

# Brands Emotions Judge





# INTRODUCTION

- ❖ A system that can judge the emotions expressed in tweets would have numerous advantages and potential uses to companies.
- ❖ Companies understand customer preferences and emotions related to different products or features impacting product development and innovation.
- ❖ The system would give insights of how the companies are doing against their competitors in the eyes of the public therefore assisting marketing strategies.
- ❖ In addition, it would monitor and assess the emotional reactions of the public during a crisis.



# PROBLEM STATEMENT

- ❖ Social media contains a lot of unstructured data availed by members of the public freely. This data contains important information.
- ❖ On the other hand, companies require lots of information to help them make critical decisions about their products and services.
- ❖ Our system is looking in to capturing information about people's emotion on companies' products and services. It looks into making the information on social media platform useful and available to companies benefiting both the social media and target companies.



# DATA UNDERSTANDING

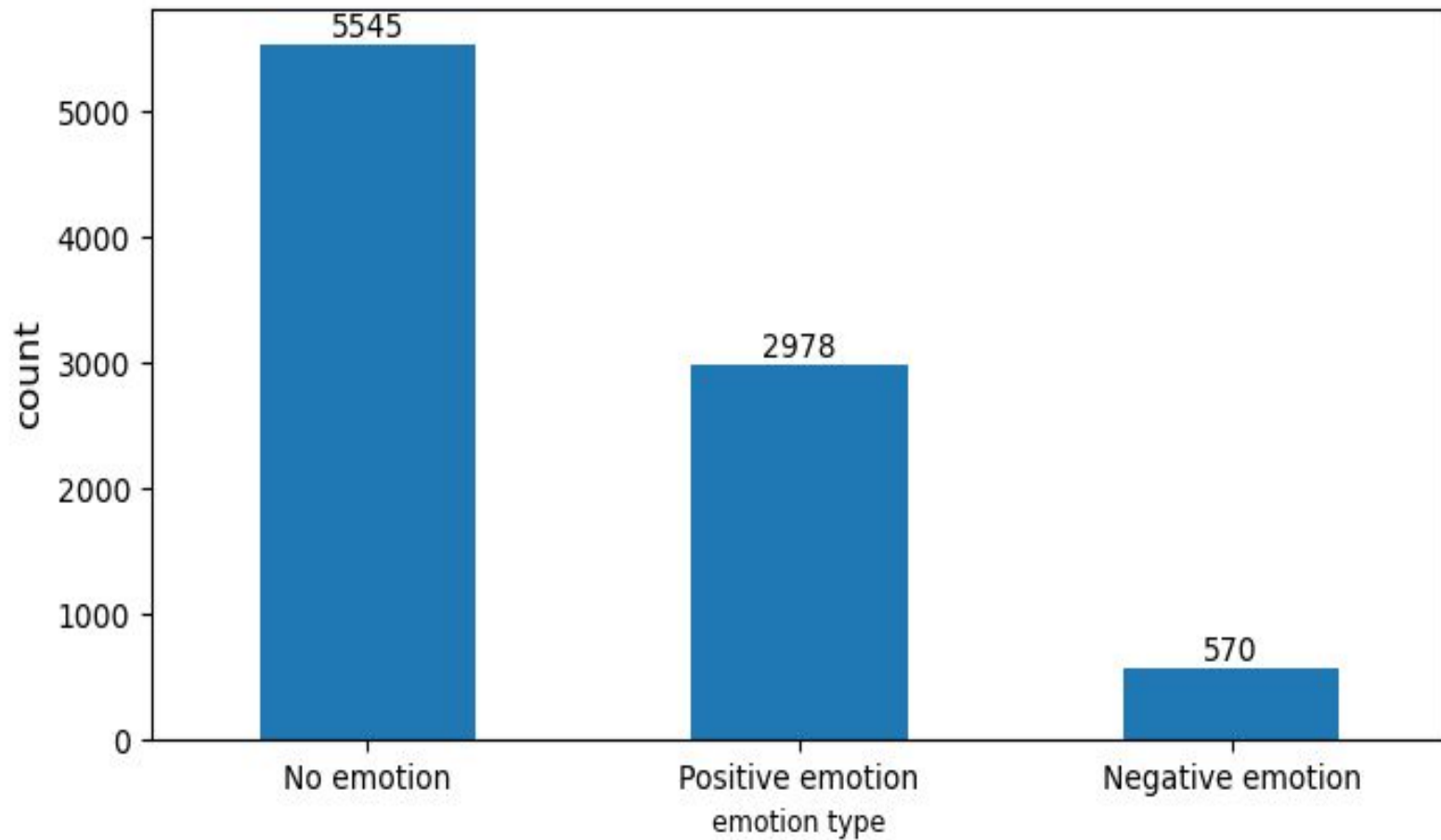
- ❖ The data used in this project was sourced from Brands and Product Emotions.
- ❖ This data was made by giving tweets to a crowd and they gave feedback on what emotion was expressed on the tweet, either positive, negative or no emotions towards the brand or product. If some emotion was expressed they were also asked to give which brand or product was the target of that emotion.
- ❖ The data comprise of 3 columns and 9094 rows. The rows represent different tweets.



