

# Short Text Emotion Detection Using Multiclass Support Vector Machine



Presente by

A. V. V. Sairaju

18021F0025

MCA

Supervised by

Smt. SSSN Usha Devi


Assistant Professor

Department of CSE



# Abstract

Emotions capture the essence of the communication process between people and electronic communication systems.



**Detecting Emotions such as joy, anger, sadness, fear, and the like improves the computer-generated response process for users.**



# Challenges

## **Classifying multiple Emotions**

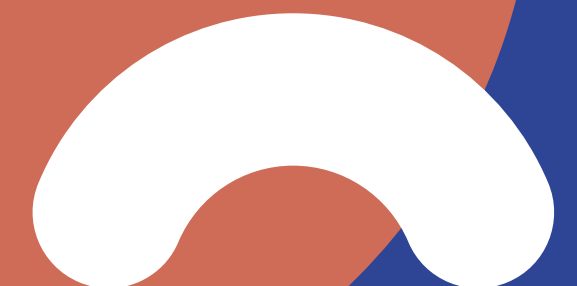
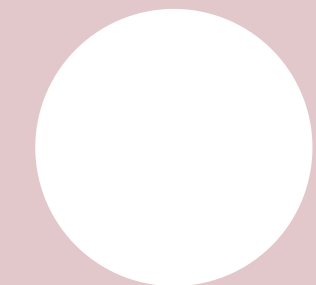
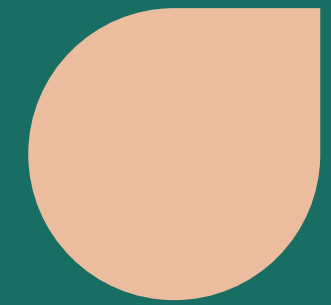
Emotion detection is a multi-class classification problem.

## **Text might be related to multiple emotions**

Here, the sentence may contain the co-occurrence of words related to more than one emotion

# Existing System

**Two  
basic supervised  
topic models are used  
to associate emotions  
with topics**



# Topic Models

1. Weighted Labelled Topic Model (WLTm)
2. X-term Emotion Topic Model (XETM)

# Sampling Algorithms

1. Gibbs Sampling Algorithm
2. Alias Method.

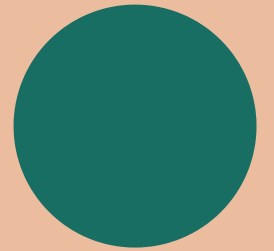
**Topic Modeling**

**Emotion Detection**

**Supervised  
Learning Model**

**Short Text**


**Sampling  
Algorithms**





# Proposed System

The proposed system uses an SVM classifier to map the short text statement to a specific emotion.





# Why SVM?

**We found SVM has a better solution to deal with multi-class data.**

Support Vector Machine (SVM) scales relatively well on multi-class data.

# Modules

## Preprocessing

- Removing stop words
- Removing words with less than two characters
- Lemmatization
- N-grams
- Splitting data for training and testing

