



PONTIFICIA UNIVERSIDAD CATÓLICA DEL ECUADOR

SEDE ESMERALDAS

**PROGRAMA DE MAESTRÍA EN TECNOLOGÍAS DE LA
INFORMACIÓN**

LÍNEA DE INVESTIGACIÓN:

Ingeniería de Software

MODALIDAD:

Artículos profesionales de alto nivel

TÍTULO:

**CLASIFICACIÓN DE COMENTARIOS TÓXICOS EN REDES SOCIALES
MEDIANTE MACHINE LEARNING**

**PREVIO AL GRADO ACADÉMICO DE
MAGÍSTER EN TECNOLOGÍA DE LA INFORMACIÓN**

AUTOR:

Ing. Jair Oswaldo Bedoya Benavides

ASESOR:

Mgt. Xavier Quiñonez Ku

Esmeraldas, 10 de octubre del 2023





CLASIFICACIÓN DE COMENTARIOS TÓXICOS EN REDES SOCIALES MEDIANTE MACHINE LEARNING

CLASSIFICATION OF TOXIC COMMENTS ON SOCIAL NETWORKS USING MACHINE LEARNING

María Fernanda Revelo Bautista¹, Jair Oswaldo Bedoya Benavides¹, Jaime Paúl Sayago Heredia¹,
Pablo Pico Valencia¹, Xavier Quiñonez Ku¹

¹ Pontificia Universidad Católica del Ecuador, Sede Esmeraldas, Ecuador
{maria.revelo, jair.bedoya, jaime.sayago, pablo.pico,
xavier.quinonez}@pucese.edu.ec

Resumen.

En esta investigación, se aborda el problema de comentarios tóxicos en redes sociales, y cómo la inteligencia artificial (IA) y el aprendizaje automático (Machine Learning) pueden ayudar. Se presenta el desarrollo de un modelo de clasificación empleando IA con técnicas de machine learning que permita identificar comentarios tóxicos en Twitter.

El clasificador planteado, elaborado en Python, se estableció con 7 diferentes algoritmos a los que se les empleó enfoques o estrategias para la clasificación de etiquetas múltiples, preprocesamiento, limpieza y visualización de datos. Este modelo fue entrenado con un total de 159571 comentarios del conjunto de datos del repositorio Kaggle denominado Jigsaw, el cual tiene los comentarios clasificados con diversas características. Posteriormente del entrenamiento, evaluación y comparación del modelo creado se consiguió como resultado un clasificador capaz de identificar palabras o comentarios tóxicos y ofensivos cuya precisión fue del 92.16%

Palabras Clave: análisis de los sentimientos, clasificación de texto, comentarios tóxicos, machine learning, tweets, Twitter.

¹ Maestría en Tecnologías de la Información, PUCE Sede Esmeraldas, ORCID: <https://orcid.org/0009-0002-8611-1379>, <https://orcid.org/0009-0001-8437-9681>, <https://orcid.org/0000-0003-3657-5407>, <https://orcid.org/0000-0003-3518-3313>, <https://orcid.org/0000-0002-8653-6214>





Abstract.

This research addresses the problem of toxic comments in social networks, and how artificial intelligence (AI) and machine learning (Machine Learning) can help. It presents the development of a classification model using AI with machine learning techniques to identify toxic comments on Twitter.

The proposed classifier, developed in Python, was established with 7 different algorithms using approaches or strategies for multi-label classification, preprocessing, cleaning and data visualization. This model was trained with a total of 159571 comments from the Kaggle repository dataset called Jigsaw, which has the comments classified with various features. After the training, evaluation and comparison of the model created, the result was a classifier capable of identifying toxic and offensive words or comments with an accuracy of 92.16%.

Keywords: sentiment analysis; text classification; toxic comments; machine learning; tweets; twitter.

AUTOR DE CORRESPONDENCIA: Xavier Quiñonez Ku

- **Nombre de la revista científica:** Proceedings, Communications in Computer and Information Science
- **Enlace (URL) de la revista:**
<https://link.springer.com/conference/icate>
- **ISSN de la revista:** 978-3-030-42517-3 (Online)
- **Medio(s) de indexación:**
 - Scopus,
 - SCImago SJR: Q3
 - DBLP
 - Google Scholar
 - EI-Compendex
 - Mathematical Reviews
 - Web of Science
 - Doi, Crossref
- **Nombre del editor de la revista:**
Miguel Botto-Tobar, Marcelo Zambrano Vizueté, Sergio Montes León, Pablo Torres-Carrión, Benjamin Durakovic
- **Correo electrónico del editor de la revista:** icat2023@icat-conferences.org





- Fecha de envío del artículo a la revista: 24/09/2023

Evidencias de envío a medio científico.

