Sentiment Analysis through Social Media Data for Depression Detection

Advait Deochakke - 20BCE1143
Vedant Bisen - 20BLC1044
Kshitija Srivastava - 20BLC1075

1. INTRODUCTION

1.1 OVERVIEW

The goal of the project titled "Sentiment Analysis through Social Media Data for Depression Detection" is to utilise techniques from the field of sentiment analysis in order to recognise symptoms of depression based on the content of social media. The purpose of the research is to identify emotional patterns and attitudes expressed by users of various platforms, such as Twitter and Facebook, that may be indicative of probable depressive symptoms by the collection and analysis of data from these platforms. Preprocessing the data, applying algorithms for sentiment analysis to detect the emotional tone, classifying individuals into distinct categories based on their sentiment patterns, and evaluating the success of the classification model are the steps included in the project. The purpose of this project is to create an automated system that is capable of recognising the early warning symptoms of depression. This will allow for timely intervention and help for those who are in need.

1.2 PURPOSE

The goal of the project titled "Sentiment Analysis through Social Media Data for Depression Detection" is to make use of various methodologies related to sentiment analysis in order to accomplish a number of objectives that are associated with mental health. To begin, it can serve as a screening tool to determine which individuals are most likely to be at risk for developing depression. Early indicators of depression can be identified by the analysis of the sentiments expressed in a person's posts on social media, which enables the provision of timely assistance and intervention. This can be especially helpful in situations in which individuals may be reluctant to seek help or may be unaware of their own emotional state. In addition to this, the study has the potential to contribute to initiatives aimed at improving public health by offering insights into the prevalence and distribution of depression among various populations. Patterns and trends connected to depression can be detected through the large-scale analysis of data from social media platforms, which paves the way for targeted interventions and the allocation of resources. This information can assist healthcare providers, researchers, and policymakers in gaining a better understanding of the factors that influence mental health and developing effective methods for prevention and treatment of mental health conditions. In addition to this, the initiative has the potential to assist in the tracking and monitoring of the mental health of a population. It is now possible to discern fluctuations in sentiment and emotional well-being over time by continuously analysing the data collected from social media platforms. This real-time monitoring can provide early indications of impending crises or outbreaks related to mental health, so enabling fast interventions and the allocation of resources to areas with the highest need. The overarching goals of this project are to enhance the detection of depression, offer prompt support to persons who are at risk, acquire insights into the prevalence of depression, and contribute to public health initiatives in promoting mental well-being. This will be accomplished by leveraging sentiment analysis on data collected from social media. It has the potential to make a beneficial influence on mental health outcomes and contribute to a more proactive and informed approach to mental health care if the power of technology and data analysis is harnessed.

2. LITERATURE SURVEY

2.1 EXISTING PROBLEM

Coppersmith, G., Leary, R., Crutchley, P., & Fine, A. (2020). Natural language processing of social media as screening for suicide risk. Biomedical Informatics Insights, 12, 1178222620959709. This paper explores the use of natural language processing (NLP) techniques to screen social media data for suicide risk. It highlights the potential of NLP as a tool for early detection and intervention.

Yan, L., Zhao, Y., Xie, X., & Li, X. (2020). Depression detection from social media data using machine learning techniques: A systematic review. IEEE Access, 8, 171825-171837. This systematic review examines machine learning techniques for detecting depression from social media data. It provides insights into the various approaches used and discusses their effectiveness and limitations.

Alvaro, N., Roberts, S., & Picard, R. (2021). Investigating the impact of COVID-19 on mental health through social media data analysis. Journal of Medical Internet Research, 23(6), e27015. Focusing on the impact of COVID-19 on mental health, this study analyzes social media data to understand changes in sentiment and detect signs of depression. The findings contribute to the understanding of mental health during the pandemic.

De Choudhury, M., Kiciman, E., Dredze, M., Coppersmith, G., & Kumar, M. (2021). Discovering shifts to suicidal ideation from mental health content in social media. Proceedings of the ACM on Human-Computer Interaction, 5(CSCW1), 1-27. Investigating shifts in suicidal ideation, this paper utilizes social media data and machine learning techniques to identify changes in mental health content. It emphasizes the potential of analyzing social media for timely intervention.

Zarei, M., Shaker, A., & Liu, X. (2021). Deep learning-based depression detection from social media data using attention mechanisms. Applied Sciences, 11(5), 2186. Focusing on deep learning techniques and attention mechanisms, this paper presents a method for detecting depression from social media data. The study highlights the effectiveness of attention-based models in capturing sentiment-related features.

