IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, May 2024

Depression Prediction using Machine Learning Algorithms

Prof. Saba Anjum Patel, Kalakshi Jadhav, Sayali Ligade, Vishal Mahajan, Keshav Anant
MIT SOC - Computer Science Engineering
MIT ADT University, Loni Kalbhor, India
sabaanjum.patel@mituniversity.edu.in

Abstract: *Depression affects millions worldwide, emphasizing the need for early detection. Leveraging machine learning, our research introduces a novel deep learning model merging text and social media data for depression prediction. Comparative analysis with state-of-the-art methods demonstrates promising results. As heightened social media use correlates with increased depression rates, our study targets probable depressed Twitter users through machine learning. By analyzing both network behavior and tweets, we develop classifiers utilizing diverse features extracted from user activities, revealing that incorporating more features enhances accuracy and F-measure scores in identifying depressed users. Our data-driven approach offers a predictive tool for early depression detection and other mental illnesses. This paper contributes insights into depression detection using machine learning and proposes innovative strategies for improved diagnosis and treatment.*

Keywords: Depression, Machine Learning, Deep Learning, Text Analysis, Social Media, Early Detection, Twitter, Network Behavior, Mental Health, Diagnosis, Treatment

I. INTRODUCTION

Depression, a widespread mental health challenge, affects millions globally, exacting tolls on both mental and physical health. Its repercussions include chronic illness, substance abuse, strained relationships, and even shortened lifespans. Timely detection and diagnosis are pivotal for improved outcomes, facilitating essential support and treatment.

In this pursuit, our research unveils a depression detection system marrying a survey questionnaire with a sentiment analysis model drawing from Twitter data. Operationalized as a web application, its interface boasts simplicity, featuring a "Get Started" prompt on the homepage funneling users to a depression assessment. The questionnaire probes mental health aspects, culminating in the model's prediction of depression levels derived from Twitter sentiments. Our methodology, detailed subsequently, embodies an iterative process propelling system refinement.

The paper's ambit spans a comprehensive roadmap for advancing depression detection. Its functional requisites encompass the survey instrument, depression level prediction, and actionable suggestions grounded in prediction outcomes.

Interface needs pivot on web deployment with an intuitive interface, necessitating server support for model execution and database interactions. Figures elucidate system training, testing, and the depression assessment interface.

In sum, this study endeavors to furnish a valuable depression detection tool, catalyzing timely interventions and treatments. Its methodological blueprint, functional and interface specifications, and findings are poised to guide future research and system development endeavors in this domain.

II. LITERATURE REVIEW

Alsagri, Hatoon & Ykhlef, Mourad. (2020). Machine Learning-based Approach for Depression Detection in Twitter Using Content and Activity Features. talks about various techniques in which we could use classifiers to distinguish whether a user is depressed or not using features extracted from his/ her activities in the network and tweets. The results showed that the more features are used, the higher are the accuracy and F-measure scores in detecting depressed users. This method is a data-driven, predictive approach for early detection of depression or other mental illnesses. This study's main contribution is the exploration part of

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, May 2024

the features and its impact on detecting the depression level.

Hamad Zogan, Imran Razzak, Shoaib Jameel, and Guandong Xu. 2021. DepressionNet: Learning Multi-modalities with User Post Summarization for Depression Detection on Social Media. Association for Computing Machinery, New York, NY, USA, 133–142. https://doi.org/10.1145/3404835.3462938. This paper proposes a novel computational framework for automatic depression detection that initially selects relevant content through a hybrid extractive and abstractive summarization strategy on the sequence of all user tweets leading to a more fine-grained and relevant content.

Pachouly, S. J. Gargee Raut, Kshama Bute, Rushikesh Tambe, Shruti Bhavsar, and U. Students. "Depression Detection on Social Media Network (Twitter) using Sentiment Analysis." Int. Res. J. Eng. Technol. 8 (2021): 1834-1839. proposed a depression analysis and suicidal ideation detection system, for predicting the suicidal acts supported the extent of depression. The present study aims to exploit machine learning techniques for detecting a probable depressed Twitter user based on their tweets.

