

Prueba técnica

Arrate Landia Iglesias



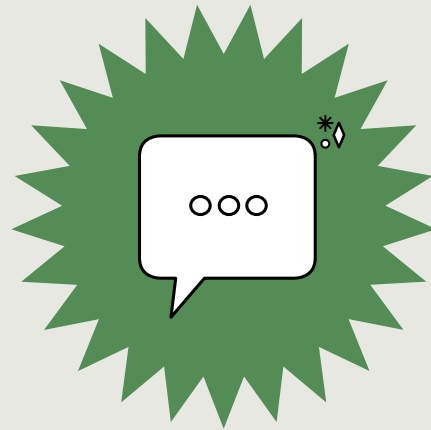
Recolección de datos

Snsrape es una biblioteca de Python que proporciona una interfaz para extraer datos de la plataforma de Twitter.

Contiene las clases y funciones relacionadas con la extracción de datos de Twitter

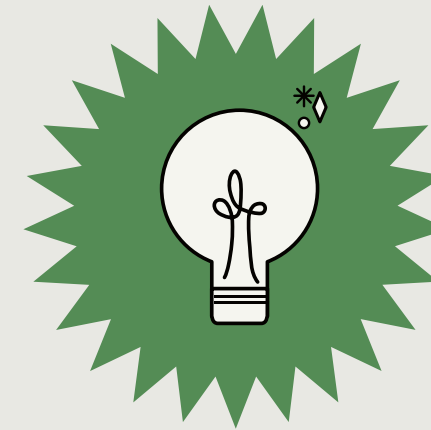


Base de datos



tweets

id INTEGER PRIMARY KEY,
text TEXT,
date TEXT,
author_id INTEGER,
retweets INTEGER,
replies INTEGER,
likes INTEGER,
quotes INTEGER



usuarios

id INTEGER PRIMARY KEY,
name TEXT,
username TEXT

Análisis

Tweet con mayor repercusión social

Se han tenido en cuenta -> Retweets, Replies, Likes, Quotes

Para reactivarse!!! Tenemos Meetup este martes 11 de abril a las 19:00 en @TheBridge_Tech @dev7ba nos hablará sobre la mempool!! Charla para todo el mundo, sepas o no sobre #Bitcoin @GaboHBeaumont @kilrau @libertariadora @olya_borderless <https://t.co/bHzpDduIRoSe> agradece RT

