Advanced Analysis of Depression in Women Using e-LSTM and All-Lexicon LR Classifiers

Jayaprakash B
School of Sciences Computer Science & IT,JAIN (Deemed to be University)
Bangalore,Karnataka, India
b.jayaprakash@jainuniversity.ac.in

Mandar Diwakar
AI & Ds Vishwakarma Institute of Technology
Pune , India
mandar.diwakar@viit.ac.in

A. Krishna Kumar
Department of Civil Engineering Karpagam Academy of Higher Education
Coimbatore- 641021, India
krishnakumar.a.kahe@kahedu.edu.in

Wamika Goyal
Centre of Research Impact and Outcome,Chitkara University
Rajpura- 140417, Punjab, India
wamika.goyal.orp@chitkara.edu.in

Divya N
Department of Civil, Prince Shri Venkateshwara Padmavathy Engineering College
Chennai – 127, India
divya.n_civil@psvpec.in

Bharat Bhushan
Chitkara Centre for Research and Development, Chitkara University
Himachal Pradesh-174103 India
bharat.bhushan.orp@chitkara.edu.in

2024 International Conference on Recent Advances in Science and Engineering Technology (ICRASET) | 979-8-3503-8860-2/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICRASET63057.2024.10895471

***Abstract* —Mental illness like depression is frequent. It is extended dissatisfaction, declining pleasure, or apathy to once-pleasurable activities. Not all mood swings are depression. It might influence family and neighbor ties. It may cause work or school issues. Depression affects everyone. Depression is increased by trauma like abuse or catastrophic loss. Depression affects women more than males. Survey data analysis may predict anxiety and depression. Self-harm may decrease with mental health diagnosis and prediction. Depression is emotional. It might be feelings. affect daily life, including wrath, sorrow, and despair. Also, common. According to CDC Trusted Source, 18.5% of Americans had depressed symptoms every two weeks in 2019. Depression is distinct from melancholy after a catastrophe or family loss, despite its similarities. Lose loved one. Grief should not make us feel bad or lose ourselves. Low self-esteem is common in depression. Many happy events accompany emotional distress. Remembering good times and pain. The following circumstances may aggravate depression: RA, asthma, CVD. Metabolic syndrome, obesity Understanding depression's nature is key. Everyone feels sad and miserable. The ensemble binary classifier beat all baseline strategies in all trials and metrics. Processed depression dataset trained proposed ensemble model. Study links women's expected quality of life markers to depression. The Ensemble model performed best with 75.64% accuracy and 0.7595 F1 Score. Integration of numerous models improves efficiency and generalizability. From 64.25% accuracy and 0.6595 F1 Score, the LSTM model performs 52.48% and 0.5513. Even the All Lexicons LR model, with an F1 result of 0.7296 and accurateness of 74.15%, behind the Ensemble model.**

Keywords- depression, postpartum depression (ppd), machine learning (ml), mental health, anxiety, self-harm, ensemble classifier.

I. INTRODUCTION

At the very least, the growth in the incidence of depression may be attributed, at least in part, to the fact that people's lives have changed over the course of the last century. Despite the fact that the rates of diagnosis have grown, there are still a large number of cases of mental illness that have not been diagnosed. For the purpose of locating individuals who are depressed or who are at risk for developing depression, it would be beneficial to make use of automated detection methods [1]. A depressive episode might leave them empty, resentful, and hopeless. Activity interest may diminish. Normal mood fluctuations are not depressive episodes. They remain most days for two weeks. More individuals, especially women, are worried, depressed, and anxious. Women's economic suffering may have diverse societal effects. PPD moms frequently experience physical and mental health issues.

