NLP: se prueban estategias de NLP (text2emotion, textblob) –input: “elonmusk.csv”

Scrapea\_tweets: Se prueban funciones para scrapear tweets y además funciones para generar db con los tweets scrapeados que se van

a dividir para ser procesados (cloud y local)

Load\_and\_clean\_scraped\_tweets: toma todos los tweets scrapeados y los combina y exporta un sólo archivo-- output: “df\_completa.parquet”

NLPK\_V2: toma archivo del tipo "df\_completa.parquet" se prueban funciones para la limpieza de tweets y funciones para matchear las columnas datetime de la bd de tweets y

la bd de cotizaciones de binance.

Se prepara una función que construye dataframes con esta información.

Compute\_strategies: lee los datos de los tweets combinados "df\_completa.parquet" realiza el analisis de sentimiento y arroja

la archivos de la forma “tweets\_split\_\*\_processed.parquet"

Timeframe\_dataframe: Lee archivos "tweets\_split\_\*\_processed.parquet" y los concatena.Tiene una función refinada para generar dataframes con

datos de cotización de binance y tweets.

Binance\_scrapper: realiza el scrappeado de binance y arroja archivos como"historical\_doge\_REAL\_price\_1Hour.parquet"

Generador\_timeshifts: toma la función generadora de timeshifts de la notebook "Timeframe\_dataframe: ". Toma los tweets\_completo\_testing.parquet los graficos de binance como "historical\_doge\_REAL\_price\_1Hour.parquet" y

genera los dataframes de timeshifts combinados del tipo "dataframe\_shifts\_horas\_real.parquet"

Modelos\_ML\_horas\_DOGE: toma archivos del tipo 'dataframe\_shifts\_horas\_real.parquet' y realiza el análisis de modelos de machine learning.