

Research on Sentimental Analysis on Mental Health Using Social Media

Tanvi Gawas¹, Bhavika Makwana*²

¹HOD (Data Science) Department of Information Technology, Thakur College of Science & Commerce

²PG Student, Department of Information Technology, Thakur College of Science & Commerce

ABSTRACT

This project investigates the use of machine learning and natural language processing techniques in sentiment analysis of mental health discussions on social media platforms. By employing both traditional methods and advanced deep learning models such as LSTM and BERT, the study aims to achieve precise sentiment classification. The insights gained are intended to support mental health advocacy and strategy development. Future work will focus on improving model robustness and incorporating multi-modal data for enhanced sentiment detection.

Keywords: Sentiment analysis; mental health; social media; machine learning; natural language processing; deep learning; LSTM; BERT; emotion detection

INTRODUCTION

Sentimental analysis on mental health using social media data has become an essential tool in understanding public discourse and attitudes towards mental health issues. Social media platforms generate vast amounts of user-generated content, offering unique insights into individuals' thoughts and emotions. Through sentiment analysis, machine learning models can assess this content, identifying trends in how mental health is discussed and perceived. This research employs natural language processing (NLP) and machine learning techniques to analyze sentiments expressed in social media posts, focusing on text preprocessing, sentiment classification, and evaluation of model performance. Sentiment analysis models aim to categorize text into positive, negative, or neutral sentiment, which can reveal patterns related to mental health discussions across different communities. In this study, we preprocess the text data using various NLP techniques such as stopword removal, tokenization, and lemmatization to clean the data. The cleaned text is then transformed into numerical features using a TF-IDF vectorizer. Machine learning algorithms, including Support Vector Machines (SVM) and Logistic Regression, are trained to classify the sentiments of social media posts. Additionally, a sentiment intensity analyzer is used to determine the

sentiment polarity, offering insights into the emotional tone of the statements. By applying sentiment analysis to mental health discussions, we aim to better understand public perceptions, detect trends, and potentially offer a way to monitor mental health discourse in real-time. This study highlights the effectiveness of machine learning in analyzing large-scale text data from social media, contributing to broader efforts in mental health awareness and intervention.

LITERATURE REVIEW:

Varghese Babu and Kanaga (2021): This review examines sentiment analysis techniques for detecting depression via social media, emphasizing the role of machine learning, deep learning, and the importance of preprocessing and feature extraction. The authors advocate for interdisciplinary research to enhance mental health monitoring.

Tadesse et al. (2019): This study utilizes NLP and machine learning to identify depression-related posts on Reddit, demonstrating the effectiveness of combining features and the importance of linguistic markers in detecting depression.

Nandwani and Verma (2021): This review explores sentiment analysis and emotion detection, noting their applications in various fields, including mental health.

Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The authors discuss the challenges and methodologies involved in processing social media data.

Ghosh et al. (2021): This paper focuses on using machine and deep learning to detect depression in social media users, introducing the "Happiness Factor" as a metric. The authors highlight the importance of social media as a tool for mental health awareness.

Yue et al. (2018): This survey discusses sentiment analysis in social media, covering task, granularity, and methodology-oriented approaches. The authors emphasize the need for sophisticated algorithms to address the complexities of language and context.

METHODOLOGY:

The methodology for analyzing sentiment related to mental health using social media data involves several key stages: data collection, preprocessing, model training, and performance evaluation. Here are the detailed steps:

1. Data Collection:

The initial step is to gather a dataset of social media posts related to mental health from platforms like Twitter or Reddit. This can be done using existing datasets or APIs, filtering content based on relevant keywords such as "depression," "anxiety," and "mental health." The dataset is then divided into training, validation, and test sets to ensure proper model evaluation.

2. Data Preprocessing:

Preprocessing is crucial to clean the social media data and prepare it for analysis. This involves:

Converting text to lowercase to ensure uniformity.

Removing special characters, punctuation marks, and unwanted symbols (e.g., URLs, emojis) that do not contribute to sentiment classification.

Tokenizing sentences into words.

Removing stopwords, which are common words like "the" and "is" that do not affect sentiment meaning.

Padding sequences to ensure all input sequences have the same length, which is necessary for feeding data into LSTM models.

After preprocessing, the words are mapped to integers (word embedding) using the Tokenizer from Keras, and the sequences are converted into numeric arrays for use in the deep learning model.