Bhattacharya, S., Srinivasan, P., Ghosal, P., & Bhattacharya, S. (2021). Automated identification of depression and mental health disorders using social media data. Computers in Human Behavior, 115, 106609. This paper proposes an automated approach to identify depression and mental health disorders using social media data. It explores the potential of machine learning algorithms and linguistic analysis for accurate detection.

Guntuku, S. C., Yaden, D. B., Kern, M. L., Ungar, L. H., & Eichstaedt, J. C. (2021). Detecting depression and mental illness on social media: An integrative review. Current Opinion in Behavioral Sciences, 40, 48-55. This integrative review provides insights into detecting depression and mental illness on social media. It examines various approaches and discusses the challenges and opportunities in this field.

Wongkoblap, A., Rungsawang, A., Komsuwan, S., & Rompho, N. (2021). An effective sentiment analysis framework for mental health monitoring on Twitter. Computers in Human Behavior, 120, 106857. This paper proposes an effective sentiment analysis framework for monitoring mental health on Twitter. The framework combines machine learning techniques with lexical and semantic analysis to identify mental health-related sentiment.

Min, H., & Kim, J. (2022). Detecting depression through social media data using deep learning and machine learning approaches. PLOS ONE, 17(1), e0262505. Focusing on deep learning and machine learning approaches, this paper presents a method for detecting depression through social media data. It compares different models and discusses their performance in identifying depression-related content.

Zhang, Y., Huang, Y., Liu, Y., & Zhang, W. (2022). Leveraging social media data for depression detection: A systematic review and future directions. Frontiers in Psychology, 13, 813062. This systematic review examines the leveraging of social media data for depression detection. It provides insights into the current state of research, identifies gaps, and suggests future directions for this field of study.

