

AFYAMIND EMOTION DETECTION MODEL

By: Neural Ninjas

MEET THE TEAM



Simon Irungu
Scrum master
Deployment



John Nyangoya
Tableau



Eugene Mallah
Dataset + EDA



Bob Lewis
Data Report



Sophie Muchiri
Notebook
Slides

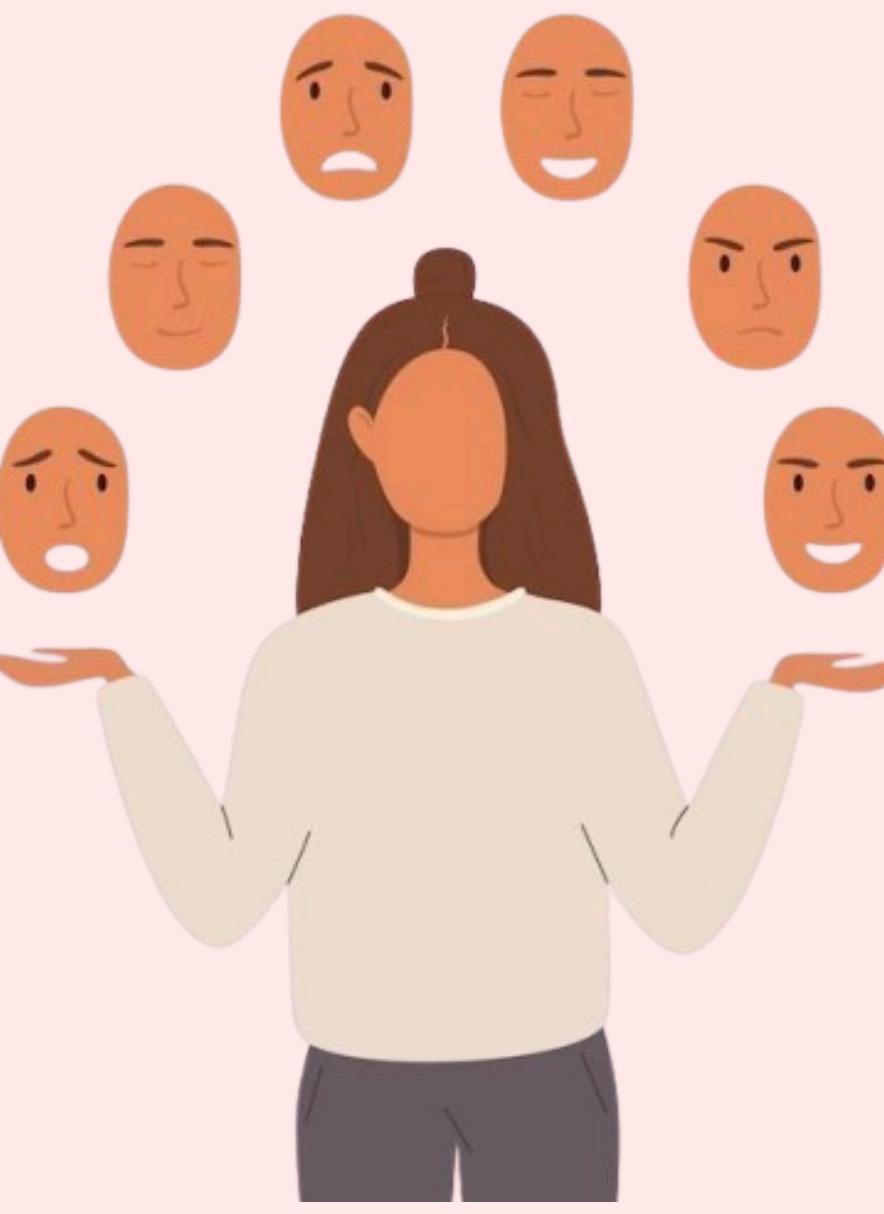


INTRODUCTION

In the digital era, understanding emotions expressed in text is essential for enhancing user interactions, improving customer support, and analyzing public sentiment.

This project aims to create a model capable of classifying user text into 6 emotion categories or Neutral using the GoEmotions dataset.

This technology offers a powerful tool for businesses to anticipate issues, personalize responses, and drive customer satisfaction in the competitive and fast-growing gig economy.