Usuario con la mayor cantidad de tweets

TheBridge_Tech con 45 tweets

Mes con mayor número de tweets

Marzo

Palabras más frecuentes

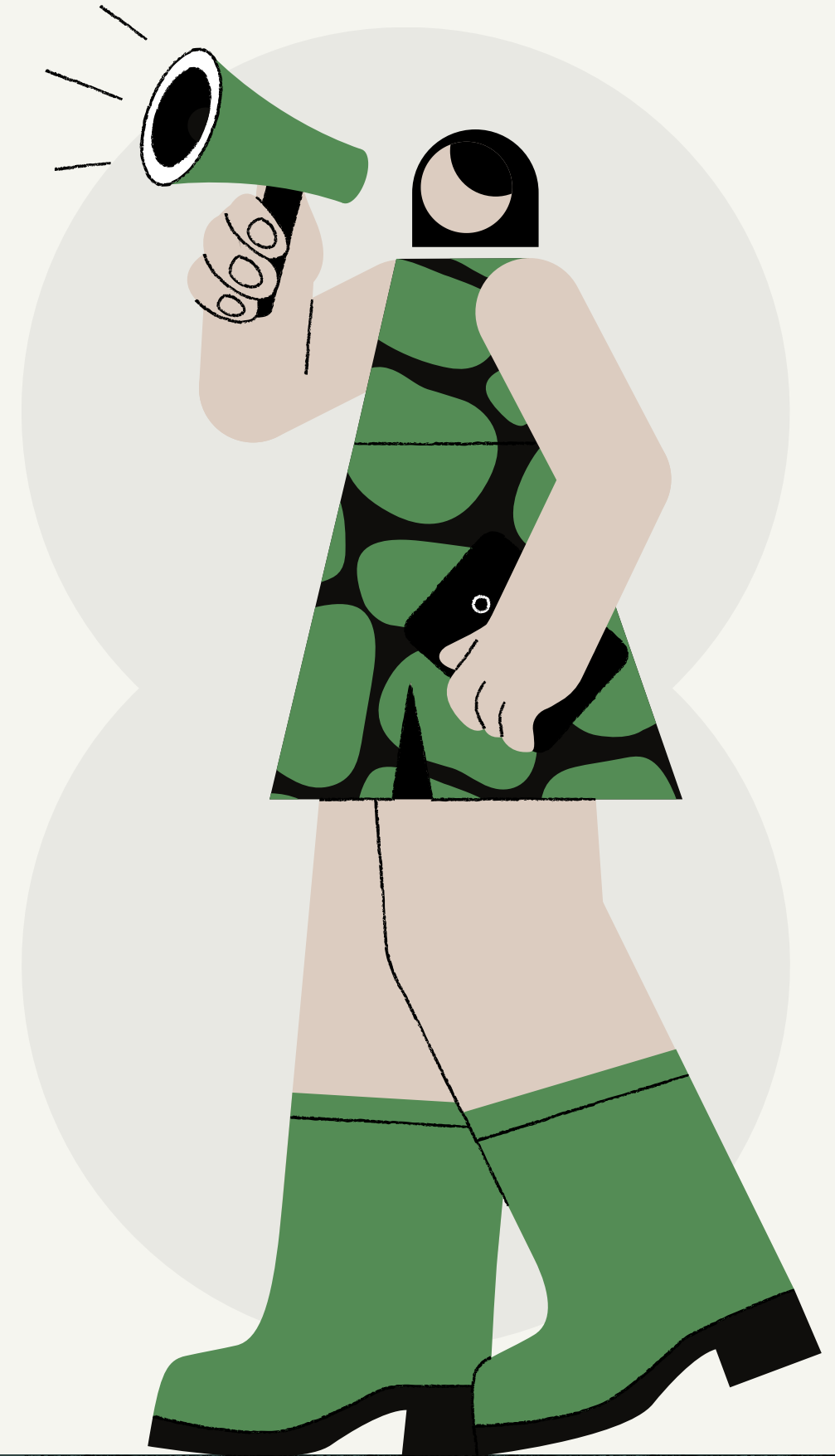
sngular: 27

gracias: 20

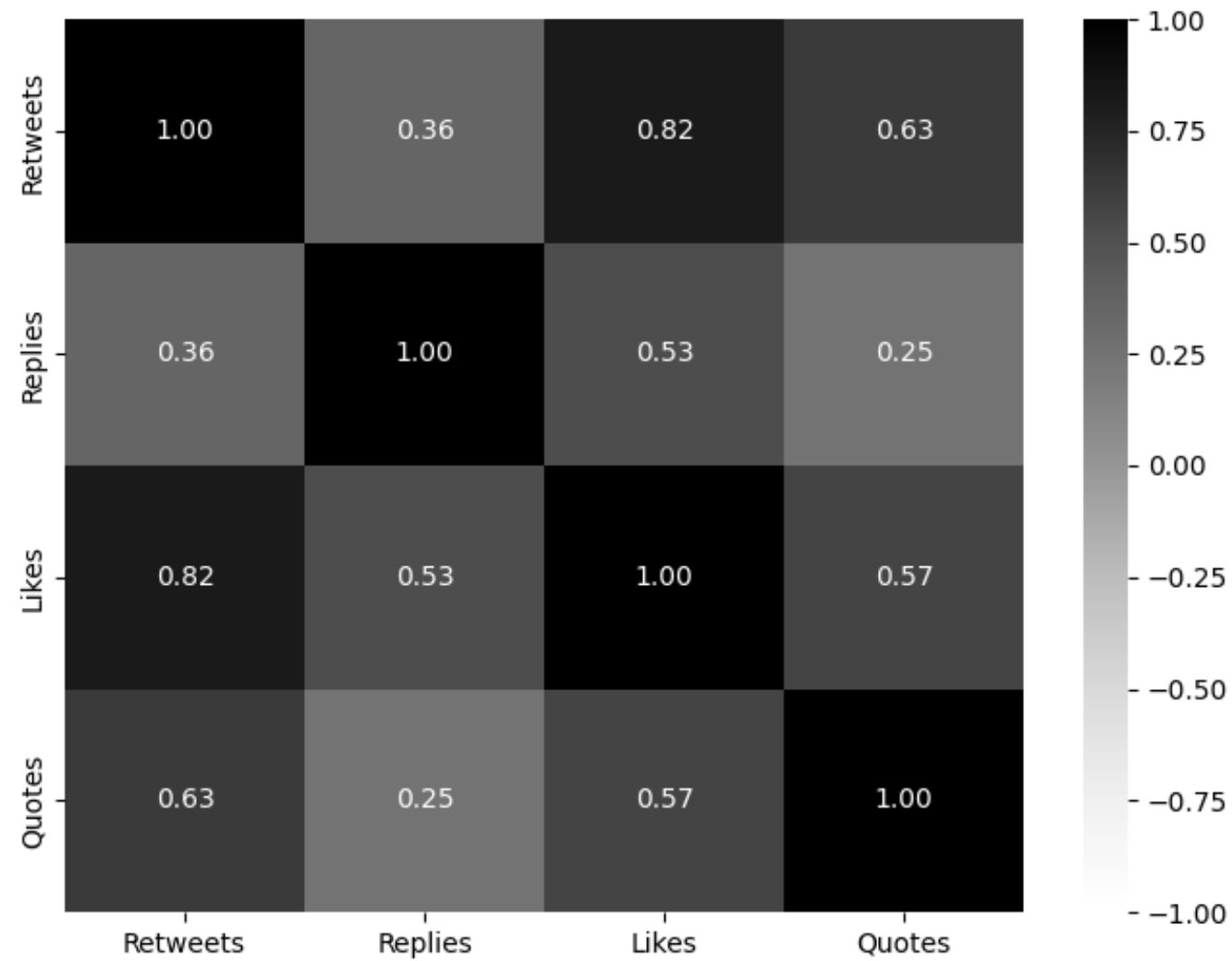
digital: 20

entradas: 19

bootcamp: 17



Correlación matemática entre las métricas públicas



Conclusión

Mi conclusión final sobre el análisis es que twitter no es muy útil para llegar a la gente.

No hay mucha interacción.



Determinación de sentimiento de los tweets

	Tweet	Sentiment
0	Ayer visitamos a los alumnos de @TheBridge_Tec...	1
1	¿Qué hacían nuestros alumni antes del #bootcam...	0
2	#MLMonitoring and #Observability, mar, 23 may ...	1
3	Nos colamos en el campus de #Bilbao para descu...	0
4	Así es el #bootcamp de #UXUIProductDesign que ...	0
...
138	¿Los colores tienen género?¿podemos hacer una ...	1
139	@BitcoinTuesday_ ¡Un placer!	0
140	Gracias a nuestros amigos @TheBridge_Tech por ...	0
141	¡Nos vamos al WorkinLan Workshop en Bilbao jun...	0
142	Les compartimos un mini resumen del 1º encuent...	1

Variables más importantes del modelo

Palabras clave (odio)
Polaridad de las palabras (bueno)
Intensificadores y negaciones
Emoticonos y emojis
Hashtags
Menciones de usuarios

Mejora del modelo

Recopilación de datos adicionales
Limpieza y preprocesamiento de datos
Selección de características relevantes
Ajuste de hiperparámetros
Utilización de modelos de lenguaje preentrenados
Experimentar con diferentes algoritmos y arquitecturas de modelos
Validación cruzada
Análisis de errores y retroalimentación del modelo

Otras oportunidades de modelos de ML

Detección de tendencias
Recomendación de usuarios y contenido
Detección de noticias falsas
Clasificación de contenido
Detección de eventos y emergencias

Despliegue del modelo pre entrenado



← → ↻ 🏠 127.0.0.1:5000/predict

Ingrese el texto del tweet:

Predecir

← → ↻ 🏠 127.0.0.1:5000/predict

La predicción de sentimiento para el tweet 'hola que tal!' es: Sentimiento negativo o nulo

El texto del tweet se ha ingresado en la base de datos.

[Ingresar otro texto](#)