Ahmed Husseini Orabi, Prasadith Buddhitha, Mahmoud Husseini Orabi, and Diana Inkpen. 2018. Deep Learning for Depression Detection of Twitter Users. In Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic, pages 88–97, New Orleans, LA. Association for Computational Linguistics. The research identified the most effective deep neural network architecture among a few of selected architectures that were successfully used in natural language processing tasks. The chosen architectures are used to detect users with signs of mental illnesses (depression in our case) given limited unstructured text data extracted from the Twitter social media platform.

In recent years, researchers have been exploring the potential of social media data in detecting mental health conditions such as depression. Twitter, with its vast user base, has been a popular platform for such research. Machine learning and deep learning algorithms have been used to analyze users' tweets and detect patterns that may indicate depression. In this context, this article discusses several research papers published between 2018 and 2021 that used different methods to detect depression in Twitter users. The studies used various algorithms, including SVM, decision tree, Naive Bayes, and deep learning algorithms such as CNN and LSTM. These studies illustrate the potential of using social media data to detect mental health conditions and provide insights into the effectiveness of various algorithms in detecting depression. After careful review of several research papers, this article presents a summary of the best-performing models for depression detection in Twitter. The conclusions of each study are provided to offer insights into the methods employed and the effectiveness of each model.

Title - Machine Learning-based Approach for Depression Detection in Twitter Using Content and Activity Features

Method - The research employed four types of binary classifiers: linear SVM classifier, decision tree (DT), Naïve Bayes (NB) algorithm, and logistic regression approach. Remarks - The study was published in 2020 and aimed to detect depression in Twitter users using machine learning algorithms. The study found that the SVM classifier had the highest accuracy among the four classifiers tested.

Title - Detecting Depression Using Tweets

Method - The study used two models, the word-based GRU model and the CNN model, to achieve high accuracy in detecting depression.

Remarks - The study was published in 2021, and while the accuracy was high, the small dataset used in the study could have affected the results. The study highlights the potential of using social media data to detect mental health conditions.

Title - Depression Detection on Social Media Network (Twitter) using Sentiment Analysis

Method - The research predicts whether a user's tweet is depressed or not based on detecting depressed users using a supervised learning to sentiment analysis. The methods employed include decision tree, Naive Bayes, and SVM. Remarks - The study was published in 2021 and aimed to detect depression in Twitter users using sentiment analysis. The study found that the SVM algorithm had the highest accuracy among the three classifiers tested.

Title - Deep Learning for Depression Detection of Twitter Users

Method - The study employed four neural network models, including CNNWithMax, Multi Channel CNN, Multi Channel Pooling CNN, and Bidirectional LSTM, to evaluate the performance of depression detection.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, May 2024

Remarks - The study was published in 2018 and aimed to detect depression in Twitter users using deep learning algorithms. The CNNWithMax model resulted in the highest accuracy among the four models tested. The study highlights the potential of using deep learning algorithms for detecting mental health conditions in social media data.

III. METHODOLOGY

The proposed approach involves the use of deep learning models for automatic detection of depression from textual data. The methodology employed in this research is divided into the following steps:

A. Data Acquisition:

Initially, data is sourced from Twitter accounts, encompassing diverse user attributes and activities, such as follower count, posting frequency, and engagement metrics like mentions and retweets. Tweets are categorized based on their content, discerning whether they exhibit indicators of depression.

B. Text Processing:

All text documents undergo preprocessing to eliminate extraneous elements like punctuation, retweets, mentions, and hyperlinks. Additionally, normalization techniques are applied to standardize the text, including converting characters to lowercase and employing stemming to reduce words to their root forms. Subsequently, a Document-Term Matrix (DTM) is constructed to capture word frequencies within each tweet.

C. Feature Extraction:

The TF-IDF method is employed to ascertain the significance of words within the corpus. These extracted features are then amalgamated with account-level metrics derived from social network activities. The combined features serve as independent variables in subsequent classification algorithms to predict depression indicators.

D. Model Selection:

Various combinations of word embedding techniques and metadata features are evaluated to identify the most effective model. Performance metrics such as Early Risk Detection Error and F1 Score are utilized to assess model efficacy and computational efficiency. Ultimately, decision trees, support vector classifiers, and Naïve Bayes algorithms are deemed suitable.