OBJECTIVES

Main Objective: To build and evaluate a machine learning model that can automatically classify text into emotional categories using the GoEmotions dataset.

Build and compare baseline models with deep learning models



Deploy the model on Streamlit and integrate it with an existing model to provide intervention.

Explore and understand the distribution of emotions in the GoEmotions dataset.

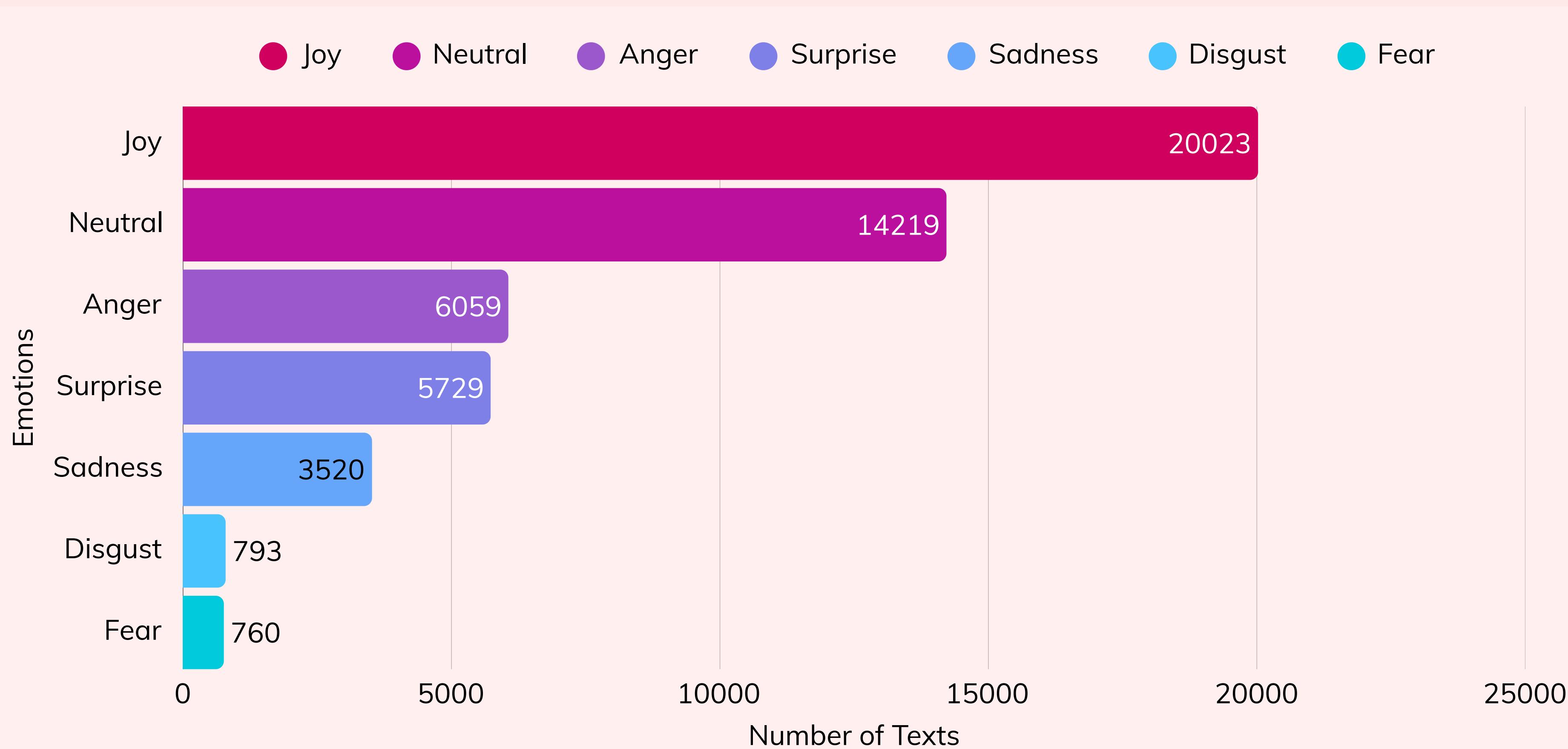
Evaluate models using metrics appropriate for multi-label classification