The max postpartum depression severe. Even without depression-specific forecasting methods, ML may expose this often-obscured illness. In order to have a comprehensive knowledge of a diagnosis of depression, it is necessary to represent and investigate the qualities of language within the context of the diagnosis. This research focusses primarily on the detection of depressed symptoms via the utilization of text classifiers as its primary strategy. With the goal of improving depression identification, the goal of this explore is to compare and analysis the performance of hybrid and ensemble techniques using the idea of strengthening depression detection [2]. Ensemble models are considered to be more successful than hybrid models when it comes to finding solutions to classification problems. Selecting the features that provide the most benefits and then multiplying the total number of those features by a factor of two is one method for improving performance. A number of techniques, including ensemble methods, deep neural networks, emotion lexicons, and the identification of sadness

II. RELATED WORKS

More and more studies show that depression is becoming more common in today's culture; in fact, by 2030, experts expect that it will be major reasons of sickness universal. According to research, many people consider depression to be "a disease of modernity" because of the strong correlation between the disorder and contemporary ways of living [3-4]. Effective identification and intervention measures are crucial in tackling the growing incidence of depression, which is a significant public health problem.

Not only is depression stigmatized in society, but it is also infamously hard to identify and treat effectively. Misdiagnosis of depression is common, according to the authors, which makes therapy more difficult and has serious repercussions, such as increased risk of suicidal ideation in untreated patients [5]. It is crucial for the well-being of individuals and society as a whole to recognize mental illness early and get proper treatment since the social stigma associated with it makes these problems even worse.

A number of linguistic and cognitive science investigations have shown that depressed people use language differently. A non-invasive and easily accessible method of diagnosing mental health concerns, researchers have shown that these language signals may be used to diagnose depression [6]. The development of instruments that analyze language to determine the possibility of depression in people has been made possible by this finding.

Research into mental health, and especially depression, has shifted its focus to online social material. Revisions have shown that public mass media sites may be secondhand to detect indicators of depression by analyzing user-generated content such as thoughts and emotions expressed on these platforms. Social media research has the potential to shed light on people's feelings and choices to seek professional treatment, as authors have investigated ways to exploit user behavior on social networking sites to identify mental health concerns. references [7-8].

Now days the accuracy of depression identification is getting easier because to the merging of AI and NLP methods. Although sub symbolic AI methods are useful for uncovering statistical correlations and word frequencies, they are unable to understand conversation systems used for sentiment analysis, according to research [9]. A potential strategy for automated depression diagnosis has been found, however, by integrating sub symbolic approaches with symbolic learning procedures. This boosts the capacity of NLP systems to predict depressed symptoms more correctly.

A number of prediction tests have shown the efficacy of ensemble approaches, which integrate several learning processes. To address various prediction challenges, researchers have extensively used ensemble techniques. Continuing from previous developments, the present research improves depression detection as a text classification problem, thereby expanding our knowledge of automated depression identification. Each of the three datasets underwent eight analyses, with the goal of improving text categorization using ensemble approaches [10-13] by merging lexicon-based models with deep learning (DL) techniques. The research on automated depression diagnosis has been greatly advanced by hybrid approaches that use attention-LSTM and LSTM technique.

The article delves into the topic of automated depression identification via the use of text categorization algorithms derived from natural language processing (NLP). It presents the outcome of the preliminary statistics examination and summaries the information obtained from the experiments. The essay does more than just compare two models; it also discusses their weaknesses and suggests directions for future studies [14-16]. In order to evaluate textual content and determine users' linguistic preferences, natural language processing (NLP) algorithms have been created. Different machine learning methods are evaluated according to the effect they have on certain properties.

Methods like TF-IDF, N-gram, and LDA have been used in the pursuit of better classification accuracy via feature combination. When it came to performance, models based on Convolutional Neural Networks (CNNs) were determined to be superior than RNNs. This was especially true when CNN models were combined with optimized embeddings, which allowed for more generalization power. The efficacy of integrating feature-based approaches with sophisticated machine learning models for text categorization tasks is shown here [17].

Reusing language feature needs in contexts that are semantically relevant has been achieved via the use of ml and dl approaches [18]. Text categorization is approached from several angles, with characteristics based on relatedness and semantic similarity being among the most prominent. Several approaches may be used to examine latent semantics, including neural embedding, topic modelling, vector space modelling, along with neural networks. Discovering statistical characteristics in texts is also achieved via feature engineering and statistical analysis; these features are vital for improving the efficacy and precision of text categorization algorithms.