# DATA PREPARATION

- ❖ "I can't tell" and 'No emotion toward brand or product' were merged to form a single column as they looked the same
- ❖ Row with missing value on tweet\_text column were removed as they would not make sense to this project
- ❖ Our target column was label encoded to numeric figures .
- ❖ The text on tweets was tokenized and transformed to numeric form for processing
- ❖ The data was separated into train and test data.

Emotion count on tweets

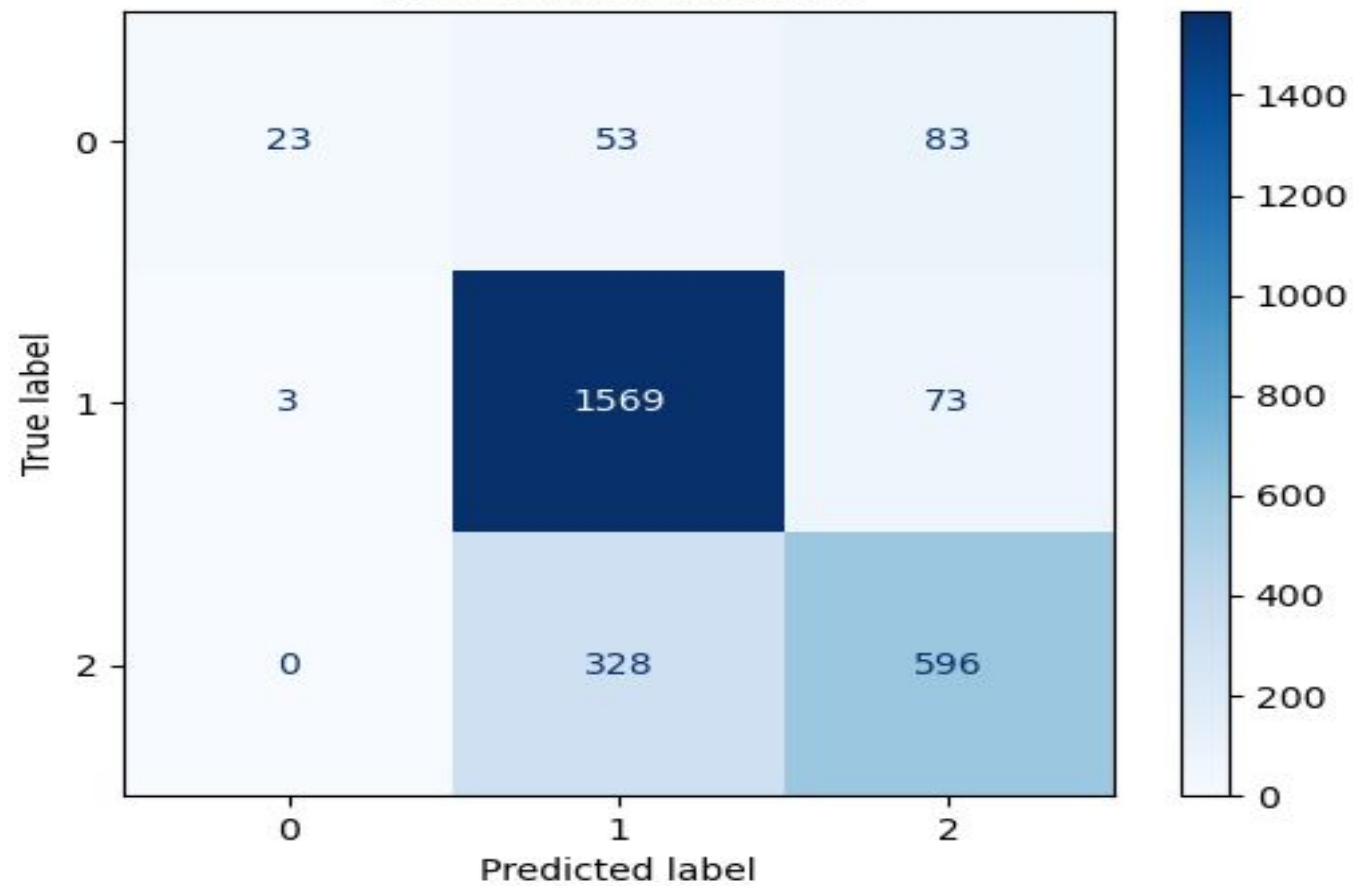




# MODELING

- ❖ Random forest classifier model was trained on the training data and tested on the test sets.
- ❖ Our main aim was to check for accuracy of the model on the test set.
- ❖ The other objective was that the model was supposed to detect negative emotions which were the minority in our data.
- ❖ The first model did not perform to our expectations and therefore need for improvement.