## Build Model

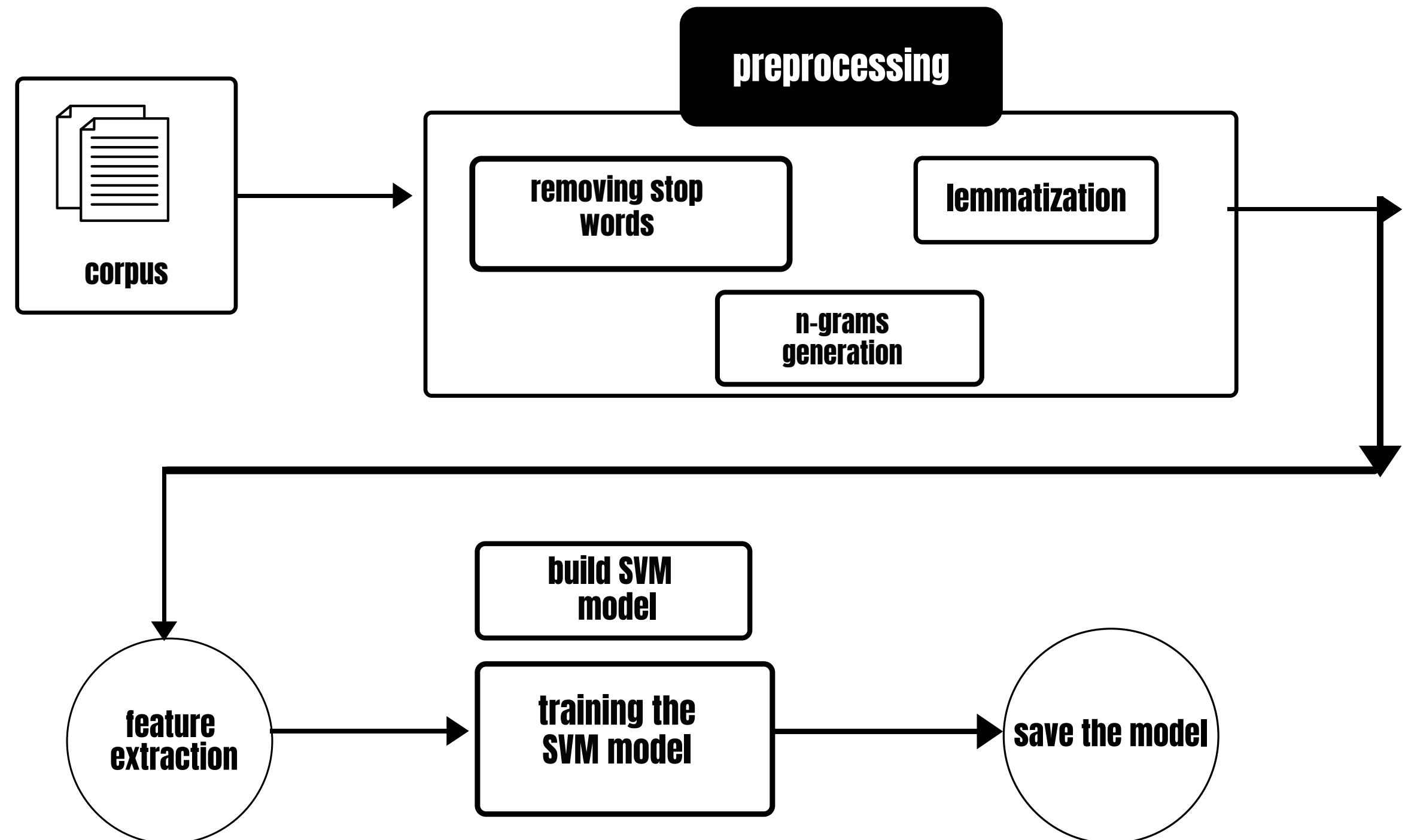
- Training SVM classifier

## Model Evaluation

- Testing the Model with testing data
- Evaluating results

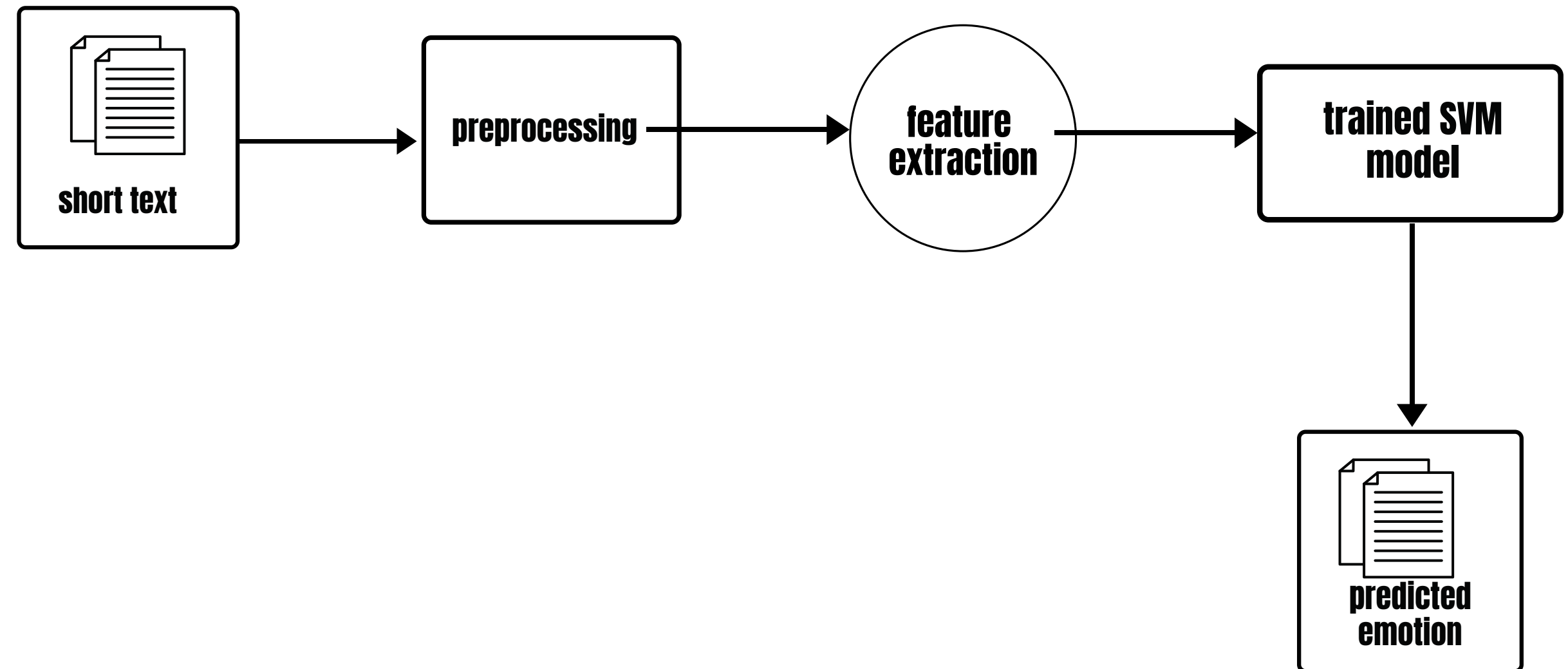
# Architecture

## Training Phase

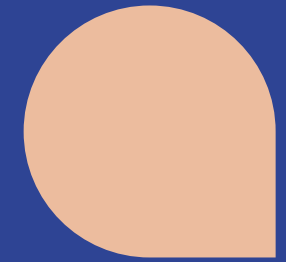


# Architecture

Testing Phase



# Module Implementation



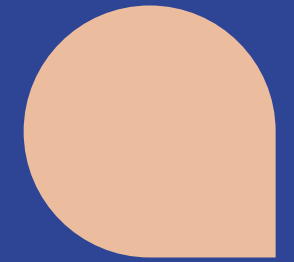
## Programming Language

### Python

Python combines the power of general-purpose programming languages with the ease of use of domain-specific scripting languages.

Python has libraries for data loading, visualization, statistics, natural language processing, image processing, and more.

# Module Implementation



## Libraries

### Scikit learn

Scikit-learn provides a range of supervised and unsupervised learning algorithms via a consistent interface in Python.

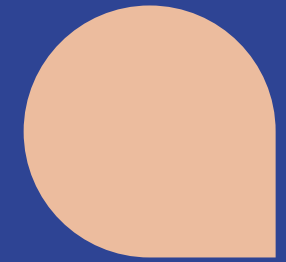
### Numpy

NumPy is one of the fundamental packages for scientific computing in Python

### Pandas

it offers powerful, expressive and flexible data structures that make data manipulation and analysis easy

# Module Implementation



## Libraries

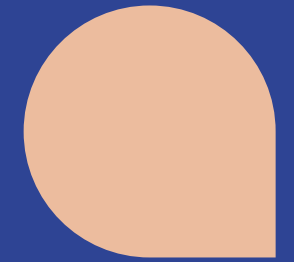
### Matplotlib

Matplotlib is a Python library used for plotting beautiful and attractive Graphs. With the help of this library, we can plot 2D and 3D graphs

### NLTK

The Natural Language Toolkit (NLTK) is a platform used for building Python programs that work with human language data for applying in statistical Natural Language Processing

# Module Implementation



**Programming  
tool**

## Jupyter Notebook

The Jupyter Notebook is an interactive environment for running code in the browser. It is a great tool for exploratory data analysis and is widely used by data scientists.

