DEPRESSION DETECTION FROM SOCIAL MEDIA DATA USING CNN AND LINGUISTIC METADATA FEATURES

J. RACHANA
Assistant Professor
Department of CSE
rachanajakkula@tkrec.ac.in
Teegala Krishna Reddy Engineering
College, Hyderabad

PULLABOINA SNEHA
Department of CSE
snehapullaboina@gmail.com
Teegala Krishna Reddy Engineering
College, Hyderabad

SATHYAM LAXMI VENKATA PRASAD
Department of CSE
prasadslv075@gmail.com
Teegala Krishna Reddy Engineering
College, Hyderabad

RESU SAI CHAND GOUD
Department of CSE
saichandresu2000@gmail.com
Teegala Krishna Reddy Engineering
College, Hyderabad

ABSTRACT: This project mainly concentrates on the comparison of various machine learning algorithms in predicting depression at an early stage. We take a popular reddit dataset as an input and train various models to find the best one among them. Programming language used is python and hybrid algorithms are helpful for detecting the most accurate algorithm that can be used to predict depression using various linguistic metadata features. Depression is a prevalent mental illness characterized by a depressed mood, stressful life experiences, and a sense of despair. It has an impact on your mood and capacity to operate, and it has the potential to lead to suicide. Depression is a substantial contributor to the worldwide burden of mental diseases and is a main cause of disability. According to studies, women are more likely than males to suffer from depression. Around 700,000 people commit suicide each year. Suicide is the fourth leading cause of mortality among people between the ages of 15 and 29. Depression is a common illness that affects 3.8 % of the world's population, with 5.0 % and 5.7 % of people over the age of 60 suffering from it. We target the early diagnosis of sadness in this study by applying several Machine Learning algorithms based on messages and posts on social media networks. Based on word embedding methods WORD2VEC, GLOVE, the Machine Learning algorithm LOGISTIC REGRESSION and the neural network algorithm CONVOLUTIONAL NEURAL NETWORK(CNN) are trained and compared to a classification-based user-level Linguistic metadata

INDEX TERMS: –— CNN (Convolutional Neural Network), Logistic Regression, Glove, Linguistic Metadata.

I. INTRODUCTION:

According to World Health Organization (WHO), more than 300 million people worldwide are suffering from depression, which equals about 4.4 percent of the global population. While forms of depression are more common among females (5.1 percent) than males (3.6 percent) and prevalence differ between regions of the world, it occurs in any age group and is not limited to any specific life situation. Latest results from the 2016 National Survey on Drug Use and Health in the United States report that, during the year 2016, 12.8 percent of adolescents between 12 and 17 years old and 6.7 percent of adults had suffered a major depressive episode (MDE). Precisely defining depression is not an easy task, not only because several sub-types have been described and changed in the past, but also because the term "being depressed" has become frequently used in everyday language. In general, depression can be described to lead to an altered mood and may also be accompanied, for example, by a negative self-image, wishes to escape or hide,

vegetative changes, and a lowered overall activity level. The symptoms experienced by depressed individuals can severely impact their ability to cope with any situation in daily life and therefore differ drastically from normal mood variations that anyone experiences. At the worst, depression can lead to suicide. They further speculate that the fear of discrimination might be relatively unimportant in their study because people hope to keep their psychiatric treatment secret. While depression and other mental illnesses may lead to social withdrawal and isolation, it was found that social media platforms are indeed increasingly used by affected individuals to connect with others, share experiences, and support each other. Based on these findings, peer-to-peer communities on social media can be able to challenge stigma, increase the likelihood to seek professional help, and directly offer help online to people with mental illness. All this emphasizes the importance of research toward ways to assist depressed individuals on social media platforms and on the internet in general. This paper is therefore focused on ways to classify indications of depression in written texts as early as possible based on machine learning methods

II LITERATURE SURVEY:

"Utilizing Neural Networks and Linguistic Metadata for Early Detection of Depression Indications in Text Sequences" As the detailed look at the current ERDE metric has shown, one priority of future work in this area should be to agree on a new metric for early detection tasks like Erick. Concerning the models presented in this work, additional experiments will be necessary to find better ways to integrate the metadata features directly into the neural network. On the other hand, utilizing ensembles of more than just two models and calibrating the resulting probabilities seems promising. Combining word embeddings of two models in a single neural network has also not been evaluated yet. Another possible improvement would be to use recently published language modelling methods like BERT as input for the network and to compare a self-trained model using this approach to the fast Text word embeddings of this work. "Universal language model fine-tuning for text classification" Authors have proposed ULM Fit, an effective and extremely sample-efficient transfer learning method that can be applied to any NLP task. We have also proposed several novel finetuning techniques that in conjunction prevent catastrophic forgetting and enable robust learning across a diverse range of tasks. Our method significantly outperformed existing transfer learning techniques and the state-of-the-art on six representative text classification tasks. We hope that our results will catalyze new developments in transfer learning for NLP. "A sensitivity analysis of convolutional neural networks for sentence classification" In this author used a neural network model for text classification. The author proposed Convolutional Neural Network model for sentence classification. The author concluded that CNNs are used to explore the effect of each part of the architecture on the performance. "Enriching word vectors with sub word information" The author said the usage of NLP to map words from vocabulary to a corresponding vector, used to find word predictions. Word vector algorithms like Word2Vec, Glove and many other algorithms used in NLP convert the text to the vector forms. "In an Absolute State: Elevated Use of Absolutist Words Is a Marker Specific to Anxiety, Depression, and Suicidal Ideation" Absolutist thinking is considered a cognitive distortion by most cognitive therapies for anxiety and depression. Yet, there is little empirical evidence of its prevalence or specificity. Authors predicted and found that anxiety, depression, and suicidal ideation forums contained more absolutist words than control forums. Suicidal ideation forums also contained more absolutist words than anxiety and depression forums. Authors show that these differences are more reflective of absolutist thinking than psychological distress. It is interesting that absolutist words tracked the severity of affective disorder forums more faithfully than negative emotion words. Finally, Authors found elevated levels of absolutist words in depression recovery forums. This suggests that absolutist thinking may be a vulnerability factor. "Detecting early risk of depression from social media user-generated content" Detecting depression at early stage from social media data is an important task. In this paper the author proposed that using supervised learning and information retrieval methods we can able to detect the depression at an early stage. "Basic concepts of Depression" Author said that the concepts involved in depression are complex. They have evolved over the years, and often, as is common in psychiatry, the meanings have changed subtly in the process. The core elements, and workable definitions for the disorder and its boundaries, are now well established. Some aspects of classification remain problematic, but the separation of bipolar and unipolar disorder was a major advance. Depressions are the most common disorders in psychiatry, both for psychiatrists and for general practitioners, so that understanding of their elements is important. "Delay and failure in treatment seeking

