

Aplicación Machine Learning para la detección de violencia en las redes sociales

Luis David Marin Gallego
Ingeniería de sistemas y computación
Universidad de Caldas, Manizales Colombia
luis.1701624364@ucaldas.edu.co



FACULTAD DE INGENIERÍAS

Problema



Del año 2020 al 2021 se presentó un aumento en los usuarios de redes sociales enorme, pasó de **3.71** mil millones a **4.20** mil millones, un **+13.2%**. (Kemp, 2021)



En 2015 el **14.5%** de los estudiantes entre **12 y 18 años** reportaron discursos de odio en la escuela durante el año escolar, entre los que destaca el discurso de odio racial con un **6.3%** y relacionados con su etnia **3.5%**. (Zhang et al., 2017)



En los meses Enero y Febrero del 2017 **twitter** solo eliminó el **1%** de los post's reportados como discurso de odio por los usuarios. (Michael Knigge, 2017).

Objetivo

Desarrollar un sistema machine learning que detecte el discurso de odio en textos, blogs y publicaciones en redes sociales, que alerte al usuario sobre el contenido que está por ver o por publicar.

Pregunta

¿De qué maneras podría, el uso de Machine Learning, ayudar en la disminución del acoso y el discurso de odio en las redes sociales, alertando al usuario de lo que está por ver o por publicar?

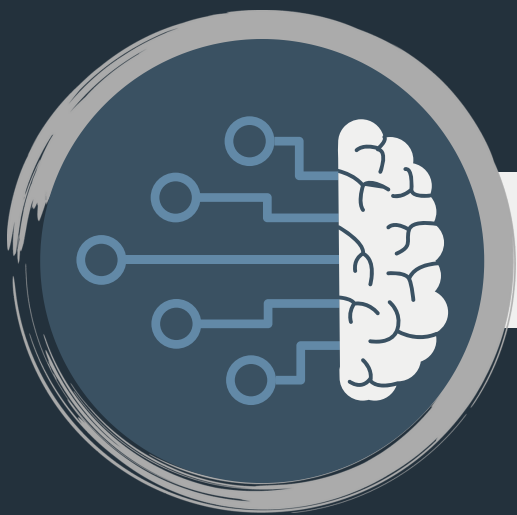


Metodología



Recopilación de datos

- Tweets
- Posts
- Libros
- Construidas



Machine Larning

- Entrenamiento
- Pruebas
- Comparación
- Elección



Desarrollo prueba piloto

- Requerimientos
- Análisis y diseño
- Implementación
- Integración y pruebas

Resultados esperados

Resultado	Descripción	Indicador de logro
Generación de nuevo conocimiento.	Comprender y conocer el estado actual del discurso de odio en redes sociales.	1 articulo cientifico enviado a revista indexada. Trámite de registro de software.
Fortalecimiento de la comunidad cientifica colombiana.	Formación de investigadores.	Formación de estudiantes 1 de maestría 1 de pregrado.
Apropiación social del conocimiento.	Participación en eventos científicos presentando los resultados de la investigación.	1 ponencia del evento en el cual se participó.

Conclusión

Se debe hacer algo para tratar de frenar la proliferación de odio en redes sociales, con este proyecto se espera tener una mejor comprensión a medio tiempo de las carencias que tienen actualmente las redes sociales en su lucha contra el discurso de odio, y a la larga se espera que las personas tengan mejor conocimiento de las repercusiones de sus actos, incluso si son por internet.

Referencias

- Kemp, S. (2021). Digital 2021: Overview report. <https://datareportal.com/reports/digital-2021-global-overview-report>
- Zhang, A., Wang, K., Zhang, J., & Oudekerk, B. A. (2017). Indicators of School Crime and Safety: 2016. NCES 2017-064/NCJ 250650. National Center for Education Statistics, 1–261. <https://nces.ed.gov/pubs2017/2017064.pdf>
- Michael Knigge (2017). Hatespeechcurbgoesbeyond Facebook, Twitter. DW.COM. Recuperado 7 de abril de 2021, de <https://www.dw.com/en/hate-speech-curb-should-look-beyond-facebook-twitter/a-39550114>