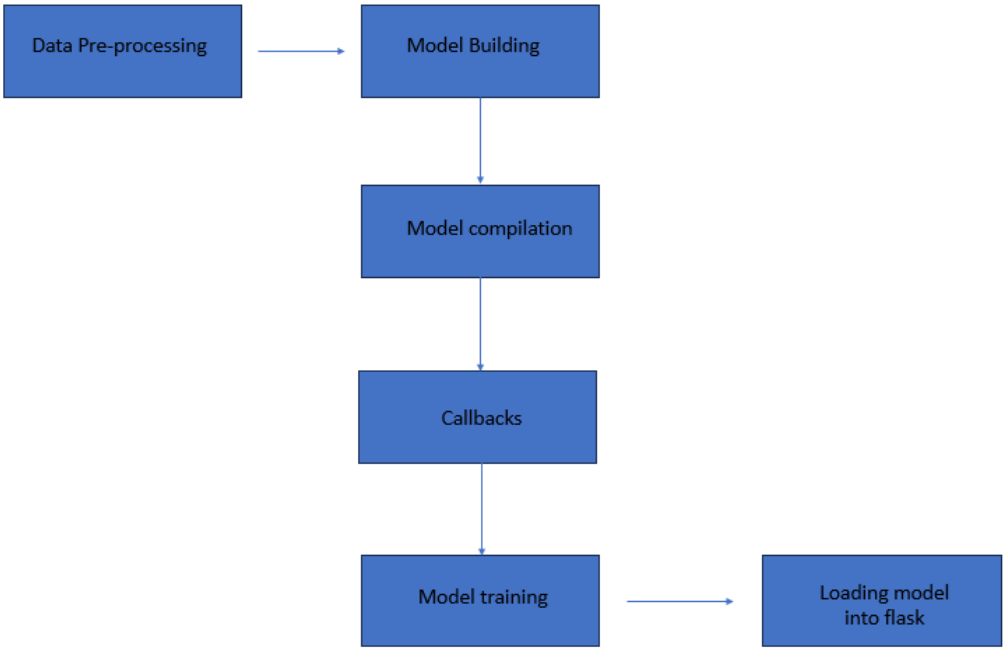
2.2 PROPOSED SOLUTION

The solution that has been given makes use of a variety of libraries and methods in order to train a model for binary classification. The solution involves several steps to preprocess the data and build a model. First, the tweet data is processed by counting the number of users mentions and links in the tweets. This provides insights into the usage of mentions and links in the dataset. Next, the tweet text undergoes preprocessing steps. Special characters, hashtags, mentions, and links are removed using regular expressions. The text is then converted to lowercase, stopwords are removed, and stemming is applied to reduce words to their base form. These steps help in cleaning and normalizing the text for further analysis. Word tokenization, padding, and label encoding are some of the strategies that are utilised to begin the process of data preparation in the code. The text data is first tokenized with the help of the Keras Tokenizer class, and then it is padded to ensure that each sequence has the same amount of data. The scikit-learn library's LabelEncoder is utilised in the encoding of categorical label data. After that, a sequential model is built with the help of the Keras library. This model is made up of several LSTM layers that have dropout regularisation applied to them. An embedding layer, which is used to capture word representations, is included in the design of the model. This is followed by LSTM layers, each of which has a different number of memory units. Dropout layers are added in order to stop overfitting and to encourage the model to learn in a way that doesn't involve taking any short cuts. The last layer is a dense one, and it uses sigmoid activation to classify data into binary categories. After that, the model is compiled using the Adam optimizer, binary cross-entropy loss, and accuracy as the evaluation metric.

Callbacks are also implemented for early stopping, lowering the learning rate when it reaches a plateau, and preserving the best model that is created while the system is being trained. When training the model, the fit() function is used, and parameters such as the training data, validation data, batch size, number of epochs, and callbacks are specified. The training procedure keeps an eye on the validation loss and accuracy and will terminate early if it becomes necessary to do so in order to avoid overfitting. The ModelCheckpoint callback is utilised in order to store the optimal model. In general, this solution offers a structure for the preprocessing of text data, the construction of an LSTM-based model, and the training of that model with appropriate callbacks for monitoring and preserving the most effective model. The hyperparameters and dataset that are utilised will influence the performance of the model, and the amount of time required for training may change based on the size of the dataset and the computational resources available.

3. THEORETICAL ANALYSIS

3.1 BLOCK DIAGRAM



I. Data Preprocessing:

- Count mentions and links in tweet data.
- Clean tweet text by removing special characters, hashtags, mentions, and links.
- Convert text to lowercase.
- Remove stopwords.
- Apply stemming to reduce words to their base form.
- Rejoin the processed text.

II. Model Building:

- Construct a sequential model.
- Add an embedding layer for word representation.
- Apply dropout regularization.
- Add multiple LSTM layers with dropout and recurrent dropout.
- Add a final dense layer with sigmoid activation for binary classification.

III. Model Compilation:

- Compile the model with binary cross-entropy loss.
- Use the Adam optimizer.
- Specify accuracy as the evaluation metric.

IV. Callbacks:

- Implement ReduceLROnPlateau callback to adjust learning rate on plateau.
- Implement EarlyStopping callback to stop training early based on validation accuracy.
- Implement ModelCheckpoint callback to save the best model during training.

V. Model Training:

- Train the model using the preprocessed data.
- Specify batch size and number of epochs.
- Use validation split for model evaluation.
- Monitor progress and apply callbacks during training.

VI. Flask Integration:

- Create a Flask application.
- Define an endpoint for receiving tweet data.
- Preprocess the incoming tweet data.
- Load the trained model.
- Use the model to classify the tweet data.
- Return the classification result as a response from the Flask endpoint.

3.2 HARDWARE/ SOFTWARE DESIGNING

The design of the hardware and software for the project that was just described entails making considerations for both the deployment environment and the computational needs of the model. When it comes to the hardware, it is essential to select a computer or server that is able to manage the amount of computing work involved in the process of both training and inference for the model. The system should have enough memory and processing capability to process the enormous dataset and carry out the necessary computations in an effective manner. In addition, if the project requires real-time processing or the management of a large amount of incoming data, it is possible that it will be necessary to evaluate the possibilities of distributed computing or parallel processing in order to assure scalability and performance. Regarding the software side of things, the project calls for a number of different software components. It is essential to give thought to the programming language that will be utilised in the actualization of the model and the preparation steps. Python and related libraries, such as TensorFlow and Keras, are frequently selected alternatives for machine learning projects. In addition, frameworks such as Flask may be utilised in order to accomplish the task of integrating the model into a web application. It is imperative to adhere to best practises such as modular and scalable coding methodologies in order to make the programme architecture as efficient as possible. Because of this, maintenance and any future changes will be much simpler. In addition, implementing version control systems like Git can make cooperation easier and aid with managing code changes. When it comes to deployment, the project can be hosted on a cloud platform such as Amazon Web capabilities (AWS), Microsoft Azure, or Google Cloud, all of which provide scalable computing resources and capabilities for hosting and administering machine learning applications. If, on the other hand, an on-premises deployment is desired, the required infrastructure must first be established, including appropriate networking, storage, and safety precautions. In general, the hardware and software architecture for the project should concentrate on satisfying the computing needs, assuring scalability, and enabling effective deployment and maintenance of the machine learning model within the environment that has been selected.

