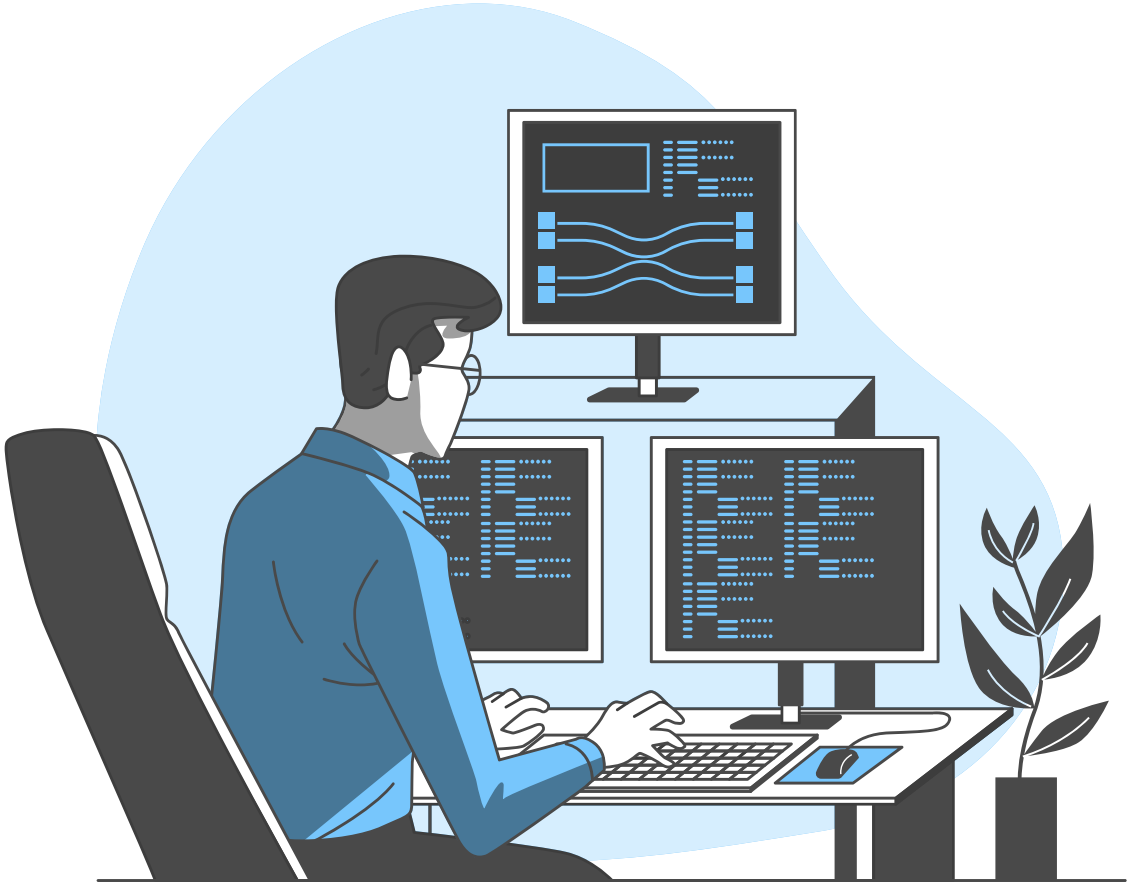


# Filtering News Articles


Group 9

Angela Tseng
Yogesh Boricha
Abdul Shaik
Sadaf Davre

# Goals

## Fetch Articles

Fetch 5 news articles each regarding the hashtags

## Fetch Hashtags

Fetch top 50 hashtags from an area based on WOEID

## Assign 0/1

Assign 0 or 1 based on validity of news article

# Work Division

01

## Angela Tseng

Project Manager

- Writing, organizing notes, meetings, guidance

02

## Abdul Shaik

Programming Lead

- Programming, guidance

03

## Yogesh Boricha

Content Designer

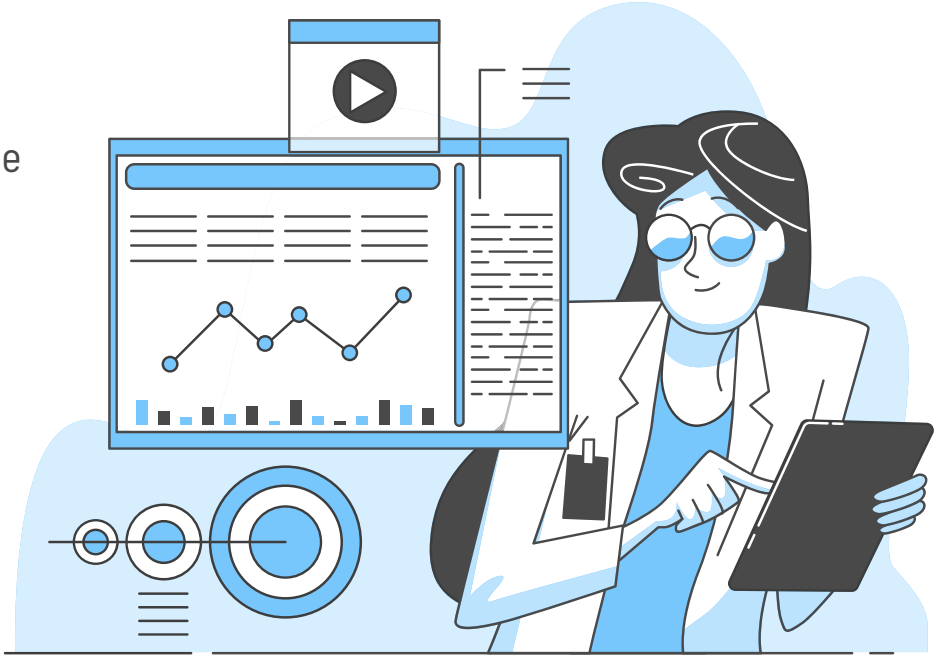
- Research

04

## Sadaf Davre

Programmer

- Programming



# Demo

# Implementation Details

## 01

### TF-IDF

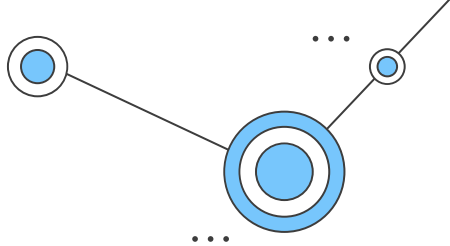
- Passive Aggressive Classifier
- Simple and easy to use

## 02

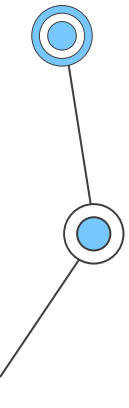
### BERT

- Transformer based technique
- Bidirectional training of the attention model
- Deeper sense of language context

# Implementation Details



- Integrate automatic processing content from Twitter with automatic processing of downloaded news sources and automatic text classification
- Use Twitter's elevated access status and various keys to fetch the top 50 Twitter trending topics in a specified area
- Install packages read in search results and store the title, link, content, and date of the articles into a csv
- Train the algorithm to filter true and false news articles
- Clean the training data
- Train and classify the data before the accuracy score of the trained model is printed
- Pickle package loads a model and vectorize, tokenize the data, and predict the reliability of the article
- Output a list of links with accompanying 0's and 1's on the left



# Limitations

## Unfamiliarity

Machine learning and data extraction are new concepts that required research to use.

...

## Accuracy

The accuracy of detection is not as accurate as desired, whether it be due to incorrectness or bias.

...

## Scope

The scope had to be adjusted 2 or 3 times, creating time limitations on the project functionalities.

...

# Lessons Learned

## Time Management

Training a model, research, and implementation

## Topic Proficiency

Programming proficiency, ML and classifier knowledge

## Scope

Scope creep and project adjustments

# Recommendations for Future Work

- Test plan and evaluation
- User interface
- Classifiers