

A dark-themed background featuring a tablet in the foreground. The tablet screen shows the Twitter login interface with a blue background, the Twitter bird logo, and the text 'Welcome to Twitter', 'See what's happening in the world right now', and buttons for 'Sign up' and 'Log in'. Overlaid on the tablet screen is the title 'Sentiment Twitter Analysis' in a large, white, sans-serif font.

Sentiment Twitter Analysis

Julián Niebieskikwiat Godoy

What?

- People nowadays express their feelings via social networks.
- To classify a Tweet (text) as a positive or negative Tweet sentiment.



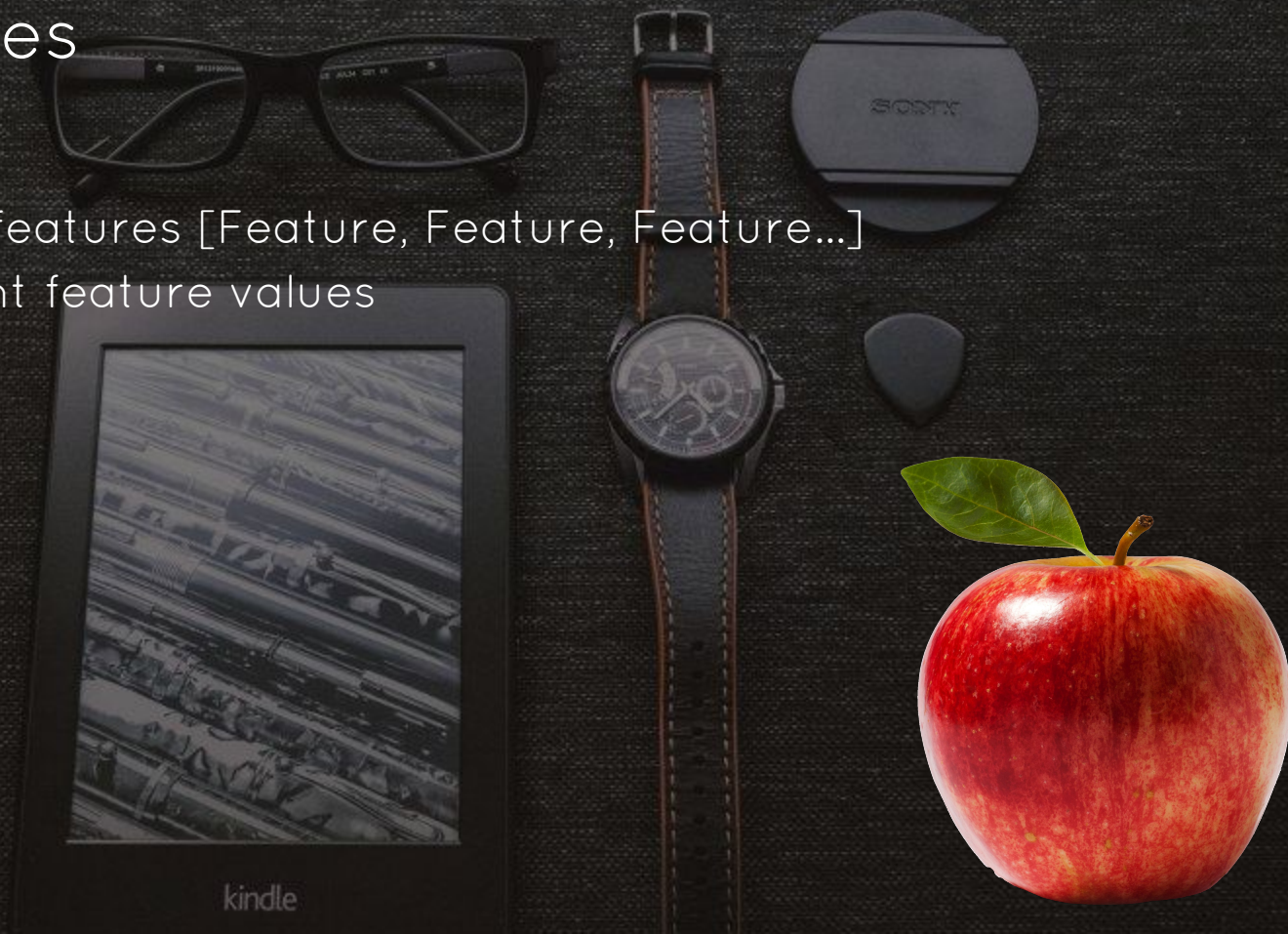
How?

- Labeled Dataset (supervised learning). Source: <https://github.com/caesar0301/awesome-public-datasets>
- Clean data
- Train algorithm (NLTK Naive Bayes Classifier) and dump it
- Predict
- Accuracy
- Twitter connection

```
Terminal
/home/julian/Documents/ITESM/7moSemestre/IA/FinalProject/SentimentalTwitterAnaly
sis> python test_accuracy.py
0.7395
/home/julian/Documents/ITESM/7moSemestre/IA/FinalProject/SentimentalTwitterAnaly
sis> 
```


Naive Bayes

- Classes
- Vectors of features [Feature, Feature, Feature...]
- Independent feature values



What worked?

- 73.9% accuracy
- Twitter connection
- Real time charts

What didn't work?

- Train algorithm with a lot of data
- “Tricky” tweets like: “I love being sad”

What can be improved?

- Porter Stemming
- Charts
- More than just “Positive” or “Negative” classes

What for?

- Get statistics from the opinion of people about an specific topic
- Get real time charts
- Different classifications (not just positive or negative)