DATA UNDERSTANDING

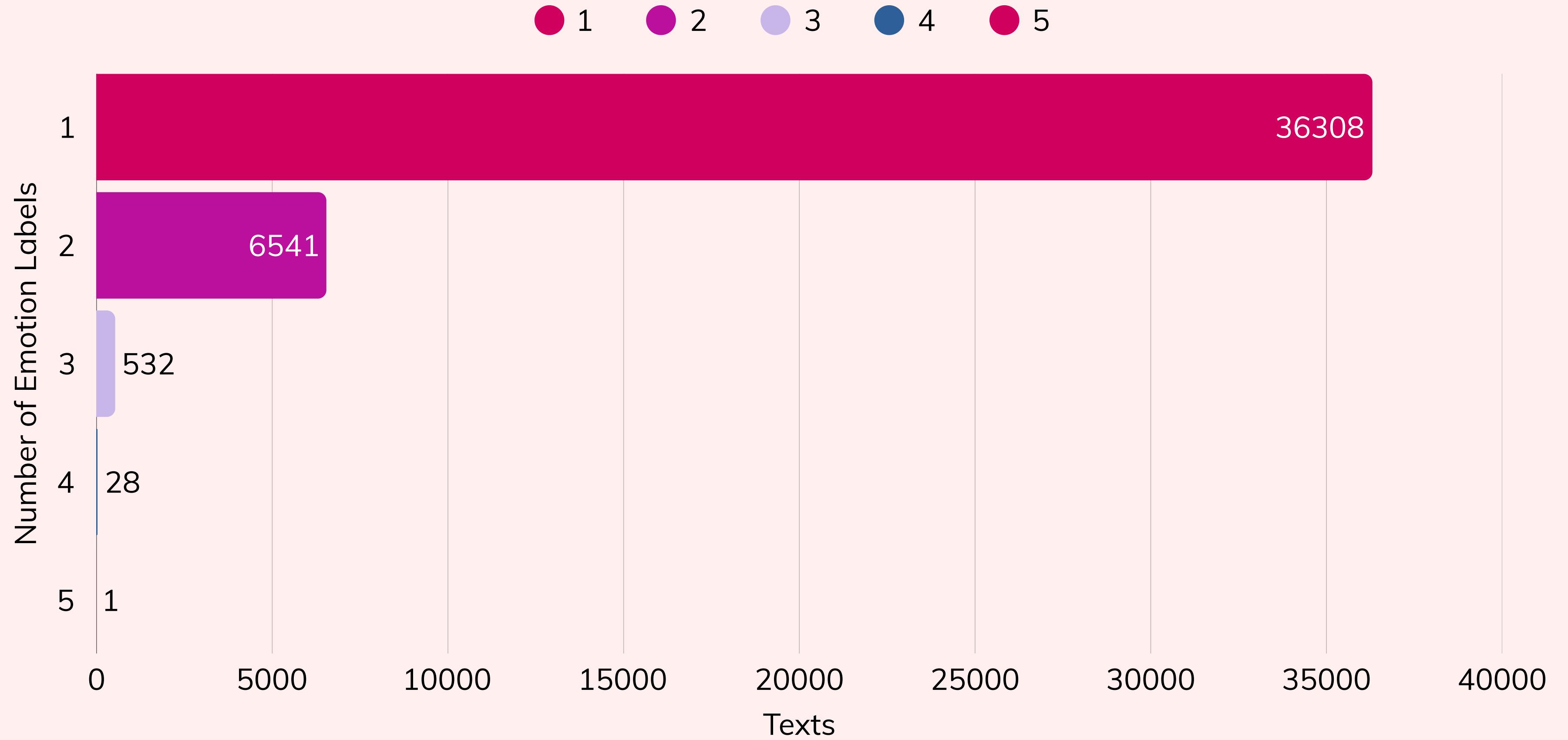
- The dataset had 27 emotions (eg: amusement, realization)
- Using Ekman's theory, we were able to map them into the basic, universal emotions.
- Namely: joy, disgust, fear, anger, sadness, surprise and neutral