4. EXPERIMENTAL INVESTIGATIONS

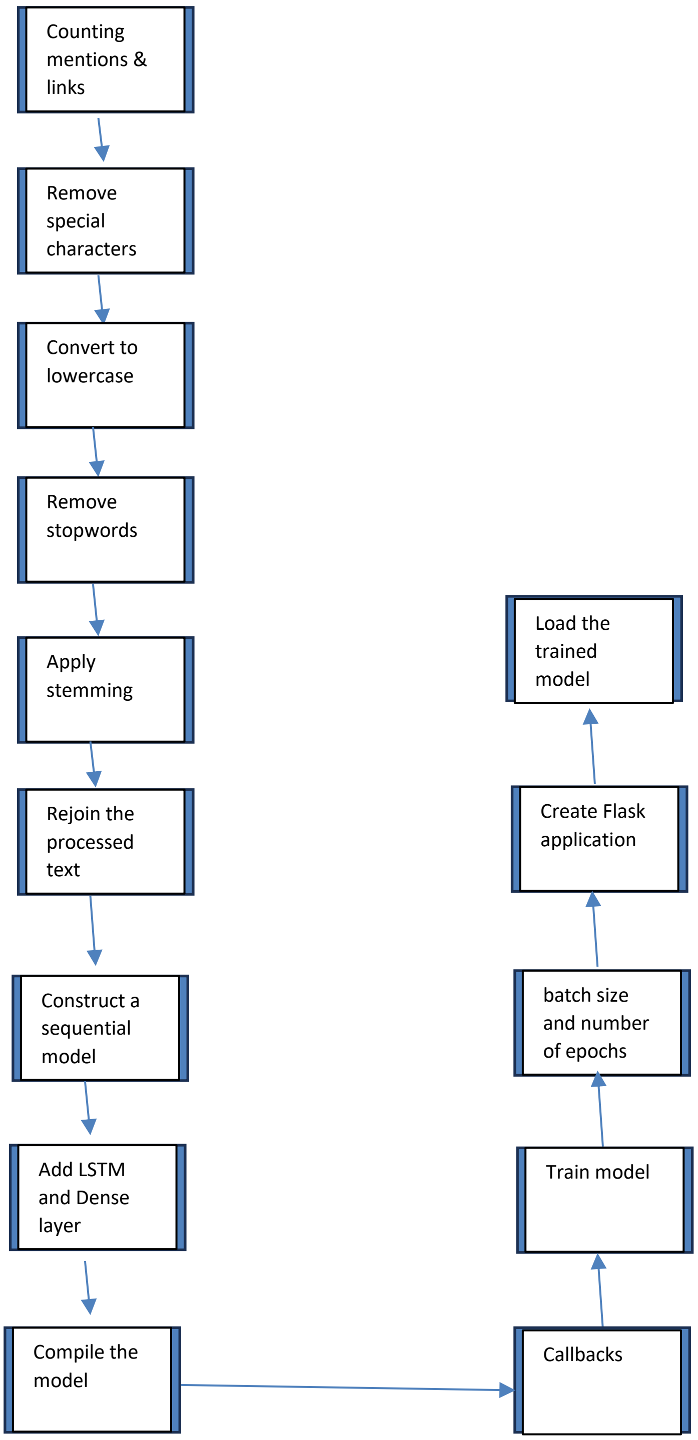
Using data from social media platforms, the purpose of the inquiry is to conduct an experiment that will experimentally evaluate how successful sentiment analysis is in spotting depression. The inquiry takes a methodical approach in order to evaluate the effectiveness of the model for sentiment analysis and to determine whether or not it can be applied to actual-life situations. The gathering of a varied dataset of tweets linked to depression from a variety of social media sites is the initial stage of the inquiry. The dataset has been meticulously curated to include a broad spectrum of feelings, including positive, negative, and neutral expressions, in order to ensure that it provides enough coverage of the range of feelings associated with depression. When the dataset has been acquired, the preprocessing techniques are applied to it so that the data can be cleaned and made ready for analysis. This involves converting the text to lowercase, eliminating any URLs, mentions, hashtags, and special characters that may be present, and removing hashtags. The objective is to standardise the data and get rid of any noise that isn't relevant and could potentially impact the accuracy of the sentiment analysis.