first Confusion Matrix



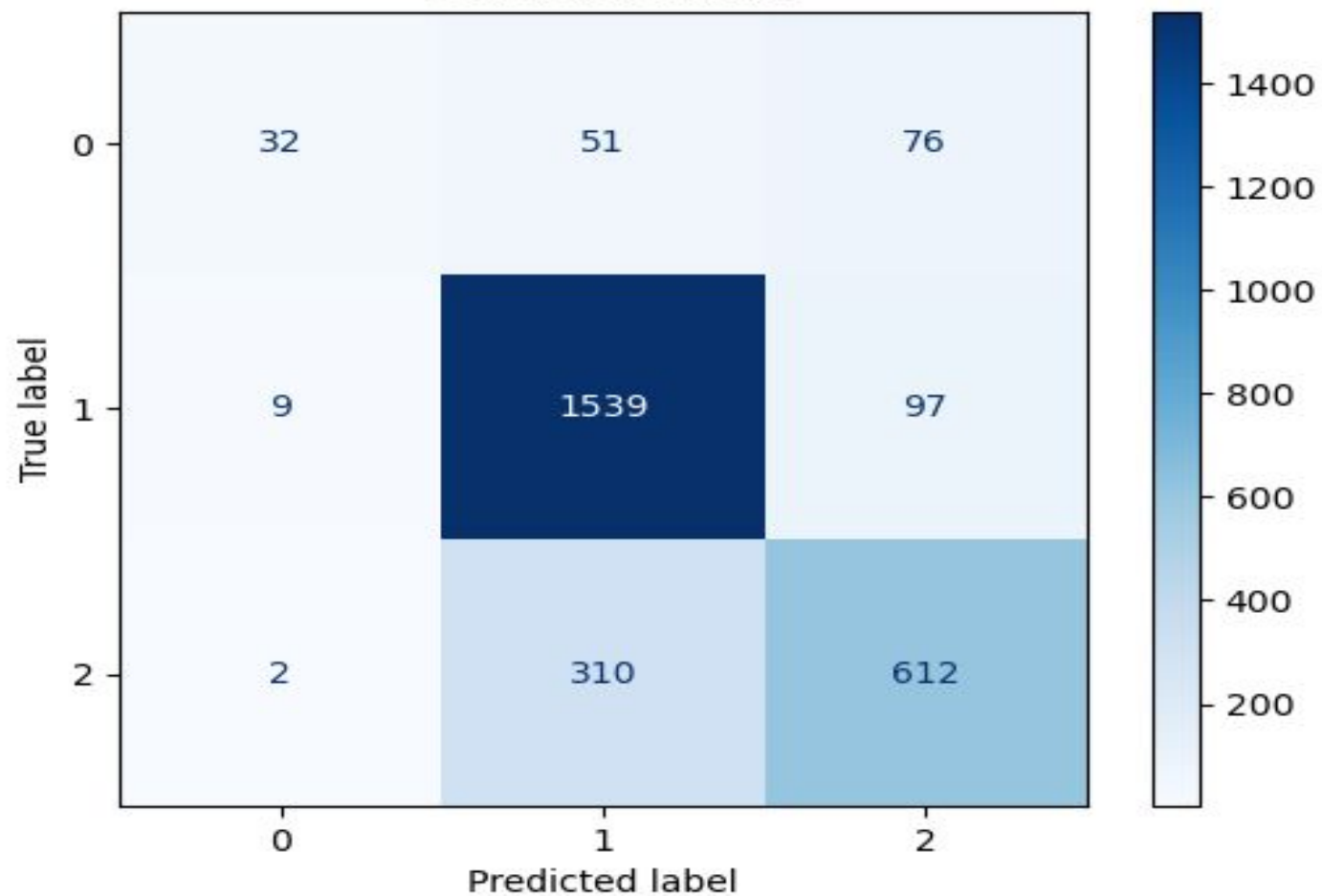




# Model improvement

- ❖ The first model showed signs of overfitting which leads to poor predicting results.
- ❖ It also showed poor detection of negative feelings. This was as a result of unproportional training data set.
- ❖ To improve the model the training data was improved by oversampling and the new data was used to train the model.
- ❖ Hyper parameter tuning was also done to improve our model predictions

Confusion Matrix





# Conclusions

- ❖ The final model has an accuracy of around 80% on test data which is fairly good and can have many several uses.
- ❖ The model is still biased on capturing negative emotions even after improvement.
- ❖ The model still can't be reliable on sensitive task such as customer support as it is still poor on detecting negative emotions and a score 80% accuracy is still low for sensitive cases.



# Recommendations

- ❖ It is recommended that our model should be incorporate in tech companies' system to help in detecting public emotions
- ❖ We recommend that the model should not be used for sensitive feeling cases such as customer support as it is having a lower power in detecting negative feelings.
- ❖ More data is required for further training of the model to improve its accuracy and detection of negative feelings.