***Sentiment Analysis Dashboard on Social Media Presence***

*Submitted by*

NIMIT KWATRA [Reg No: RA2111003030213]

DEVESH TOMAR[Reg No: RA2111003030208]

SHRIKANT CHOUDHARY[Reg No: RA2111003030224]

SUNIDHI SINGH [Reg No: RA2111003030219]

*Under the guidance of*

Dr. AMIT KUKKER

(Professor, Department of Computer Science & Engineering)

*in partial fulfilment for the award of the degree*

*of*

BACHELOR OF TECHNOLOGY

in

COMPUTER SCIENCE & ENGINEERING

of

FACULTY OF ENGINEERING AND TECHNOLOGY



SRM INSTITUTE OF SCIENCE & TECHNOLOGY, NCR CAMPUS

MAY 2021

SRM INSTITUTE OF SCIENCE & TECHNOLOGY

(Under Section 3 of UGC Act, 1956)

BONAFIDE CERTIFICATE

Certified that this project report titled “Sentiment Analysis Dashboard On Social Media Presence***”*** is the Bonafide work of “NIMIT KWATRA[RA2111003030213], SUNIDHI SINGH[RA211100 ”, who carried out the project work under my supervision. Certified further, that to the best of my knowledge the work reported herein does not form any other project report or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

Dr. AMIT KUKKAR

GUIDE

Professor

Dept. of Computer Science & Engi-

neering

Signature of the Internal Examiner SIGNATURE

Dr. R.P.Mahapatra

HEAD OF THE DEPARTMENT

Dept. of Computer Science & Engineering

Signature of the External Examiner

# 

# ABSTRACT

Mental health is a crucial aspect of people’s lives, and its neglect in some industries can impact work performance.

Many people are dissatisfied with their work environment, which can affect the company’s performance if it’s not comfortable or causes problems.

The dashboard will visually display the sentiment of customer feedback on social media, showing the percentage of positive, negative, and neutral sentiment, and highlighting common topics and keywords.

This technique can analyze a person’s emotions and opinions, helping to detect their mental health status or the problems they’re facing.

Sentiment analysis can be used to evaluate various types of content, including customer messages, online reviews, and social media posts.

The project will utilize Natural Language Processing (NLP) and deep learning techniques such as LSTM model to achieve maximum accuracy in sentiment analysis.

# ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my guide, Dr. AMIT KUKKAR for his valuable guidance, consistent encouragement, personal caring, timely help and providing me with an excellent atmosphere for doing research. All through the work, in spite of his busy schedule, he has extended cheerful and cordial support to me for completing this research work.

Author: Devesh Tomar, Nimit Kwatra, Shrikant Choudhary, Sunidhi Singh

TABLE OF CONTENTS

[ABSTRACT iii](#_Toc13813)

[ACKNOWLEDGEMENTS iv](#_Toc13814)

[LIST OF FIGURES vi](#_Toc13816)

[1 INTRODUCTION vii](#_Toc13819)

[2 LITERATURE SURVEY viii](#_Toc13822)

[3 Problem Statement ix](#_Toc13823)

[4 Proposed solution xi](#_Toc13824)

[5 Objective xii](#_Toc13825)

[6 Implementation Design xiii](#_Toc13826)

[6.1 Design Diagram xvi](#_Toc13827)

[7 Conclusion xvii](#_Toc13828)

[8 Future Enhancement xviii](#_Toc13830)

[9 Limitations xix](#_Toc13831)

[Appendix A xx](#_Toc13832)

[References xxi](#_Toc13833)

# LIST OF FIGURES

Fig 1. Design Diagram/ Block diagram -------------------------------------------------------------------- xvi

Fig 2. Data flow diagram ------------------------------------------------------------------------------------- xvii

Fig 3. Use case diagram -------------------------------------------------------------------------------------- xviii

CHAPTER 1

# INTRODUCTION

With the rapid rise of social media as a primary channel for communication, people across the world are now able to connect, share opinions, and express a wide range of emotions in real time. This ever-growing digital presence generates a massive volume of user-generated content, encompassing comments, posts, hashtags, and reactions that reflect public sentiment. However, the sheer volume and diversity of this data make it challenging to gain meaningful insights. This project, titled "Sentiment Analysis for Social Media," aims to address this challenge by developing an intelligent system capable of interpreting the emotions embedded in social media content.