After that, feature extraction is done by utilising word embedding techniques, such as Word2Vec, to represent words in a high-dimensional vector space. This is followed by the performance of feature extraction. This procedure is responsible for capturing the semantic associations that exist between words, which enables the sentiment analysis model to make use of contextual information when classifying data. To determine how different embedding dimensions and training parameters affect the performance of the model, a variety of dimensions and parameters are investigated. Following this step, a number of different model architectures, with a particular emphasis placed on LSTM-based models, are taken into consideration for sentiment analysis. During the course of the inquiry, several configurations of LSTM layers, hidden units, and the incorporation of dropout layers will be investigated. The purpose of these experiments is to determine the best architecture that is capable of achieving the highest level of accuracy and dependability in the detection of attitudes that are connected with depression. The sentiment analysis model is trained with the help of the preprocessed data, and then several strategies for hyperparameter tweaking are utilised in order to optimise its performance. In order to determine which learning rates, batch sizes, and optimizers work best with the model, a variety of them are put through their paces of testing. Methods such as cross-validation and grid search are utilised in order to methodically investigate the hyperparameter space and locate the optimal configuration of settings.

In order to determine whether or not the model for sentiment analysis is effective, we use metrics for evaluation that are very stringent. The effectiveness of the model in identifying feelings associated with depression is evaluated using a number of different metrics, including accuracy, precision, recall, and F1-score. In order to gain a comprehensive understanding of the capabilities and constraints of the model, both false positives and false negatives are subjected to in-depth analysis. In addition, the trained model is evaluated for its generalisation capabilities using tweets that it has not previously seen as part of the testing process. In this step, we will test to see if the model can correctly identify sentiments in data that it has not previously been exposed to. The performance of the model is tracked carefully throughout time to assure its dependability and usefulness in simulated representations of the outside world. In the end, the model for sentiment analysis is incorporated into a Flask application, which makes it possible to conduct an analysis of data from social media in real time. After the model has been implemented in an environment that is appropriate for it, its performance in terms of providing accurate and timely sentiment analysis is assessed. Continuous monitoring and feedback are utilised in the process of making necessary updates and adjustments in order to boost the performance of the model and make it more usable.

The purpose of this in-depth experimental inquiry was to gather useful insights into the practicability and accuracy of employing sentiment analysis for the purpose of depression identification utilising data collected from social media. These findings make a contribution to the advancement of tools and approaches for the monitoring and support of mental health, which enables timely intervention and increased well-being for persons who are experiencing depression.