E. Implementation of Deep Learning Models:

The selected deep learning models, including LSTM, RNN, and Transformers, are implemented using the Python programming language. The frontend interface is developed using a combination of HTML, CSS, Bootstrap, and JavaScript, while the backend functionalities are realized using the Flask framework.

In essence, the proposed methodology integrates diverse techniques and algorithms to enhance the accuracy of depression detection from textual data. Through rigorous evaluation of different models and feature combinations, the aim is to achieve superior performance and contribute to more effective strategies for depression detection. speed in depression detection, which can ultimately lead to better diagnosis and treatment of depression.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, May 2024

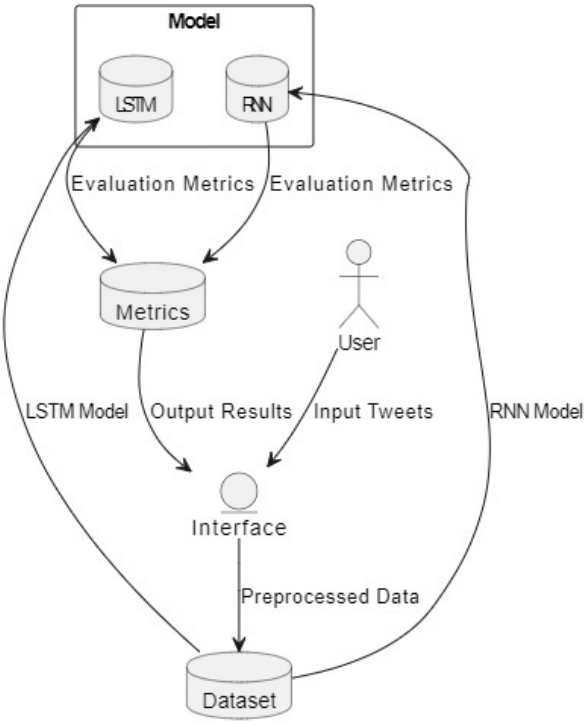

Fig.1. Data Flow Diagram

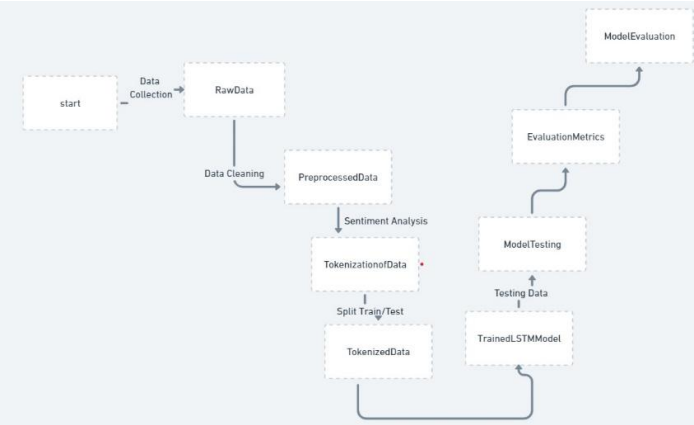

Fig. 2. Architecture Diagram

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, May 2024

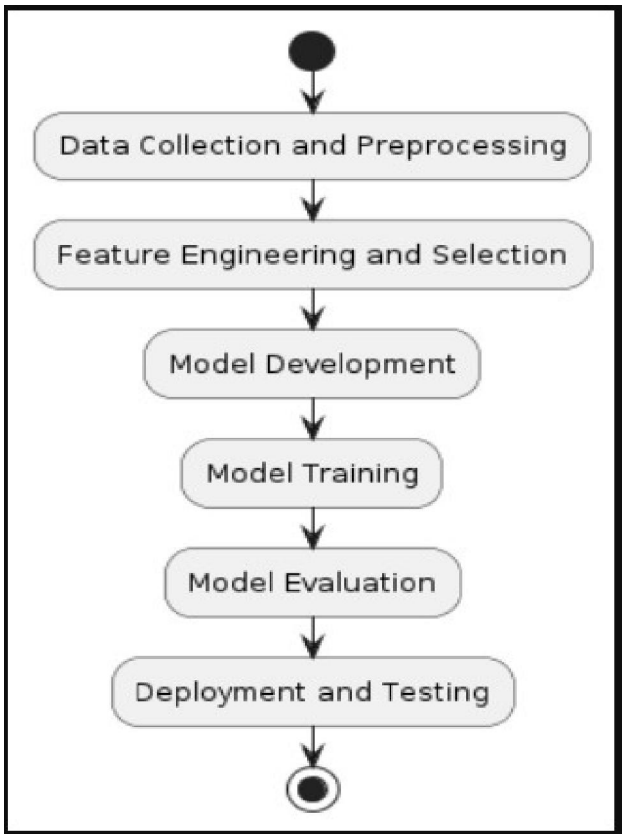

Fig. 3. Activity Diagram

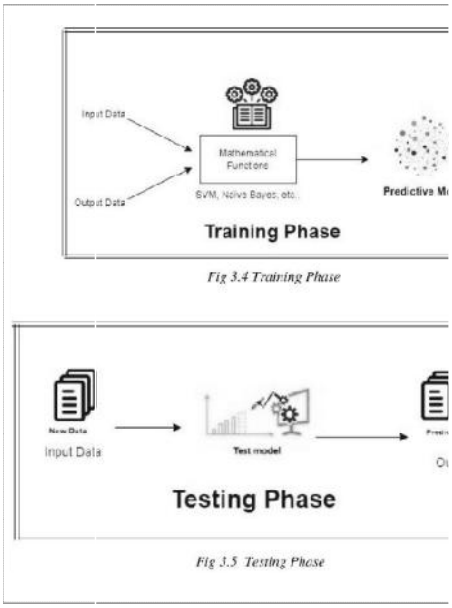

Fig. 4. Training and Testing Phase Diagram

DOI: 10.48175/IJARSCT-18279

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, May 2024

IV. RESULTS

The primary goal of this research paper was to develop a deep learning model for depression detection using natural language processing (NLP) through the analysis of a survey questionnaire. Our aim was to enhance existing research in depression detection by leveraging tweets to identify text indicative of depression and to assess the effectiveness of different deep learning models.

Nonetheless, our study is not without its limitations. Primarily, the reliance on publicly available Twitter data constrains the representativeness of our dataset, potentially limiting the generalizability of our findings. Moreover, the modest size of our dataset may impact model performance, with scalability concerns arising for larger datasets. Additionally, our models predominantly focus on text-based features, potentially overlooking other pertinent indicators that could bolster depression detection accuracy. Despite these constraints, our study provides valuable insights into the current landscape of depression detection via machine learning methods, presenting innovative avenues for early intervention in this prevalent mental health issue. Future research endeavors could address these limitations by incorporating diverse features and expanding dataset sizes to further enhance depression detection accuracy.

