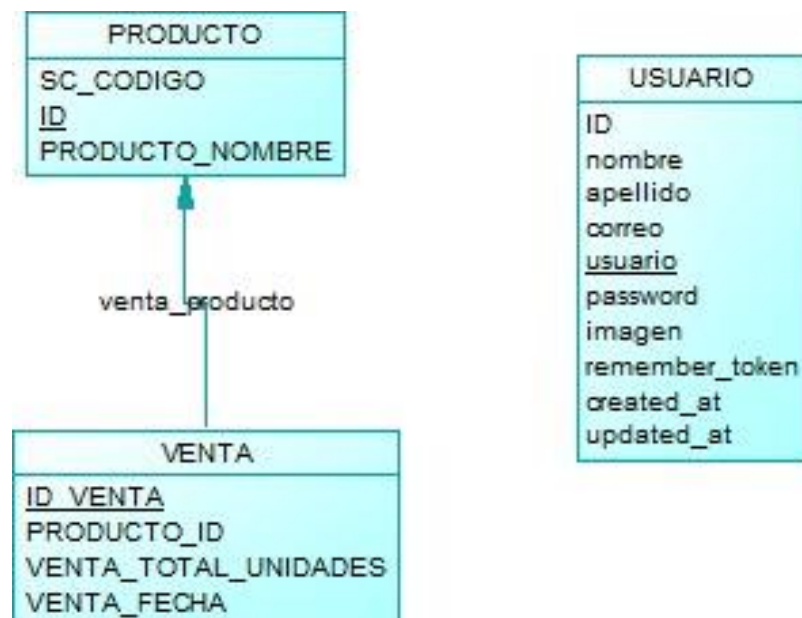
- Modelo Lógico



- Modelo Conceptual



- Modelo Físico



4.3. Implementación

Una vez que se tiene bien definido el diseño del sistema web, se tiene que realizar la implementación, en donde se desarrolla netamente todo el sistema.

Se tiene que tener claro cómo se va a llevar el proceso de desarrollo y ejecución del sistema, es por eso que se debe entender los principales modelos y métodos existentes, todo este proceso coordinado que se debe realizar ayuda a analizar y ver los recursos, los diferentes enfoques y las competencias con las cuales se pueda trabajar con el sistema, así también siguiendo un proceso se puede llegar a estandarizar y realizar un buen trabajo que pueda ser adaptable, confiable y seguro con el fin de que las empresas lo usen sin ningún inconveniente.

Se tiene que saber que al realizar un proceso ordenado en el desarrollo de software se puede lograr incrementar la productividad y hacer un buen manejo de la tecnología aplicando nuevos recursos y haciendo mejores prácticas para la industria, el proceso de desarrollo de software también permite mejorar esfuerzos, dar soporte para mejoras continuas y permite ejecutar un excelente producto de software.

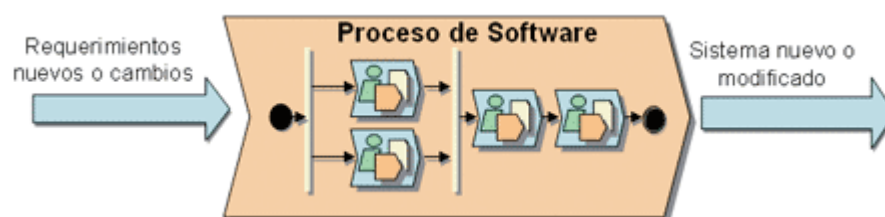


Ilustración 10. Ejemplo de proceso de software.

(2019). Proceso de software [Imagen]. Recuperado de: <https://sg.com.mx/revista/1/procesos-software>

Para llevar a cabo la implementación de este proyecto se decidió usar el método RUP y el modelo en cascada, se tomó la decisión de escoger estos modelos para el proceso de ingeniería ya que RUP usa metodologías que se adaptan a las necesidades de las personas o de la organización y el modelo Cascada porque es un proceso secuencial que permite representar el proyecto en base a la continuidad de sus fases mientras se realiza toda la implementación, es decir este modelo sigue un orden que ayuda a realizar y culminar las fases que necesita el desarrollo.

El proceso de ingeniería RUP permite tener un enfoque diferente para poder cumplir con todas las etapas del desarrollo de software, esta metodología se basa en la realización de los casos de uso y manejo de la arquitectura como tal para el sistema web ya que todo se junta para culminar con éxito el proyecto.

El ciclo de vida del modelo en cascada consiste en la parte de: Análisis, Diseño, Implementación, Pruebas y Documentación en donde en cada etapa se tiene que

definir bien sus conceptos para así lograr cumplir el objetivo que es desarrollar el sistema web totalmente funcional.

Para el análisis se hizo un estudio de todos los requerimientos que el sistema iba a usar, así mismo se determinó como se obtuvo la información viendo y analizando el tipo de información que se tiene y su estructura, se detalló los diferentes modelos de investigación que se usaron los cuales ayudan a que se pueda entender mejor la información y poder estructurar el sistema web, en esta fase se ocuparon 2 meses ya que el análisis que se necesita hacer debe ser cauteloso para que todo se realice correctamente.

El diseño es en donde se detalla toda la parte que se tiene que realizar el filtrado y la limpieza de datos ya que la información siempre tiene una estructura diferente que hace que no se adapte al modelo de base de datos que se realiza para este sistema web, además se tiene que saber cómo se va a realizar la carga de toda la información a la base y como se la va a manipular, aquí también se diagraman todos los casos de uso de las funcionalidades del sistema web, se hacen los diagramas de secuencia, diagramas de actividad, se especifica el modelo de datos y se detallan todos los componentes que tendrá el sistema web para la parte del diseño, esta fase se realizó en dos meses y 15 días debido a que se debe dibujar los diagramas y pensar bien como estará el diseño del sitio web.

La parte principal para que se logre realizar el sistema web es la fase de implementación en donde se tiene que mostrar cómo se desarrolló el sistema, se debe destacar todos los modelos matemáticos que en este proyecto se usó, también se detalla el proceso de ingeniería que se usó y como este se fue implementando en cada parte, aquí se tuvo contrariedades con la implementación de los algoritmos ya que fueron complejos de comprender, ver e investigar para lograr encontrar uno que se adapte a la forma de los datos, aquí se tomó tres meses y quince días para poder culminar toda esta fase.