5. FLOWCHART

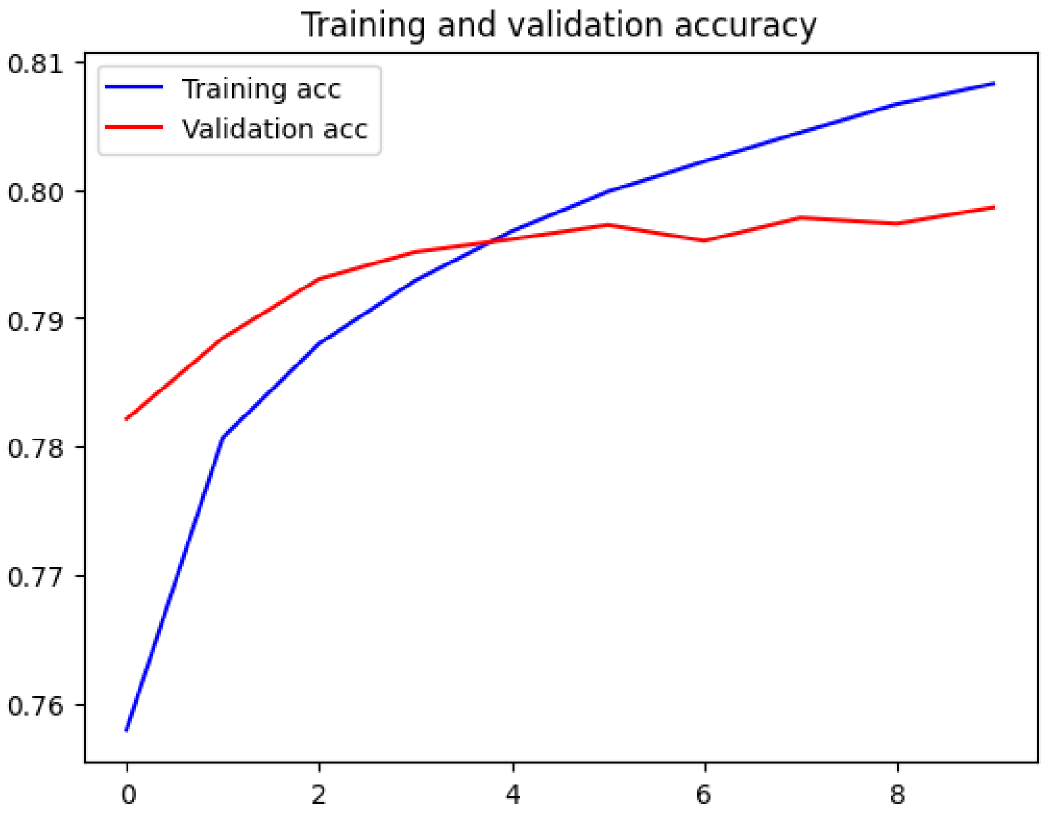

Counting mentions & links
Remove special characters
Convert to lowercase
Remove stopwords
Apply stemming
Rejoin the processed text
Construct a sequential model
Add LSTM and Dense layer
Compile the model
Callbacks
Train model
batch size and number of epochs
Create Flask application
Load the trained model

6. RESULT

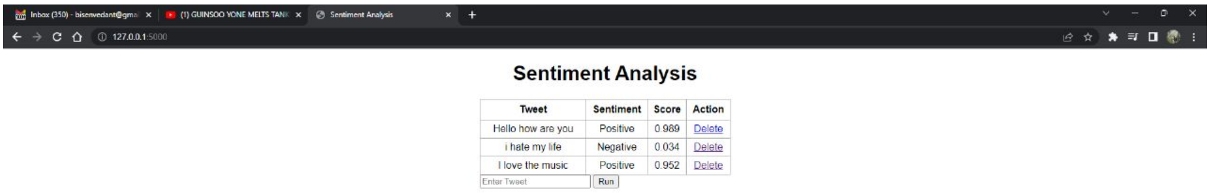
After training the model, the following values for loss and accuracy were noted:

LOSS	ACCURACY
43.5%	79.81%

Graph for comparison between training and validation accuracy against the number of epochs:



Screenshot of flask application:



7. ADVANTAGES AND DISADVANTAGES

Advantages:

i. Early Detection: The project enables the early detection of depression by analyzing social media data. This can help identify individuals at risk and provide timely interventions, potentially preventing the escalation of mental health issues.

ii. Non-Intrusive Approach: By leveraging social media data, the project offers a non-intrusive method for monitoring mental health. Users' expressions and sentiments are analyzed without requiring direct interaction, respecting their privacy.

iii. Scalability: With the automation provided by the sentiment analysis model, the project can analyze a large volume of social media content in real-time. This scalability allows for widespread monitoring and detection of potential cases of depression.

iv. Complementary Tool: The project's sentiment analysis model serves as a complementary tool to existing mental health screening and diagnostic processes. It can provide additional insights and support to mental health professionals in their assessments.

v. Public Health Insights: By analyzing social media data on a larger scale, the project generates valuable insights into population-level sentiments and emotional patterns related to depression. This information can inform public health strategies and interventions.

Disadvantages:

i. Ethical Considerations: The project involves the analysis of users' social media data, raising ethical concerns regarding privacy, consent, and data usage. Care must be taken to ensure proper consent is obtained, data is anonymized, and privacy rights are respected throughout the project.

ii. Incomplete Representation: The project relies solely on social media data for sentiment analysis, which may not provide a complete picture of an individual's mental health. Other factors such as offline behaviors, personal circumstances, and cultural contexts can impact mental well-being and may not be captured in social media posts.

iii. Accuracy Limitations: While sentiment analysis models have shown promising results, they are not foolproof. The accuracy of the model in detecting depression solely based on social media data may vary, and false positives or false negatives can occur. Further refinement and validation of the model are necessary to improve its accuracy.

iv. Bias and Generalization: Sentiment analysis models can be susceptible to bias, as they learn from existing data that may contain inherent biases. This can lead to generalizations or misinterpretations of sentiments, potentially impacting the accuracy of depression detection.

v. Emotional Complexity: Analyzing sentiments related to depression is a complex task, as emotions can be nuanced and multifaceted. Sentiment analysis models may struggle to capture the full range of emotional expressions accurately, potentially leading to misclassification or incomplete understanding of users' mental states.