after first onset of mental disorders in the World Health Organization's World Mental Health Survey Initiative" The author has concluded that, failure and delays in treatment seeking were generally greater in developing countries, older cohorts, men, and cases with earlier ages of onset. These results show that failure and delays in initial help seeking are pervasive problems worldwide. Interventions to ensure prompt initial treatment contacts are needed to reduce the global burdens and hazards of untreated mental disorders. 6 "Understanding and Improving Convolutional Neural Networks via Concatenated Rectified Linear Units" In this paper, author aim to provide insight on the property of convolutional neural networks, as well as a generic method to improve the performance of many CNN architectures. Specifically, author first examine existing CNN models and observe an intriguing property that the filters in the lower layers form pairs (i.e., filters with opposite phase). Inspired by their observation, author propose a novel, simple yet effective activation scheme called Concatenated REL and theoretically analyses its reconstruction property in CNNs. "Stigma, agency and recovery amongst people with severe mental illness" This study is unique in which author examined reactions and responses to stigma. It allowed us to explore how far the intensive delivery of services, in combination with recovery housing, affected stigma. When interacting with this outside world, author noted that participants did not report high levels of directly experienced stigma. However, they engaged in conscious social and psychological strategies which revolved around being, acting and looking 'normal' to avoid and pre-empt both internal (felt) and external (enacted) stigma. "Distributed representations of words and phrases and their compositionality" Author proposed that recently introduced continuous Skip-gram model is an efficient method for learning high-quality distributed vector representations that capture a large number of precise syntactic and semantic word relationships. In this paper author present several extensions that improve both the quality of the vectors and the training speed. By subsampling of the frequent words author obtain significant speedup and also learn more regular word representations. "Naturally occurring peer support through social media: the experiences of individuals with severe mental illness using YouTube" At the time of this study, among the individuals who self-identified as having a SMI and who uploaded videos or posted comments to YouTube, author observed a sense of reward emerging from their interactions, mutual learning, and offered peer support. Of importance is our finding that peer support is happening naturally among individuals with highly stigmatized psychiatric illnesses within an unmonitored and public online platform. "The future of mental health care: Peer-to-peer support and social media" The author concluded that people with serious mental illness are increasingly turning to popular social media, including Facebook, Twitter or YouTube, to share their illness experiences or seek advice from others with similar health conditions. This emerging form of unsolicited communication among self-forming online communities of patients and individuals with diverse health concerns is referred to as peer-to-peer support. "Internet use and stigmatized illness" In this study the author, examined internet use and showed significant and consistent patterns of increased internet use among people with stigmatized illness when compared to those with non-stigmatized illnesses. These associations were stable across differing levels of comorbidities. These findings indicate that the internet may be a useful tool for contacting and communicating with people with a psychiatric stigmatized illness. "The Language of Depression" The author has concluded that the persons of different ages will have different languages when they are depressed. The author also said that people above 50 years have a different way of communication with the other people when they are depressed. This kind of depression is very dangerous and may lead to suicidal attempts. Also, the persons 35-40 have different language when they are depressed and 90% of the people can be easily cured with the medications.

III. EXISTING SYSTEM

The existing system uses Support vector machine (SVM) machine learning algorithm to train the model. The goal of the SVM algorithm is to create the best decision boundary that can segregate n-dimensional space. This system does not contain linguistic metadata features which can be used for early detection of depression.

Disadvantages: Less prediction, Lower accuracy

IV. PROPOSED SYSTEM:

In this project, we propose the model using WORD2VEC, Glove algorithm with logistic regression and CNN (Conventional Neural Networks). The proposed system uses linguistic metadata features that can be used as input to predict early depression from user's social media posts. Word embedding vector and Linguistic

metadata features will be concatenated and then input to CNN and Logistic regression model to build a depression detection model