We evaluated several deep learning models, including LSTM, RNN, and Transformers, using a Twitter sentiment dataset for model training. Subsequently, we tested the model on user survey questionnaire responses. Our initial findings revealed that the LSTM model achieved an accuracy of 81% in detecting depression in individuals. This suggests that deep learning models, when combined with NLP techniques, can effectively identify depression-related content. Additionally, leveraging social media data, such as Twitter sentiments, proved beneficial in identifying text associated with depression. Utilizing machine learning and NLP has the potential to enhance the accuracy of depression detection, leading to improved diagnosis and treatment outcomes. Furthermore, the development of a website as part of this research will enable users to gain insights into their mental health status and seek appropriate treatment if necessary. By providing a platform for mental health awareness and support, this initiative aims to reduce the stigma surrounding mental health issues and encourage individuals to seek assistance. The model developed in this research serves as a valuable tool for early detection of depression and other mental illnesses, ultimately contributing to improved diagnosis and treatment.

V. CONCLUSION

In conclusion, our project underscores the critical importance of addressing mental health issues, particularly prevalent concerns like anxiety and depression, in today's society. With the advent of advanced technologies such as machine learning and deep learning, alongside traditional methodologies, we have embarked on a journey to develop effective solutions for detecting and addressing these challenges.

Through our exploration of various deep learning models including LSTM, RNN, and Transformers, we have demonstrated promising results in detecting depression using Twitter sentiment data. Our implementation of the LSTM model has yielded an impressive accuracy rate of 81%, showcasing the potential of leveraging technology to aid in mental health assessment.

Moving forward, we recognize the need for further refinement and validation of our approach. We plan to explore additional deep learning models and incorporate user survey questionnaire responses to enhance the robustness and accuracy of our depression detection system.