- Certificado de Aprobación por el asesor



Esmeraldas, 13 de octubre del 2023

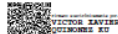
CERTIFICADO

Mediante la presente comunicación, deseo certificar que los distinguidos maestrantes: la Ing. María Fernanda Revelo y el Ing. Jair Bedoya Benavides, inscritos en el programa de Maestría en Tecnologías de la Información, completaron exitosamente la redacción de su artículo científico. Han atendido diligentemente las observaciones efectuadas por los revisores y su asesor.

Por consiguiente, me complace informar que el documento ha alcanzado el nivel de calidad requerido para su evaluación y calificación. Adjunto al presente, se encuentra el informe de análisis realizado a través del sistema Turnitin, el cual revela un nivel de similitud del 9%.

Quedo a disposición para cualquier consulta o requerimiento adicional.

Cordialmente,



Xavier Quiñonez Ku
Docente y Asesor Académico





MARIA REVELO - JAIR BEDOYA - VF

INFORME DE ORIGINALIDAD

9%

INDICE DE SIMILITUD

6%

FUENTES DE INTERNET

5%

PUBLICACIONES

3%

TRABAJOS DEL
ESTUDIANTE

FUENTES PRIMARIAS

1	vijaygadepalli.medium.com Fuente de Internet	1%
2	www.researchsquare.com Fuente de Internet	1%
3	Devi, P.R. Suganya, R. Baskaran, and S. Abirami. "Multi-Label Learning with Class-Based Features Using Extended Centroid-Based Classification Technique (CCBF)", <i>Procedia Computer Science</i> , 2015. Publicación	<1%
4	www.researchgate.net Fuente de Internet	<1%
5	Submitted to Cork Institute of Technology Trabajo del estudiante	<1%
6	Submitted to The University of the West of Scotland Trabajo del estudiante	<1%
7	dspace.daffodilvarsity.edu.bd:8080 Fuente de Internet	<1%





- Capturas del envío

ICAT 2023 (author) Help / Log out

New Submission Submission 25 Conference News EasyChair

My Submissions for ICAT 2023

Using the submission author environment you can view or manage your submissions to ICAT 2023. You can make new submissions or update your previous submissions.

To **make a new submission** click on "New Submission".

To **view or update your existing submission** click on the corresponding "view" icon.

Additional information about submission to ICAT 2023 can be found at the [ICAT 2023 Web page](#).

Questions about submissions can be sent to the conference contact email mbotto@ieee.org.

Please note that if you do nothing (not even click on the menu) for more than two hours, your session will expire and you will have to log in again.

#	Authors	Title	View	Paper	Program
25	María Revelo-Bautista, Jair Bedoya-Benavides, Jaime Sayago-Heredia, Pablo Pico-Valencia and Xavier Quiñonez-Ku	CLASSIFICATION OF TOXIC COMMENTS ON SOCIAL NETWORKS USING MACHINE LEARNING			

ICAT 2023 (author) Help / Log out

New Submission Submission 25 Conference News EasyChair

ICAT 2023 Submission 25

If you want to **change any information** about your paper, use links in the upper right corner.

For all questions related to processing your submission you should contact the conference organizers. [Click here to see information about this conference](#).

[Update information](#)
[Update authors](#)
[Update file](#)
[Withdraw](#)

Submission 25

Title	CLASSIFICATION OF TOXIC COMMENTS ON SOCIAL NETWORKS USING MACHINE LEARNING
Paper:	(Sep 24, 15:32 GMT)
Author keywords	sentiment analysis text classification toxic comments machine learning tweets twitter
Abstract	This research addresses the problem of toxic comments in social networks, and how artificial intelligence (AI) and machine learning (Machine Learning) can help. It presents the development of a classification model using AI with machine learning techniques to identify toxic comments on Twitter. The proposed classifier, developed in Python, was established with 7 different algorithms using approaches or strategies for multi-label classification, preprocessing, cleaning and data visualization. This model was trained with a total of 159571 comments from the Kaggle repository dataset called Jigsaw, which has the comments classified with various features. After the training, evaluation and comparison of the model created, the result was a classifier capable of identifying toxic and offensive words or comments with an accuracy of 92.16%.
Submitted	Sep 24, 15:32 GMT
Last update	