Advantages: Higher prediction, Accuracy is more

V.SYSTEM ACHITECTURE

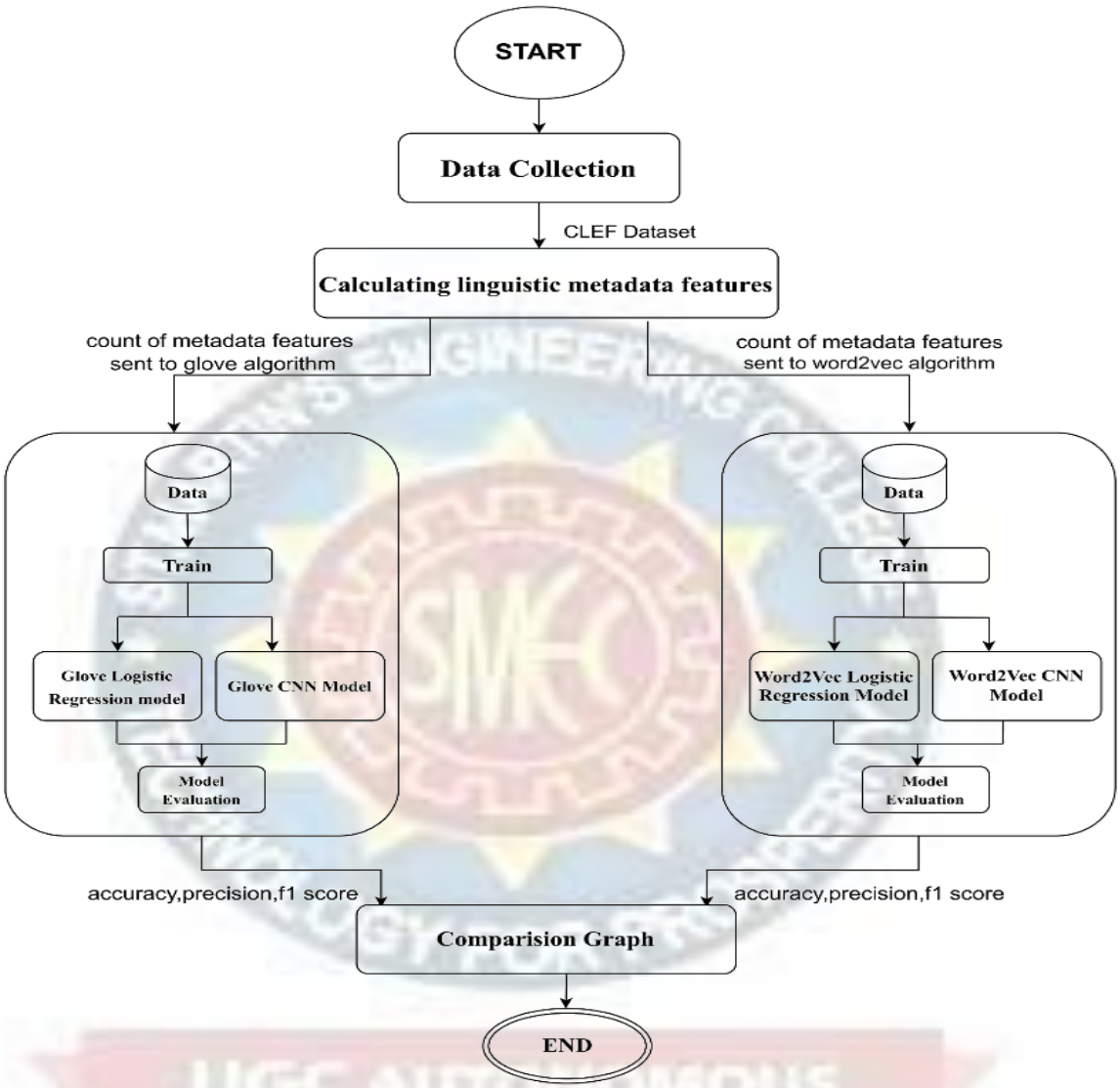
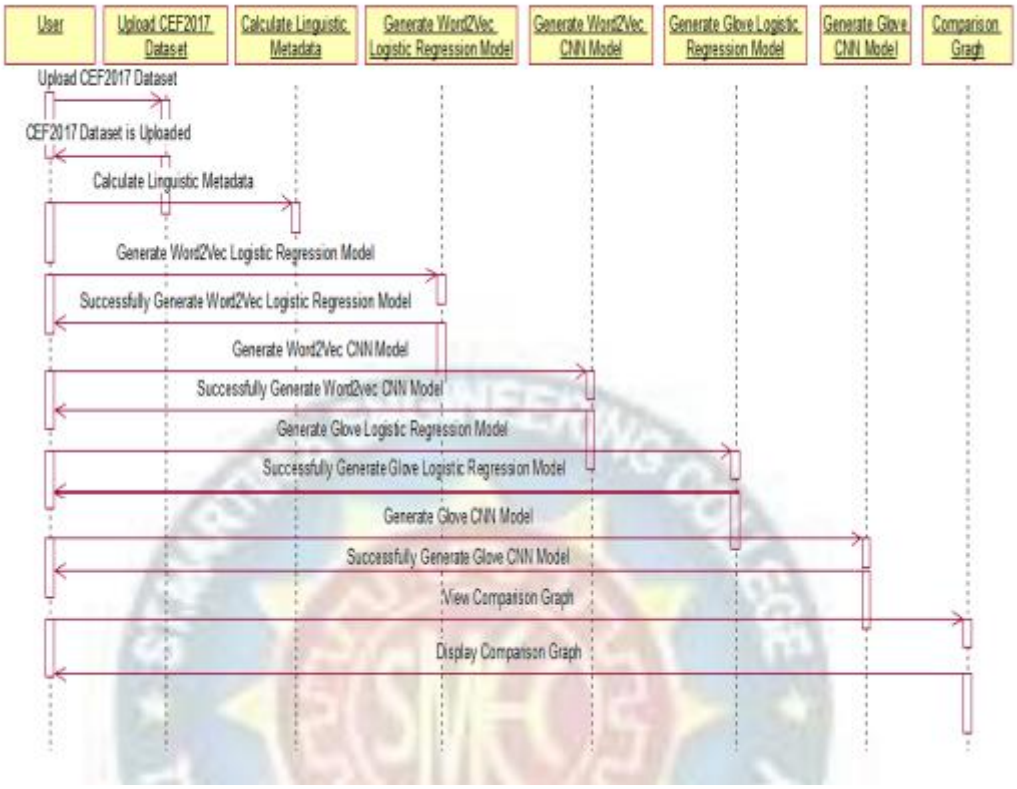