Para las pruebas se necesitó hacer pruebas de estrés, progresivamente se fueron cargando los datos al sistema lo que hizo que se pudiera observar si es que se volvía lento o si tenía alguna complicación pero esta prueba fue superada con éxito ya que la forma en la cual está estructurado el sistema no permite que esté tenga alguna caída en su ejecución, luego se debió hacer pruebas para ver cómo se comportan los algoritmos y saber cuál es la mejor predicción que se tendrá al momento de ejecutar el sistema esta parte de pruebas se las realizó en quince días un tiempo razonable para determinar el buen funcionamiento del sistema web.

La última fase corresponde a la documentación en donde se detalla todos los aspectos y versiones que se utilizan para ejecutar el sistema ya que todas las herramientas usadas tienen compatibilidad para que se puedan comunicar entre sí, así mismo se realiza aquí el manual de usuario para que todos puedan saber cómo

se ejecuta este sistema web, aquí se necesitó quince días para poder realizar el manual y ver las versiones que tiene este proyecto.

Para ello primero se definió que se va a trabajar con los lenguajes de programación PHP y Python, con el framework LARAVEL, el servidor de páginas web Apache, el gestor de base de Datos de Microsoft SQL Server, el editor de código visual Studio Code, el IDE Spyder y la herramienta de modelado PowerDesigner.

Una vez definidas todas las herramientas para elaborar el proyecto se empezó con la instalación del servidor de páginas web Apache, y de PHP en su versión más actual. Luego se instaló el gestor de dependencias para programar en PHP llamado COMPOSER y posteriormente se creó nuestro proyecto en LARAVEL.

Para el almacenamiento de la información se instaló el gestor de base de datos SQL Server de Microsoft en su versión developer ya que ésta es una edición gratuita y cuenta con todas las características necesarias para el desarrollo del proyecto.

Luego se instaló el editor de código fuente Visual Studio Code el cual fue fundamental para la elaboración de código con los lenguajes PHP y Python, sin embargo, también fue necesario instalar el IDE Spyder ya que está más orientado a la programación científica en lenguaje Python y fue más manejable la elaboración de los scripts que se necesitaba para realizar las proyecciones. Y, por último, para generar los modelos de datos (conceptual, lógico y físico) fue necesario el uso de la herramienta de modelado PowerDesigner.

Una vez instalado todo lo necesario para la creación del proyecto, se empezó a trabajar con el modelo de datos, definiendo atributos y claves necesarias en cada tabla de nuestro modelo, para posteriormente crearlas en el gestor de base de datos.

Luego como se iba a trabajar bajo la arquitectura MVC, se empezó a trabajar con la lógica del sistema web, es decir con la parte del Backend. Primero se configuró la conexión con la base de datos desde nuestro proyecto creado en Laravel, y a continuación, se crearon los modelos y controladores que se utilizarían en el desarrollo del proyecto. Así, para el desarrollo de los algoritmos matemáticos se decidió crear un controlador por cada algoritmo que se implementaría, es decir, se creó un controlador para la Regresión Polinomial y un controlador para Series Temporales con esto se trabajaría de manera más ordenada y controlada.

Sin embargo, antes de desarrollar e implementar los algoritmos matemáticos de cada método de proyección fue necesario comprender cómo estos funcionan, para lo cual se acudió con expertos en el área matemática, en donde se definió que las variables principales que se utilizarían para realizar las proyecciones serían el número de unidades vendidas del producto y la fecha en la que fue vendido.

Teniendo en cuenta esto, primero se estudió el algoritmo de Regresión Polinomial, que consiste en encontrar un polinomio de grado n que se ajuste a los datos que se quiere proyectar. Así, se debe formar un sistema lineal de ecuaciones en forma matricial de la siguiente manera:

$$\mathbf{A} \cdot \mathbf{X} = \mathbf{B}$$

Dado así la expresión matricial del sistema es:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

Donde:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \text{ es la matriz de coeficientes del sistema.}$$

$$\mathbf{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ es la matriz de incógnitas.}$$

$$\mathbf{B} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \text{ es la matriz de términos independientes.}$$

Para resolver este sistema de ecuaciones es necesario poner el sistema en forma matricial y la solución será el producto de la inversa de la matriz de coeficientes (A) por la matriz de términos independientes (B), como se indica a continuación:

$$\mathbf{X} = \mathbf{A}^{-1} \cdot \mathbf{B}'' \text{ (Ejercicios de Matematicas, 2019)}$$

En nuestro proyecto para determinar este sistema lineal matricial en primer lugar, fue necesario agrupar las unidades vendidas del producto que se quiere proyectar en un periodo de tiempo determinado puede ser semanal, quincenal, mensual, trimestral, etc. en este caso debido a la cantidad de datos se agrupó quincenalmente.

Luego se formó una matriz de 2 columnas y n filas (a la cual llamaremos matriz **T**), en la cual la primera columna representaría el número de la quincena y la segunda columna las unidades vendidas en ese periodo de tiempo. Una vez obtenida la matriz **T** se realizó el cálculo de la matriz **A** y la matriz **B** del sistema lineal matricial.

Para cada elemento de la matriz **A** se realizó la sumatoria de las n quincenas elevando cada quincena desde 0 hasta el grado del polinomio buscado, y la dimensión de la matriz dependería del grado del polinomio que se desea proyectar más 1, por

ejemplo, si se desea proyectar un polinomio de grado 2 la dimensión de la matriz sería de 3x3, así para un polinomio de grado 3 sería de 4x4, de grado 4 sería de 5x5.

Para cada elemento de la matriz **B** del sistema lineal matricial se realizó la sumatoria de la multiplicación de la quincena elevada desde 0 hasta el grado del polinomio a proyectar por las unidades vendidas. Y por último el resultado sería la matriz **X**, en la cual sus elementos serán los coeficientes del Polinomio buscado.

A continuación, para entender mejor cómo se realizó este proceso se realizará un ejemplo con los datos reales de la empresa Dysur de los meses de agosto, septiembre y octubre del 2016, en donde se agruparon los datos en 5 quincenas. Este ejemplo se lo realizó para un polinomio de grado 2:

Matriz T

Quincena (X)	Unidades vendidas (Y)
1	275
2	277
3	429
4	333
5	77

Cálculo de la Matriz **A** y Matriz **B**

$$\sum x = 1 + 2 + 3 + 4 + 5 = 15$$

$$\sum y = 275 + 277 + 429 + 333 + 77 = 1391$$

$$\sum xy = 1 \times 275 + 2 \times 277 + 3 \times 429 + 4 \times 333 + 5 \times 77 = 3833$$

$$\sum x^2y = 1^2 \times 275 + 2^2 \times 277 + 3^2 \times 429 + 4^2 \times 333 + 5^2 \times 77 = 12497$$

$$\sum x^2 = 1^2 + 2^2 + 3^2 + 4^2 + 5^2 = 55$$

$$\sum x^3 = 1^3 + 2^3 + 3^3 + 4^3 + 5^3 = 225$$

$$\sum x^4 = 1^4 + 2^4 + 3^4 + 4^4 + 5^4 = 979$$

$$A = \begin{pmatrix} 5 & 15 & 55 \\ 15 & 55 & 225 \\ 55 & 225 & 979 \end{pmatrix}; \vec{b} = \begin{pmatrix} 1391 \\ 3833 \\ 12497 \end{pmatrix}; \vec{X} = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}$$

$$A = \begin{pmatrix} 5 & 15 & 55 \\ 15 & 55 & 225 \\ 55 & 225 & 979 \end{pmatrix}, \text{ inverse: } \begin{pmatrix} \frac{23}{5} & -\frac{33}{10} & \frac{1}{2} \\ -\frac{33}{10} & \frac{187}{70} & -\frac{3}{7} \\ \frac{1}{2} & -\frac{3}{7} & \frac{1}{14} \end{pmatrix}$$

$$\rightarrow A^{-1} = \begin{pmatrix} \frac{23}{5} & -\frac{33}{10} & \frac{1}{2} \\ -\frac{33}{10} & \frac{187}{70} & -\frac{3}{7} \\ \frac{1}{2} & -\frac{3}{7} & \frac{1}{14} \end{pmatrix}$$

$$\vec{X} = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = A^{-1} \vec{b} = \begin{pmatrix} \frac{23}{5} & -\frac{33}{10} & \frac{1}{2} \\ -\frac{33}{10} & \frac{187}{70} & -\frac{3}{7} \\ \frac{1}{2} & -\frac{3}{7} & \frac{1}{14} \end{pmatrix} \begin{pmatrix} 1391 \\ 3833 \\ 12497 \end{pmatrix} = \begin{pmatrix} -\frac{9}{5} \\ \frac{2054}{7} \\ -\frac{382}{7} \end{pmatrix}$$

Luego

$$P2(x) = -\frac{9}{5} + \frac{2054}{7}x - \frac{382}{7}x^2$$

Y así una vez entendido el funcionamiento del algoritmo de Regresión Polinomial se pudo implementar en el lenguaje PHP. Para lo cual en el controlador se implementó una función la cual recibe como parámetros el código del producto y el grado del polinomio a ser proyectado, en el mismo se realiza la consulta a la base de datos del producto y las unidades vendidas respectivas al mismo. Luego se generó la matriz **T** con la agrupación de los datos en un periodo quincenal, para posteriormente construir la matriz **A** y la matriz **B**.

Sin embargo, cuando se realizó este algoritmo se tuvo inconvenientes debido a que para obtener el polinomio es necesario calcular la inversa de las matrices involucradas, es por eso que se vio la necesidad de trabajar en conjunto con el lenguaje Python generando un script que se ejecutaría realizando una llamada desde la función del controlador, este script ayudaría al cálculo de la inversa y multiplicación

de matrices, con esto se simplificaría de gran manera el desarrollo e implementación de este algoritmo.

Finalmente, luego de haber resuelto el sistema de ecuaciones matricial se obtuvo los coeficientes del polinomio buscado, con los cuales se pudo obtener una gráfica representativa del mismo y a su vez se pudo calcular la proyección de las ventas en base al polinomio generado. Y para saber que polinomio es el que más se ajusta a los datos fue necesario calcular el error que existía entre los datos proyectados y los datos reales. Para lo cual se definió la siguiente fórmula:

$$\sum_{i=0}^n \frac{|UnidadesVendidasReales_i - UnidadesProyectadas_i|}{|UnidadesProyectadas_i|}$$

Donde:

n= es el número de filas de la matriz **T**.

Y así, mediante este valor calculado se puede determinar que polinomio es el óptimo para estos datos.

Sin embargo, luego de haber implementado el método de Regresión Polinomial se notó que las proyecciones que generaba el polinomio obtenido no eran de lo más acertadas, por lo que se decidió investigar otro método de pronóstico el cual fue el método de Series Temporales.

Este método consiste básicamente en agrupar los datos ordenados cronológicamente en el tiempo, ya que a partir de esto se puede pronosticar en base a hechos ocurridos en periodos de tiempo pasados, es decir que las ventas de un producto en el mes de agosto pueden estar relacionadas con las ventas del mes de agosto del año pasado. Para realizar el pronóstico es necesario emplear métodos estadísticos especializados, en este proyecto se empleó el modelo SARIMA.

El modelo SARIMA es utilizado en series temporales no estacionarias, esto significa que los datos no oscilan alrededor de la misma media, varianza y covarianza. La importancia de este método es que se puede identificar tendencia y estacionalidad en los datos y además se compone de otros modelos estadísticos como el modelo ARIMA, lo cual lo hace más robusto para el pronóstico de datos.

El modelo ARIMA usa un paso de diferenciación I (d) para eliminar la no estacionariedad lo cual permite que el método admita series temporales con tendencia. Sin embargo, este modelo no identifica la estacionalidad y a partir de este hecho se tomó la decisión de utilizar el modelo SARIMA, que tiene una correlación estacional y puede identificar la estacionalidad de las series de tiempo. Cabe recalcar

que se debe realizar pronósticos para tiempos futuros no muy lejanos, ya que a fechas más lejanas existirá menor precisión en los pronósticos.

SARIMA al componerse de ARIMA utiliza los parámetros (p, d, q) no estacionales de éste para calcular el mejor modelo para los datos que se desea proyectar. Por lo que, la notación para este modelo es la siguiente:

$$\text{“SARIMA}(p, d, q).(P, D, Q)m$$

Donde:

(p, d, q) : *explican la estacionalidad, tendencia y ruido en los datos respectivamente y p va a tomar el valor de 0 ya que proviene del modelo ARIMA y este modelo se utiliza para datos no estacionales.*

(P, D, Q) : *explican la estacionalidad, tendencia y ruido en los datos respectivamente.*

m : *es el periodo en el que se presenta la estacionalidad de los datos.”*

En un principio se trató de implementar el modelo SARIMA de forma parametrizable, es decir de calcular los mejores valores para **(p, d, q)** y **(P, D, Q, m)** y esto no era lo óptimo al momento de realizar las proyecciones. Por lo que se implementó este modelo de una forma que calculase estos valores automáticamente. La implementación se la realizó con Python, ya que permite utilizar librerías y módulos muy potentes para facilitar el desarrollo de este algoritmo.

