

ANÁLISIS DE DATOS: TWITTER

Javier Castro M.

Universidad De Chile

May 29, 2021

Motivación: El principio del proyecto es entender cómo los datos impactan en el resultado de nuestro modelo. O bien, entender que los posibles sesgos que presenten nuestros datos se traspasaran a los resultados.

TWINT: TWITTER INTELLIGENCE TOOL

Desde el **github** de **twint**:

“Twint is an advanced Twitter scraping tool written in Python that allows for scraping Tweets from Twitter profiles without using Twitter’s API.”

No hay una documentación detallada de twint, pero típicamente se sigue la siguiente estructura para obtener tweets:

```
c = twint.Config()
c.Hide_output = True
c.Geo = "-33.45776872061534,-70.66448325142338,10km"
c.Since = '2019-10-16'
c.Until = '2019-10-22'
c.Store_csv = True
c.Output = 'data/santiago_estallido.csv'
twint.run.Search(c)
```

FIGURE: Uso típico de twint.

Existen otros campos que se pueden modificar:

```
Username: Optional[str] = None
User_id: Optional[str] = None
Search: Optional[str] = None
Username: Optional[str] = None
User_id: Optional[str] = None
Search: Optional[str] = None
```

```
Username: Optional[str] = None
User_id: Optional[str] = None
Search: Optional[str] = None
Username: Optional[str] = None
User_id: Optional[str] = None
Search: Optional[str] = None
```

FIGURE: Three simple graphs

TWITTER EN EL ESTALLIDO SOCIAL

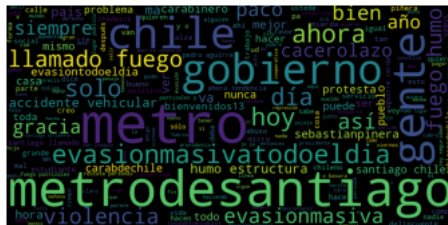


FIGURE: Wordcloud de twitter el 18 de octubre del 2019 con coordenadas en Beaucheff y un radio de 10km.

TWITTER EN EL DÍA DE LA MUJER



FIGURE: Wordcloud de twitter el 8 de marzo del 2020 con coordenadas en Beaucheff y un radio de $10km$.

Muchas Gracias!