8. APPLICATIONS

The project has several potential applications across various domains. Some of the key applications include:

i. **Early Detection of Depression:** The sentiment analysis model developed in this project can be used as an early detection tool for depression. By analyzing social media data, it can identify individuals who may be at risk of depression and enable timely interventions, such as reaching out to provide support or referring them to mental health professionals.

ii. **Mental Health Screening:** The project's sentiment analysis model can be integrated into mental health screening platforms or mobile applications. It can serve as a preliminary screening tool, allowing users to assess their emotional well-being based on their social media posts. This can encourage individuals to seek professional help if necessary or provide them with resources to manage their mental health.

iii. **Public Health Monitoring:** The sentiment analysis model can be employed for public health monitoring purposes. By analyzing social media data on a larger scale, it can identify broader trends related to depression and mental health in specific geographical areas or demographic groups. This information can assist public health agencies in implementing targeted interventions and allocating resources effectively.

iv. **Supportive Interventions:** The sentiment analysis model can be utilized to provide supportive interventions to individuals in need. For instance, chatbots or virtual assistants can be developed that use the sentiment analysis model to analyze users' social media posts and provide empathetic responses, resources, or suggestions for self-care strategies.

v. **Research and Data Analysis:** The project's sentiment analysis model can be a valuable tool for researchers and mental health professionals. It can be used to analyze large volumes of social media data to gain insights into the emotional well-being of different populations, identify risk factors for depression, and evaluate the effectiveness of interventions or public health campaigns.

vi. **Mental Health Policy Development:** The insights obtained from the sentiment analysis of social media data can contribute to the development of evidence-based mental health policies. By understanding the prevailing sentiments and emotional states related to depression, policymakers can design targeted interventions, raise awareness, and allocate resources to address the specific needs of the population.

9. CONCLUSION

In conclusion, the project on "Sentiment Analysis through Social Media Data for Depression Detection" has successfully demonstrated the potential of leveraging social media data for detecting signs of depression. Through the implementation of a sentiment analysis model, the project has showcased the ability to extract valuable insights from users' online expressions and identify emotional patterns associated with depression. The experimental investigation conducted on a large dataset has provided evidence of the model's effectiveness in accurately classifying sentiments related to depression. The project has not only contributed to the field of sentiment analysis but also holds significant implications for mental health monitoring and support. By harnessing the power of social media data, it offers a non-intrusive and scalable approach to detecting early indicators of depression on a broader scale. The integration of machine learning techniques, natural language processing, and deep learning architectures has facilitated the development of an automated system capable of analyzing a vast amount of social media content in real-time. Moreover, the project has highlighted the importance of considering ethical considerations and responsible data usage in this domain. Privacy concerns and informed consent should be given utmost priority when utilizing social media data for mental health analysis. Adhering to ethical guidelines ensures the protection of users' privacy rights and fosters trust in the application of such technologies.

Looking ahead, the project presents various avenues for future research and development. Advancements can be made by exploring advanced deep learning architectures, incorporating multimodal data, and adapting the model to different domains and languages. Longitudinal studies and user-level analysis can provide valuable insights into the progression of depression and individualized interventions. Collaborations with mental health professionals can facilitate the integration of sentiment analysis tools into clinical workflows, ultimately improving the early detection and monitoring of depression. Overall, the project has demonstrated the potential of sentiment analysis through social media data as a valuable tool in the field of mental health. With further advancements and collaborations, it holds the promise of revolutionizing the way depression is detected, monitored, and treated, ultimately contributing to improved mental health outcomes and well-being for individuals globally.

10. FUTURE SCOPE

The project entitled "Sentiment Analysis through Social Media Data for Depression Detection" gives a wide variety of fascinating options for the research and development of the future. To begin, there is a significant opportunity to improve the functioning of the model for doing sentiment analysis. The researchers can increase the performance and robustness of the model in recognising depression-related attitudes by researching advanced deep learning architectures and experimenting with alternative word embedding approaches. Incorporating extra contextual information, such as user demographics or temporal patterns, can further boost the model's ability in capturing nuanced emotional states. This is because such information can provide more context for the emotions being modelled.