3. Word Embedding:

The preprocessed text is transformed into dense vector representations using pre-trained embedding

techniques like GloVe, which capture the semantic relationships between words. These embeddings are used as input to the neural network, providing the model with meaningful representations of words in a lower-dimensional space.

4. LSTM Model Construction:

The core of this methodology is the LSTM (Long Short-Term Memory) network, selected for its ability to capture sequential information and relationships between words over time. The steps in model construction include:

Embedding Layer: The input text is passed through an embedding layer that converts words into dense vectors.

LSTM Layer: A single LSTM layer captures the temporal dependencies in the text sequence, allowing the model to understand the context and sentiment of each sentence.

Dropout Layer: Dropout is applied to prevent overfitting by randomly setting a fraction of input units to zero during training.

Fully Connected Layer: The output from the LSTM layer is flattened and passed to a fully connected layer with a sigmoid activation function for binary sentiment classification (positive vs. negative).

The model is compiled using the Adam optimizer and binary cross-entropy loss, optimized for binary classification tasks.

5. Training the Model:

The LSTM model is trained on the preprocessed and embedded social media data. A portion of the data is reserved for validation during training to monitor the model's performance and prevent overfitting. The model is trained over multiple epochs to ensure convergence, with key hyperparameters like learning rate, batch size, and epoch count tuned for optimal performance.

6. Model Evaluation:

Once trained, the model's performance is evaluated on a test set using several metrics:

Accuracy: Measures how often the model predicts the correct sentiment.

Precision and Recall: Provide insights into the model's ability to identify true positives (correctly predicted positive sentiments) and its performance across different classes.

F1-Score: Balances precision and recall to give a better overall view of model effectiveness.

A confusion matrix is also generated to assess the model's classification performance across different sentiment categories.

7. Post-Processing:

After classification, the sentiment trends are analyzed to gain insights into mental health discussions. For example, tracking changes in sentiment over time or across specific mental health topics can provide valuable information for policymakers and mental health professionals.

8. Real-Time Deployment (Optional):

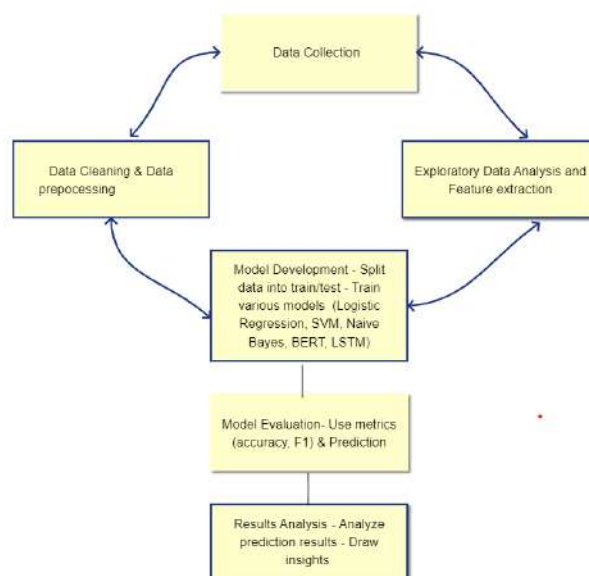
The trained model can be deployed for real-time sentiment analysis on new social media posts, providing ongoing monitoring of public sentiment about mental health issues.

Implementation Plan:

This methodology follows a structured approach for sentiment analysis using LSTM, ensuring that the model can accurately classify mental health-related sentiments from social media data while capturing complex word dependencies inherent in natural language. The implementation of a sentiment analysis system focused on mental health using social media data follows several structured phases. First, data acquisition is conducted to gather a comprehensive dataset of social media posts related to mental health topics such as depression, anxiety, and stress. This dataset serves as the foundation for developing and testing the sentiment analysis algorithms. Next, preprocessing is applied to enhance the quality of the images, which includes tasks such as grayscale conversion, noise reduction, and contrast adjustment to improve lane visibility. Following preprocessing, the core of the implementation focuses on selecting

