June ,2025

Sentiment Analysis (Emotion Prediction)

Guided By:
sambit Subhasis Sahu
Presented By :
Lipa Prusty

Institute Name : IGIT, Sarang
OCAC Training Centre

ABSTRACT :

This project focuses on emotion prediction from textual input using machine learning techniques. By leveraging Natural Language Processing (NLP) methods such as TF-IDF for text vectorization and Logistic Regression for classification, the system can accurately identify emotional states like Happy, Sad, and Angry from user messages. The model is trained on labeled emotion datasets and aims to help machines better understand human emotions, enabling applications in sentiment analysis, mental health support, and intelligent chat systems.

Today's Agenda

1. Introduction to the Session
2. Objective
3. Tools And Technology
4. OutPut
5. Outcome
6. Real Life Problem & My Project
7. Future Scope
8. Conclusion

Introduction

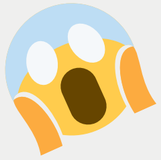
Emotion prediction is an NLP task focused on identifying human emotions from text, such as happy, sad, angry, or fear. It's widely used in areas like sentiment analysis, chatbots, and mental health monitoring.

In this project, we use a labeled dataset of sentences to train a logistic regression model that predicts the emotion behind a message. Text is processed using TF-IDF vectorization, turning it into numerical features for machine learning.

This approach offers a simple, fast, and effective solution for understanding emotional tone in user-generated content.

OBJECTIVE

How are you feeling today?



The objective of this project is to build a machine learning model that can analyze a given text message and accurately predict the underlying emotion expressed, such as Happy, Sad, Angry, etc. Using natural language processing (NLP) techniques and a logistic regression classifier trained on labeled sentence-emotion data, the system aims to understand emotional tone in written communication, which can be useful for applications like sentiment analysis, mental health monitoring, and conversational AI.

Technical Specification

TOOLS:

- Python: Main programming language used for model development
- Jupyter Notebook / Anaconda: Development environment for running and testing code
- Pandas: For loading and preprocessing the dataset
- Scikit-learn: For machine learning models and utilities
- Matplotlib / Seaborn (optional): For plotting graphs like bar charts and confusion matrices

TECHNIQUES:

- Text Preprocessing:
  - Converting text to lowercase
  - Removing stop words and punctuation(via TF-IDF)
  - Tokenization using Tfidfvectorizer
- Feature Extraction:
  - TF-IDF Vectorization: Transforms text data into numerical features .
- Classification Algorithm:
  - Logistic Regression: A supervised learning algorithm used to predict discrete emotion labels.

Output :-

The project successfully implements an emotion prediction system that classifies user-input text into emotions such as Happy, Sad, Angry, and others using a machine learning approach. By training a Logistic Regression model on a labeled text dataset and transforming text with TF-IDF vectorization, the system is able to:

- 🔍 Analyze textual messages and detect emotional tone accurately
- 📈 Achieve good accuracy in emotion classification during model evaluation
- 💬 Allow real-time predictions from user-provided text inputs
- 🛠️ Provide a scalable foundation for chatbots, mental health analyzers, and social media sentiment tools.

Real Life Problem

In today’s digital world, people express their thoughts and emotions online through texts—on social media, chats, reviews, and emails. However, machines or systems can't inherently understand the emotional context of these messages. This leads to challenges in:

- 🧠 Mental health detection (e.g., identifying signs of depression or anxiety)
- 📞 Customer support (e.g., prioritizing angry or upset customer complaints)
- 👨‍💻 Human-computer interaction (e.g., chatbots responding with empathy)
- 📊 Social media analysis (e.g., understanding public sentiment during elections or crises)

My Project

My project focuses on Emotion Prediction from Text using Logistic Regression.

- ✅ It reads text inputs (like sentences or messages) from the user.
- ✅ Uses TF-IDF (Term Frequency-Inverse Document Frequency) to convert words into numbers.
- ✅ Applies a Logistic Regression model to classify emotions like Happy, Sad, Angry, etc.
- ✅ Helps in predicting emotions based on how the message is written.

Future Scope

Multilingual Support:

Extend the model to support multiple languages by integrating multilingual datsets and models like XLM-R or mBERT

Deep Learing Integration:

Replace Logistic Regression with more advanced models like LSTM, BERT or Transformers for higher accuracy and better contextual understanding.

Real-Time Emotion Monitoring:

Integrate the model into chatbots, mental health apps or virtual assistance to detect & respond to emotions in real time.

Mobile & Web Deployment:

Deploy the model in web or mobile applications using streamlit or TensorFlow.js to make it accessible to end=user.

Custom Emotion Categories:

Allows user or organizations to define their own emotion categories for domain specific applications (e.g., customer support, therapy, education).

Conclusion

This project successfully demonstrates how machine learning, specifically Logistic Regression, can be applied to predict human emotions from textual data. By using a combination of natural language processing techniques such as TF-IDF and supervised learning, the model is capable of identifying emotional states like happy, sad, or angry from user input. The system provides a strong foundation for real-world applications in sentiment analysis, mental health monitoring, and conversational AI. While the current model offers promising results, future enhancements using deep learning and larger, diverse datasets can further improve its performance and accuracy.

June 2025

Thank you!

Have
a good
day!