III. PROPOSED METHODOLOGY

Analysis of emotions, translation, and text summarization are just a few of the many sectors that may benefit from natural language processing technology. The secret Markov model was employed in earlier data-intensive efforts to circumvent these challenges. Lately, natural language processing methods have increasingly included DL. Thanks to advancements in machine learning (ML), which have helped with tasks like voice identification, recognition of images, and natural language processing, powerful computing platforms have recently made end-to-end training feasible. Using a bag-of-words representations of the texts, early text classification systems used ML methods without sequential word processing. LSTM neural networks rely on word embeddings; however, they have been used for text classification due to their ability to reduce the number of training samples. In order to classify texts according to their emotional content. Figure 1 shows the proposed ensemble model's flow diagram.

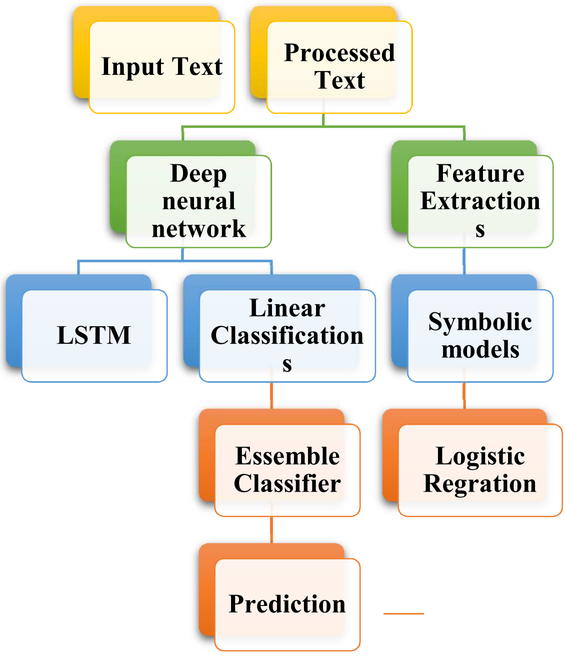

Fig. 1. Architecture of Proposed Work

A. Input data

The first step is to bring in raw text data from a variety of sources. Posts, comments, and other preprocessed text forms make up the bulk of this data set.

B. Processing the Text

Among the preprocessing stages that are applied to the raw text data are tokenization, which involves dividing text into words or tokens, stemming or lemmatization, which involves reducing words to their basic form, and changing text to lowercase. Stop words are also removed at this phase. Cleaning and preparing the text data for further analysis is a crucial part of the preprocessing.

C. Feature Extraction

Features are retrieved from text data after preprocessing. Some examples of such methods are N-grams, Word2Vec, GloVe, as well as TF-IDF, or term frequency inverse document frequency. To ensure that DL and ML models to process the text data, it has to be quantitatively represented.

D. Modelling using LR and LSTM or Attention-LSTM Deep Learning

Various models are then trained using the retrieved characteristics. For text categorization, the default ML model is Logistic Regression (LR). At the same time, the text's sequential and contextual information is captured using deep learning models such as Attention-LSTM and LSTM. It excels in deciphering word sequences with complex temporal connections.

Relevant LSTM Process Equations:

Input Gate:

$$\mathrm{j_z} = \sigma\,(\mathrm{w_z} \cdot \mathrm{x_z} + \mathrm{y_z} \cdot \mathrm{a_{z-1}} + \mathrm{b_z}) \tag{1}$$

Forget Gate:

$$\mathrm{f_z} = \sigma\,(\mathrm{w_f} \cdot \mathrm{x_z} + \mathrm{y_f} \cdot \mathrm{a_{z-1}} + \mathrm{b_f}) \tag{2}$$

Cell State Update:

$$\mathrm{G_z} = \mathrm{f_z} * \mathrm{G_{z-1}} + \mathrm{j_z} * \tanh\,(\mathrm{w_g} \cdot \mathrm{x_z} + \mathrm{y_g} \cdot \mathrm{a_{z-1}} + \mathrm{b_g}) \tag{3}$$

Output Gate:

$$o_t = \sigma\,(w_o \cdot x_t + U_o \cdot h_{t-1} + b_o) \tag{4}$$

Hidden Gate:

$$\mathrm{h_t} = \mathrm{O_t} * \tanh(\mathrm{c_t}) \tag{5}$$

E. Ensemble method

An ensemble approach is used to blend the outputs of LR, LSTM, and Attention-LSTM models. Combining the results of several models into a single more accurate forecast is the goal of techniques like bagging, boosting, and stacking. The ensemble technique takes use of each model's capabilities to improve forecast robustness.

