



Universidad del Azuay

Departamento de Posgrados

**CREACIÓN DE UN MODELO PARA LA EXTRACCIÓN
DE PALABRAS CLAVES SOBRE ACTIVIDADES DE
CIBERBULLYING EN TEXTOS CORTOS: CASO DE ESTUDIO
- TWITTER**

Maestría En Sistemas De Información Mención Inteligencia De Negocios

Autor:

Ing. William Hermel Astudillo Quituisaca

Director:

Ing. Irene Priscila Cedillo Orellana Ph. D.

Cuenca – Ecuador

2023

Dedicatoria

Este trabajo va dedicado a Dios, por darme salud, vida y valor para afrontar la vida día a día. A mi esposa Bernarda Vintimilla, quien, con su amor, siempre me ha dado valor, ánimo y fuerzas para superarme día a día. A mis hijos Carolina Maite y Matías Felipe, que día a día me dan un motivo para buscar ser mejor, y que me alegran la vida con sus ocurrencias. A mis Padres, Angelita y Julio por siempre estar pendientes y apoyarme en cada paso y etapa de mi vida.

Y a todas las personas que forman parte de mi vida y me han ayudado a alcanzar esta meta.

Agradecimiento

Agradezco en primer lugar a Dios, por permitirme estudiar esta maestría, por las bendiciones que recibo día a día de él.

Agradezco el apoyo recibido por la Ing. Priscila Cedillo PhD. Gracias por su paciencia, y las enseñanzas recibidas, en especial por la guía recibida para poder realizar este trabajo.

A todos los profesores de la maestría que me enseñaron mucho, no solo desde el punto de vista técnico, también humano.

A mi esposa Bernarda Vintimilla, por todo su apoyo incondicional y por siempre impulsarme a crecer y salir adelante, por ser mi compañera incondicional y siempre estar pendiente de mí.

Finalmente agradezco a mis padres, por el apoyo que recibo de ellos cada día, y a todas las personas que de una u otra manera aportaron para poder cumplir esta meta de mi vida, que, aunque no las nombre a todas, siempre estaré agradecido.

Resumen

El ciberbullying impacta negativamente en la sociedad debido a las consecuencias sufridas por víctimas, acosadores y espectadores. El acceso a Internet y a las redes sociales forma parte de las actividades diarias; incluyendo a usuarios más jóvenes que pueden carecer del conocimiento para manejar situaciones de ciberacoso. Educar a la sociedad actual, especialmente a la población joven, es esencial para abordar y mitigar estas consecuencias.

El presente trabajo busca contribuir a la creación de guiones de cápsulas educativas que aporten a la educación de la población. Para lograrlo, se desarrolló un modelo que permite obtener datos de la red social Twitter/X mediante un proceso de búsqueda automatizado. Se aplica un proceso de limpieza y preprocesamiento a los datos para luego procesarlos utilizando el modelo LDA para extraer las palabras clave. La generación del guion se realiza utilizando el modelo "text-davinci-003" de GPT-3 a través de la API de OpenAI, basándose en el contexto proporcionado y las palabras clave identificadas. Un experto en el dominio validó este resultado mediante un caso de estudio.

Palabras claves: Twitter, ciberbullying, análisis de datos, LDA, OpenIA, GPT-3



Firmado electrónicamente por:
**IRENE PRISCILA
CEDILLO
ORELLANA**

Ing. Irene Priscila Cedillo Orellana Ph. D.
Director del trabajo de titulación.

Ing. William H. Astudillo Q.
Autor

Abstract

Cyberbullying negatively impacts society due to the consequences suffered by victims, bullies, and bystanders. Internet and social media access are parts of daily activities, including younger users who may lack the knowledge to handle cyberbullying situations. Educating today's society, especially the younger population, is essential to address and mitigate this consequences.

The present work aims to contribute to creating educational capsules script that contribute to the population's education. To achieve this, a model was developed that allows for data from the social network Twitter/X through an automated search process. A cleaning and preprocessing process is applied to the data to then process it using the LDA model to extract keywords. the script generation is performed using the "text-davinci-003" model from GPT-3 through the OpenAI API based on provided context and the identified keywords. An expert in the domain validated this result by applying a case study.

Keywords: Twitter, cyberbullying, data analysis, LDA, OpenIA, GPT-3

Translated by



Firmado electrónicamente por:
**IRENE PRISCILA
CEDILLO
ORELLANA**



Ing. Irene Priscila Cedillo Orellana Ph. D.

Final Degree Project Director

Ing. William H. Astudillo Q.

Author