

PROYECTO 4: NLP

Planteamiento

YouTube lleva un tiempo preocupado por el aumento de los mensajes de odio entre los comentarios de sus vídeos y ha llegado a un punto donde un equipo de moderadores no da a basto y aumentar ese equipo sería prohibitivamente caro, además de que no se hace posible escalar el equipo al ritmo al que crece la plataforma y aumenta el volumen de estos mensajes.

Por ese motivo han decidido subcontratar a una consultora, donde trabajáis, en busca de una solución para poder detectar este tipo de mensajes de forma automática y para así poder eliminarlos, banear al usuario o tomar las acciones necesarias..

Han hecho hincapié en que es importante la implementación de la solución que encontréis, quieren una solución práctica por encima de una herramienta precisa.

Plazos

Vuestro jefe quiere ver las propuestas de solución el día **19 de Noviembre** (el 20 se presenta)

Condiciones de entrega

Para el día de la reunión, será necesario entregar:

- Presentación de 30 min (negocio + tecnica)
- Proyecto en github

Tecnologías que se pueden aplicar

- Scikit-learn
- Spacy / NLTK
- Pandas
- Huggingface
- Streamlit / Dash / FastAPI / Django / Flask
- BeautifulSoup / requests / Scrapy
- Trello / Jira ...

Niveles de Entrega

Nivel Esencial:

Un modelo de ML que reconozca los mensajes de odio

Overfitting inferior al 5%

Una solución que productivice el modelo (una interfaz, API o lo que se os ocurra, que permita a un usuario consultar si un mensaje es o no de odio)

Nivel Medio:

Un modelo de ML con técnicas de ensemble que reconozca mensajes de odio Una solución que permita reconocer los posibles mensajes de odio dado un enlace a un vídeo en concreto

Incluir tests unitarios

Nivel Avanzado:

Un modelo que implemente redes neuronales y mejore significativamente los resultados frente a una solución de Machine Learning (RRN o LSTM)

Una solución que permita introducir la url de un vídeo concreto reconocer mensajes de odio haciendo seguimiento del video en tiempo real

Dockerizar la aplicación

Nivel Experto:

Utilizar un modelo basado en transformers

Guardar en base de datos los resultados de las predicciones.

Trackear los experimentos realizados con MLFlow.

Datos



[Youtube Comments](#)