

CONTINUIDAD**PROYECTO “DETECCIÓN DEL ACOSO EN REDES SOCIALES APLICANDO *EARLY DETECTION*”**RESOLUCIÓN VIDI Nº **031/2024**UNIDAD ACADÉMICA **FACULTAD DE INGENIERÍA**DIRECTORA **Dra. CAGNINA, Leticia Cecilia**EQUIPO DE INVESTIGACIÓN **Dr. ERRECALDE, Marcelo Luis (Co-Director)**
Mg. Lic. TALAMÉ, María LorenaCAMPO DE APLICACIÓN **ESTRUCTURAS, PROCESOS Y SISTEMAS POLÍTICOS Y SOCIALES**DISCIPLINA GENERAL **CIENCIAS EXACTAS Y NATURALES: Informática**PALABRAS CLAVE **CYBERBULLYING – DEEP LEARNING – EARLY DETECTION – SOCIAL NETWORKS**FINANCIAMIENTO **CONSEJO DE INVESTIGACIONES****PROYECTO DE INVESTIGACIÓN CONVOCATORIA RRNº 1265/2019**
CONTINUIDAD RRNº 325/2020 con FINALIZACIÓN RRNº 564/2021
CONTINUIDAD RRNº 575/2021 con FINALIZACIÓN RRNº 665/2022
CONTINUIDAD RRNº 672/2021 con FINALIZACIÓN RVIDiNº 061/2023**RESUMEN**

Los niños y los adolescentes cada vez más utilizan las redes sociales para comunicarse. Sin embargo, el hecho de una falsa identidad digital hace que sean víctimas de acosadores. El cyberbullying es un flagelo al que se enfrentan que ocasiona cambios de actitud, de comportamiento y hasta a veces, llega a provocar el suicidio. Es por esto que es necesario contar con un filtro de mensajes agresivos para intentar terminar con el ciberacoso y alertar a los usuarios al detectar conversaciones con características de acoso antes de que éste efectivamente suceda

En este trabajo, a través de técnicas de deep learning y considerando la detección anticipada o early detection, se desarrollará un modelo capaz de predecir conversaciones agresivas. Esto permitirá establecer una barrera entre el acosador y su potencial víctima evitando las consecuencias indeseables del ciberacoso

Para ello se aplicarán algoritmos de early detection y modelos de lenguajes pre-entrenados buscando un modelo que tenga un alto grado de confiabilidad en su predicción. Se hará mediante la recopilación de conversaciones de una red social en un período de tiempo establecido. Luego se prepararán esos textos para su posterior análisis. Se diseñarán distintos modelos que se evaluarán con métricas acordes al objetivo propuesto. Por último, se validará el modelo seleccionado.

ABSTRACT

Children and adolescents increasingly use social networks to communicate. However, the fact of a false digital identity makes them victims of bullies. Cyberbullying is a scourge they face that causes changes in attitude, behavior and sometimes even leads to suicide. This is why it is necessary to have a filter of aggressive messages to try to end cyberbullying and alert users to detect conversations with characteristics of harassment before it actually happens.

In this work, through deep learning techniques and considering early detection, a model capable of predicting aggressive conversations will be developed. This will allow the establishment of a barrier between the harasser and the potential victim, avoiding the undesirable consequences of cyberbullying.

For this purpose, early detection algorithms and pre-trained language models will be applied in search of a model that has a high degree of reliability in its prediction. This will be done by collecting conversations from a social network in an established period of time. These texts will then be prepared for further analysis. Different models will be designed and evaluated with metrics according to the proposed objective. Finally, the selected model will be validated.