## Artificial

## Intelligence and Machine Learning

Project Report

Semester-IV (Batch-2022)

**Social Media Sentiment Analysis Using Twitter**

A red and white sign

Description automatically generated with low confidence

**Supervised By: Submitted By:**

Rajeev Bhardwaj Teena Goyal

2210990910

G13

**Department of Computer Science and Engineering**

## Chitkara University Institute of Engineering & Technology,

## Chitkara University, Punjab

**ABSTRACT**

In today's digitally connected world, social media platforms like Twitter have become invaluable sources of real-time data reflecting public opinions, emotions, and sentiments. This project aims to leverage the power of machine learning techniques to analyze sentiment trends within a Twitter dataset. The objective is to develop a robust sentiment analysis model capable of accurately classifying tweets into positive, negative, or neutral categories.

The project begins with the collection of a sizable Twitter dataset using appropriate APIs or web scraping techniques. Preprocessing steps such as text normalization, tokenization, and noise removal are applied to clean the raw tweet data. Feature extraction methods, including bag-of-words, TF-IDF, or word embeddings, are employed to represent the textual data in a format suitable for machine learning algorithms.

Next, various machine learning models are explored and trained on the preprocessed dataset. These models may include traditional algorithms such as Naive Bayes, Support Vector Machines (SVM), as well as more advanced techniques like Recurrent Neural Networks (RNNs) or Transformers. The models are evaluated based on metrics such as accuracy, precision, recall, and F1-score to determine their performance in sentiment classification.

Furthermore, techniques for model optimization and fine-tuning are employed to enhance the overall accuracy and generalization capability of the sentiment analysis system. Hyperparameter tuning, cross-validation, and ensemble methods may be utilized to achieve optimal results.

In addition to model development, visualization techniques are employed to gain insights into sentiment trends over time and across different topics or events. Graphical representations such as word clouds, sentiment distributions, and time-series analyses provide a comprehensive understanding of the underlying sentiment dynamics within the Twitter dataset.

The proposed sentiment analysis system holds significant implications for various applications, including brand monitoring, market research, and public opinion analysis. By accurately gauging the sentiment of Twitter users, businesses, policymakers, and researchers can make informed decisions and gain valuable insights into public perception and sentiment towards specific topics, products, or events.

In conclusion, this project demonstrates the efficacy of machine learning techniques in analyzing social media sentiment using Twitter data. By harnessing the vast amount of user-generated content on Twitter, the developed sentiment analysis model offers a powerful tool for understanding and interpreting public sentiment in the digital age.