Fig.3.5: Architecture diagram for project planning

Block Diagram



Sequence Diagram

VI. MODULE IMPLEMENTATION:

Module 1: Glove Logistic Regression Model After calculating Linguistic metadata features, the output is sent to the Glove algorithm where it translates the text into respective vector forms and then applied Logistic regression Algorithm which is used for classification purpose. Therefore, it is a hybrid algorithm as it is combined of two algorithms.

Module 2: Glove CNN model the output of linguistic metadata features is given as input to the Glove algorithm which translates the text into the respective vector forms with the help of co-occurrence matrix then applied CNN which is used for Text classification purpose. Therefore, it is a hybrid algorithm as it is combined of two algorithms CNN and Glove.

Module 3: Word2Vec Logistic Regression Calculated Linguistic Features are given to word embedding Algorithm called Word2Vec which converts the text to vector representation and we apply Logistic regression on it to classify the text.

Module 4: Word2Vec CNN Count of Metadata Features are given as input to the Word2Vec that converts the English text to the vector forms and then neural Network model named CNN is applied to classify the Text

VII. CONCLUSION:

This work has been used to examine the currently popular ERDE metric for early detection tasks in detail and has shown that especially ERDE5 is not a meaningful metric for the described shared task. Only the correct prediction of few positive samples has an effect on this score and the best results can therefore

often be obtained by only minimizing false positives. A modification of this metric, namely ERDE%, has been proposed that is better interpretable in the case of shared tasks that require information to be read in chunks. Exemplary scores using this score have been shown in comparison to ERDE scores for the experiments in this work. The analysis of the resulting word vectors has shown that the model has learnt some features specific to this domain and is viable for general syntactic questions in the English language as shown based on the standard word analogy task. The five submitted predictions achieved the best F1 and ERDE50 scores in both tasks and the CNN without metadata in particular achieved the best results in the new anorexia task. The same working notes has also been used to evaluate the modified ERDE% metric for all and again shows how especially the original ERDE5 metric favors systems that correctly predict test users with only few documents in total regardless of their overall performance