Para la implementación fue necesario la instalación de las librerías numpy (trabajo con matrices y vectores), pandas (manipulación y análisis de datos), matplotlib.pyplot (generación de gráficos) y pmdarima (envuelve los modelos de estadísticos). Luego se desarrolló el script que se usaría para realizar las proyecciones, el cual tiene dos parámetros de entrada, los cuales son el código del producto y el número de periodos a ser proyectados. Desde PHP se realiza la llamada a este script y se envían los parámetros requeridos. Una vez que el script recibe los parámetros filtra los datos de acuerdo con el código de producto, los agrupa en un periodo quincenal y los aloja en un DataFrame (**Frame1**) el cual no es más que una Tabla similar a la de SQL.

Luego para aplicar el modelo SARIMA fue necesario separar **Frame1** en dos grupos (**FrameTrain** y **FrameTest**) con la intención de obtener un conjunto de datos para entrenar al modelo y otro conjunto para comparar la información pronosticada, ya que con esto se puede determinar qué tan confiable es el modelo de pronóstico.

Ya con estos dos Frames fue posible la implementación del modelo SARIMA, esto se lo realizó utilizando la función `auto_arima()` de la librería pmdarima, en la cual se debe especificar los siguientes parámetros: **Datos de entrenamiento**, los valores iniciales de **p, d, q, P, D, Q** y **m** .

Esta función va a calcular el modelo que mejor se ajuste (**SarimaModel**) a los datos proporcionados generando diferentes combinaciones de los parámetros de estacionalidad, tendencia y ruido. Al haber obtenido **SarimaModel** ya se pudo proyectar la información utilizando la función predict() propia de este método, la que recibe como parámetros el número de periodos a proyectar y un valor alpha, el cual no es más que un porcentaje de error que aceptará o rechazará los pronósticos generados.

Finalmente, cuando se obtuvo los valores pronosticados no fue más que hacer uso de la librería matplotlib.pyplot para generar los gráficos en los cuales se pueda observar los datos proyectados y los datos reales.

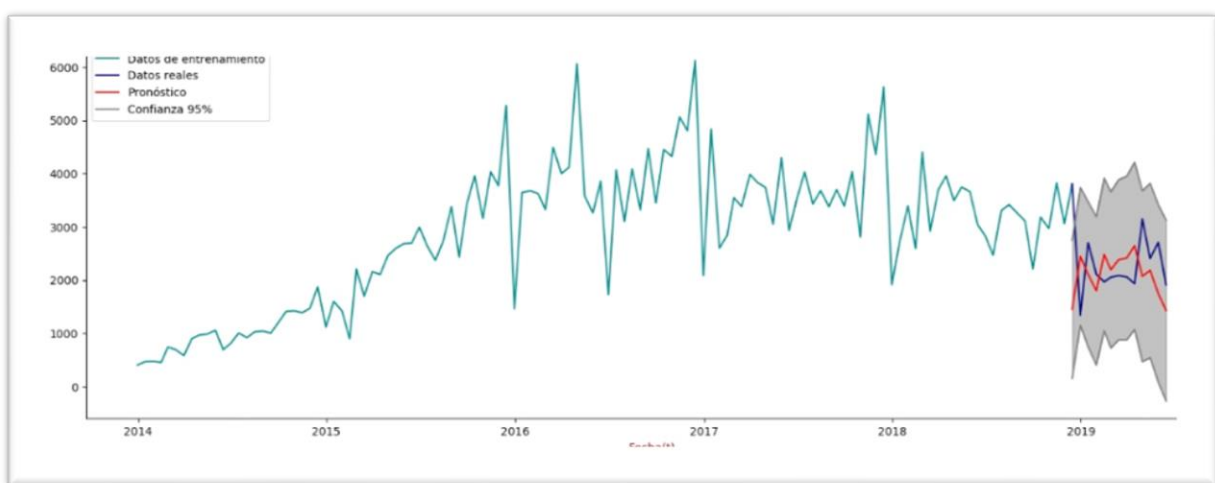


Ilustración 11. Resultado de la proyección con series temporales.

(2019). Resultado de la proyección [Captura de imagen]

4.4. Pruebas

Se realizó pruebas de estrés ya que es un sistema web que funciona en base a una gran cantidad de datos. Primero se comenzó con pocos datos corriendo el sistema con datos históricos de 3 años dando como resultado que el sistema se ejecutó con total normalidad en un tiempo de 2.5 segundos, dando así la pauta para continuar con la carga de los demás datos para su ejecución, la carga de datos se realizó de año en año, la última prueba que se realizó fue con el histórico de datos de cinco años dando un resultado positivo después de su ejecución, la cual tuvo un tiempo de respuesta de 3.5 segundos ya que el sistema se toma sus minutos en procesar la información pero dando buenos resultados lo que quiere decir que el programa trabaja bien con toda la información cargada, y en el caso de que llegara a colapsar se realizarán las correcciones necesarias.

Así mismo se hicieron pruebas de carga en donde se tuvo que verificar si los datos que se iban subiendo en la base se ingresaban de forma correcta ya que se tenía que verificar el formato de fechas y el tipo de escritura de los datos que se tenía en la información, así se hicieron las debidas correcciones en el formato para que la carga sea correcta en la base.

También se probarán diferentes escenarios en lo que se refiere a conexión de base de datos con el lenguaje PHP, ya que es posible que los tiempos de respuesta difieran uno del otro considerablemente. Luego dependiendo del resultado se determinará con que base de datos el sistema responde óptimamente y se aconsejará utilizar este gestor para un mayor rendimiento del sistema.

Por último, se realizará pruebas de predicción, esto significa que se probará el sistema para verificar si los algoritmos empleados están proyectando datos coherentes; dado un ejemplo, se simulará una predicción de un año ya conocido, es decir un año que ya existan registros reales y dependiendo del resultado se verificará si está dentro de un rango de error aceptable.

4.5. Documentación

4.5.1. Requisitos

Se necesita tener en cuenta cierta documentación o aspectos para poder instalar y ejecutar los frameworks, editores de texto y lenguajes que se usan en el proyecto, a continuación, se detalla lo que cada uno de estos necesitan para tener un correcto funcionamiento:

Laravel: Si se lo va a realizar localmente se debe tener instalado PHP >= 7.2.0, en el proyecto se instaló el intérprete de lenguajes de SCRIPT XAMPP el que ayuda para poder ejecutar PHP, es de origen público y actúa como un servidor web libre, fácil de usar y capaz de interpretar páginas dinámicas.

