

The News Classifier

Final Project

Ulises Aquino
and Javier Alvarez Jiménez

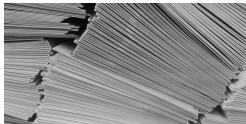
IronHack
Mexico

December 2019



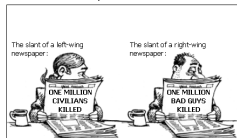
The News

- Nowadays, we have a lot of information, news included.



- The information can be biased.

Skewed News by Eric Perlman



Steps

- Scrap news

Steps

- Scrap news
- Sentiment analysis

Steps

- Scrap news
- Sentiment analysis
- Clusterize

Steps

- Scrap news
- Sentiment analysis
- Clusterize
- Summarize and analyze the clusters

Steps

- Scrap news
- Sentiment analysis
- Clusterize
- Summarize and analyze the clusters
- Upload

Text Summarization¹

Algorithm:

- Text \rightarrow sentences

¹https:

//stackabuse.com/text-summarization-with-nltk-in-python/

Text Summarization¹

Algorithm:

- Text \rightarrow sentences
- Bag of words

¹[https:](https://stackabuse.com/text-summarization-with-nltk-in-python/)

Text Summarization¹

Algorithm:

- Text \rightarrow sentences
- Bag of words
- Scores the words

¹https:

//stackabuse.com/text-summarization-with-nltk-in-python/

Text Summarization¹

Algorithm:

- Text \rightarrow sentences
- Bag of words
- Scores the words
- Rate the sentences

¹[https:](https://stackabuse.com/text-summarization-with-nltk-in-python/)

Text Summarization¹

Algorithm:

- Text \rightarrow sentences
- Bag of words
- Scores the words
- Rate the sentences
- Keep the best sentences

¹[https:](https://stackabuse.com/text-summarization-with-nltk-in-python/)

Final work: <https://bias.netlify.com/>

Recommendation System

The key point is to get scores from the users.

- +1 if the user read the article

Recommendation System

The key point is to get scores from the users.

- +1 if the user read the article
- +1 if the user saves the article

Recommendation System

The key point is to get scores from the users.

- +1 if the user read the article
- +1 if the user saves the article
- If the user comments the article, we make a sentiment analysis and we add the compound value

Improvements

The clustering system

```
{'title': "Rvan's World is ranked number one for the second year in a row.",  
'summary': "As President Donald Trump deals with impeachment, here's a look at the people, teams and orga  
in sports who should face impeachment. A case linked to a $670 million acquisition shows that there may b  
to the authorities' abuse of law enforcement to advance corrupt business interests. Sarah, 30, says she c  
paid for a holiday with the money she blew on one expensive night out. A year after Cuba's government all  
n the island, we follow three activists using the internet to drive rapid change. Rvan's World is ranked  
e for the second year in a row. A website will feature some of the beloved comic strip's classics and, La
```

Figure: Bad cluster.

- Track the news
 - Classification according to the clusters
- Ideas after use it
 - Correct the titles
- Suggestions...