Mr. Umesh A. Patil, et. al. International Journal of Engineering Research and Applications
www.ijera.com
ISSN: 2248-9622, Vol. 12, Issue 4, (Series-III) April 2022, pp. 23-25

RESEARCH ARTICLE OPEN ACCESS

Review on Predicting Depression Level of the Person from Social Media Posts

Mr. Umesh A. Patil (Assistant Professor), Miss Jyoti J. Patil (Student), Miss Mayuri M. Patil (Student), Mr.Akshay B. Huddar (Student), Miss Komal B. Ghadge (Student)
(Department of Computer Science And Engineering, D.Y.Patil Technical Campus Talsande, Kolhapur.)

ABSTRACT
The use of Social Networking Sitesis increasing nowadays, especially by the new generations. The availability of SNS allows users to express their interests, feelings and share daily routines. Many researchers prove that using user-generated content correctly may help determine people's mental health. Mining the UGC could help to predict mental health levels and depression. Depression is a serious mental illness. However, from the user profile, we can collect all the information that relates to a person's mood. In this research, we decided to search how users' posts can help classify users according to mental health levels. We propose a system that uses networking sites as a source of data and screening tool to classify the user using artificial intelligence according to the UGC on SNS. We created a model that classifies the UGC using two different classifiers: Support Vector Machine (SVM), Naïve Bayes,and Transformer.
***Keywords*–**Analyze, Detection, Depressed, Prediction

Date of Submission: 09-04-2022 Date of Acceptance: 26-04-2022

I. INTRODUCTION

Mental illness has been prevalent in the world, depression is one of the most common psychological problems we know and we would like to help as much as we can.Especially in this pandemic depression rate has increased due to people stressing a lot about their career and family. Many factors can lead to depression, for example, personal issues, such as hopelessness, severe anxiety, or impulsivity; social factors, like social isolation andoverexposure to deaths; or negative life events, etc. Nowadays people use different social media such as Facebook, Instagram, and Twitter. Theyshare their thoughts, emotions, and inner feelings. We can know aboutthe daily activity, mood, and opinion of the individual. By analyzing a user's activity (data) and applying machine learning algorithms to it, we can predict the depression level of that individual.

II.RELEVANCE

By using this software, we can define the stresslevel of the user. By using stress level results we can guide the user to avoid depression. As result, they can give the main reason for depression. This software can be a help to overcome this depression problem.

III.LITERATURE REVIEW

https://ieeexplore.ieee.org/abstract/document/9065101

•Nafiz Al Asad 2019 IEEE International Conference on Signal Processing, Information, Communication Systems (SPICSCON)

•https://www.researchgate.net/publication/338158822_Detection_of_Suicide_Ideation_in_Social_Media_Forums_Using_Deep_Learning

•Michael Mesfin Tadesse, Hongfei Lin, Bo Xu, and Liang Yang ○ Received: 17 November 2019; Accepted: 20 December 2019; Published: 24 December 2019

•https://ieeexplore.ieee.org/document/9395943

•Keshu Malviya, Dr. Bholanath Roy, Dr. Saritha SK ○ International Conference on Artificial Intelligence and Smart Systems (ICAIS-2021)

IV.OBJECTIVES

- To remove noise from the data.
- To separate and group similar words from the data.
- To train a model using a dataset.
- Classify a given text as depressed or non-depressed

V.RESEARCH GAP

Sr.no.	Name	Details	Limitation
1.	Paper 2020	This paper proposed a model that takes a username and analyses the user's social media posts to determine the levels of vulnerability to depression. The machine learning model is trained to classify the depression criteria in six ranges. The verdict is depressed when the percentage is above borderline (above 55%).	This proposed system has low accuracy for the prediction of depression levels**.**
2.	Paper 2020	In the previous system due to the lack of accurate predictions were unreliable. In this system, the LSTM-CNN hybrid model considerably improves the accuracy of text classification.	Data deficiency is one of the most critical issues of this proposed system, mainly applied supervised learning techniques. They usually require.

