

Ejemplo de informe

Antonio afgs@ucm.es

Otro otro@ejemplo.com

Introducción

Este es un informe de ejemplo para ver cómo usar markdown.

El informe está dividido en dos: `index.md` y `content.md`. Para compilarlo, ejecutar `pandoc index.md content.md -o informe.pdf`.

Metodología

Datos

Se usa el dataset sobre tuits de Elon Musk, `elon_tweets_2012-2022.csv`. Se investigan los términos “twitter” y “tesla”, por ser compañías relevantes. Esto se ha decidido tras hacer un análisis de distintos términos. Los resultados están en la sección Resultados.

Código

```
import matplotlib.pyplot as plt
import nltk
import pandas as pd
from textblob import TextBlob

tknizr = nltk.tokenize.TweetTokenizer()

def preprocess(text):
    tokens = [t.lower() for t in tknizr.tokenize(text)]
    return tokens

df = pd.read_csv("elon_tweets_2012-2022.csv")
df["tokens"] = df["tweet"].apply(preprocess)
df["created_at"] = pd.to_datetime(df["created_at"]).dt.tz_localize(None)
df["sentiment"] = df["tweet"].apply(lambda x: TextBlob(x).sentiment.polarity)
df["month"] = df["created_at"].dt.to_period("M")

def select_token(token):
    return df[df["tokens"].apply(lambda x: token in x)]
```

```
def plot_time_density(df, keyword):  
    twits = select_token(keyword)  
    time_density = twits["month"].value_counts()  
    time_density.sort_index().plot(kind="line")  
  
plot_time_density(df, "twitter")  
plot_time_density(df, "tesla")  
  
plt.legend()  
plt.savefig("tesla_vs_twitter.png")
```

Resultados

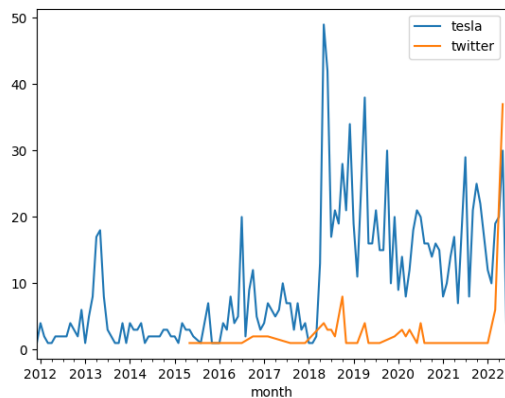


Figura 1: En esta gráfica se ve la evolución de menciones a “tesla” frente a “twitter”

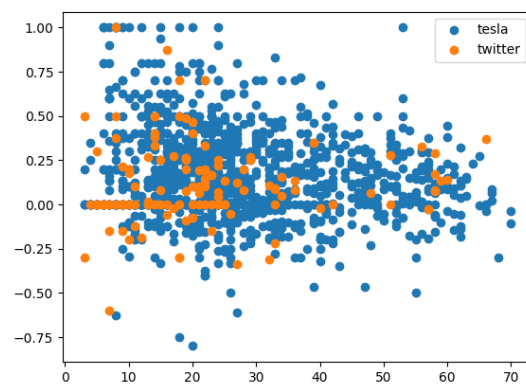


Figura 2: En esta gráfica se ve la polaridad del sentimiento frente a la longitud del tuit