Proyecto Análisis y Predicción de mercado de Bitcoin Basado en Sentimientos de Twitter.



Diagrama de Flujo Data Engineer

procesos





tareas Automatización del Almacenamiento de los datos SAAS como Airflow o cronjobs. una base de datos como PostgreSQL o MySQL Es necesario investigar cual es el sistema de almacenamiento más beneficioso para el proyecto a nivel de scalabilidad.

Desarrollo de un esquema de versionamiento v de forma que se puedan revertir cambios <u>o trabajar</u> control de calidad Este versionamiento es a Implementación de nivel interno del proyeto procesos de control de calidad para asegurar que los datos extraídos sean para que todos esten enterados de que se han realizado cambios.

versionamiento

Proyecto Análisis y Predicción de mercado de Bitcoin Basado en Sentimientos de Twitter.



Tweet_BTC

timestamp	Date
user_id	String
screen_name	String
full_name	String
text	String
hashtags	Array
mentions	Array
Like	Number
retweet_count	Number
favorite_count	Number
tweet_sentiment	Number

Estructura de Datos

timestamp:

Fecha y hora en que se publicó el tweet.

user_id:

Identificación única del usuario que publicó el tweet.

screen_name:

Nombre de usuario del usuario que publicó el tweet.

full_name:

Nombre completo del usuario que publicó el tweet (si está disponible).

text:

Texto del tweet.

hashtags:

Hashtags incluidos en el tweet (si existen).

mentions:

Usuarios mencionados en el tweet (si existen).

retweet_count:

Número de veces que el tweet ha sido retwitteado.

favorite_count:

Número de veces que el tweet ha sido marcado como favorito.

tweet_sentiment:

Sentimiento del tweet después de análisis (Positivo, Negativo, Neutro)