

# Twitter Text Classification



Jose Yamas

# **Project Objective:**

**Can we predict gender from a tweet?**

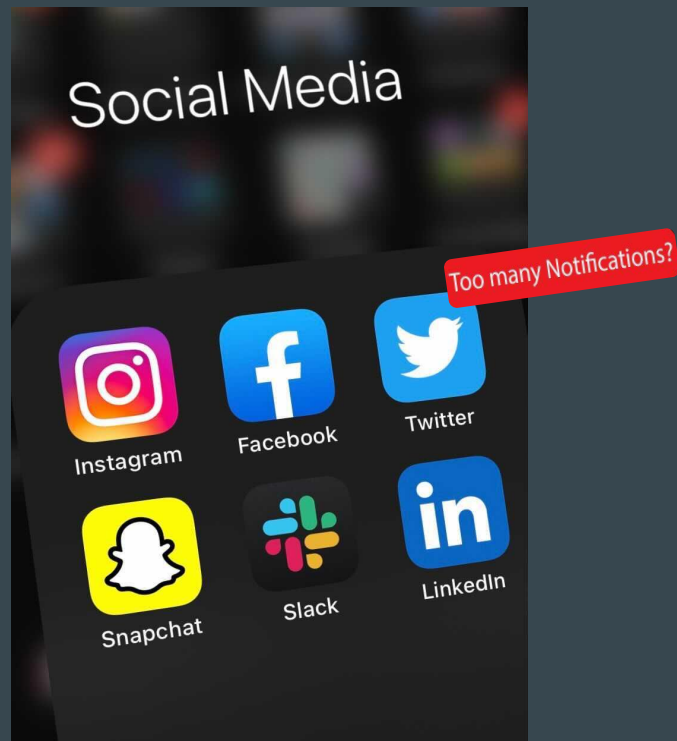
# Overview

## Why Twitter?

- Twitter is Unique

## Why Should We Care?

- Imbalanced Data
  - Brand: Female: Male: Unknown Ratio
  - About 1:1:1:0.25 Ratio
- Forensics Use



# Tech Stack Tools



# Text Classification Process

## Clean Data

- Remove:
  - Punctuation
  - @'s and Links
  - Numbers
  - Whitespaces\_
  - Emojis 😊
  - Stop Words
- Lower Case Text

## NLP Pipeline

- Train/Test Split
- CountVectorizer()
- TfidfTransformer()
- RandomForestClassifier()
  - Fit Model
  - Predict
- Calculate Performance
  - Analyze Performance
- *Tune Model\**

## Visualization

- Confusion Matrix
  - Predicted/Actual
- Seaborn Heatmap
- Visualize Feature Importance

# Understanding the Data

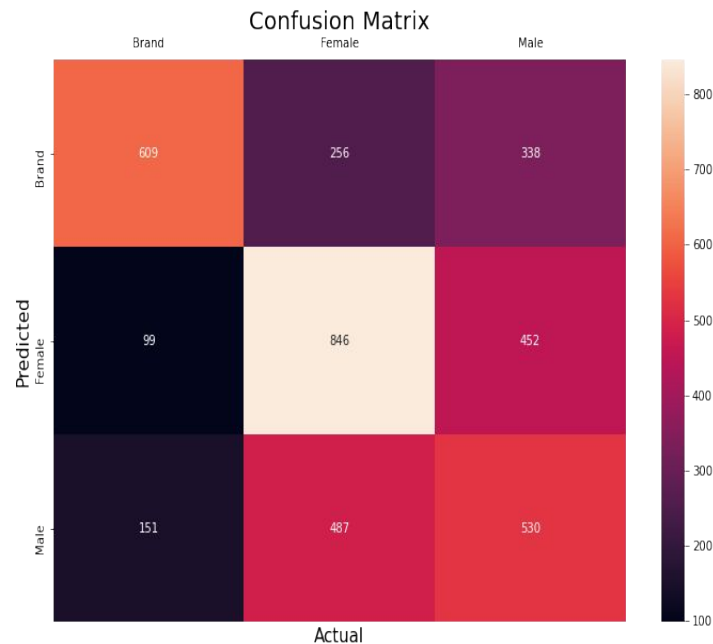
# Data

```
Accuracy = 0.5262738853503185
          precision    recall  f1-score   support

     0       0.70      0.50      0.58      1203
     1       0.53      0.61      0.57      1397
     2       0.40      0.45      0.43      1168

 accuracy                   0.53      3768
 macro avg              0.55      0.52      0.53      3768
 weighted avg           0.55      0.53      0.53      3768
```