VIII. FUTURE SCOPE:

In this project, we compare all the algorithms using comparison graph based on accuracy, precision, recall and f1score.We will pick the best algorithm that has more accuracy among all algorithms to detect depression. Also, we can enhance this model by utilizing ensembles of more than just two models and calibrating the resulting probabilities seems promising. Combining word embeddings of two models in a single neural network has also not been evaluated yet. Another possible improvement would be to use recently published language modelling methods as input for the network as BERT algorithm and to compare a self-trained model using this approach to the Fast Text word embeddings of this work. This could be very helpful for the faster execution and more accurate values can be produced which increases the performance of the model

IX REFERENCE:

[1]. Marcel Trotsky, Sven Ketika, and Christoph M. Friedrich, "Utilizing Neural Networks and Linguistic Metadata for Early Detection of Depression Indications in Text Sequences" vol.32,2020 pp 3-14.
[2]. J. Howard and S. Ruder, "Universal language model fine-tuning for text classification," in Proc. 56th Anu. Meeting Assoc. Computer. Linguistics (Vol. 1: Long Papers), 2018, vol. 1, pp. 328–339.
[3]. M. Al-Mousawi and T. Johnstone, "In an absolute state: Elevated use of absolutist words is a marker specific to anxiety, depression, and suicidal ideation," Clinical Psychological Sci., vol. 6, no. 4, pp. 529–542, 2018.
[4]. P. Bojanowski, E. Grave, A. Joplin, and T. Manolov, "Enriching word vectors with subworld information," Trans. Assoc. Computer. Linguistics, vol. 5, pp. 135–146, 2017.
[5]. Y. Zhang and B. Wallace, "A sensitivity analysis of convolutional neural networks for sentence classification," in Proc. 8th Int. Joint Conf. Natural Lang. Process. (Vol. 1: Long Papers), 2017, pp. 253–26
[6]. H. Almeida, A. Briand, and M.-J. Mears, "Detecting early risk of depression from social media user-generated content," in Proc. Conf. Labs Eval. Forum, 2017. [Online]. Available: http://ceur-ws. org/Vol-1866/paper_127.pdf
[7]. E. S. Paykel, "Basic concepts of depression," Dialogues Clinical Neuroscience, vol. 10, no. 3, pp. 279–289, 2017.
[8]. P. S. Wang, M. Anger Meyer, G. Borges. Chihuly. Huang, et al., "Delay and failure in treatment seeking after first onset of mental disorders in the WHO," vol. 6, no. 3, 2016, Art. no. 177.
[9]. W. Shang, K. Sohn, D. Almeida, and H. Lee, "Understanding and improving convolutional neural networks via concatenated rectified linear units," in Proc. 33rd Int. Conf. Mach. Learn., 2016, vol. 48, pp. 2217–2225.
[10]. R. Whitley and R. D. Campbell, "Stigma, agency and recovery amongst people with severe mental illness," Social Sci. Med., vol. 107, pp. 1–8, 2015
[11]. T. Manolov, I. Subsieve, K. Chen, G. S. Corrido, and J. Dean, "Distributed representations of words and phrases and their compositionality," in Proc. Neural Inf. Process. Syst., 2014, pp. 3111–3119.
[12]. J. A. Nayland, S. W. Grande, K. A. Asch Brenner, and G. Elwyn, "Naturally occurring peer support through social media: the experiences of individuals with severe mental illness using YouTube," PLOS One, vol. 9, no. 10, 2014, Art. no. e110171.
[13]. A. Nayland, K. A. Asch Brenner, L. A. Marsh, and S. J. Bartels, "The future of mental health care: Peer-to-peer support and social media," Epidemiology Psychiatric Sci., vol. 25, no. 2, pp. 113–122, 2014.
[14]. M. Berger, T. H. Wagner, and L. C. Baker, "Internet use and stigmatized illness," Social Sci. Med., vol. 61, no. 8, pp. 1821–1827, 2014. 59
[15]. W. Bucci and N. Freedman, "The language of depression," Bulletin Menninger Clinic, vol. 45, no. 4, pp. 334–358, 2013