Sentiment analysis, a branch of artificial intelligence, offers a solution by detecting and interpreting emotions within text. By analysing large amounts of social media data, the project seeks to uncover sentiment trends and nuances, distinguishing between positive, negative, and neutral tones and even identifying specific emotions like joy, anger, or surprise. The project provides a sentiment analysis dashboard designed to visualize these insights in an intuitive way, making them accessible to users across various sectors.

Understanding sentiment on social media is vital for multiple reasons. In the business sector, it allows companies to gain insights into customer attitudes toward products, services, and brand image, directly impacting reputation management and marketing strategies. In the public sphere, organizations and policymakers can gauge public opinion on key issues and events, helping them to stay informed of societal trends. Additionally, individuals and influencers can benefit from a deeper understanding of their audiences' emotional responses, enabling more meaningful communication.

The sentiment analysis dashboard empowers users by transforming complex, large-scale data into actionable insights. This project ultimately aims to simplify social media analysis, making it easier for users to make data-driven decisions and confidently navigate the fast-paced world of social media communication.

# CHAPTER 2

# LITERATURE SURVEY

## 2.1 Sentiment Analysis of Consumer Reviews Using Deep Learning.

Deep learning-based LSTM models show promising results for sentiment analysis of consumer reviews, offering comparable or superior performance metrics such as accuracy, precision, recall, and F1-score.

This project will be using sentiment analysis helps businesses understand customer reviews and improve products and services

Existing literature lacks exploration into alternative deep learning architectures like transformers for sentiment analysis of consumer reviews.

Amjad Iqbal , Rashid Amin , Javed Iqbal , Roobaea Alroobaea, Ahmed Binmahfoudh and Mudassar Hussain(2022)

## 2.2 Sentiment Analysis of Persian Movie Reviews Using Deep Learning

Deep learning excels over traditional methods in Persian sentiment analysis, with stacked-bidirectional-LSTM and 2D-CNN achieving high accuracies of up to 95.61% and 89.76% respectively

A similar type of dataset has been used, which helps to compare the performance and accuracy of our model.

Limited research exists on deep learning-driven Persian sentiment analysis, indicating a need for further exploration

Kia Dashtipour , Mandar Gogate , Ahsan Adeel , Hadi Larijani and Amir Hussain(2021)

2.3 A Deep Neural Network-Based Approach for Sentiment Analysis of Movie Reviews.

This model has implemented a deep neural network with seven layers on movie review data. The model achieves accuracy of 91.18%, recall of 92.53%, F1-Score of 91.94%, and precision of 91.79%

. A similar type of dataset has been used, which helps to compare the performance and accuracy of our model. It is possible to increase the classification accuracy of the model by adding more layers to the model and testing it with different datasets.

By Kifayat Ullah, Anwar Rashad, Muzammil Khan, Yazeed Ghadi, Hanan Aljuaid,

Zubair Nawaz

1.4 Analysis Sentiment based on IMDB aspects from movie reviews using SVM

Based on the results obtained in this study, the SVM classification model used in this study is also able to produce 79% accuracy, 75% precision, and 87% recall. SVM is very suitable for use in text data problems. A similar type of dataset has been used, which helps to compare the performance and accuracy of our model. For further research, additional aspects such as film genre can be added to the research paper.

By Nur Ghaniaviyanto Ramadhan, Teguh Ikhlas Ramadhan

# PROBLEM STATEMENT

In the digital age, social media platforms such as Twitter, Facebook, Instagram, and others have emerged as a significant medium through which the public expresses opinions, shares experiences, and provides feedback. For organizations, these platforms serve as a rich source of information about customer perceptions, public sentiment, and brand reputation. However, the challenge lies in effectively harnessing this immense volume of unstructured data to derive meaningful insights, as sifting through millions of posts, comments, and reactions manually is virtually impossible.

Organizations seek actionable insights from this data to understand how their products, services, or overall brand are being received by the public. However, without a structured approach to analyzing this information, critical trends, issues, and opportunities may go unnoticed. This underscores the need for an automated solution that can interpret and classify the sentiment of social media content, identifying whether it reflects positive, negative, or neutral attitudes and even pinpointing specific emotions such as satisfaction, frustration, or excitement.