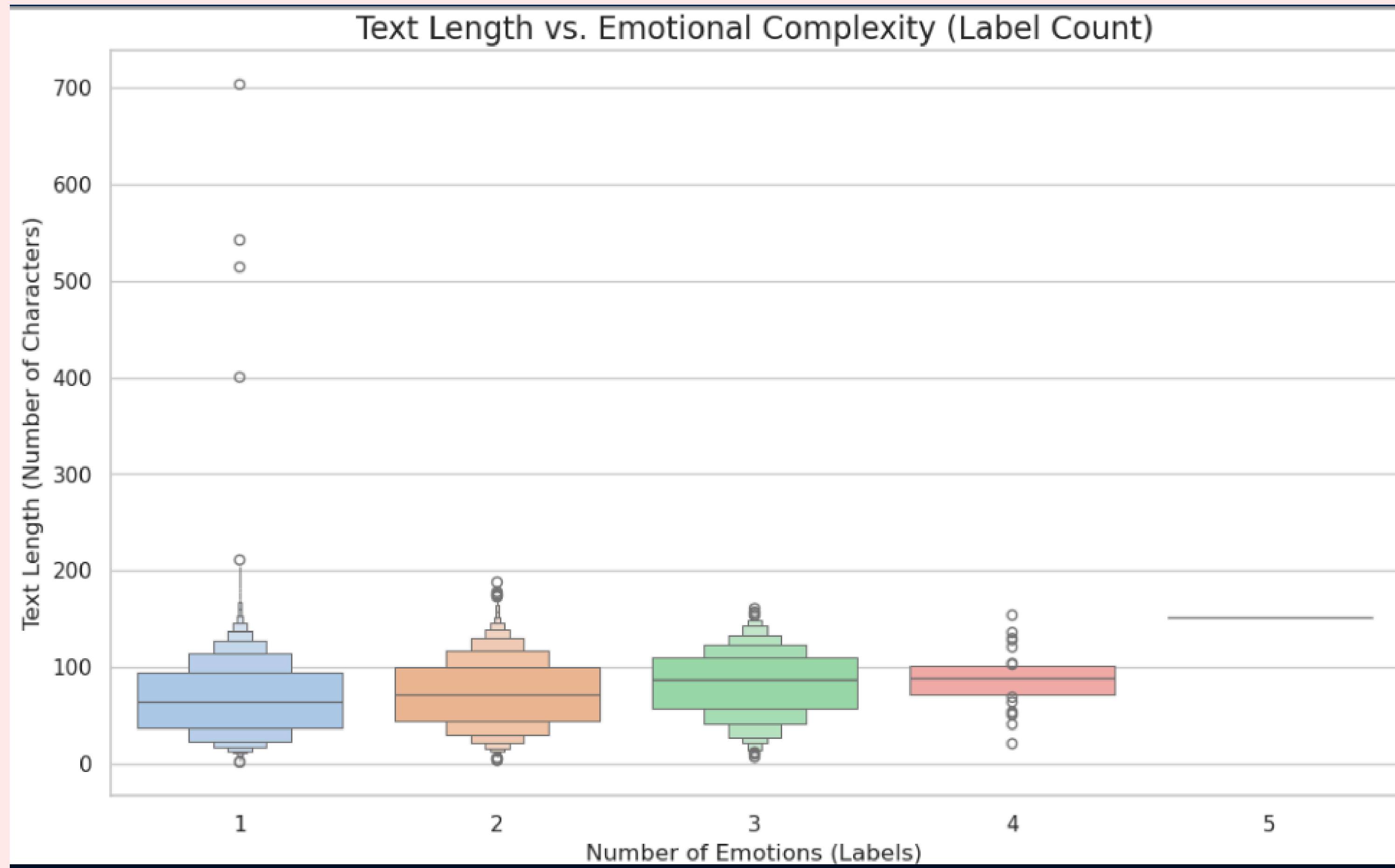
EMOTION DISTRIBUTION



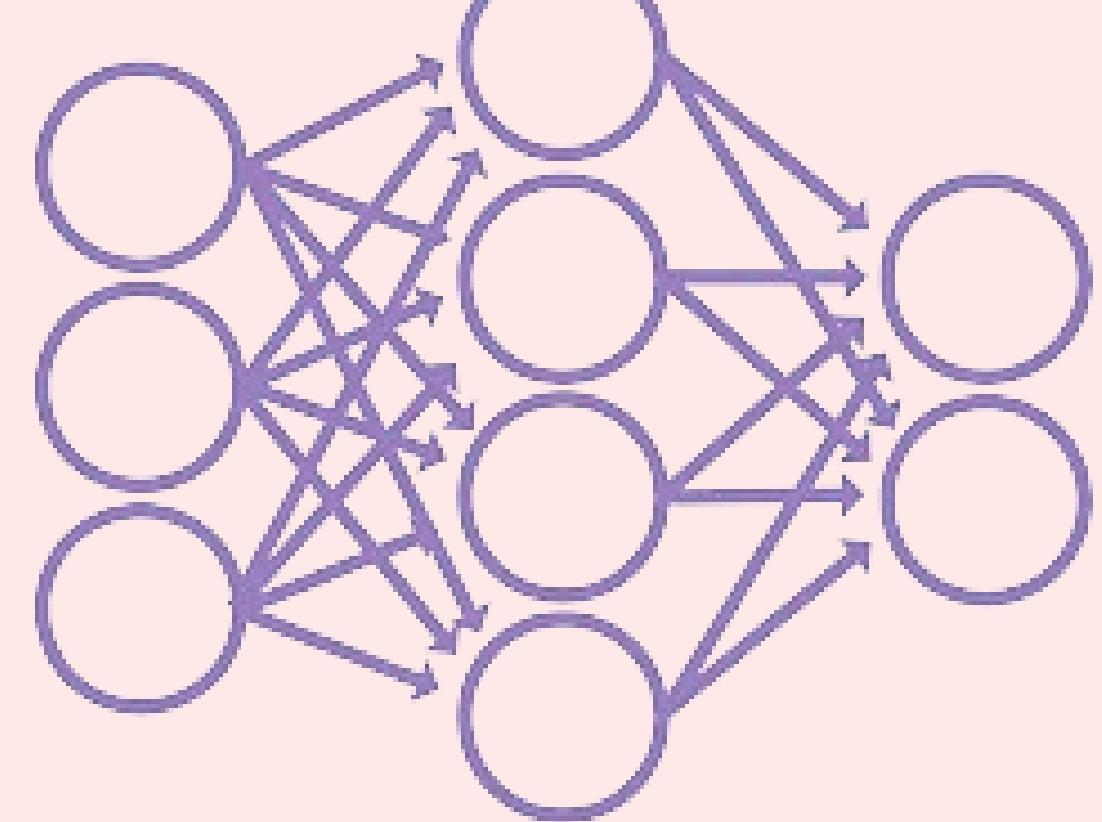
DISTRIBUTION OF TEXTS VS NO OF EMOTIONS



TEXT LENGTH VS EMOTION COMPLEXITY



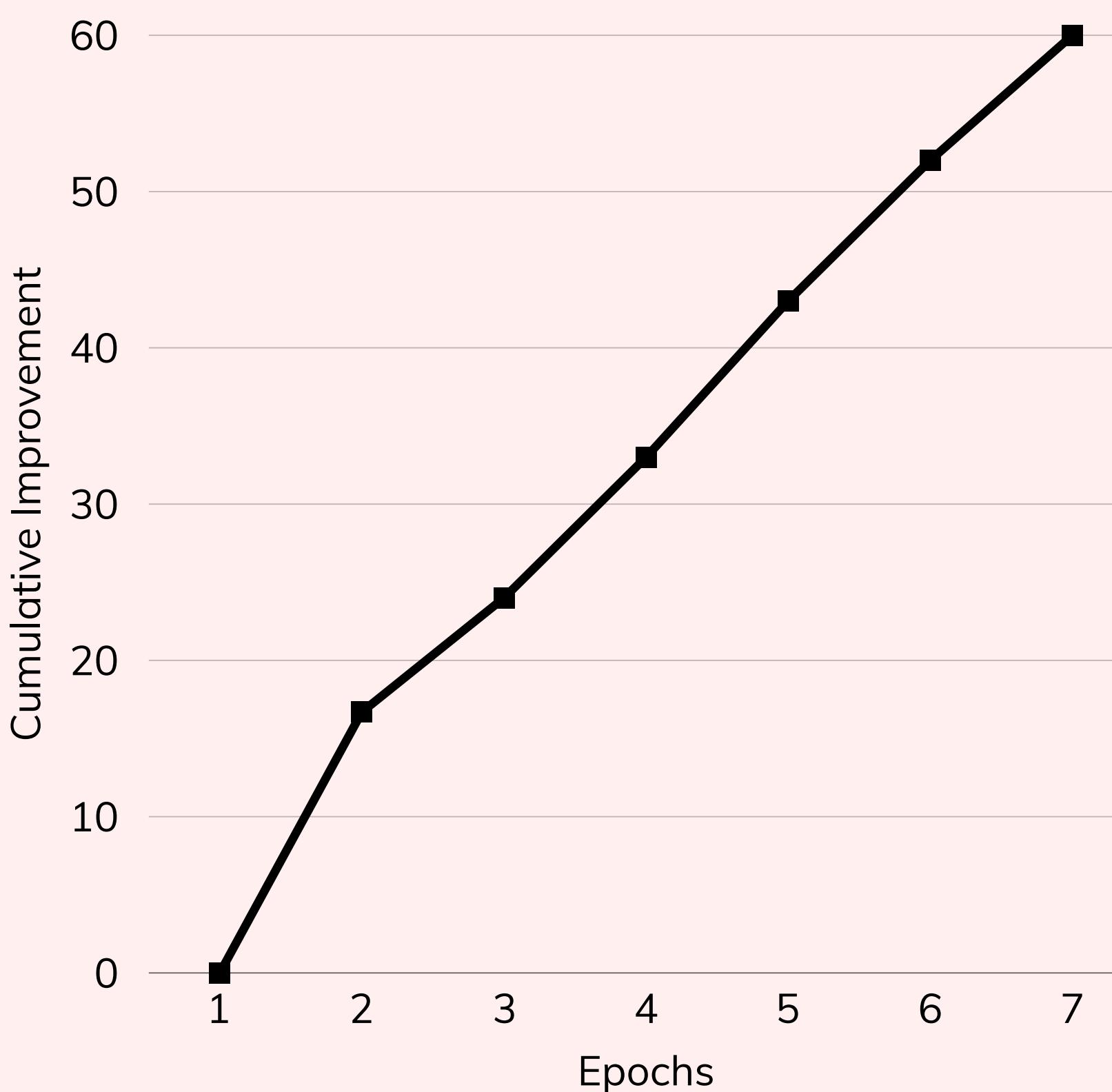