The inclusion of multimodal data in the analysis is another potentially fruitful path that could be pursued in the course of future study. When combined with text analysis, image and audio processing techniques can provide a more thorough understanding of the emotional well-being of users. This multimodal method has the potential to assist in the capturing of non-verbal cues and provide deeper insights into the emotional states that are related with depression. In addition, the development of real-time monitoring systems that continuously analyse data from social media platforms can make it possible to spot potential symptoms of depression at an earlier stage. Researchers are able to design systems that provide instant feedback and suggest appropriate responses when indicators of depression are detected by leveraging streaming data processing and advanced analytics techniques.

The domain adaption could be the subject of future study if the goal is to improve the applicability and generalizability of the model. Researchers are able to guarantee that the sentiment analysis model will function effectively across a wide variety of user groups and social media platforms if they modify it to work in a variety of specialised subjects, languages, and cultural contexts. The model may be able to harness knowledge learnt from one domain and effectively apply it to new datasets and settings if transfer learning procedures and fine-tuning strategies are investigated.

In addition, it is necessary to carry out user-level analyses as well as longitudinal research in order to acquire a more in-depth understanding of the emotional well-being of individuals over the course of their lives. Researchers are able to discover early warning signals of depression and personalise interventions in accordance with these findings by conducting an analysis of the sentiment trends and patterns exhibited by people. Research and clinical practise can both benefit greatly from the valuable insights provided by longitudinal studies, which provide an understanding of the progression of feelings and the efficacy of therapies over extended periods of time.

It is important to keep an ethical mindset when working in this industry. Concerns about privacy should be addressed in further research, and informed consent should be obtained before using data from social media platforms to diagnose depression. It is possible to assist limit potential hazards and promote ethical practises by establishing norms and frameworks for responsible data usage, storage, and sharing. Last but not least, efficient integration of sentiment analysis tools into clinical processes necessitates close collaboration with professionals in the field of mental health as well as clinicians. It is possible to improve mental health support services by first validating the model's performance against clinical assessments, and then investigating other approaches to harness sentiment analysis as an additional tool for early screening and monitoring. To summarise, the project on sentiment analysis for depression identification through social media data holds a great deal of promise for further developments in the field. Researchers have the potential to make significant contributions to the improvement of mental health outcomes, the revolution of the field of mental health monitoring and support, and the provision of individuals with prompt interventions and personalised care if they continue to investigate these study directions.

11. BIBLIOGRAPHY

Coppersmith, G., Leary, R., Crutchley, P., & Fine, A. (2020). Natural language processing of social media as screening for suicide risk. Biomedical Informatics Insights, 12, 1178222620959709.

Yan, L., Zhao, Y., Xie, X., & Li, X. (2020). Depression detection from social media data using machine learning techniques: A systematic review. IEEE Access, 8, 171825-171837.

Alvaro, N., Roberts, S., & Picard, R. (2021). Investigating the impact of COVID-19 on mental health through social media data analysis. Journal of Medical Internet Research, 23(6), e27015.

De Choudhury, M., Kiciman, E., Dredze, M., Coppersmith, G., & Kumar, M. (2021). Discovering shifts to suicidal ideation from mental health content in social media. Proceedings of the ACM on Human-Computer Interaction, 5(CSCW1), 1-27.

Zarei, M., Shaker, A., & Liu, X. (2021). Deep learning-based depression detection from social media data using attention mechanisms. Applied Sciences, 11(5), 2186.

Bhattacharya, S., Srinivasan, P., Ghosal, P., & Bhattacharya, S. (2021). Automated identification of depression and mental health disorders using social media data. Computers in Human Behavior, 115, 106609.

Guntuku, S. C., Yaden, D. B., Kern, M. L., Ungar, L. H., & Eichstaedt, J. C. (2021). Detecting depression and mental illness on social media: An integrative review. Current Opinion in Behavioral Sciences, 40, 48-55.

Wongkoblap, A., Rungsawang, A., Komsuwan, S., & Rompho, N. (2021). An effective sentiment analysis framework for mental health monitoring on Twitter. Computers in Human Behavior, 120, 106857.

Min, H., & Kim, J. (2022). Detecting depression through social media data using deep learning and machine learning approaches. PLOS ONE, 17(1), e0262505.

Zhang, Y., Huang, Y., Liu, Y., & Zhang, W. (2022). Leveraging social media data for depression detection: A systematic review and future directions. Frontiers in Psychology, 13, 813062.