Moreover, a solution that offers a user-friendly interface for visualizing sentiment trends is essential for stakeholders who may not have technical expertise but require quick and accessible insights. An effective sentiment analysis system should thus not only analyze social media data but also present the findings in an intuitive, easily interpretable manner. This would empower businesses, marketers, and decision-makers to make data-driven decisions that enhance customer experience, improve brand management, and respond promptly to public feedback, ultimately strengthening their connection with their audience.

CHAPTER 4

PROPOSED SOLUTION

1. **Data Preparation**:
   1. Load dataset.
   2. Split into training and testing data.
2. **Model Architecture**:
   1. Define Sequential model.
   2. Add Embedding layer.
   3. Add Dropout layer for overfitting prevention.
   4. Add LSTM layer.
   5. Add another Dropout layer.
   6. Add Dense output layer with sigmoid activation.
3. **Model Compilation**:
   1. Compile model with appropriate function, optimization algorithm, and accuracy metric.
4. **Model Training**:
   1. Train model on training data, validate with test data.
5. **Model Evaluation**:
   1. Evaluate model performance on test data.

CHAPTER 5

OBJECTIVE

1. **Understanding Human Sentiments:** The main goal is to recognize and characterize people’s moods and feelings with various forms of data: face and voice, text and biometric data.

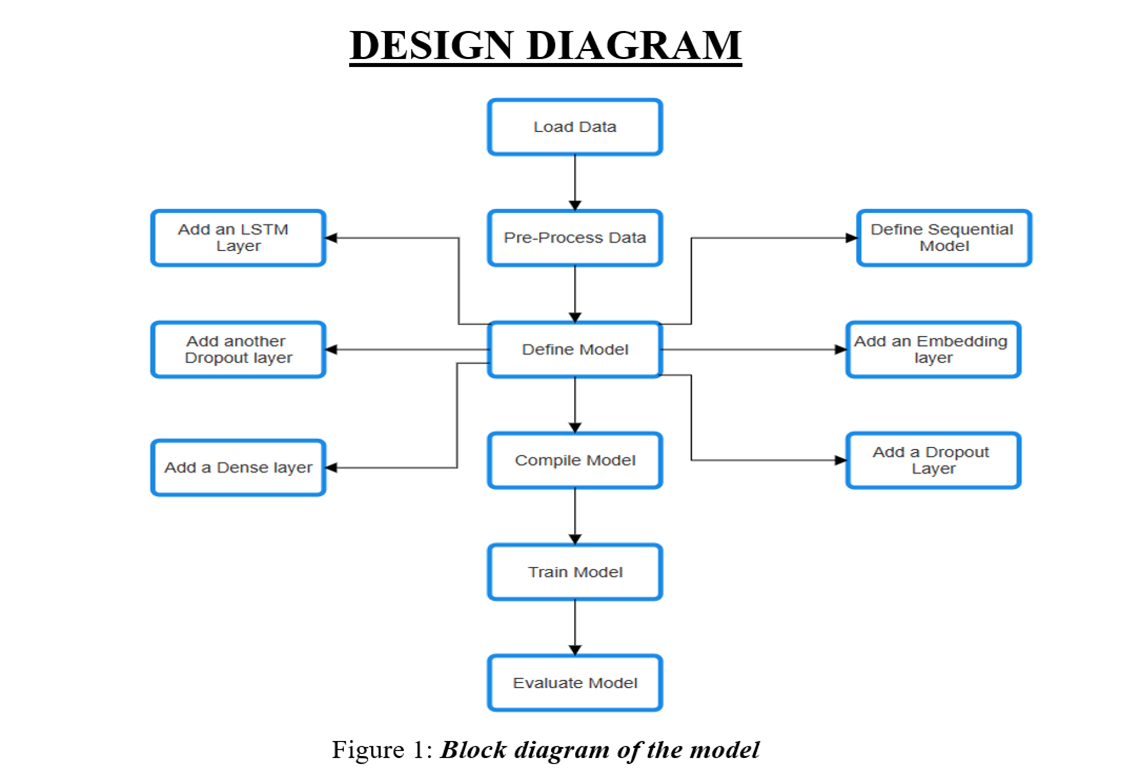
2.**Enhancing Human-Computer Interaction:** Emotion detection is a means of enhancing the human-an environment interface introducing capability for capturing the mood of a user.