F. Prediction and Evaluation

A forecast of the likelihood of experiencing depression was proven to be dependent on the text that was examined and this was the final result that was obtained. This output is compared to conventional measures such as recall, accuracy, precision, and F1-score in order to get an evaluation of the performance of the model. In order to determine how well the model performed, this comparison is being carried out.

G. Result Interpretation:

The efficiency of the model is understood by interpreting the forecasts. In order to prove that the ensemble model is better at detecting depression, it is compared to individual models. To make text-based depression diagnosis more accurate and resilient, the proposed ensemble model makes good use of several ML and DL approaches. Combining ensemble approaches with LSTM and Attention-LSTM makes for a more accurate and dependable prediction system.

IV. Result and Discussion

To name just a few of the numerous fields that could potentially profit from the application of natural language processing technology, we can include the analysis of emotions, translation, and text summarization. In prior data-intensive efforts to bypass these issues, the secret Markov model was utilized as they were being implemented. For some time now, natural language processing techniques have been gradually using DL.

The purpose of this research is to establish the feasibility of the combination of models that was provided is successful in the automated identification of depression. One of the criteria that is discussed in this study is a number of distinct criteria. Listed below, in addition to the parameters themselves, are the formulae that correspond to each of the parameters.

Accuracy: Relative to the total number of occurrences, the accuracy is the proportion of instances that are correctly classified.

$$\text{Accurancy} = \frac{\text{TP} + \text{TN}}{\text{TP+TN+FP+FN}} \tag{6}$$

Precision: When compared to the total number of optimistic observations that were anticipated, precision describes the degree to which a certain number of observations were successfully predicted.

$$\mathrm{P} = \frac{TP}{TP + FP} \tag{7}$$

Sensitivity: The proportion of genuine principles that were properly identified is referred to as the recall, and it is equal to the sensitivity.

$$Recall = \frac{TP}{TP + FN} \tag{8}$$

Through the process of comparing the outcomes of several models on a wide range of datasets, it is possible to identify patterns in the accuracy and efficiency of categorization. Carrying out this operation is a possibility. As can be seen in Table 1 and Figure 2, the attention-focused LSTM model obtains great results that are superior to those of its rivals across all metrics. This is the case on all fronts. It is true that this is the situation. The accuracy of this instrument is 0.6469, its threshold is 0.6648, and the F1 result is 0.6662. All of these specifications are accurate. It has a reliability of 67.88%, according to this instrument. Both Table 2 and Figure 2 demonstrate that the Ensemble model achieves the highest metrics, with an F1 result of 0.7595 and an accurate prediction rate of 75.64% respectively.

TABLE I. Comparative Analysis of Performance

Models	Accuracy	Precision	Recall	F1 Score
AFINN	0.5910	0.6436	0.6435	0.6431
NRCSA	0.5924	0.6439	0.6440	0.6437
LSTM	0.6425	0.6458	0.6584	0.6595
Attention LSTM	0.6788	0.6469	0.6648	0.6662

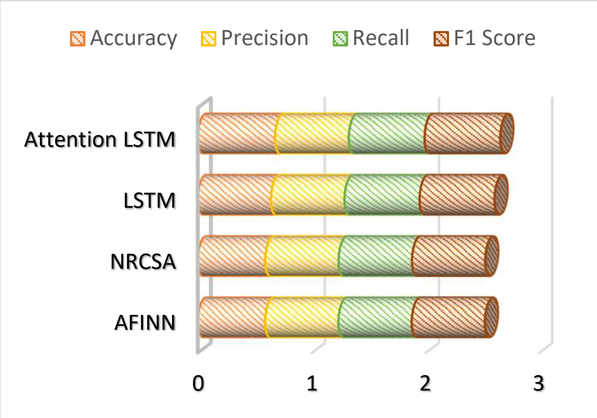

Fig. 2. Comparative Analysis of Performance Chart

TABLE II. A COMPARISON OF MULTIPLE DATASETS

Methods	Accuracy	Precision	Recall	F1 Score
LSTM	0.5248	0.5334	0.5119	0.5513
All Lexicons LR	0.7415	0.7415	0.7495	0.7296
Ensemble	0.7564	0.8125	0.7526	0.7595

