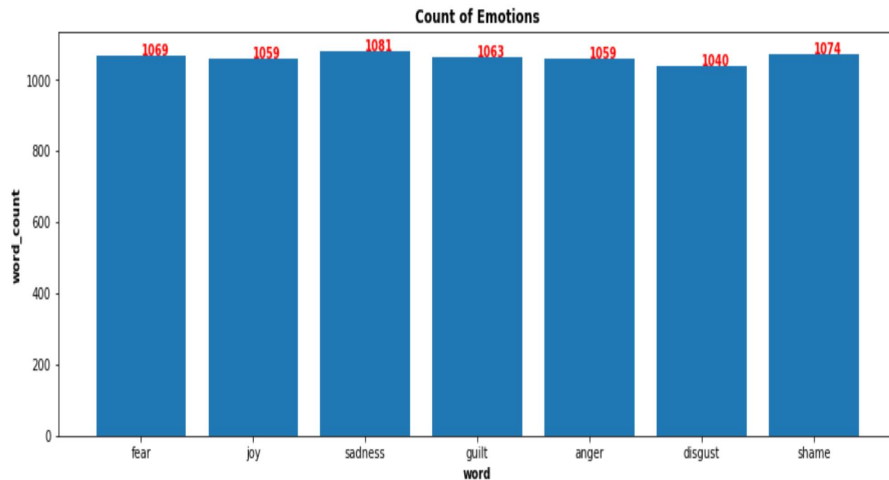# Emotion Classification
Final Progress Report
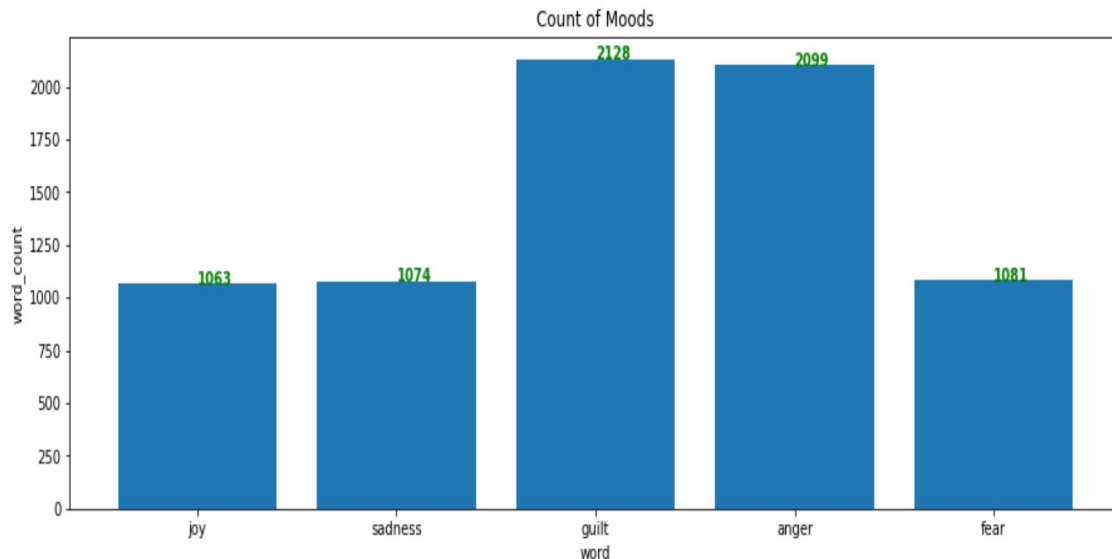
DataSet Overview: Part 1

➢ **Data Format** - Text

➢ **Emotion Categories** - Seven

   (Fear, joy, sadness, guilt, anger, disgust, shame)

