Sarcasm Detection Using Natural Language Processing

By Adam Pell

Outline

Business Problem

Data Understanding

Methodology

Results

Recommendations

Sarcasm is Hard

Tone and context dependent

Sarcasm relies heavily on tone of voice and context, which can be difficult to pick up in written text.

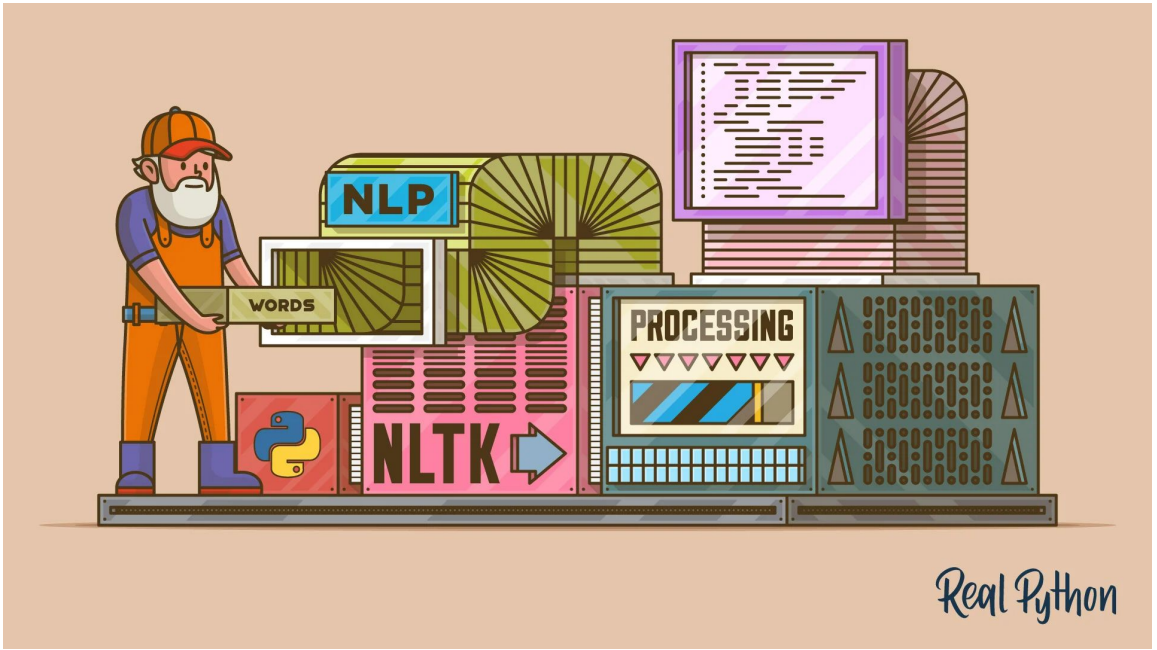
Requires understanding nuance

Detecting sarcasm requires understanding subtle nuances in language and meaning.

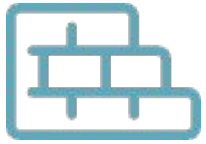
Difficult to detect in written text

Without vocal cues , sarcasm can be very difficult to identify in written communication.

The Solution: Natural Language Processing



Applications



Bullying Detection

Differentiate harmful behavior, resulting in bans and suspensions

Inappropriate Behavior

Detect genuine threats or concerning language

Algorithm Enhancement

Deliver similar content to people who interact heavily with sarcastic posts

Understanding sarcasm opens up a world of possibilities

Dataset Overview

- 1m+ entries
- Comments and parent posts
- Labeled for sarcasm
- Balanced

Methodology

Supervised Learning

Unsupervised Learning

Target Metric: **Recall**

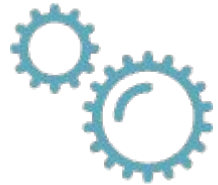
Results: Final Model Key Trends

Most Consistent

Least variation between precision and recall, highest accuracy

Best Sarcasm Detection

Achieves the highest recall for Class 1 (sarcasm)



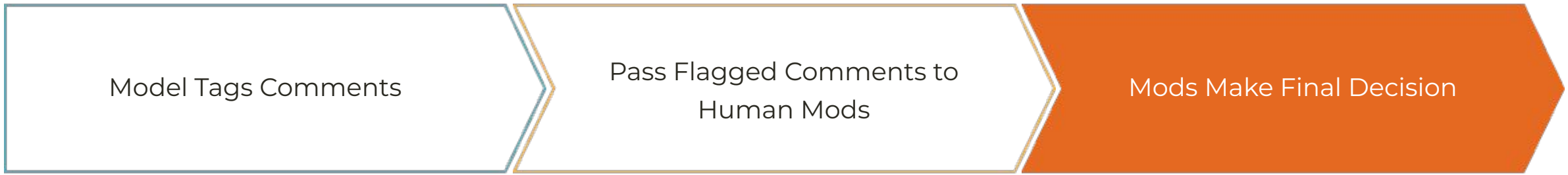
Further Training Needed

Sarcasm detection still proves difficult

Model Performance Compared

Model	Recall (Class 0)	Recall (Class 1)	Accuracy	Time to Run
Naive Bayes	74%	55%	64%	>1s
XGBoost	94%	21%	58%	5min.
Neural Network	75%	57%	66%	21min.
Recurrent Neural Network	91%	16%	54%	31min.

Recommendation: A Hybrid Approach



Our model could tag messages as sarcastic or not, with speed and scale

The model will flag problematic content to moderators

Moderators look at flagged text and decide whether to take action

Next Steps

- Data quality and quantity
  - Better labeling practices
- Transfer learning
  - Fine-tune using large, pre-trained model
- Optimize models
  - Longer, more intensive training and experimentation

Thank You!

Contact Information

- **Email:** apell7591@gmail.com
- **Phone:** 917-434-6615
- **Github:** apell74
- **LinkedIn:** Adam Pell

Relevant Links and Sources

- Project Repository
- Google Colab Notebook
- Kaggle Dataset
- Further Reading