MODELLING



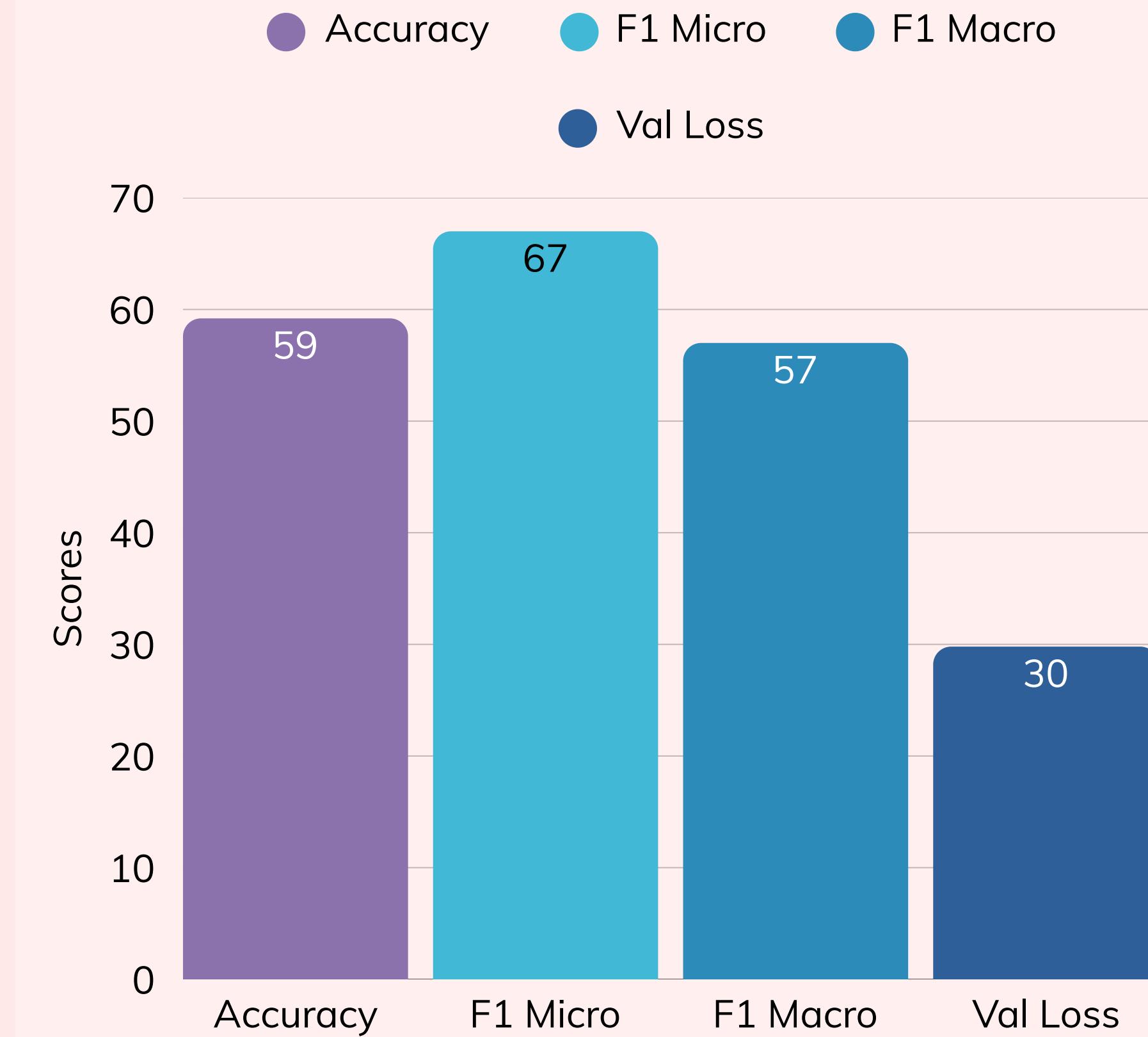
- After preprocessing the text data, we trained the Traditional models: Logistic Regression, Multinomial NB and Linear SVC.
- We used F1 score as our main metric of success
- Linear SVC performed best even after applying class weights with an F1 score of: 0.5530
- We then implemented Deep learning model: Roberta which achieved an F1(micro) of 0.66