➢ **Label Distribution** - Nearly equal number in each label
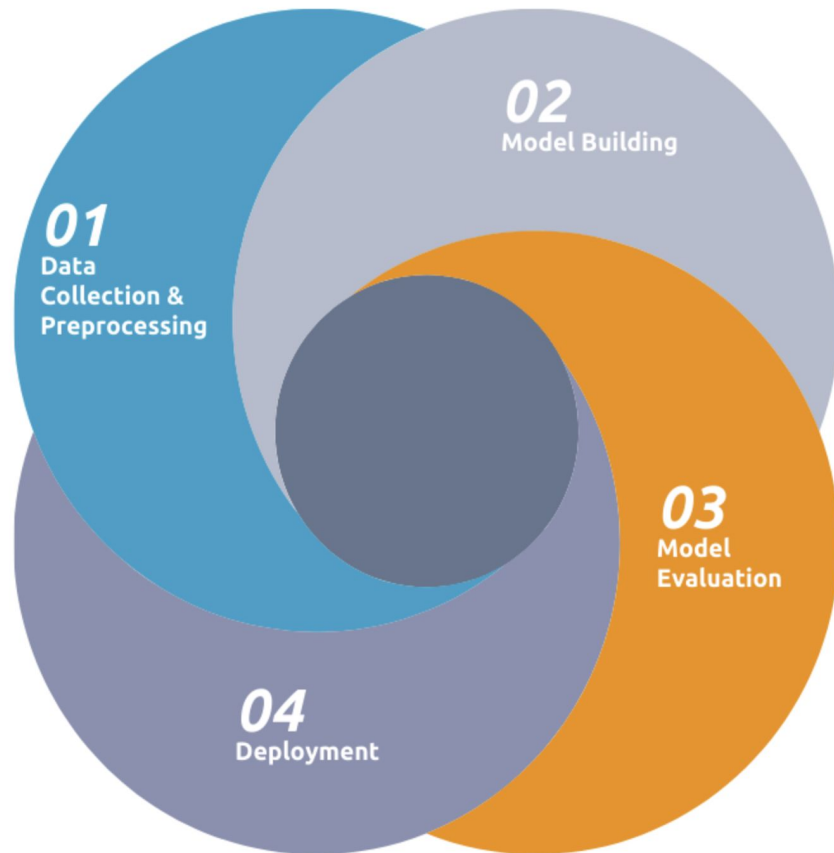

Count of Emotions

Dataset Overview: **Part 2**

➢ **Data Format** - Text

➢ **Emotion Categories** - *Five*

- fear
- joy
- sadness
- guilt/shame - guilt
- anger/disgust - anger

Dataset Overview: Word Cloud

Fig: Words in Dataset

Workflow of System

➢ **Data Collection & Pre-processing**(assembling

data, tokenization, lemmatization)

➢ **Model Building**(Classification algorithm)

➢ **Model Evaluation**(Confusion Matrix, Accuracy)

➢ **Deployment**(Flask, MongoDB)

