Negativity in Online Communities: →

Sentiment Analysis & Topic Classification

ML

By: Tessa Wu & Lingwen Deng

Problem Statement

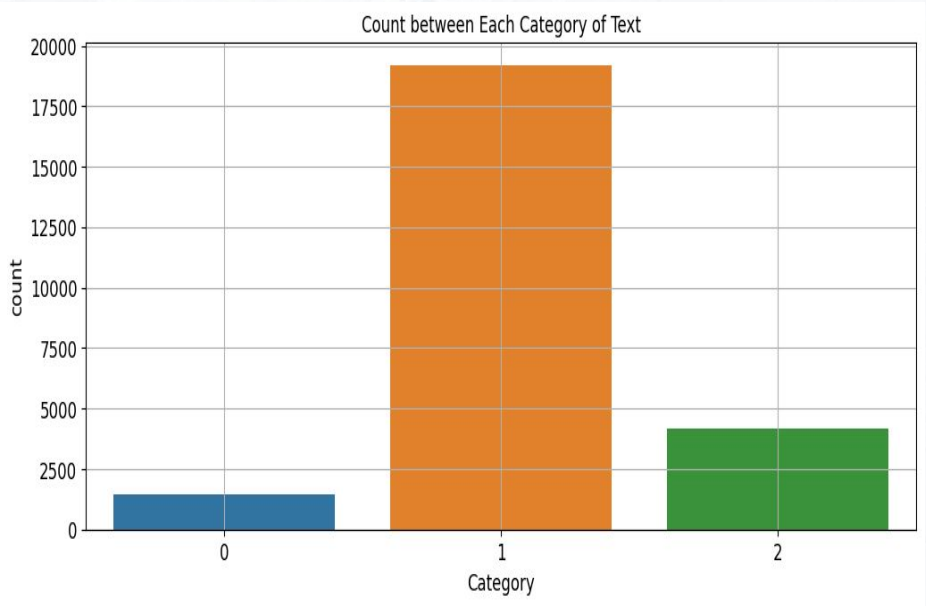
Background

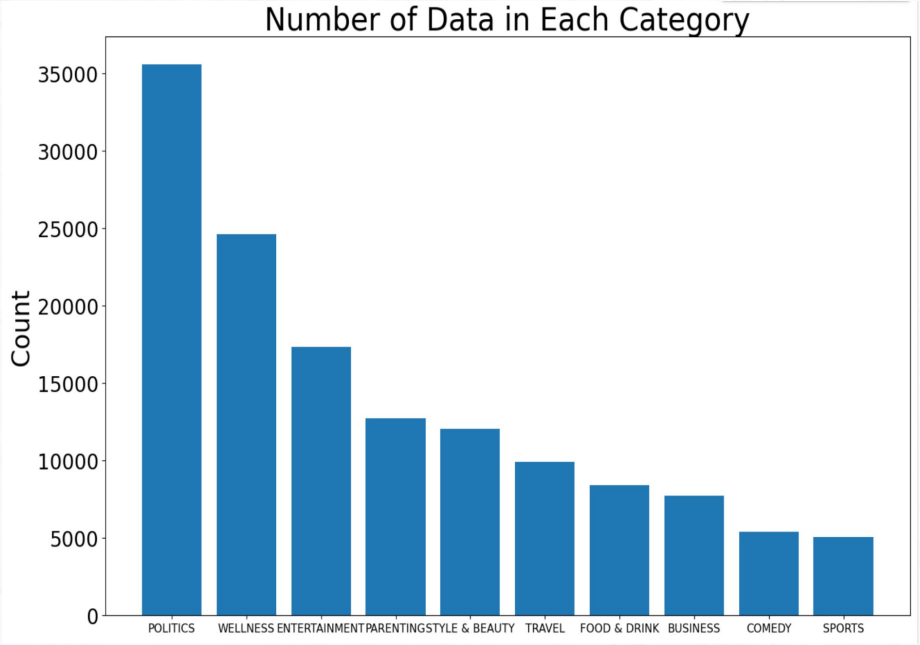
Growth of online communication has given rise to a concerning escalation in the use of hate speech and offensive language, posing a significant challenge in fostering a healthy digital community. Tackling this issue requires a nuanced understanding of textual data.