0 - 'Brand', 1 - 'Female', 2 - 'Male'



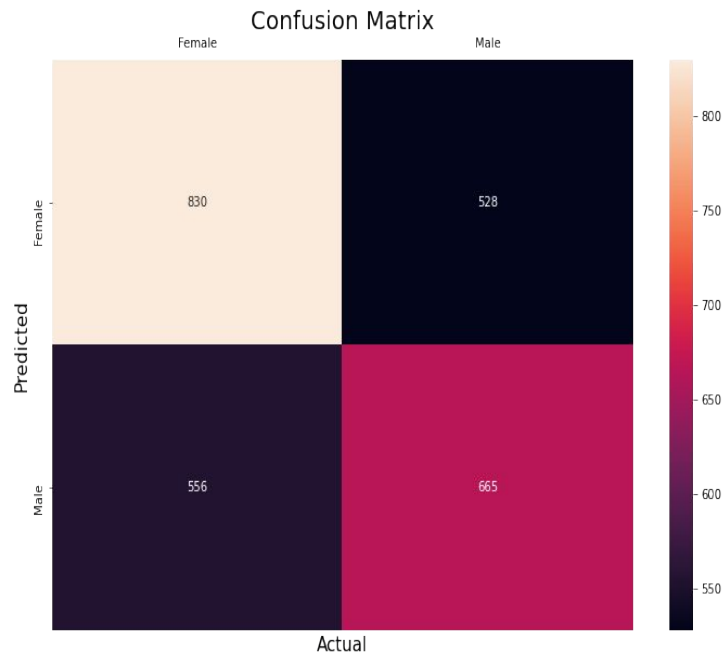
# Data

```
Accuracy = 0.5800697944939899
          precision    recall  f1-score   support

     0       0.60       0.61       0.60       1358
     1       0.56       0.55       0.55       1221

 accuracy                   0.58       2579
 macro avg       0.58       0.58       0.58       2579
 weighted avg    0.58       0.58       0.58       2579
```

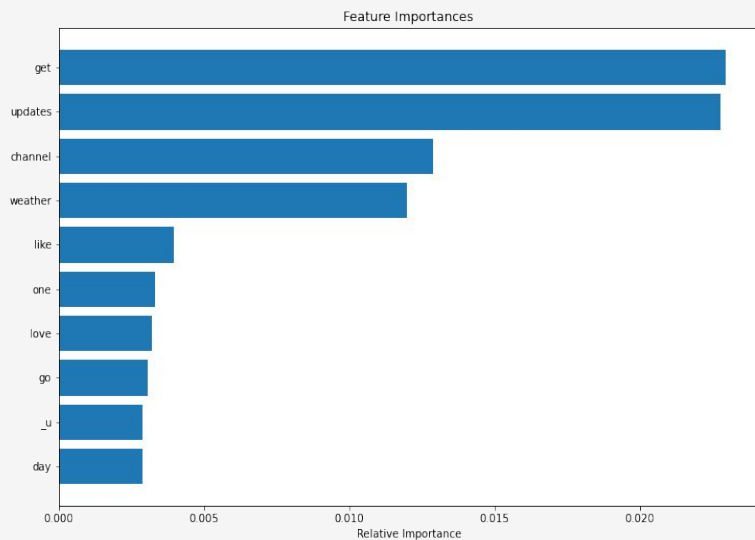
0 - 'Female', 1 - 'Male'