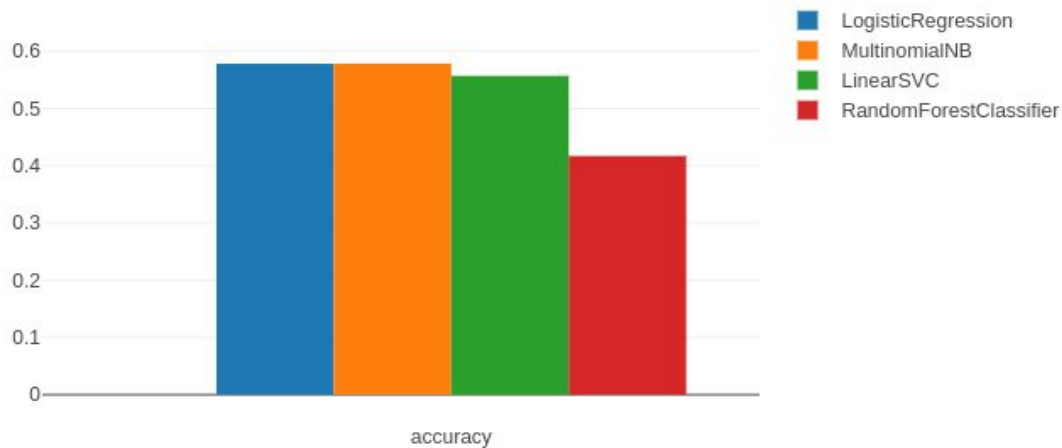# Preprocessing & Feature Extraction

➢ **Text Preprocessing:**

  ○ Lower case

  ○ Tokenize the sentence

  ○ Remove stop words
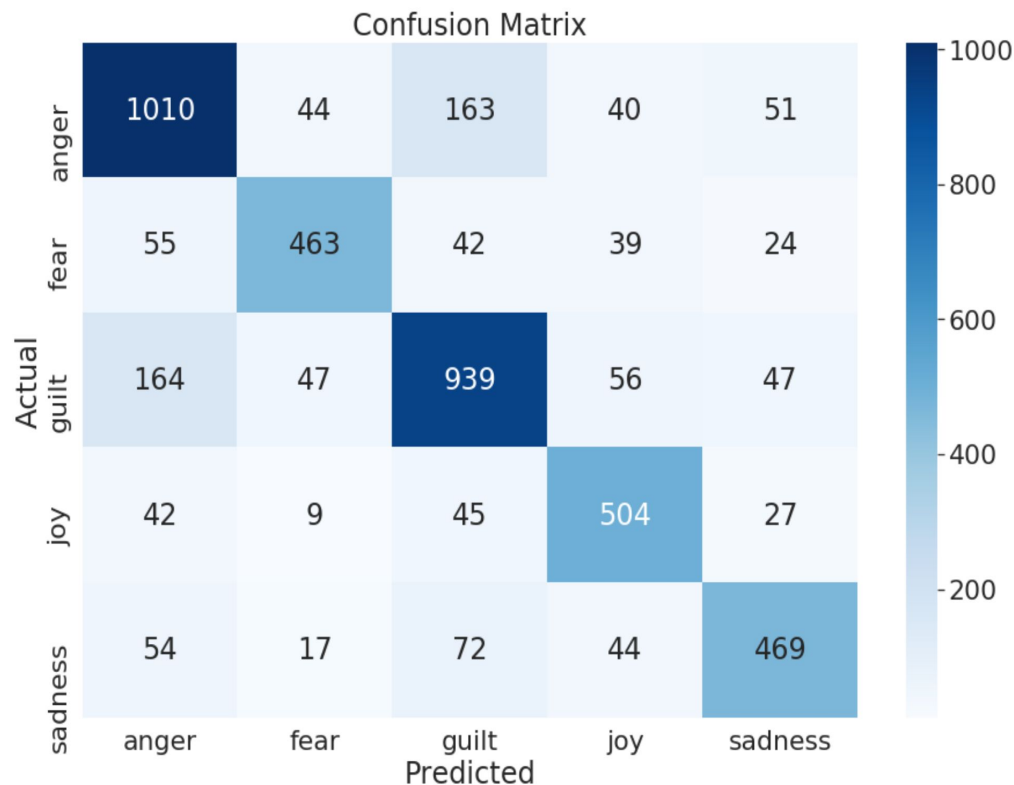
  ○ Lemmatization

➢ **Feature Extraction**

  ○ Countvectorizer

Model Building

➢ Multinomial Naive Bayes
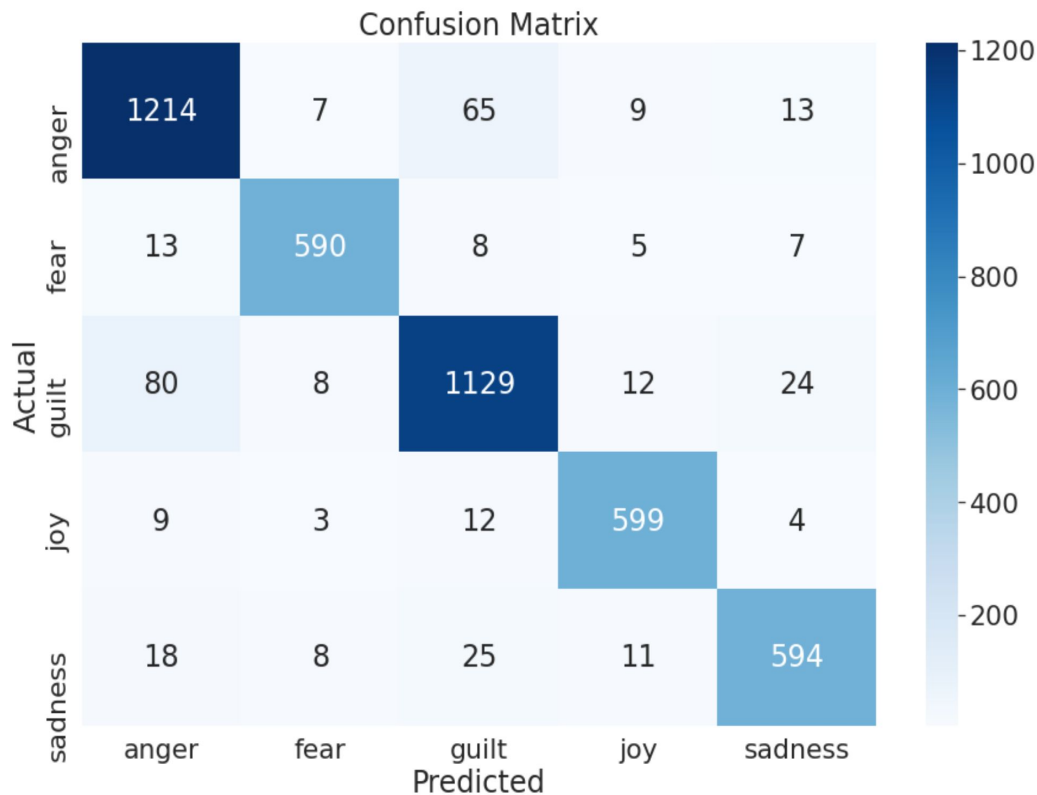
➢ Logistic Regression