1. **Applications in Mental Health:** Emotion detection can help to track the condition of the patient’s psyche to work out the material for a therapist, as well as help identify the likelihood of an emotional breakdown.
2. **Personalization:** In the areas such as marketing, customer relations, and entertainment it is possible to create an individual environment because the technology can analyse the emotional response of the user.
3. **Social Robotics and AI:** Introducing emotion detection feature to robots and AI will improve their empathetic communication and can be effective in sectors such as: care giving, teaching and therapy.
4. **Research and Development:** Emotion detection helps to enrich the existing knowledge of the emotional expression and perception field in psychology, neuroscience and artificial intelligent systems.
5. **Ethical Considerations:** Some of the areas where one needs to know the ethical issues to guide the development of the Emotion detection technology includes.

CHAPTER 6

IMPLEMENTATION DETAILS

* **Algorithm 1: Data Preparation**
* Input: dataset D.
* Output: Pre-processed dataset D’.
* Step 0: Start.
* Step 1: Import necessary libraries.
* Step 2: Load the dataset D from csv file using pd.read\_csv().
* Step 2.1: Drop unnecessary columns from the dataframe.
* Step 3: Tokenize the text into individual words by word\_tokenize().
* Step 3.1: Convert all the words to lowercase using word.lower().
* Step 4: Create a corpus of words from the tokenized text num\_words.
* Step 5: Split the dataset into training (80%) and testing (20%) set train\_size as int(shape[0]\*0.8).
* Step 5.1: Create X\_train  by selecting the first train\_size.
* Step 5.2: Create Y\_train  by selecting the first train\_size.
* Step 5.3: Create X\_test  by selecting the all rows from train\_size.
* Step 5.4: Create X\_test  by selecting the all rows from train\_size.
* Step 6: Tokenize the words and pad sequences for equal input dimensions in both training and testing sets.
* Step 6.1: Fit the tokenizer on the texts and Transform each text in X\_train to a sequence of integers.
* Step 6.2: Pad sequences to the maxlen. Adding zeros at the end of each sequence until it is of length maxlen. If a sequence > maxlen, it will be truncated.
* Step 6.3: Repeat Steps 6.1 and 6.2 for X\_test
* Step 7: Encode the sentiment labels in the training and testing sets using LabelEncoder().
* Step 8: Store the Preprocessed data as D'.
* Step 9: Stop
* **Algorithm 2: Building the model**
* Input: Pre-processed dataset D
* Output: Trained LSTM model.
* Step 0: Start.
* Step 1: Initialize a Sequential model:
* Step 2: Add an Embedding layer as the input layer:
* model. Add (Embedding())
* Step 3: Add a Dropout layer to prevent overfitting:
* model. Add (Dropout())
* Step 4: Add an LSTM layer:
* model. Add (LSTM())
* Step 5: Add another Dropout layer:
* model. Add (Dropout ())
* Step 6: Add a Dense output layer with a sigmoid activation function:
* model.add(Dense(activation as 'sigmoid'))
* Step 7: Stop
* **Algorithm 3: Model Training and Evaluation**
* Input: Trained LSTM model
* Output: Accuracy on the Testing data.
* Step 0: Start.
* Step 1: Compile the model using binary cross-entropy loss and Adam optimizer:
* model.compile(loss as 'binary\_crossentropy', optimizer as 'adam').
* Step 2: Print the model summary.
* Step 3: Train the model on the training data.
* Step 3.1: Set epoch.
* Step 3.2: Set batch size.
* Step 4: Evaluate the model on the test data and print the accuracy: scores as model.evaluate(X\_test, y\_test).
* Step 5: Stop.