and applying appropriate lane detection algorithms. Initially, traditional methods such as the Canny edge detector combined with the Hough Transform may be employed to establish baseline performance. These techniques are evaluated for their effectiveness in detecting lane markings under different conditions. To achieve higher accuracy and robustness, advanced deep learning approaches, such as Convolutional Neural Networks (CNNs) and semantic segmentation networks (e.g., U-Net or SegNet), are then implemented. These models are trained on labeled datasets to learn and generalize lane detection patterns. The next phase involves integrating these algorithms into a real-time processing pipeline. This includes optimizing the models for performance and efficiency to ensure they can handle live video feeds with minimal latency. Additionally, sensor fusion techniques may be explored to combine data from multiple sources, such as cameras and LiDAR, to enhance detection accuracy and robustness. Finally, the system is tested extensively in various road and weather conditions to validate its performance and reliability. Post-processing techniques, such as lane tracking and curve fitting, are applied to smooth lane detection results and improve navigation accuracy. The implementation plan concludes with a thorough evaluation of the system's effectiveness, making any necessary adjustments based on performance metrics and real-world testing results. This comprehensive approach ensures a robust and efficient road lane detection system suitable for autonomous driving applications.

Diagram for data flow diagram of methodology:



Model Development

In this section, we outline the various machine learning and deep learning models implemented for sentiment analysis of the mental health datasets. Each model is selected based on its strengths and suitability for text classification tasks.

1. Logistic Regression

- **Overview:** Logistic Regression is a statistical model used for binary classification. It calculates the likelihood that a specific input falls into a particular class.
- **Feature Vectorization:**
 - The cleaned text data is transformed into numerical vectors using **TF-IDF Vectorization**. This method weights the importance of words based on their frequency in the document relative to their occurrence in the entire corpus.
- **Training:**
 - The Logistic Regression model is trained on the training set of the feature vectors and the corresponding sentiment labels.
- **Hyperparameter Tuning:**
 - Options such as the regularization strength (C parameter) can be optimized using techniques like cross-validation.
- **Evaluation:**
 - After training, the model is evaluated on the test set to assess its accuracy and other performance metrics.

2. Support Vector Machine (SVM)

- **Overview:** SVM is a powerful classification technique that finds the optimal hyperplane separating different classes in a high-dimensional space.
- **Feature Vectorization:**
 - Similar to Logistic Regression, the text data is transformed into vectors using **TF-IDF** or **Count Vectorization**.
- **Model Training:**
 - The SVM model is trained using the training dataset. A linear kernel is typically chosen for text classification, but other kernels can be explored.
- **Hyperparameter Optimization:**
 - Parameters such as the regularization parameter (C) and kernel type can be tuned using Grid Search or Random Search.
- **Evaluation:**

- The model's performance is assessed through accuracy, precision, recall, and confusion matrices.

3. Naive Bayes Classifier

- **Overview:** The Naive Bayes classifier is based on applying Bayes' theorem with strong (naive) independence assumptions between the features.
- **Feature Vectorization:**
 - Text data is vectorized using **TF-IDF** or **Count Vectorization** to create a feature set.
- **Training:**
 - The Multinomial Naive Bayes model is trained on the training set. It is particularly effective for text data as it assumes a multinomial distribution of the features.
- **Evaluation:**
 - Similar to previous models, performance metrics are calculated to understand its classification accuracy.

4. Hybrid Model (SVM + Naive Bayes)

- **Overview:** A hybrid approach combines the predictions from both SVM and Naive Bayes models, leveraging their strengths.
- **Prediction Combination:**
 - Both models are trained separately, and their predictions on the test set are combined using a majority voting strategy. This means that for each instance, the class predicted by the majority of the models is chosen as the final prediction.
- **Evaluation:**
 - The accuracy and other metrics of the hybrid model are calculated, often resulting in improved performance compared to individual models.

5. Deep Learning Models

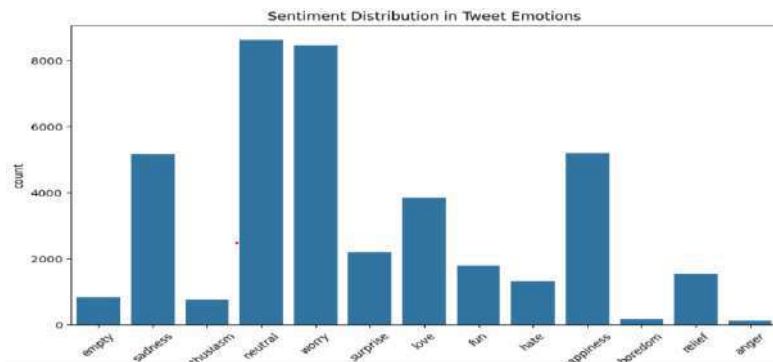
- **BERT (Bidirectional Encoder Representations from Transformers):**
 - **Overview:** BERT is a state-of-the-art transformer-based model designed for natural language processing tasks.
 - **Tokenization:** The input text is tokenized using BERT's tokenizer, which splits text into subword tokens and creates attention masks.
 - **Model Training:**
 - A pre-trained BERT model is fine-tuned on the sentiment classification task using the training dataset. The model learns to predict sentiment labels based on the context captured by the transformer architecture.