Python: Es un lenguaje de programación muy potente que se usa para crear sitios web, juegos, softwares de todo tipo, gráficos y más, originado en 1980 y el objetivo es ser más entendible para los humanos, se tiene que ver cuál es la versión más estable en este caso se usa Python 3.4, se tiene que tener en cuenta de instalar todos los módulos como "Add Python 3.6 to PATH" o "Add Python to your environment variables" que ayudan a la ejecución del programa.

PHP: Ya que es un lenguaje diseñado para el preprocesado de texto plano su interfaz se ejecuta mediante phpMyAdmin en el navegador así se necesitará un servidor web en este caso se usa XAMPP, se necesitará usar PHP >= 7.1.3 y las diferentes librerías que apoyan al programa, también es compatible con SQL SERVER para la base de datos con la cual se trabajó.

Visual Studio Code: es un editor de código fuente desarrollado por Microsoft, es una herramienta muy ligera y de fácil ejecución tiene un peso menor a 100mb lo cual hace que no ocupe mucho espacio en disco, basta con tener un procesador $\geq 1.6\text{GHz}$ y memoria RAM $\geq 1\text{GB}$, se lo puede utilizar en diferentes sistemas operativos entre ellos Windows 7 o superior como es el caso de este proyecto.

SQL Server: para el proyecto como motor de base de datos se usó SLQ Server Management Studio el cual necesita de la instalación de .NET ≥ 4.6 el cual se instalará juntamente con el instalador del SQL y el SQL Server Native Client, en la parte de hardware se requiere un mínimo de 6GB de disco, 4GB de memoria y un procesador $\geq 2.0\text{GHz}$ para que se pueda garantizar el correcto funcionamiento de la herramienta.

Hosting: Se debe tener en cuenta ciertos aspectos que den fiabilidad al momento de elegir un Host como: lugar del hosting, tiempo de carga, servicio al cliente, copias de seguridad y mantenimiento, en la parte de hardware se debe tener en cuenta el espacio en disco, transferencia de datos, migración y base de datos que dependiendo del costo se puede obtener mejores características.

CAPITULO 5 CONCLUSIONES Y RECOMENDACIONES

5.1. Conclusiones

- Enfocarse en empresas que venden productos masivamente facilitó el estudio realizado en el presente proyecto.
- El uso de los algoritmos matemáticos para la proyección de ventas implica el procesamiento de grandes volúmenes de información para que sus resultados sean más aproximados a la realidad.
- De esta manera se aplicó la metodología en cascada por ser un método clásico de desarrollo de software el cual se orienta mejor a la ejecución de proyectos de corto plazo que estén bien definidos y detallados, dando lugar al desarrollo de software ordenado siguiendo sus etapas definidas.
- Debido a la necesidad de un enfoque técnico para la asignación de tareas y responsabilidades en el desarrollo del software se aplicó la metodología RUP.
- El proceso de ingeniería RUP permitió prever los cambios que se pueden tener acorde a los requerimientos durante las etapas del desarrollo para de esta forma optimizar y mejorar el sistema.
- Entre los modelos adoptados en este proyecto, el que más sobresale y mejor se adaptó a los datos que se usó fue el SARIMA, lo cual se pudo notar en los resultados obtenidos de forma gráfica al momento de probar el sistema.
- De acuerdo con el contexto que se analizó en el modelamiento puede servir una diversidad de variables, sin embargo, en nuestro caso las variables más significativas y con las cuales se realizó casi todo el proceso fueron dos, la variable fecha y la variable precio.
- Para realizar los scripts de Python fue necesario instalar el entorno de desarrollo multiplataforma Spyder ya que sin la ayuda de este IDE no se hubiera podido desarrollar los scripts porque visual Studio Code no es potente para visualizar los scripts en Python.
- En los resultados gráficos que se obtuvieron se observó que mientras mayor sean los periodos que se desee proyectar existirá menor precisión en las estimaciones, debido a que los datos a través del tiempo tienden a comportarse de diferente manera y esto ocasionaría que el nivel de confianza de los resultados disminuya.

- Después de haber realizado todo el proceso de desarrollo del sistema y de haber culminado las pruebas planteadas se puede dar solución a la hipótesis planteada ya que a través de modelos matemáticos conjunto con la inteligencia artificial se puede mejorar el análisis de datos.

5.2. Recomendaciones

- Las proyecciones que se realicen no deben ser en periodos muy alejados ya que el nivel de confianza de los resultados disminuiría.
- Para usar las librerías que se requieren en los scripts de Python es importante instalarlos en el directorio del proyecto, ya que si se las instala en el directorio de Python al ejecutar los scripts el programa no reconocerá las importaciones de las librerías
- Se tiene que tomar en cuenta el lenguaje, el framework y el motor de base de datos con el cual se pueda trabajar todo.
- Para el caso fue necesario tomar una empresa estándar de venta de productos tangibles para que proporcione los datos con un histórico de al menos tres años.
- Para la conexión entre Laravel y SQL Server es necesario descargar las bibliotecas de enlaces dinámicos (.dll) `php_pdo_sqlsrv_72_ts` y `php_sqlsrv_72_ts` y agregarlas a las extensiones de PHP
- También es necesario agregar las bibliotecas de enlaces dinámicos (.dll) descargadas al archivo `php.ini` para cargarlos en tiempo de ejecución y evitar inconvenientes en la conexión entre el motor de base de datos SQL Server y el framework Laravel.
- Los protocolos TCP/IP del SQL Server deben estar habilitados para poder realizar la conexión. Entre la base de datos y laravel
- Antes de instalar los frameworks e IDE's necesarios, se debe verificar que las versiones sean compatibles para no tener problemas al momento de incorporarlos todos.
- Para todo desarrollo de aplicaciones es recomendable adoptar una metodología de desarrollo ya que permite tener organización y seguir un procedimiento.