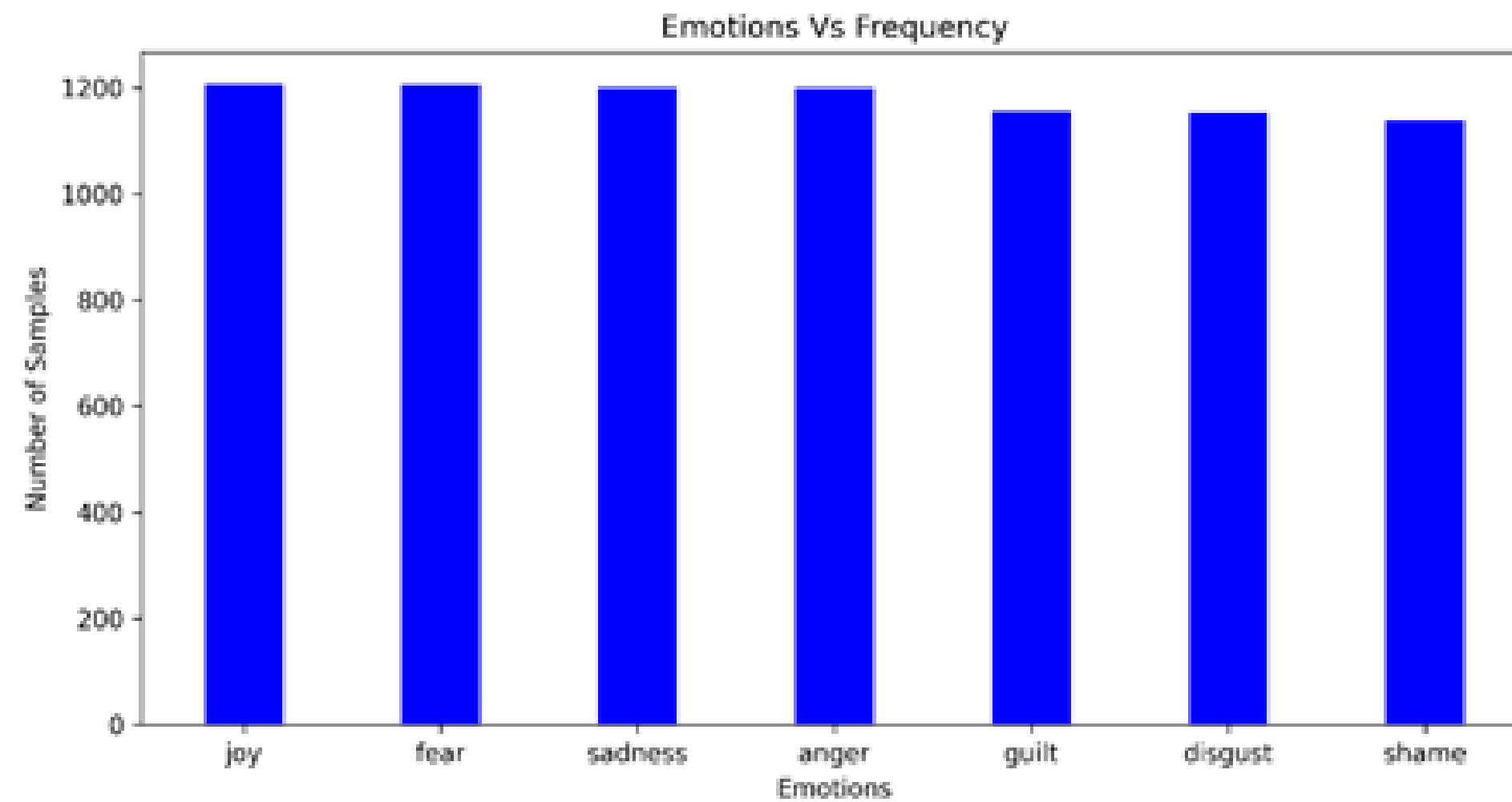
While the Jupyter Notebook supports many programming languages, we only need Python support. The Jupyter Notebook makes it easy to incorporate code, text, and images.