➢ Linear SVM

➢ Random Forest

Model Evaluation : Confusion Matrix (SVM)
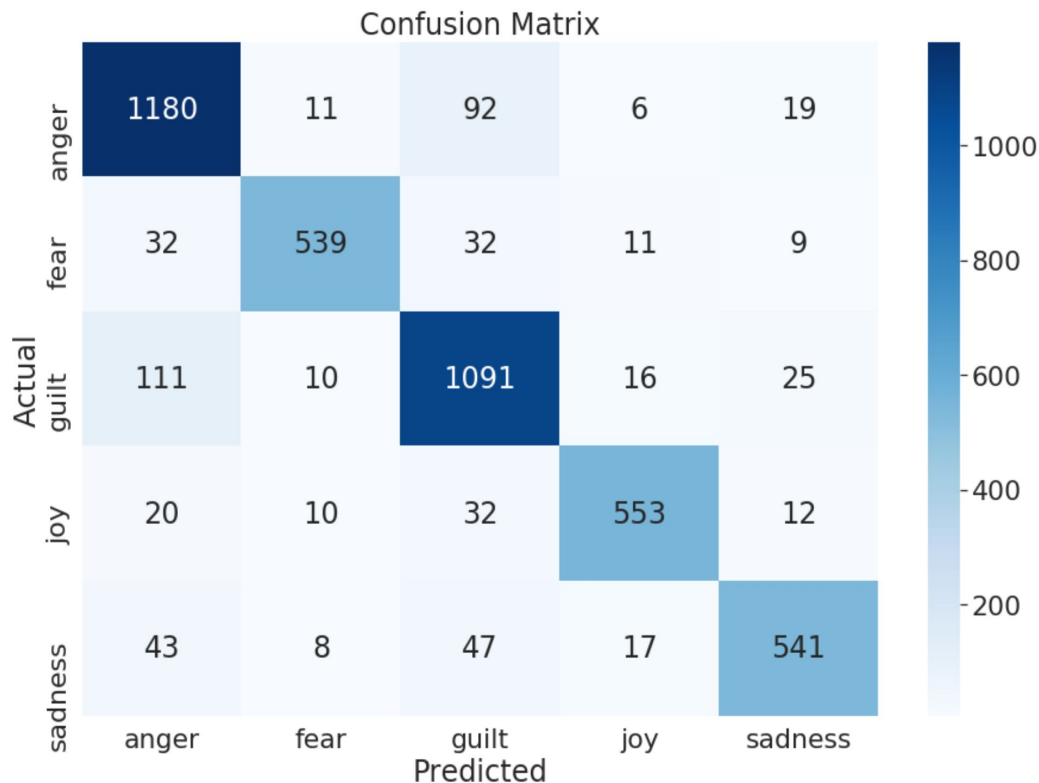