REFERENCIAS BIBLIOGRÁFICAS

- (30 de octubre de 2017). Obtenido de Recuperado de <https://openclassrooms.com/en/courses/4309151-gestiona-tu-proyecto-de-desarrollo/4538221-en-que-consiste-el-modelo-en-cascada>
- (octubre de 2019). Obtenido de https://www.google.com.ec/search?hl=es-419&biw=1366&bih=625&tbm=isch&sxsrf=ACYBGNQOpCOdljyTrFbss8R_NtfQcsPmTQ%3A1570067218928&sa=1&ei=EIOVXZyaOMfn5gLVup3wAg&q=derivadas+parciales+grafica&oq=derivadas+parciales+gr&gs_l=img.1.0.0j0i8i30.1398.2682..3709...0
- Bagnato, J. I. (26 de Febrero de 2019). *Pronóstico de series temporales en Python*. Obtenido de Aprende Machine Learning: <https://www.aprendemachinelearning.com/pronostico-de-series-temporales-con-redes-neuronales-en-python/>
- Ejercicios de Matemáticas*. (2019). Obtenido de <https://www.ematematicas.net/matrices.php?tipo=7>
- gestiopolis*. (2019). Obtenido de <https://www.gestiopolis.com/metodos-y-tecnicas-de-investigacion-cientifica/>
- Ingeniería de Software, U. e. (2019). *Ecured*. Obtenido de https://www.ecured.cu/Modelo_en_cascada
- Ley Organica de Telecomunicaciones. Art 94 Capitulo I*. (s.f.). Obtenido de <https://www.gobiernoelectronico.gob.ec/wp-content/uploads/2018/10/Ley-Org%C3%A1nica-de-Telecomunicaciones.pdf>
- López, B. S. (2016). *INGENIERIAINDUSTRIALONLINE.COM*. Obtenido de <https://www.ingenieriaindustrialonline.com/herramientas-para-el-ingeniero-industrial/pron%C3%B3stico-de-ventas/>
- Maelec. (25 de Marzo de 2013). *2011–2019 ClubEnsayos.com*. Obtenido de <https://www.clubensayos.com/Espa%C3%B1ol/AJUSTES-DE-CURVAS/631050.html>
- Market, L. y. (24 de 05 de 2017). *Comunicae.es*. Obtenido de <https://www.comunicae.es/nota/en-mas-del-80-de-las-empresas-los-procesos-de-1187101/>
- Norma Técnica para la Aplicación de la Ley Orgánica de*. (2019). Obtenido de <https://www.uta.edu.ec/v3.2/uta/reglamentosexternos/normaLOTAIP.pdf>
- Prabhakaran, S. (febrero de 2019). *machine learning plus*. Obtenido de https://www.machinelearningplus.com/time-series/arima-model-time-series-forecasting-python/amp/?utm_source=mybridge&utm_medium=blog&utm_campaign=read_more
- Raffino, M. E. (18 de Enero de 2019). *Método inductivo*. Obtenido de <https://concepto.de/metodo-inductivo/>
- Romero, J. (2015). *Que es la Metodología orientada a objetos*. Obtenido de https://www.academia.edu/15943563/Que_es_la_Metodolog%C3%AD_a_orientada_a_objetos_-_Introducci%C3%B3n_a_UML
- software recopilation*. (2019). Obtenido de <https://softwarerecopilation.wordpress.com/modelo-rup/>

Telégrafo, E. (17 de junio de 2013). *Ley Orgánica de Comunicación*. Obtenido de <https://www.eltelegrafo.com.ec/images/eltelegrafo/banners/2013/17-6-13-Ley-de-comunicacion.pdf>

Wacker, J. G., & Lummus, R. (2002). *Sales Forecasting for Strategic resource planning*. Arizona.

APÉNDICE

MANUAL DE USUARIO

Contenido del sistema web

- *Introducción*
- *Proceso*
 - o *Capturas de pantalla del funcionamiento del sistema web.*

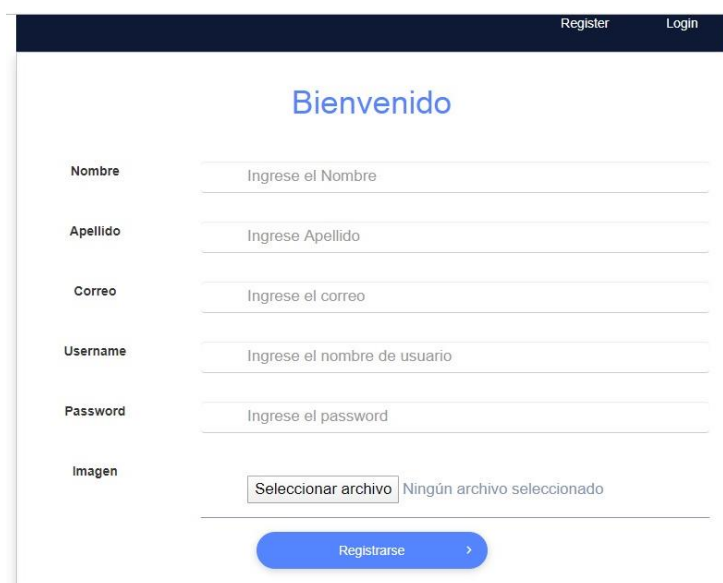
❖ Introducción

Este sistema de predicción de ventas se basa en la implementación de varios algoritmos matemáticos que permiten realizar un proceso bastante complejo mediante el uso de polinomios, matrices, series temporales y el modelo SARIMA lo que se llega a obtener resultados de forma gráfica para así tomar decisiones en cuanto a ventas dentro de la empresa.

Este sistema web se desarrolla en el framework Laravel con los lenguajes de PHP y Python, así mismo se realiza la interfaz gráfica con HTML lo que permite tener implementado un sistema muy amigable al usuario.

❖ Proceso

1. Para poder acceder al sistema y utilizar este servicio se tiene que realizar el **registro de usuario** en donde se debe completar un formulario con los datos básicos para que así el sistema pueda tener el registro de cada usuario, al finalizar el llenado de datos se tiene que presionar el botón de “REGISTRARSE”

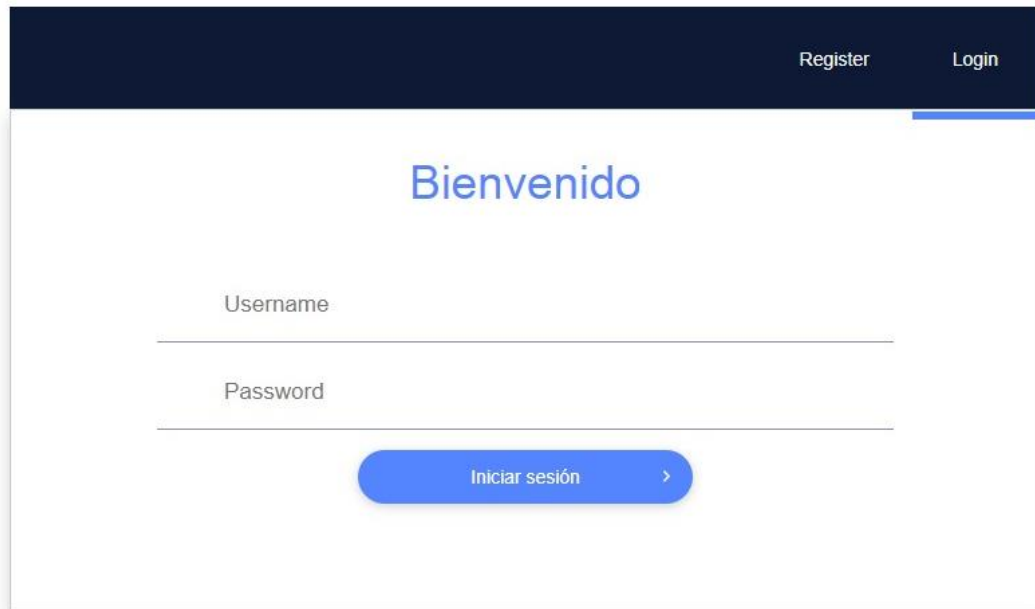


La imagen muestra una interfaz web de registro de usuario. En la parte superior, hay una barra de navegación con los enlaces "Register" y "Login". El título principal de la sección es "Bienvenido". El formulario de registro contiene los siguientes campos:

- Nombre:** Un campo de texto con el placeholder "Ingrese el Nombre".
- Apellido:** Un campo de texto con el placeholder "Ingrese Apellido".
- Correo:** Un campo de texto con el placeholder "Ingrese el correo".
- Username:** Un campo de texto con el placeholder "Ingrese el nombre de usuario".
- Password:** Un campo de texto con el placeholder "Ingrese el password".
- Imagen:** Un campo de carga de archivos con el botón "Seleccionar archivo" y el texto "Ningún archivo seleccionado".

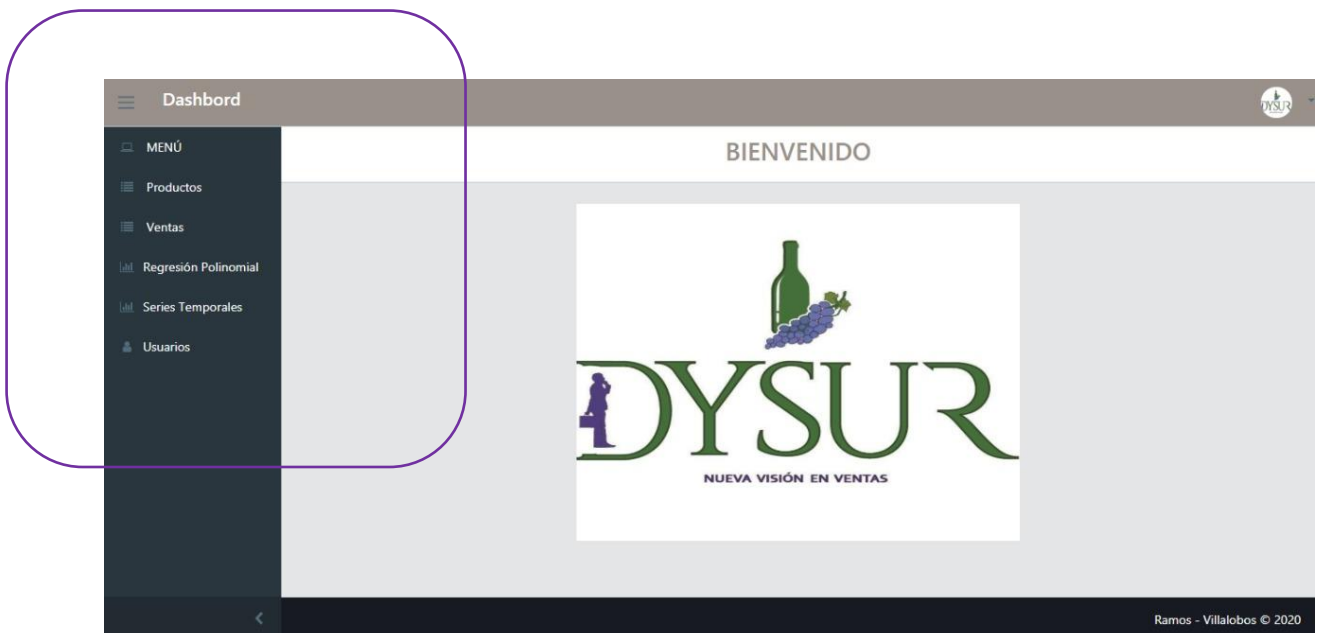
En la parte inferior del formulario, hay un botón azul con el texto "Registrarse" y una flecha de dirección a la derecha.

2. Una vez completo el registro para ingresar al sistema web se tiene que realizar el **Login** de usuario, así de esta forma se podrá acceder al perfil que cada usuario tiene registrado en el sistema web y manipular únicamente los datos asociados a ese perfil.



The image shows a login form with a dark blue header containing 'Register' and 'Login' links. The main content area is white and features the word 'Bienvenido' in blue. Below this, there are two input fields: 'Username' and 'Password'. At the bottom, there is a blue button labeled 'Iniciar sesión' with a right-pointing arrow.

3. Cuando ya se ingresa al sistema con el usuario respectivo se presenta la **Página de Inicio** en donde se puede observar el menú de opciones que tiene el sistema web como se puede observar en el gráfico siguiente.



4. Dentro de las opciones del dashboard, se tiene el botón de **Productos** el cual una vez seleccionado muestra la vista de los productos aquí se presenta la tabla con todos los productos que se encuentren en la base de datos y adicional se puede realizar la búsqueda por el código del producto.

Dashboard

MENÚ

- Productos
- Ventas
- Regresión Polinomial
- Series Temporales
- Usuarios

SISTEMA DE PROYECCIÓN DE VENTAS

Listado de Productos

Ingrese código del producto

ID PRODUCTO	CODIGO PRODUCTO	NOMBRE PRODUCTO	COSTO UNITARIO
1	134	RON PON PON	22.6
2	111	NORTEÑO ANIZADO	36.299999

5. Otra opción que se puede escoger es la parte de **Ventas** en donde igual se puede visualizar el listado de las ventas con sus respectivos registros y permite realizar la búsqueda mediante el ID de Venta.

Dashboard

MENÚ

- Productos
- Ventas
- Regresión Polinomial
- Series Temporales
- Usuarios

SISTEMA DE PROYECCIÓN DE VENTAS

Listado de Ventas

Ingrese nombre de producto...