Objective

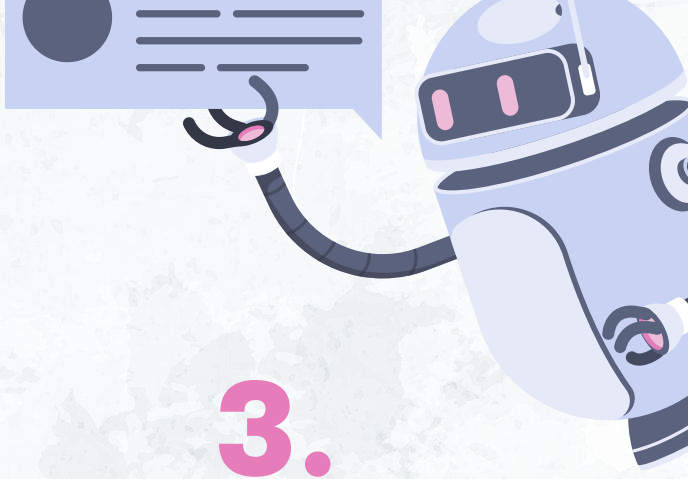
Develop a robust classification system that not only identifies hate speech and offensive language in textual data but also categorizes such instances into relevant topics.

Exploratory Data Analysis (EDA)





Method



1.

Data Acquisition & Preprocess:

- Kaggle Text Data
- Text Cleaning (stopwords removal, spelling corrections, lemmatization, tokenization, etc.)
- Word embedding

2.

Sentiment Analysis:

- Logistic regression
- Bernoulli Naives Bayes
- Multinomial Naive Bayes
- Neural Network

3.

Topic Classification:

- Naives Bayes
- RNN
- LSTM
- Fine-Tuning BERT

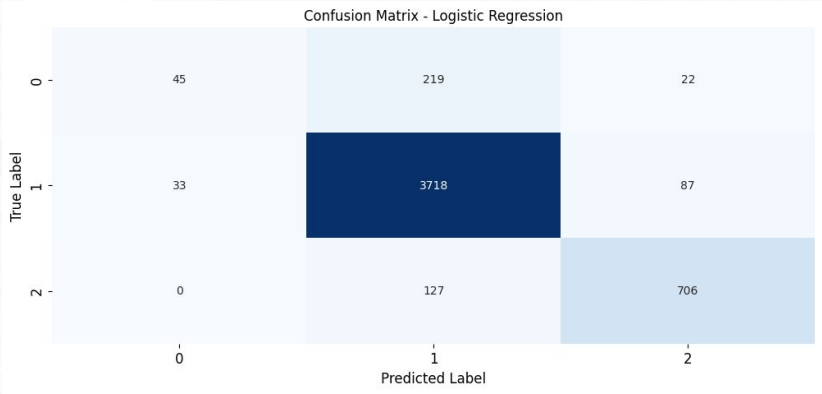
Challenges

- Run time inefficiency
- Type Mismatch: string, lists, vectorization, sparse matrix
- Dataset pertaining to both classification
- Insufficient Data

- **Insufficient Data**
  - Data are distributed unevenly between categories
  - Combine data of similar categories
  - Keep the top 10 categories with most data

Results: Sentiment Analysis

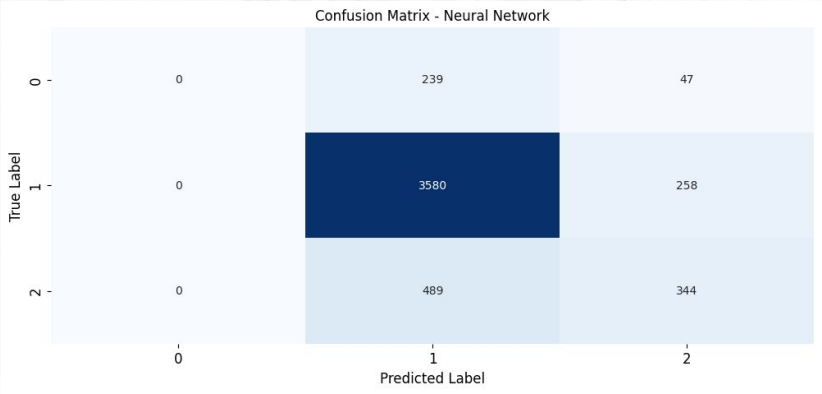
Logistic Regression



- Accuracy Score:

  90.15%

Neural Networks



- Accuracy Score:

  79.16%

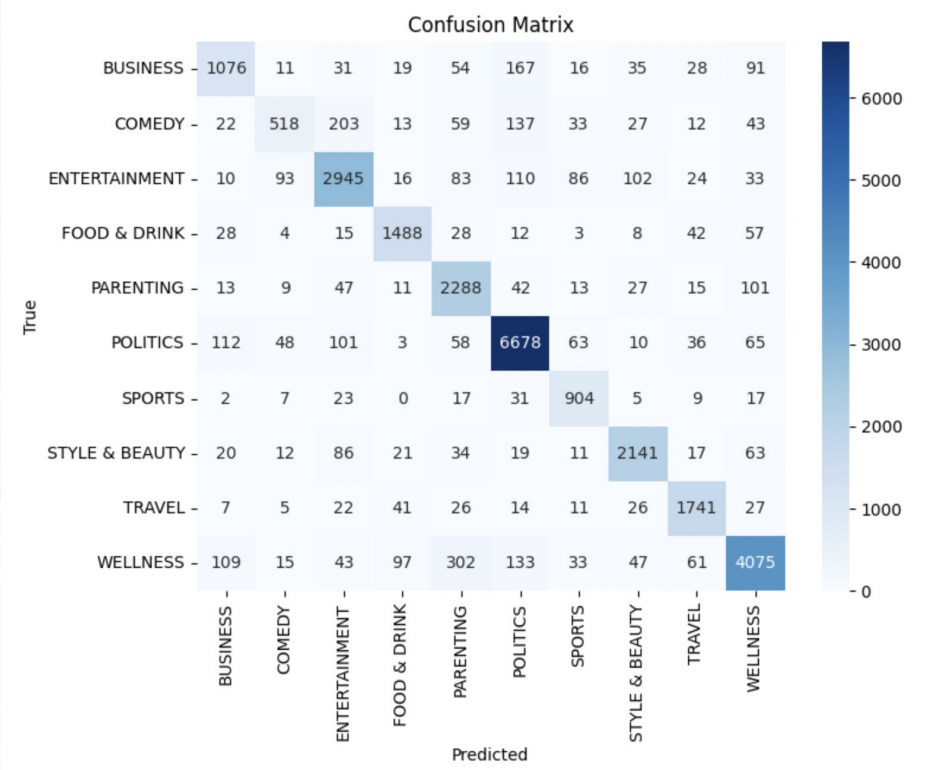
Results: Topic Classification

Fine-Tuning BERT

- Accuracy Score:

  85.82%

	precision	recall	f1-score	support
BUSINESS	0.77	0.70	0.74	1528
COMEDY	0.72	0.49	0.58	1067
ENTERTAINMENT	0.84	0.84	0.84	3502
FOOD & DRINK	0.87	0.88	0.88	1685
PARENTING	0.78	0.89	0.83	2566
POLITICS	0.91	0.93	0.92	7174
SPORTS	0.77	0.89	0.83	1015
STYLE & BEAUTY	0.88	0.88	0.88	2424
TRAVEL	0.88	0.91	0.89	1920
WELLNESS	0.89	0.83	0.86	4915
accuracy			0.86	27796
macro avg	0.83	0.82	0.82	27796
weighted avg	0.86	0.86	0.86	27796



Conclusion

Our Project

- Level of complexity in textual preprocessing for NLP: tokenization, stopword removal, one-hot encoding, etc.

Bigger Picture:

- Importance of not only identifying negative speech but also recognizing the relevant topic it pertains to. Reveals the diverse landscape these languages can permeate highlighting the intricate layers of online communities.

Fun Fact →

Researchers are starting to consider emoji as a part of textual analysis as it can provide more nuanced understanding of online communication.

😍: might indicate positive language

🤬: might indicate negative language