METRICS OF THE BEST MODEL

Cumulative Improvement Over Time



Final Model Metrics



DEPLOYMENT

We deployed our model using Streamlit

Our model can analyze single and multiple texts into different emotional categories.

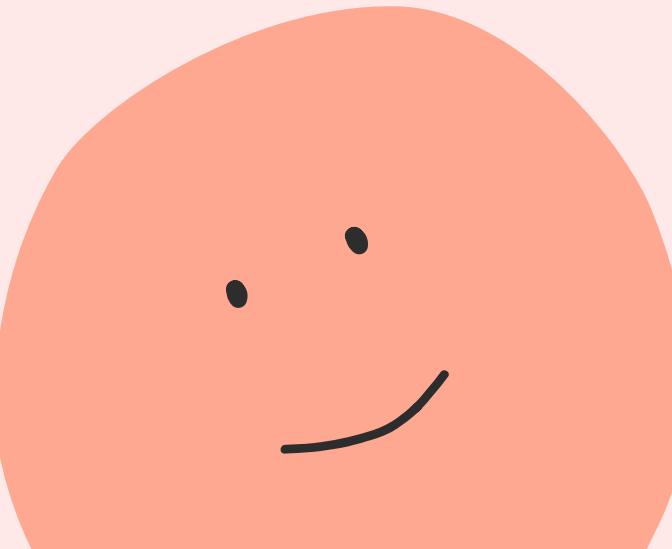
Our model can also take in CSV files with the text column and give an emotion for each of the entries. A user can also save it as a CSV file for future individual analysis.





CONCLUSION

- By using the mapped Ekman emotions a business can move past generic "positive/negative" scoring to understand the quality of customer feeling.
- The high frequency of Neutral comments suggests the system effectively identifies purely informational transactions.
- Most text express a single dominant emotion, allowing businesses to focus on one primary emotion per comment without complex multi-label analysis.
- The Afyamind model was successfully integrated with Deepseek's LLM to create an emotion sensitive chatbot.



NEXT STEPS



- Improve Model Robustness: Focus on techniques to boost the performance of minority classes
- Dataset Enhancement: Future work should consider improving the overall quality of the dataset, perhaps by focusing on samples that express more complex or subtle emotional language.
- Establish a feedback loop by anonymously logging text inputs and model predictions in Streamlit. This real-world data helps uncover weaknesses missed during validation.



THANK YOU!