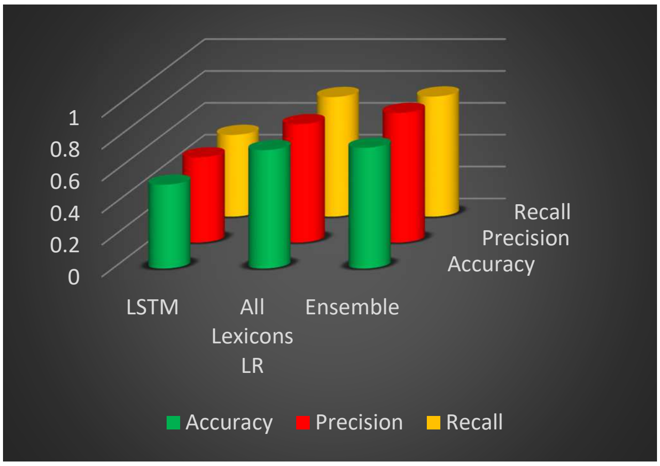

Fig. 3. Performance Analysis of Multiple Datasets

In sharp contrast to the statement that came before this one, this one is. These results represent, in comparison to the one that was achieved before them, a significant improvement. Even though its approach has a competitive reliability of 74.15% and an F1 result of 0.7296, the All The Lexicon definitions LR model does not have a degree of precision that is appropriate in respect to the overall performance of the ensemble. This is the case despite the fact that its methodology has a competing reliability.

This suggests that ensemble approaches often provide superior overall outcomes. Hybrid models, including those using LSTM and lexicon-based methods, may improve accuracy and F1 Scores by mixing varied features and methodologies, as this performance discrepancy shows. The LSTM model's efficacy seems to be affected by differences in datasets, as its performance decreases across various datasets (as seen in Table 2 and Figure 3). The ensemble technique, on the other hand, is resilient and generalizable since it consistently performs well across datasets. Ensemble approaches often provide better results, especially in difficult or diverse datasets, however lexicon-based models do provide useful insights.

V. CONCLUSION

This research demonstrates that the use of many social media datasets for the goal of diagnosing depression provides essential insights into the effectiveness of different text classification techniques. The outcomes of the study indicate that traditional models, such as Logistic Regression (LR), may still be superior than deep learning (DL) models if they make use of sentiment lexicons. This is in spite of the fact that deep learning models, particularly those that make use of complex architectures, have made substantial progress in the last several years. To be more specific, as compared to hybrid models that use both lexicon-based and deep learning approaches, ensemble models, which collect information from a wide variety of sources, demonstrate superior performance. Additionally, it is important to take into consideration that models that rely on a single set of lexical characteristics do less well than those that make use of many sets of lexical features. The importance of feature variety in enhancing classification accuracy is brought into focus by this particular instance. According to the findings of the study, even if the applications of feature sets improve classification performance, there is still a need for more research work. In spite of the fact that these findings have been made, this results in the following. It is possible that further text categorization models, such as transformer-based pre-trained language models and convolutional neural networks (CNNs), may be examined in future research. Additionally, in order to perhaps drive additional improvements in speed, it is possible to examine different strategies to overcome data imbalances using various methods.

REFERENCES

[1] Albert Rizzo, Russell Shilling, Eric Forbell, Stefan Scherer, Jonathan Gratch, and Louis-Philippe Morency. Autonomous virtual human agents for healthcare information support and clinical interviewing. 2016.