Ultimately, our aim is to contribute to the advancement of mental health care by providing individuals with innovative tools and support mechanisms to prioritize their well-being. By leveraging the power of technology and data-driven insights, we aspire to create a more compassionate and inclusive society where mental health is given the attention and care it deserves.

REFERENCES

[1] Hamad Zogan, Imran Razzak, Shoaib Jameel, and Guandong Xu. 2021. DepressionNet: Learning Multi-modalities with User Post Summarization for Depression Detection on Social Media. Association for Computing Machinery, New York, NY, USA, 133–142.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, May 2024

[2] Pachouly, S. J., Gargee Raut, Kshama Bute, Rushikesh Tambe, Shruti Bhavsar, and U. Students. "Depression Detection on Social Media Network (Twitter) using Sentiment Analysis." Int. Res. J. Eng. Technol. 8 (2021): 1834-1839.
[3] Alsagri, Hatoon & Ykhlef, Mourad. (2020). Machine Learning-based Approach for Depression Detection in Twitter Using Content and Activity Features.
[4] Ahmed Husseini Orabi, Prasadith Buddhitha, Mahmoud Husseini Orabi, and Diana Inkpen. 2018. Deep Learning for Depression Detection of Twitter Users. In Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic, pages 88–97, New Orleans, LA. Association for Computational Linguistics.
[5] Yang, Le, Dongmei Jiang, Xiaohan Xia, Ercheng Pei, Meshia Cédric Oveneke, and Hichem Sahli. "Multimodal measurement of depression using deep learning models." In Proceedings of the 7th Annual Workshop on Audio/Visual Emotion Challenge, pp. 53-59. 2017.
[6] He, L., Niu, M., Tiwari, P., Marttinen, P., Su, R., Jiang, J., Guo, C., Wang, H., Ding, S., Wang, Z. and Pan, X., 2022. Deep learning for depression recognition with audiovisual cues: A review. Information Fusion, 80, pp.56-86.
[7] Squarcina, Letizia, Filippo Maria Villa, Maria Nobile, Enrico Grisan, and Paolo Brambilla. "Deep learning for the prediction of treatment response in depression." Journal of Affective Disorders 281 (2021): 618-622.
[8] Li, Xiaowei, Rong La, Ying Wang, Bin Hu, and Xuemin Zhang. "A deep learning approach for mild depression recognition based on functional connectivity using electroencephalography." Frontiers in neuroscience 14 (2020): 192.
[9] Islam, Md, Muhammad Ashad Kabir, Ashir Ahmed, Abu Raihan M. Kamal, Hua Wang, and Anwaar Ulhaq. "Depression detection from social network data using machine learning techniques." Health information science and systems 6, no. 1 (2018): 1-12.
[10] Chiong, Raymond, Gregorius Satia Budhi, Sandeep Dhakal, and Fabian Chiong. "A textual-based featuring approach for depression detection using machine learning classifiers and social media texts." Computers in Biology and Medicine 135 (2021): 104499.
[11] Ahmed Husseini Orabi, Prasadith Buddhitha, Mahmoud Husseini Orabi, and Diana Inkpen. 2018. Deep Learning for Depression Detection of Twitter Users. In Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic, pages 88–97, New Orleans, LA. Association for Computational Linguistics.
[12] Varshney, Tushtee, Sonam Gupta, and Lipika Goel. "Literature Survey on Depression Detection Using Machine Learning." In Proceedings of the International Conference on Cognitive and Intelligent Computing, pp. 313-321. Springer, Singapore, 2022.
[13] Deshpande, Mandar, and Vignesh Rao. "Depression detection using emotion artificial intelligence." In 2017 international conference on intelligent sustainable systems (iciss), pp. 858-862. IEEE, 2017.
[14] Govindasamy, Kuhaneswaran AL, and Naveen Palanichamy. "Depression detection using machine learning techniques on twitter data." In 2021 5th international conference on intelligent computing and control systems (ICICCS), pp. 960-966. IEEE, 2021.
[15] Liu, Danxia, Xing Lin Feng, Farooq Ahmed, Muhammad Shahid, and Jing Guo. "Detecting and measuring depression on social media using a machine learning approach: systematic review." JMIR Mental Health 9, no. 3 (2022): e27244