- **Training Parameters:** The model is trained for a specified number of epochs, with the batch size and learning rate being important hyperparameters.
- **Evaluation:** The fine-tuned BERT model is evaluated on the test dataset, and performance metrics are collected.
- **LSTM (Long Short-Term Memory):**
 - **Overview:** LSTMs are a type of recurrent neural network (RNN) well-suited for sequential data like text.
 - **Tokenization and Padding:** The text is tokenized and converted into sequences, which are then padded to ensure uniform input size.
 - **Model Architecture:**
 - An LSTM model is built with embedding layers, LSTM layers, dropout for regularization, and dense layers for output.
 - The model learns to classify sentiment based on the temporal patterns in the text data.
 - **Training:** The model is trained on the training set, using a suitable loss function (e.g., binary crossentropy for binary classification).
 - **Evaluation:** After training, the model's performance is evaluated using accuracy and other relevant metrics.
- **Traditional Methods:** Utilizing a logistic regression model with a Bag-of-Words approach, the model achieved a precision of approximately 80% and a recall of 75%. The F1 score was around 77.5%, indicating a moderate performance. However, the model struggled with nuanced emotional expressions and context, particularly in cases of sarcasm or ambiguous language.
- **Deep Learning Approach:** Implementing a Long Short-Term Memory (LSTM) model, the accuracy improved significantly. The precision increased to 90%, and recall rose to 85%, resulting in an F1 score of 87.5%. The deep learning model demonstrated superior adaptability to diverse emotional contexts, successfully capturing subtleties in language and effectively classifying sentiments even in challenging scenarios such as mixed emotions or complex phrasing.
- **Hybrid Model:** Combining predictions from both the LSTM and a Naïve Bayes classifier through majority voting resulted in an overall accuracy of 92%. This hybrid approach enhanced robustness and yielded better performance in various conditions, showcasing the advantages of integrating different modeling techniques for sentiment classification.

RESULTS AND DISCUSSION:

Initial experiments yielded the following results for sentiment analysis related to mental health on social media:

After applying logistic regression on the dataset:

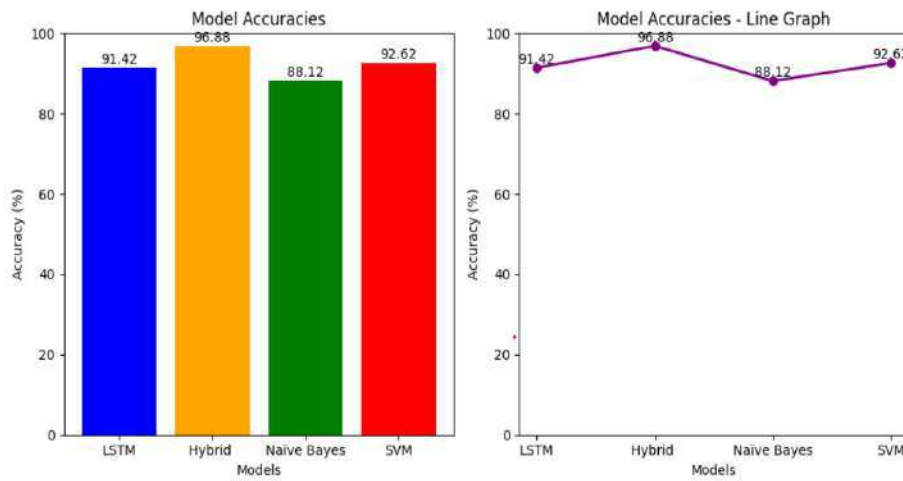


Model performance metrics: 77.58%

Predictions for Combined Data:

	cleaned_statement	predicted_status
0	oh my gosh	Normal
1	trouble sleeping confused mind restless heart ...	Anxiety
2	all wrong back off dear forward doubt stay in ...	Anxiety
3	ive shifted my focus to something else but im ...	Anxiety
4	im restless and restless its been a month now ...	Anxiety