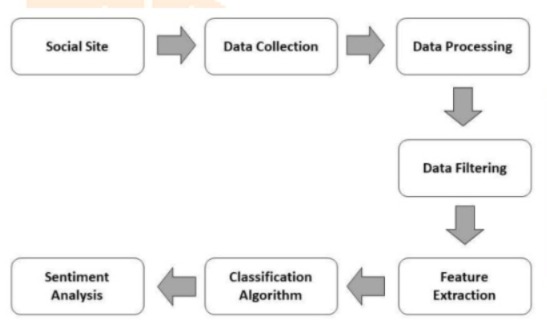


Fig 2. Data Flow diagram

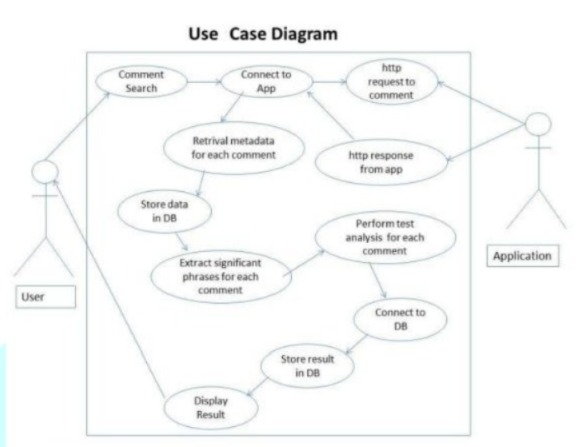


Fig 3. Use Case Diagram

CHAPTER 7

# CONCLUSION

Sentiment analysis can be defined as a rapidly developing interdisciplinary area, which covers technology aspects and psychological approaches towards people's emotional experience.

Simultaneously, the expansion of machine learning and artificial intelligence leads to understanding emotions based on different data sources including faces, tones of voices or even texts are especially promising in various areas. Some of these are in areas such as mental health, customer service, education and human computer interaction.

However, issues to do with ethical issues including privacy and consent as well as the potential for misuse of this technology are essential questions. From this progress, it will be necessary to make greater effort to promote the responsible use of these tools, especially in promoting the public interest. Therefore, it can be authorized that though the possibilities of emotion detection seem quite bright, proper consideration of benefits of the technology together with its threat can form a quite elevator and appropriate portrait to integrate the technology in everyday life.

CHAPTER 8

FUTURE ENHANCEMENT

In this project, we have demonstrated a sentiment analysis based on LSTM model for text data. Users express and share their opinions on different platforms for an organization.

Manual analysis of large amounts of such data can be difficult. So, pre-processed data from various sources can be used. Sentiment analysis gives people's sentiment toward products, services, social media, and company planning. Reviews (from sources such as IMDB) and social network posts (mostly from Twitter and Facebook) are categories of documents for sentiment analysis.

Deep Learning methods such as LSTM show better performance of sentiment classification with 86% accuracy on IMDB review dataset, 93% accuracy on Twitter review dataset and 88% on social media dataset.

Dashboard design provides a user-friendly interface that allows users to visualize sentiment of sentence either positive or negative with probability of sentence. Also, gives emoji image related to sentiment.

In future we are planning to extend this project to a larger extent where different embedding models can be considered on large variety of the datasets. We can also use better model for better accuracy avoiding overfitting

CHAPTER 9

**LIMITATIONS**

**Choosing the Right Parameters**: The choice of parameters such as the number of top words to consider, the maximum review length, the size of the word embeddings, the number of LSTM units, and the dropout rate can significantly affect the model’s performance.  
   
**Overfitting:** LSTM models are powerful and have a high capacity, which makes them prone to overfitting, especially when the amount of training data is limited. Although dropout layers are used in the model to mitigate this issue.  
   
**Training Time**: LSTM models can be slow to train, especially when the number of epochs is large. This can make the process of model selection and hyperparameter tuning quite time-consuming.  
   
**Interpretability:** While LSTM models can achieve high performance on sentiment analysis tasks, they are often considered as “black box” models.  
  
**Dependency on External Libraries**: The program relies on several external Python libraries such as Keras and NLTK. Any changes or issues with these libraries can potentially affect the program.  
  
**Hardware Requirements**: Training deep learning models like LSTM requires significant computational resources (CPU/GPU), and not all systems may be equipped to handle these requirements.