first name	last name	email	country	affiliation	Web page	corresponding?
María	Revelo-Bautista	maria.revelo@pucese.edu.ec	Ecuador	Pontificia Universidad Católica del Ecuador Sede Esmeraldas		✓
Jair	Bedoya-Benavides	jair.bedoya@pucese.edu.ec	Ecuador	Pontificia Universidad Católica del Ecuador Sede Esmeraldas		✓
Jaime	Sayago-Heredia	jaime.sayago@pucese.edu.ec	Ecuador	Pontificia Universidad Católica del Ecuador Sede Esmeraldas		
Pablo	Pico-Valencia	pablo.pico@pucese.edu.ec	Ecuador	Pontificia Universidad Católica del Ecuador Sede Esmeraldas		
Xavier	Quiñonez-Ku	xavier.quinonez@pucese.edu.ec	Ecuador	Pontificia Universidad Católica del Ecuador Sede Esmeraldas		✓

Submission 25

Title	CLASSIFICATION OF TOXIC COMMENTS ON SOCIAL NETWORKS USING MACHINE LEARNING
Paper:	(Sep 24, 15:32 GMT)
Author keywords	sentiment analysis text classification toxic comments machine learning tweets twitter
Abstract	This research addresses the problem of toxic comments in social networks, and how artificial intelligence (AI) and machine learning (Machine Learning) can help. It presents the development of a classification model using AI with machine learning techniques to identify toxic comments on Twitter. The proposed classifier, developed in Python, was established with 7 different algorithms using approaches or strategies for multi-label classification, preprocessing, cleaning and data visualization. This model was trained with a total of 159571 comments from the Kaggle repository dataset called Jigsaw, which has the comments classified with various features. After the training, evaluation and comparison of the model created, the result was a classifier capable of identifying toxic and offensive words or comments with an accuracy of 92.16%.
Submitted	Sep 24, 15:32 GMT
Last update	

first name	last name	email	country	affiliation	Web page	corresponding?
María	Revelo-Bautista	maria.revelo@pucese.edu.ec	Ecuador	Pontificia Universidad Católica del Ecuador Sede Esmeraldas		✓
Jair	Bedoya-Benavides	jair.bedoya@pucese.edu.ec	Ecuador	Pontificia Universidad Católica del Ecuador Sede Esmeraldas		✓
Jaime	Sayago-Heredia	jaime.sayago@pucese.edu.ec	Ecuador	Pontificia Universidad Católica del Ecuador Sede Esmeraldas		
Pablo	Pico-Valencia	pablo.pico@pucese.edu.ec	Ecuador	Pontificia Universidad Católica del Ecuador Sede Esmeraldas		
Xavier	Quiñonez-Ku	xavier.quinonez@pucese.edu.ec	Ecuador	Pontificia Universidad Católica del Ecuador Sede Esmeraldas		✓





Conference Information

Acronym of the event:	ICAT 2023
Name of the event:	International Conference on Applied Technologies
Web page:	https://icat-conferences.org/
Contact emails:	mbotto@ieee.org
Submission page:	https://easychair.org/conferences/?conf=icat2023

- Captura del recibido

28/9/23, 18:36

ICAT 2023 submission 25: Xavier Quiñonez - Outlook

ICAT 2023 submission 25

ICAT 2023 <icat2023@easychair.org>

Dom 24/09/2023 10:32

Para: Xavier Quiñonez <xavier.quinonez@pucese.edu.ec>

Dear authors,

We received your submission to ICAT 2023 (International Conference on Applied Technologies):

Authors : María Revelo-Bautista, Jair Bedoya-Benavides, Jaime Sayago-Heredia, Pablo Pico-Valencia and Xavier Quiñonez-Ku

Title : CLASSIFICATION OF TOXIC COMMENTS ON SOCIAL NETWORKS USING MACHINE LEARNING
Number : 25

The submission was uploaded by Xavier Quiñonez <xavier.quinonez@pucese.edu.ec>. You can access it via the ICAT 2023 EasyChair Web page

<https://easychair.org/conferences/?conf=icat2023>

Thank you for submitting to ICAT 2023.

Best regards,
EasyChair for ICAT 2023.