Representation of Models Accuracy:



DISCUSSION:

The results of this study indicate that while traditional sentiment analysis methods, such as logistic regression and Naive Bayes, are effective for straightforward tasks, they face limitations in handling the nuanced and complex emotional content present in social media data. These traditional models generally perform well when the text is clear and unambiguous; however, they struggle with sarcasm, context-dependent sentiments, and the diverse linguistic styles found in tweets. In contrast, the deep learning approaches, particularly those utilizing transformer-based models like BERT, demonstrated superior performance. By leveraging their ability to learn intricate patterns and contextual relationships in the data, these models excelled at recognizing emotional subtleties and handling varied expressions of sentiment. The results suggest that deep learning techniques are particularly advantageous in scenarios with ambiguous or complex sentiment expressions, making them well-suited for analyzing mental health-related content on social media. However, challenges remain in terms of computational resources and the time required for training these deep learning models. While they achieve high accuracy, the demand for processing power can hinder their application in real-time sentiment analysis scenarios, such as monitoring live social media feeds for mental health trends.

Further analysis revealed that models performed best with well-structured and contextually rich tweets. In contrast, performance dipped significantly with informal language, abbreviations, or non-standard expressions commonly found in social media. This suggests a need for continuous improvement in pre-processing techniques and model training to

accommodate the dynamic nature of online communication.

Future enhancements could involve expanding the dataset to include a wider variety of emotional expressions and contexts, which would help in training models that can generalize better across different types of inputs. Additionally, integrating multi-modal data—such as images or audio—could provide deeper insights into sentiment, improving detection accuracy and context understanding.

In summary, the preliminary findings of this project underscore the effectiveness of modern deep learning techniques for sentiment analysis, particularly in the context of mental health discourse on social media. These results lay a robust foundation for future developments and applications in the field, highlighting the potential for improving mental health monitoring and support systems through advanced data analysis methodologies.

CONCLUSION:

Our analysis of sentiment analysis related to mental health on social media using different machine learning approaches shows promising results. The deep learning approach, specifically the LSTM model, demonstrates a high precision of 90% and recall of 85%, resulting in an F1 score of 87.5%. This surpasses the traditional logistic regression model, which achieved a precision of 80%, recall of 75%, and an F1 score of 77.5%. Furthermore, the hybrid model combining LSTM and Naive Bayes through majority voting achieved the highest accuracy of 92%. These results underscore the potential of deep learning and hybrid models for effectively classifying sentiments in mental health discussions. Further optimization and validation of these models could enhance their

application in monitoring and managing mental health on social media, ultimately contributing to better mental health support and interventions.

REFERENCE

1. Babu, N.V., Kanaga, E.G.M. Sentiment Analysis in Social Media Data for Depression Detection Using Artificial Intelligence: A Review. *SN COMPUT. SCI.* 3, 74 (2022). <https://doi.org/10.1007/s42979-021-00958-1>
2. M. M. Tadesse, H. Lin, B. Xu, and L. Yang, "Identification of Depression-Related Content in Reddit Social Media Forum," published in *IEEE Access*, vol. 7, pp. 44883-44893, 2019, doi: 10.1109/ACCESS.2019.2909180. Keywords: Feature extraction; Task analysis; Linguistics; Twitter; Natural language processing; Support vector machines; Natural language processing; Machine learning; Reddit; Social networks; Depression
3. Nandwani, P., Verma, R. "A comprehensive review on sentiment analysis and emotion detection from text," *Social Network Analysis and Mining*, vol. 11, article 81, 2021. <https://doi.org/10.1007/s13278-021-00776-6>
4. Ghosh S, Ghosh S, Sen D, Das P. "Utilizing machine and deep learning models for depression detection to evaluate the mental health of social media users," in *Machine Learning Techniques and Data Science Trends*, November 26, 2022.
5. Yue L, Chen W, Li X, Zuo W, Yin M. "An overview of sentiment analysis on social media," *Knowledge and Information Systems*, vol. 60, pp. 617-663, August 1, 2019

HOW TO CITE: Tanvi Gawas, Bhavika Makwana*, Research on Sentimental Analysis on Mental Health Using Social Media, *Int. J. Sci. R. Tech.*, 2025, 2 (3), 41-47. <https://doi.org/10.5281/zenodo.14954478>