VI.PROPOSED ARCHITECTURE

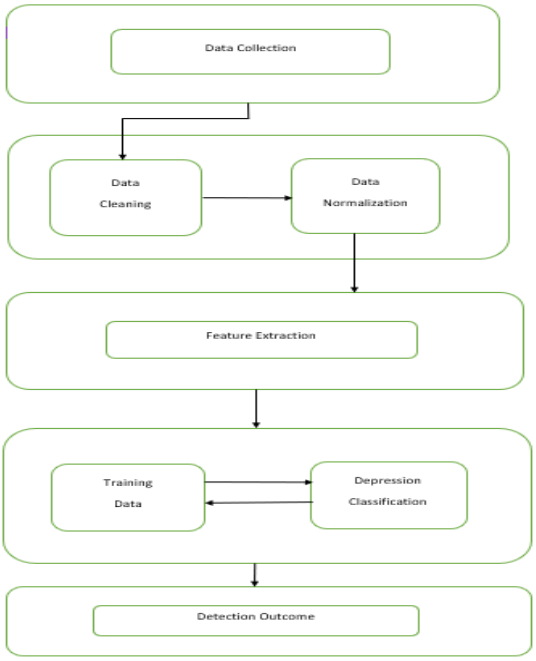

Fig. Proposed Architecture

VII.IMPLWMWNTATIONS STEPS

DATASET:

We are going to use the sentiment140 dataset. It contains 1,600,000 tweets extracted using the Twitter API. The tweets have been annotated as depressed and non-depressed and they can be used to detect sentiment.

MODULES:

1. Preprocessing:

Pre-processing involves filtering an input text to improve the accuracy of a proposed method by eliminating redundant features to process raw posts before the learning of word embedding.

2. Word Analysis:

To describe and demonstrate amongst depressive and non-depressive posts, we extract the different features given psycholinguistic measurements from the user's post. Such as it is easy to spot words that are indicative of depression: depression, treatment, suffering, crying, help, struggle, risk, hate, sad, anxiety, disorder, suicide, stress, therapy, mental health, emotional, bipolar.

Word Embeddings:

The first step is to symbolize every post as a vector. Two different word embedding schemes were chosen for getting vector representations for the posts and sentences.

Word2Vec:

In the Word2Vec approach, shallow but powerful models research words as represented by vector

space through mapping those words that occur simultaneously as well as at the same time with comparable meanings.

TF*IDF:

It is a method based on informationretrieval that is responsible for weighing the frequency of a term (TF) and the inverse frequency of the document (IDF). Every term that takes place within the textual content has been given its TF and IDF score. The TF*IDF weight of that term is calculated.

3. Classification Model:

After receiving the phrase embeddings, experiments will carry out with the following six classifiers: –

a) Naive Bayes: This classifier is a probabilistic classifier that makes use of the Bayes Theorem. It calculates the possibility of current events based on previously occurred events.

b) Logistic regression: It is a statistical model that estimates the probability of a reaction based on prediction variables.

c)Bagging model (Random Forest): An ensemble of Decision Trees where labels are predicted using a tree-like model. The outputs of all created decision trees are taken into consideration for the final output.

d)Transformer:

The task is to formulate a way to figure out if a person is depressed or not through his social media posts. In the first instance, we have planned a few relatively easy baseline models such as Linear Classifier, Boosting model, etc. Later, the usage of transformer structures and fine-tuned, unique, and pre-trained transformer models.

4. Model Evaluation:

This phase constructs model evaluation using model evaluation metrics and selecting the one with the best accuracy.

VIII. SCOPE

In future work, the limitation of language in this model will be reduced and more than one language will be considered as a sample. ○ In this system, we use text data for analysis but shortly we can determine depression levels from images also.

REFERENCES

[1]. https://ieeexplore.ieee.org/abstract/document/9065101
Nafiz Al Asad 2019 IEEE International Conference on Signal Processing, Information, Communication Systems (SPICSCON)

[2]. https://www.researchgate.net/publication/338158822_Detection_of_Suicide_Ideation_in_Socia l_Media_Forums_Using_Deep_Learning Michael Mesfin Tadesse, Hongfei Lin, Bo Xu, and Liang Yang ○ Received: 17 November 2019; Accepted: 20 December 2019; Published: 24 December 2019

[3]. https://ieeexplore.ieee.org/document/9395943 Keshu Malviya, Dr. Bholanath Roy, Dr. Saritha SK ○ International Conference on Artificial Intelligence and Smart Systems (ICAIS-2021)