Model inference using SVM

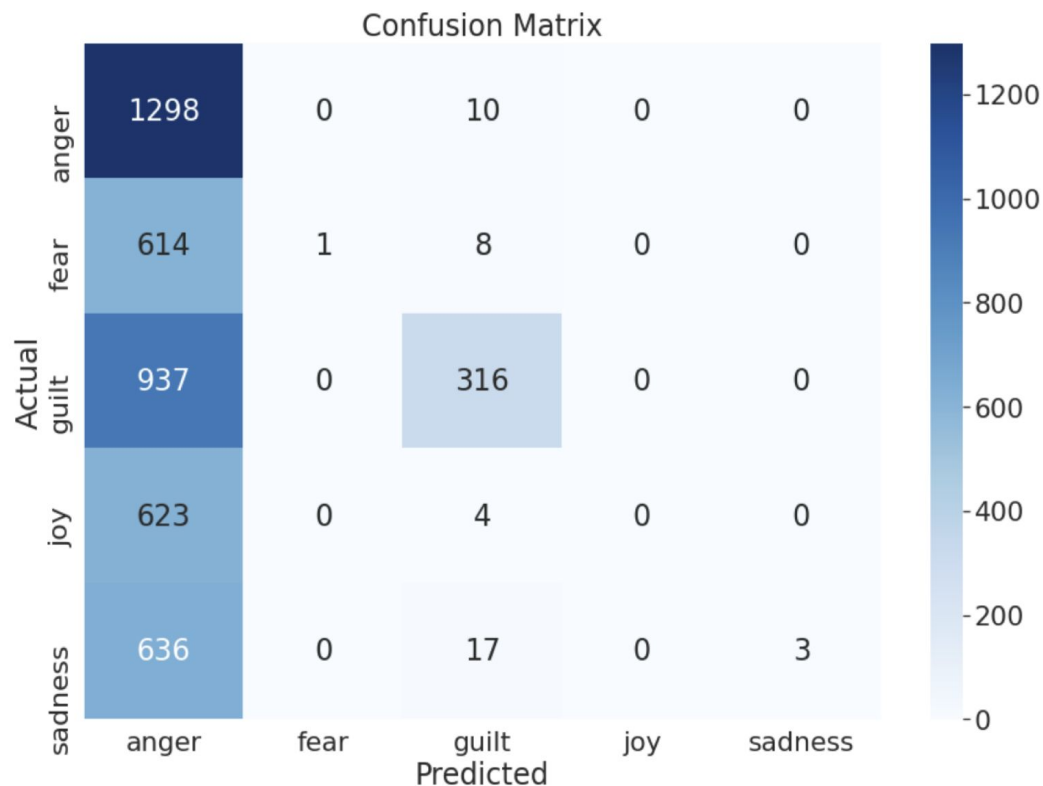# Model Evaluation : Confusion Matrix (Naive Bayes)

Model inference using Naive Bayes

Model Evaluation : Confusion Matrix (Logistic)
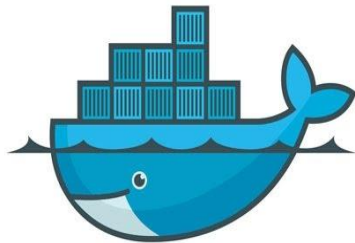

Model inference using Logistic

Model Evaluation : Confusion Matrix (Random Forest)

Model inference using Random Forest
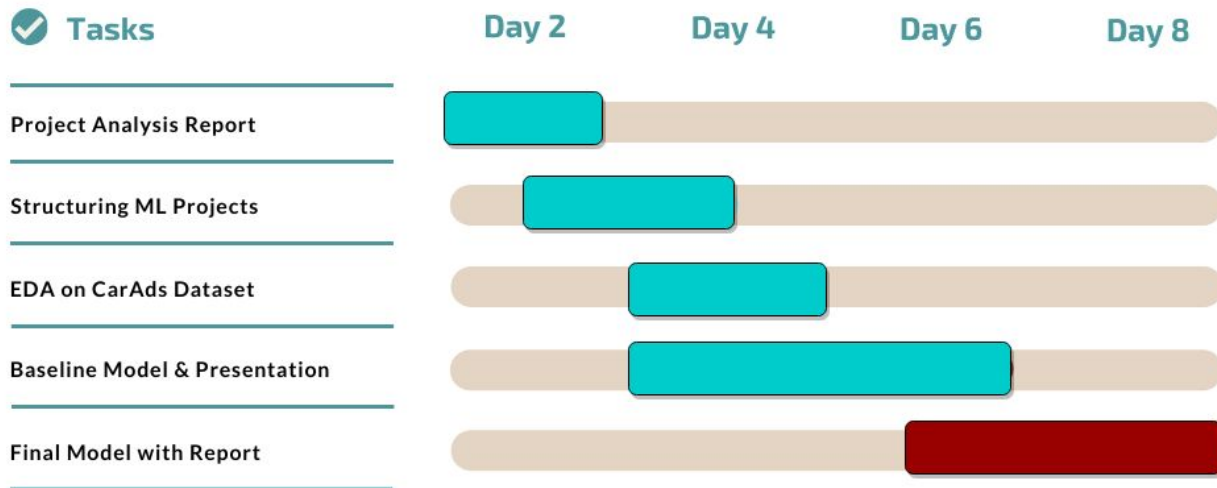
Model Deployment

PROGRESS TIMELINE

Tasks	Day 2	Day 4	Day 6	Day 8
Project Analysis Report	■			
Structuring ML Projects		■		
EDA on CarAds Dataset		■		
Baseline Model & Presentation		■		
Final Model with Report				■

DEMO