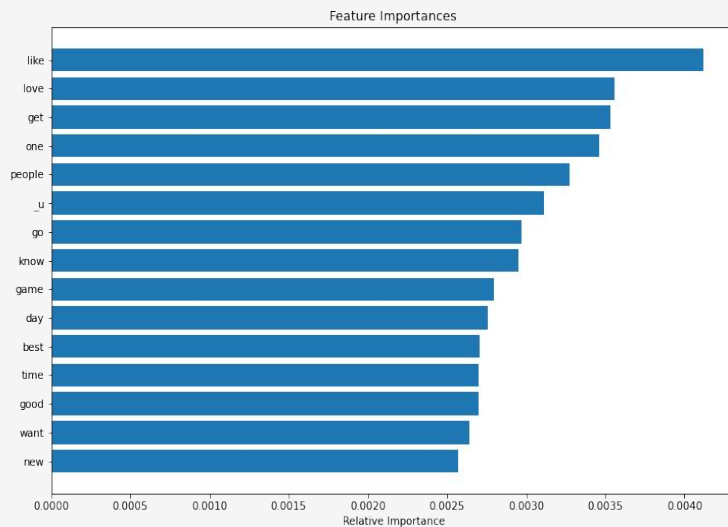


# Feature Importance

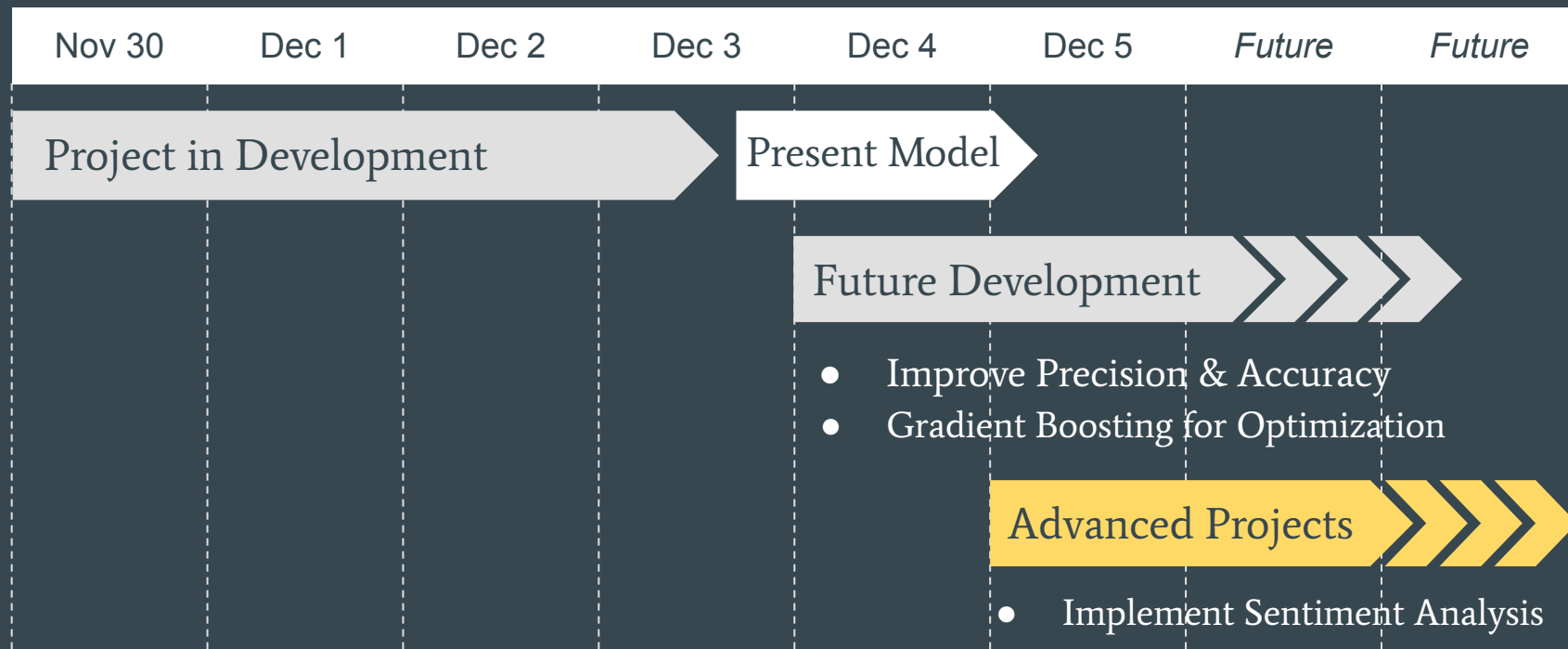
Data: ['Brand', 'Female', 'Male']



Data: ['Female', 'Male']



# Project Timeline - Future Development



# Thank You!



## Jose Yamas

Email

---

yamasjosell@gmail.com

Github

---

<https://github.com/yamasjosell>

LinkedIn

---

[https://www.linkedin.com/in/  
jose-yamas/](https://www.linkedin.com/in/jose-yamas/)