APPENDIX A

# DATASET

* Dataset 1: IMDB Dataset of 50K movie reviews from Kaggle.
* Description: This dataset contains 50,000 movie reviews from IMDb. Each review is labeled as positive or negative sentiment.
* Dataset 2: Twitter Sentiment Analysis from Kaggle.
* Description: This dataset contains tweets labelled as 0 (negative) and 1 (positive). It is used for sentiment analysis and contains text data from Twitter.
* Dataset 3: The Social Media Sentiments Analysis from Kaggle.
* Description: Dataset captures a vibrant tapestry of emotions, trends, and interactions across various social media platforms. Each sentence is labeled as positive or negative sentiment.

REFERENCES

* [1] Iqbal, A., Amin, R., Iqbal, J., Alroobaea, R., Binmahfoudh, A., & Hussain, M. (2022). Sentiment analysis of consumer reviews using deep learning. Sustainability, 14(17), 10844.
* [2] Dashtipour, K., Gogate, M., Adeel, A., Larijani, H., & Hussain, A. (2021). Sentiment analysis of persian movie reviews using deep learning. Entropy, 23(5), 596.
* [3] Ramadhan, N. G., & Ramadhan, T. I. (2021). Analysis sentiment based on IMDB aspects from movie reviews using SVM. Sinkron: jurnal dan penelitian teknik informatika, 6(1), 39-45.
* [4] Ullah, K., Rashad, A., Khan, M., Ghadi, Y., Aljuaid, H., & Nawaz, Z. (2022). A deep neural network-based approach for sentiment analysis of movie reviews. Complexity, 2022.
* [5] ‘The Power of Social Media: Connecting and Engaging in the Digital Age,’ Times of India, 2024. Available: https://timesofindia.indiatimes.com/readersblog/elrashidy-media-group/the-power-of-social-media-connecting-and-engaging-in-the-digital-age-50585/. [Accessed: 14/02/2024].
* [6] Fornacciari, P., Mordonini, M., & Tomaiuolo, M. (2015, June). A case-study for sentiment analysis on twitter. In *WOA* (pp. 53-58).
* [7] Tusar, M. T. H. K., & Islam, M. T. (2021, September). A comparative study of sentiment analysis using NLP and different machine learning techniques on US airline Twitter data. In *2021 International Conference on Electronics, Communications and Information Technology (ICECIT)* (pp. 1-4). IEEE.
* [8] Saragih, M. H., & Girsang, A. S. (2017, November). Sentiment analysis of customer engagement on social media in transport online. In *2017 International Conference on Sustainable Information Engineering and Technology (SIET)* (pp. 24-29). IEEE.
* [9] "IMDb Dataset of 50K Movie Reviews," Kaggle, 2019. [Online]. Available: https://www.kaggle.com/datasets/lakshmi25npathi/imdb-dataset-of-50k-movie-reviews/data. [Accessed: 16-03-2024]
* [10] LSTM Tutorial,’ Simplilearn, 2024. Available: https://www.simplilearn.com/tutorials/artificial-intelligence-tutorial/lstm. [Accessed: 14-03-2024]
* [11] "Twitter Sentiment Analysis: Hatred Speech," Kaggle, 2020. [Online]. Available: https://www.kaggle.com/datasets/arkhoshghalb/twitter-sentiment-analysis-hatred-speech. [Accessed: 16-03-2023].
* [12] Prastyo, P. H., Sumi, A. S., Dian, A. W., & Permanasari, A. E. (2020). Tweets responding to the Indonesian Government’s handling of COVID-19: Sentiment analysis using SVM with normalized poly kernel. J. Inf. Syst. Eng. Bus. Intell, 6(2), 112.
* [13] Ressan, M. B., & Hassan, R. F. (2022). Naive-Bayes family for sentiment analysis during COVID-19 pandemic and classification tweets. Indonesian Journal of Electrical Engineering and Computer Science, 28(1), 375.
* [14] Bello, A., Ng, S. C., & Leung, M. F. (2023). A BERT framework to sentiment analysis of tweets. Sensors, 23(1), 506.
* [15] Gaurav Umesh Awate, Rushikesh Sanjay Dhus, Sanket Rajendra Gaikwad, Rohit Ravindra Lonkar, Prof. S.R. Bhujbal(2024). Sentiment Analysis on Social Media.