

DECODING EMOTIONS THROUGH SENTIMENT ANALYSIS OF SOCIAL MEDIA CONVERSATIONS

Student Name: Loseya.V

Register Number: 511523205030

Institution: P.T.Lee Chengalvaraya Naicker College Of Engineering and Technology

Department: Information Technology

Date of Submission: 2025-04-26

1. Problem Statement

Social media platforms generate vast amounts of textual data daily, reflecting a wide range of human emotions and opinions.

Understanding these emotions is critical for businesses, policymakers, and researchers aiming to gain insights into public sentiment.

However, manually analyzing this data is infeasible due to its volume and variability. Employing sentiment analysis techniques enables

automated interpretation of emotional nuances in social media conversations, facilitating informed decision-making and trend analysis.

2. Objectives of the Project

- Develop an AI-based system to analyze sentiments expressed in social media conversations.
- Accurately classify text into sentiment categories (e.g., positive, negative, neutral).
- Identify emotional trends over time to support strategic decision-making.
- Present findings using intuitive visualizations and dashboards.

3. Scope of the Project

Features to Analyze:

- Text data from platforms like Twitter, Reddit, and Instagram.
- Contextual elements like hashtags, mentions, and emojis.
- Temporal and geographic patterns of sentiments.

Limitations:

- Focus restricted to publicly available data adhering to platform policies.
- Use of English language posts for initial analysis.
- Emphasis on text-based sentiment analysis with limited support for multimedia data.

4. Data Sources

- Dataset Source: Twitter APIs, Kaggle datasets, and web scraping tools.
- Type: Public datasets with user-generated content.
- Nature: Dynamic datasets updated periodically to capture recent trends.

5. High-Level Methodology

Data Collection:

Use APIs and web scraping tools to gather text data from social media platforms.

Data Cleaning:

Perform preprocessing to handle noise, remove irrelevant content, and standardize text formats.

Exploratory Data Analysis (EDA):

Utilize word clouds, sentiment distribution graphs, and time-series analysis to uncover patterns and trends.

Feature Engineering:

Extract features like n-grams, sentiment scores, and contextual markers (e.g., hashtags).

Model Building:

Experiment with natural language processing models like Naive Bayes, Support Vector Machines (SVM), and Transformer-based architectures (e.g., BERT) for sentiment classification.

Model Evaluation:

Evaluate models using accuracy, precision, recall, F1-score, and confusion matrix metrics.

Visualization & Interpretation:

Present sentiment trends and emotional insights through dashboards and interactive charts.

Deployment:

Develop a web-based tool or API for real-time sentiment analysis and visualization using frameworks like Flask or Streamlit.

6. Tools and Technologies

Programming Language:

Python

Notebook/IDE:

Jupyter Notebook or Google Colab

Libraries:

- Data Processing: Pandas, NumPy
- Visualization: Matplotlib, Seaborn, Plotly
- Modeling: scikit-learn, TensorFlow, Hugging Face Transformers

- Deployment: Flask, Streamlit

7. Team Members and Roles

- Handles data cleaning and EDA:Abi.D
- Leads feature engineering and model development: Keerthana.M
- Develops and deploys the web tool or API:Madhumitha.V
- Coordinates tasks and ensures timely delivery:Loseya.V