CODIGO PRODUCTO	NOMBRE PRODUCTO	ID VENTA	TOTAL UNIDADES	TOTAL (\$)	FECHA
134	RON PON PON	1	24	542.45001	2016-01-02 00:00:00.000
134	RON PON PON	2	6	135.61	2016-01-02 00:00:00.000
134	RON PON PON	3	6	135.61	2016-01-02 00:00:00.000
134	RON PON PON	4	36	813.66998	2016-01-02 00:00:00.000
134	RON PON PON	5	6	135.61	2016-01-06 00:00:00.000
134	RON PON PON	6	3	67.809998	2016-01-06 00:00:00.000
134	RON PON PON	7	36	813.66998	2016-01-08 00:00:00.000
134	RON PON PON	8	6	143.16	2016-01-12 00:00:00.000
134	RON PON PON	9	12	286.31	2016-01-12 00:00:00.000
134	RON PON PON	10	4	95.440002	2016-01-12 00:00:00.000

Cabe mencionar que estas opciones solo son informativas o de visualización de los datos ya cargados en la base.

6. En la opción de **Regresión Polinomial** ya se puede realizar la predicción de las ventas mediante los polinomios de segundo, tercer, cuarto y quinto grado para esto se tiene que seleccionar el producto del cual se va a hacer la predicción y también la estimación del polinomio es decir del grado del polinomio con el que se va a trabajar, una vez colocados esos datos se presiona el botón graficar y muestra la gráfica resultante.

Dashboard

MENÚ

- Productos
- Ventas
- Regresión Polinomial
- Series Temporales
- Usuarios

PROYECCIÓN DE VENTAS REGRESIÓN POLINOMIAL

Seleccione producto

Elija un producto...

Estimación de Polinomio

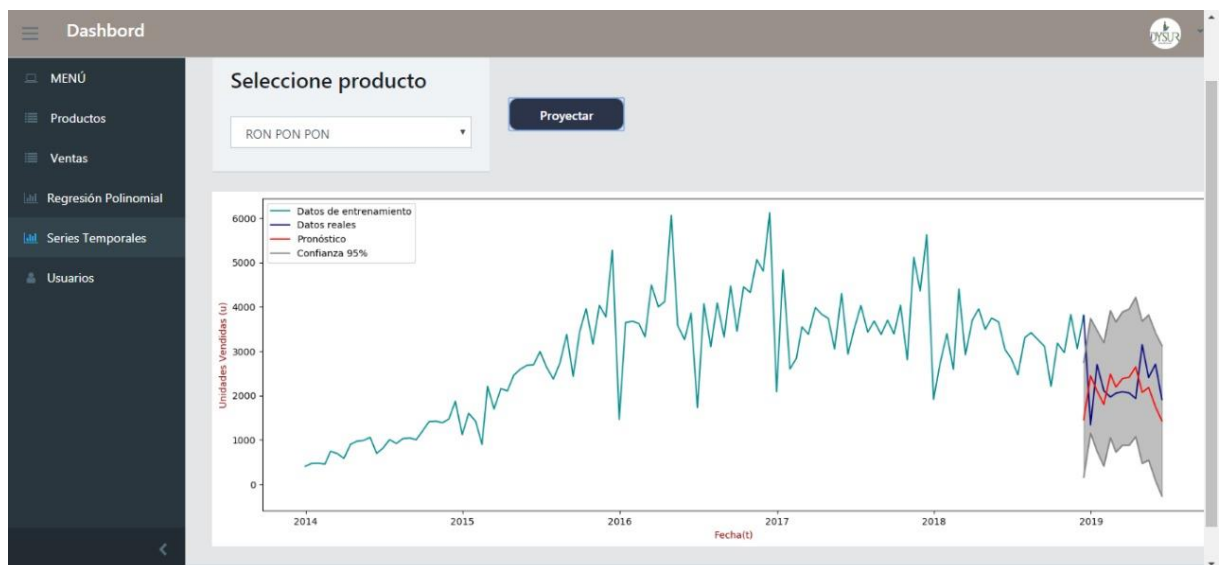
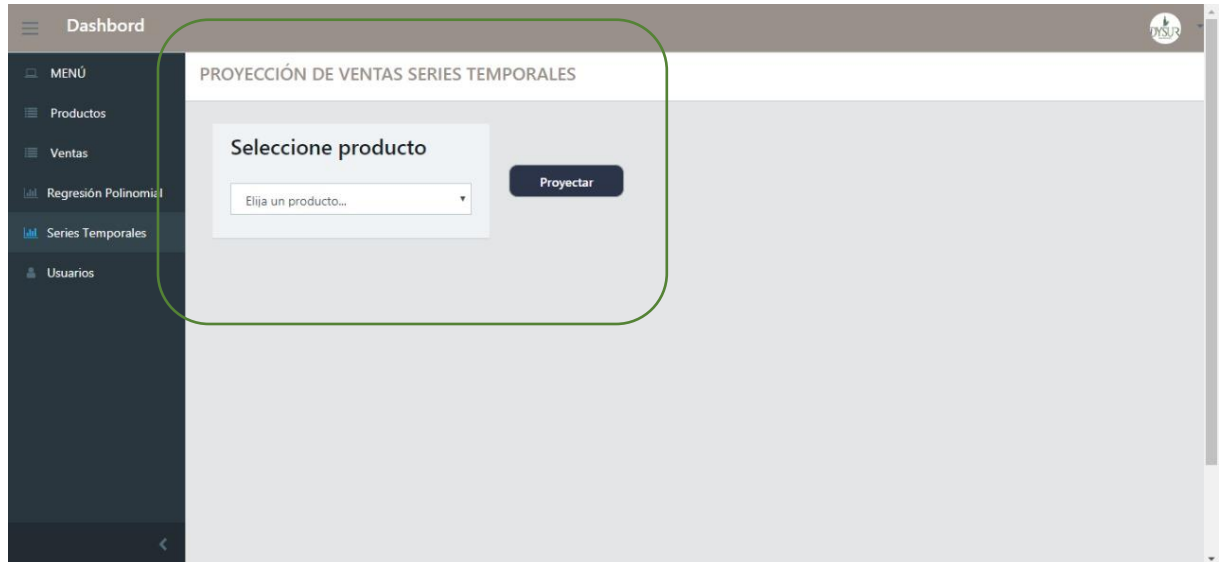
Graficar



7. Se presenta igual un cuadro de **Resumen** el cual va indicando los datos según se haya mandado a calcular el polinomio, este cuadro es dinámico lo que hace que se vaya llenando la información dependiendo de cómo se proyecta en base al polinomio.



8. La siguiente opción para realizar la proyección de ventas es mediante **Series Temporales** aquí se tiene que seleccionar el producto y se presiona el botón de proyectar.



9. La última opción es informativa y es en donde se puede visualizar todos los perfiles de los **Usuarios** cargados en el sistema web, también se tiene la opción de agregar como se ve en la figura.