# Results



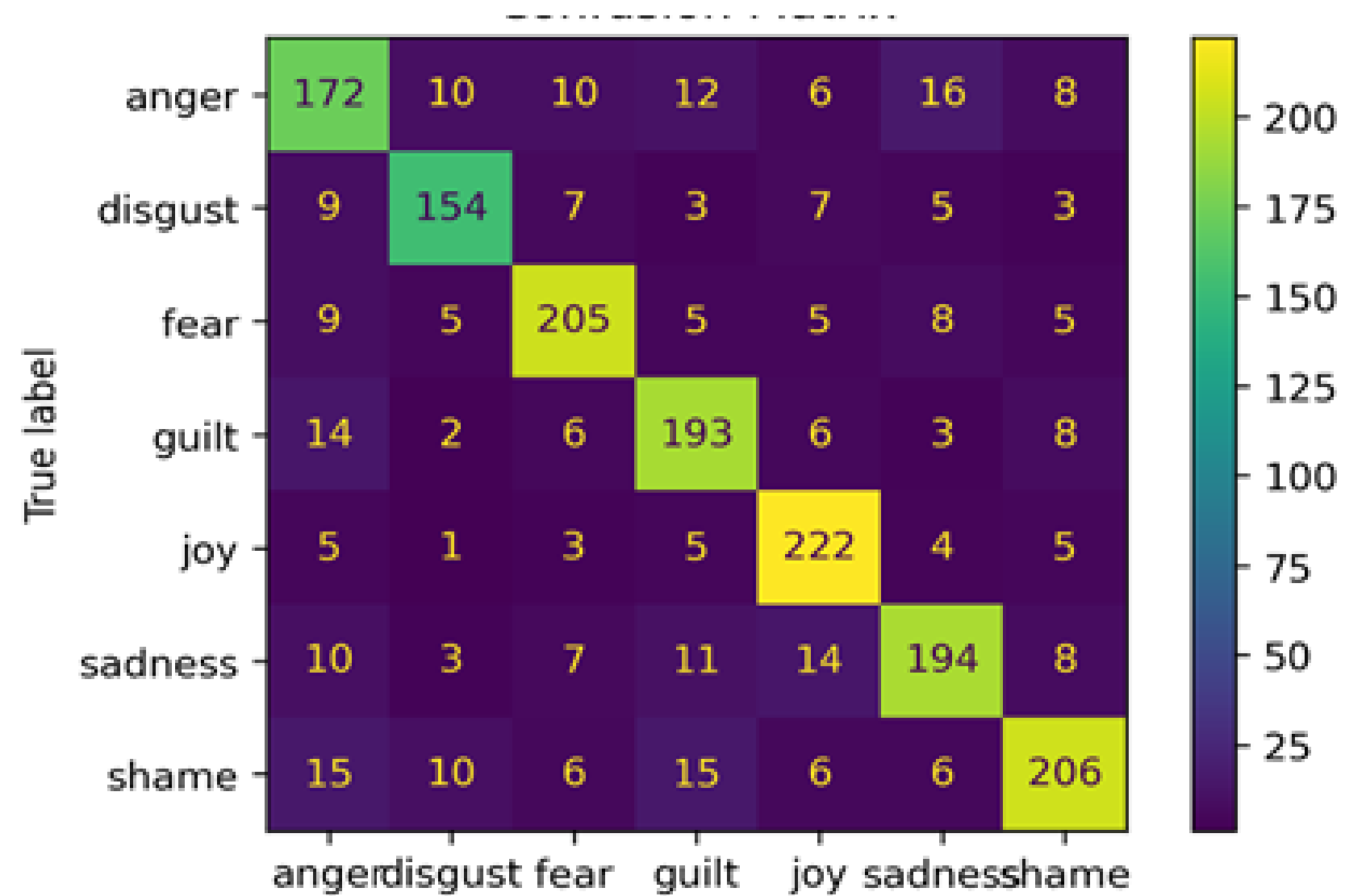
## Emotions vs Samples



# Results



## Confusion Matrix



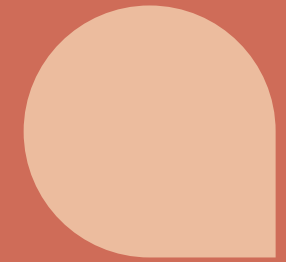
# Results



## Accuracy of models

Model Name	Accuracy
SVM with Linear Kernel	0.8147
SVM with Sigmoid Kernel	0.7455
Weighted Labeled Topic Model	0.6431
X-term Emotion Topic Model	0.6214

# Comparision



## Existing system

used the term group co-occurring in the same context to enrich the number of features

It proposed two algorithms Weighted Labeled Topic Model (WLTm) and an X-Term Emotion Topic Model (XETM).

The time complexity of the existing model is high because of the two supervised topic modeling algorithms.

## Preposed system

uses the co-occurrence of words generated using n-grams generation technique.

used the Support Vector Machine Multiclass Classifier (SVMCMC) for predicting emotion.

time complexity is better when compared to the existing system.

# Comparision



The Proposed system is performing well when compared to the existing system as it shows the better accuracy score than the existing system. The proposed system obtained 81% accuracy.

A decorative graphic on the left side of the slide. It consists of a dark green square at the top containing a solid orange circle. Below this is a dark blue shape that resembles a stylized wave or a partial circle, containing a white semi-circular cutout. The entire graphic is set against a solid orange background.

**Thank you**

A decorative graphic on the right side of the slide. It features a dark green square at the top with a white semi-circular cutout. Below this is a dark blue shape that resembles a stylized wave or a partial circle, containing a solid orange semi-circular cutout. The entire graphic is set against a solid orange background.