

Sentiment analysis on Twitter

Desarrollado con Pandas, Seaborn, SQLite3 y Heroku



Data acquisition



Twitter API



POSTMAN

Dataframe 1 con id del tweet, texto del tweet, fecha del tweet, id del autor y métricas públicas (likes, retweets, respuestas y citas)

Dataframe 2 con id del autor, nombre del autor, nombre de usuario del autor

Base de datos SQLite3

DBeaver 22.0.5 - tweets

File Edit Navigate Search SQL Editor Database Window Help

SQL Commit Rollback Auto

Database Navigator Projects

Enter a part of object name here

API_database.db

- Tables
 - tweets
 - users
- Views
- Indexes
- Sequences
- Table Triggers
- Data Types

books.db

chinook.db

df8no4dfu6vbl5 - ec2-23-23-151-191.compute-1.amazonaws.com:5432

my_database.db

tweets x users

Properties Data ER Diagram

API_database.db Tables tweets

Enter a SQL expression to filter results (use Ctrl+Space)

	author_id	text	created_at	id_tweet	retweet_count	reply_count	like_count
1	1088261225153060864	Los emprendedores y profesionales digitales del futuro	2022-07-13T14:45:06.000Z	1547230643561664515		1	[NULL]
2	1232276995796586499	Hoy hemos tenido una súper visita a nuestras oficinas @	2022-07-13T13:22:40.000Z	1547209900874506241		1	[NULL]
3	1162694149956603904	¡Descubre la experiencia de Dani! ▶ Dani estudió una #f	2022-07-12T11:22:45.000Z	1546817334882476032		2	[NULL]
4	1085383992	#VentureCapital/#PrivateEquity: @TheBridge_Tech cierr	2022-07-11T07:34:21.000Z	1546397466517540866		0	[NULL]
5	27890913	#Inversiones 💰 @TheBridge_Tech ha cerrado una r	2022-07-07T10:16:04.000Z	1544988613129261056		0	[NULL]
6	190658967	#SAVETHEDATE ▶ @TheBridge_Tech pone en marcha	2022-07-06T08:12:00.000Z	1544595001086283778		1	[NULL]
7	89141857	@ceeivalencia y @TheBridge_Tech, primera acelerad	2022-07-05T17:30:15.000Z	1544373102997655552		0	[NULL]
8	196056785	"De aquí al 2025 se van a crear 400.000 #empleos er	2022-07-05T08:07:03.000Z	1544231368896532480		2	[NULL]
9	196056785	Todo preparado en @TheBridge_Tech para la presen	2022-07-05T07:37:01.000Z	1544223809112350723		2	[NULL]
10	89141857	@ceeivalencia y @TheBridge_Tech, primera acelerad	2022-07-02T07:30:07.000Z	1543134908935200769		0	[NULL]

DBeaver 22.0.5 - users

File Edit Navigate Search SQL Editor Database Window Help

SQL Commit Rollback Auto

Database Navigator Projects

Enter a part of object name here

API_database.db

- Tables
 - tweets
 - users
- Views
- Indexes
- Sequences
- Table Triggers
- Data Types

books.db

chinook.db

df8no4dfu6vbl5 - ec2-23-23-151-191.compute-1.amazonaws.com:5432

my_database.db

users x

Properties Data ER Diagram

API_database.db Tables users

Enter a SQL expression to filter results (use Ctrl+Space)

	author_id	name	username
1	1088261225153060864	GoHub Ventures	GoHubVentures
2	1232276995796586499	Ines Calabuig	InesCalabuig2
3	1162694149956603904	The Bridge	TheBridge_Tech
4	1085383992	Webcapitalriesgo.com	WCapitalRiesgo
5	27890913	El Referente	ElReferente
6	190658967	hayCANAL.com	haycanal
7	89141857	Emprenemjunts	emprenemjunts
8	196056785	ATREVIEW	Atrevia
9	1018473762902114304	Eioua, o aprendiz	Eioua_aprendiz
10	874916978665938945	Startup Valencia	Startup_VLC
11	604485175	Berta Otero	bertinha84

Análisis de los Datos

Los 3 Tweet con más repercusión social (más likes):



Análisis de los Datos

¿Cual es el usuario que más menciona a la escuela?	HeavyMental_es, menciona a la escuela 23 veces
¿En qué mes se concentra la mayor cantidad de tweets?	En Junio con un total de 118 Tweets
¿Qué palabras son las más frecuentes?	Bootcamp, thebridge_tech, presentación, talento, digital, libro, web, becas, futuro e inteligencia
¿Qué tipo de correlación lineal encuentras entre las métricas?	Existe una correlación lineal > 0.5 entre todas las métricas. Pero las más correlacionadas son los likes y los retweets con un 69% de correlación entre ellas.

Predicciones



POSITIVO: 0.764106

NEGATIVO : 0.235894



POSITIVO: 0.981161

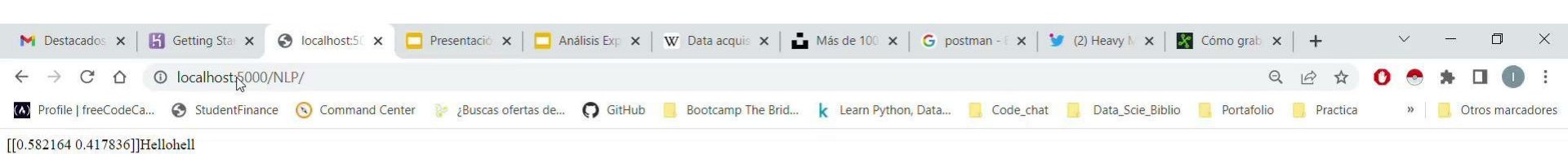
NEGATIVO : 0.018839



POSITIVO: 0.666356

NEGATIVO : 0.333644

Despliegue del modelo en Heroku





¡MUCHAS
GRACIAS!