[2] H. Aldabbas, D. Albashish, K. Khatatneh. An Architecture of IoT-Aware Healthcare Smart System by Leveraging Machine Learning. Int. Arab J. Inf. Technol. 2022.

[3] M.A Alloghani, D. Al-Jumeily, J. Mustafina, Hussain.; Aljaaf, A systematic review on supervised and unsupervised machine learning algorithms for data science. In Supervised and Unsupervised Learning for Data Science, 2020;

[4] V. Roy and S. Shukla, "Mth Order FIR Filtering for EEG Denoising Using Adaptive Recursive Least Squares Algorithm," 2015 International Conference on Computational Intelligence and Communication Networks (CICN), 2015, pp. 401-404, doi: 10.1109/CICN.2015.85.

[5] P. K. Shukla, V. Roy, P. K. Shukla, A. K. Chaturvedi, A. K. Saxena, M. Maheshwari, P. R. Pal, "An Advanced EEG Motion Artifacts Eradication Algorithm", The Computer Journal, 2021;, bxab170, https://doi.org/10.1093/comjnl/bxab170.

[6] Rubayyi Alghamdi and Khalid Alfalqi, "A Survey of Topic Modeling in Text Mining", *International Journal of Advanced Computer Science and Applications*, vol. 6, no. 1, 2015.

[7] N.; Magdi, D.A.; Dahroug, A.; Rizka, M.A. Comparative Study: Different Techniques to Detect Depression Using Social Media. In Internet of Things-Applications and Future; Springer: Singapore, 2020.

[8] M. De Choudhury, S. Counts, E. J. Horvitz, and A. Hoff, "Characterizing and predicting postpartum depression from shared Facebook data," in Proc, 2018.

[9] V. Roy. " Breast cancer Classification with Multi-Fusion Technique and Correlation Analysis" Fusion: Practice & Applications, Vol. 9, No. 2, 2023 ,PP. 48-61.

[10] B. Zohuri and M. Moghaddam, "Business Resilience System (BRS): Driven through Boolean Fuzzy logics and cloud computation: Real and

near real time analysis and decision making system", *Business Resilience System (BRS): Driven Through Boolean Fuzzy Logics and Cloud Computation: Real and Near Real Time Analysis and Decision Making System*, pp. 1-425, 2017.

[11] R. S. Michalski, J. G. Carbonell and T. M. Mitchell. Deep learning for depression detection of twitter users. In Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic, New Orleans, LA, USA, 5 June 2018.

[12] M.N. Mohammad Hossain and Sulaiman. A review on evaluation metrics for data classification evaluations. International Journal of Data Mining Knowledge Management Process, 5(2):1, 2015.

[13] A. Sangr, D. Vlachopoulos and N. Cabrera, "Building an inclusive definition of e-learning: An approach to the conceptual framework", *Int. Rev. Res. Open Distrib. Learn.*, vol. 13, no. 2, pp. 145-159, 2012.

[14] T. Ayodele, C. A. Shoniregun and G. Akmayeva , Sentiment Analysis in Social Media Data for Depression Detection Using Artificial Intelligence,2015.

[15] J. Wulf, I. Blohm, J. M. Leimeister and W. Brenner, "Massive Open Online Courses", *Business Information Systems Engineering*, vol. 6, no. 2, pp. 111-114, 2014.

[16] S. Shukla, V. Roy and A. Prakash, "Wavelet Based Empirical Approach to Mitigate the Effect of Motion Artifacts from EEG Signal," 2020 IEEE 9th International Conference on Communication Systems and Network Technologies (CSNT), Gwalior, India, 2020, pp. 323-326, doi: 10.1109/CSNT48778.2020.9115761.

[17] D. O Callaghan, D. Greene, J. Carthy and P. Cunningham, "An analysis of the coherence of descriptors in topic modeling", *Expert Syst. Appl.*, vol. 42, no. 13, pp. 5645-5657, Aug. 2015.

[18] M. L. Sein-echaluce, F. J. Garca-pealvo and M. A. Condegonzlez. Identification of imminent suicide risk among young adults using text messages. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada, 21–26 April 2018.