



## DETECTION OF DEPRESSION RELATED POSTS IN REDDIT SOCIAL MEDIA FORUM USING MACHINE LEARNING

N. Sowjanya kumari<sup>1</sup>, G. Madhuri Sai<sup>2</sup>, Ch. Sathvika<sup>3</sup>, G. Divya<sup>4</sup>, D. Amulya<sup>5</sup>, G. Pravalika<sup>6</sup>

<sup>1,2,3,4,5,6</sup> Department of Computer Science & Engineering, Vignan's Institute of Engineering For Women, Duvvada, Visakhapatnam, Andhra Pradesh, India.

### ABSTRACT

The main cause of suicide and the biggest contributor to global disability is depression. It affects how language is used and is reflected in the written content. The main goal of our investigation is to look at Reddit users' posts to see if there are any indicators that could show how pertinent online users feel about depression. To achieve this, we train the data using Natural Language Processing (NLP) methods and machine learning techniques, and then test the effectiveness of our suggested strategy. We find a vocabulary that is more prevalent in accounts of depression. The outcomes demonstrate that the performance accuracy of our suggested technique can be greatly increased. Bigram, combined with the Support Vector Machine (SVM) classifier, is the strongest single feature for detecting depression, with an accuracy rate of 80% and an F1 score of 0.80. The Multilayer Perceptron (MLP) classifier achieves the best performance for depression detection with 91% accuracy and 0.93 F1 score, best illustrating the strength and usefulness of the combined features (LIWC+LDA+bigram). Our research suggests that wise feature selections and their various feature combinations can lead to improved performance improvements.

### Keywords:

Natural Language Processing, Machine learning, Reddit, Social Networks, Depression

### INTRODUCTION

For a very long time, depression has been regarded as a single illness with a collection of diagnostic criteria. It frequently co-occurs with anxiety or other psychological and physical disorders and affects how those impacted feel and behave. The WHO study estimates that 322 million individuals worldwide—or 4.4% of the world's population—suffer from depression. The Western Pacific region, which includes China and India, and South-East Asia together account for 27% and 27% of the population at danger. Depression is still underdiagnosed and inadequately treated in many nations, which can result in negative self-perception and, in the worst cases, suicide. In addition, the stigma associated with depression stops many sufferers from getting the right kind of professional help. They consequently use more informal tools, like social media. Through online forums, micro-blogs, and tweets, people have begun to share their struggles and experiences with mental health disorders as Internet utilization has grown. Many researchers were inspired by their online activities to develop novel approaches to possible healthcare solutions and techniques for early depression detection systems. They attempted to achieve a better performance improvement by utilising various Natural Language Processing (NLP) techniques and text classification approaches. Some studies use a single collection of features to detect depression in their posts, including bag of words (BOW), N-grams, LIWC, or LDA. Other studies evaluate the effectiveness of specific features with different machine learning classifiers. Recent research looks at the effectiveness of individual features and their combos to increase accuracy, such as N-grams+LIWC or BOW+LDA and TF-IDF+LDA. They experiment with text pre-processing that is more intelligent and introduce various substitute words based on the original string's nature. For instance, Tyshchenko et al. recommended adding LIWC-like word categories and categorizing stop words as an extra feature to an already developed technique (BOW+TFIDF+LIWC). Additionally, he used a variety of feature combos with Convolutional Neural Networks (CNN), which are composed of neurons with learnable weights and vary in terms of their



layer structure, to boost performance. CNNs are the most advanced technique for text and sentence classification tasks, and they are very close to straightforward feed-forward neural networks. Multiple versions of computational linguistics depression detection tasks are compiled in a meta-analysis by Guntuku et al, Calvo et al.'s review of the taxonomy of data sources, NLP techniques, and computational methods to identify different mental health applications is another fascinating analysis of mental health support and intervention in social media. Despite this major advancement, problems still exist. Through appropriate feature selection and numerous feature combinations, this paper seeks to find a performance increase solution. In order to characterize the substance of the posts, we first select the best linguistic features used for depression identification. Second, we examine word frequency, hidden subjects, and correlation strength, content, taken from. We concentrate on the LIWC dictionary and its three feature categories (linguistic dimensions, psychological processes, and personal issues) with regard to the correlation. We select the LDA method as one of the effective features for the subject examination. By utilizing vectors based on the TF-IDF algorithm, we use unigrams and bigrams for the word frequency. Finally, using the extracted data to execute five text classification methods, we can identify depression. Based on three single feature sets and their multiple feature combinations, we compare the performance outcomes. We use information gathered from the social media website Reddit for our project. As it permits lengthier posts, it was selected as the data source. Our article targets technical methods for detection tasks and builds on the work of Calvo et al. Our study makes four distinct contributions: first, it investigates the connection between language usage and depression; second, it develops three LIWC features specifically for our research problem; third, it assesses the performance accuracy of N-gram probabilities, LIWC, and LDA as single features; and fourth, it demonstrates the predictive power of both single and combined features with suggested classification methods to achieve a higher performance in depression.

## LITERATURE SURVEY

[1] D. Maupomé & M.-J. Meurs, "Using topic extraction on social media information for the early identification of depression,". As part of the eRisk2018 shared challenge on depression, which entails the early assessment of depression risk in social media users, we build a system based on the topic extraction approach, Latent Dirichlet Allocation, and fundamental neural networks. The technique unsupervisedly extracts 30 latent topics using the frequency of unigrams, bigrams, and trigrams. The users are then given a diagnostic likelihood by a Multilayer Perceptron once it has been converted onto this feature space. Ultimately, a classification method that categorises each user is based on an absolute threshold of probability that decreases over time. The topic-document matrix that the LDA model produces next illustrates the relative relevance of each topic in each document. Lastly, a Multilayer Perceptron (MLP) is fed this representation to provide a predicted label for each user. By deleting all stop words, we limited the term-document matrix to the top 3000 n-grams with lengths 1 to 3. In our experiments, we discovered that the LDA employs strategies that enable straightforward translation of its findings to emotion analysis. We chose topic extraction since, logically, a person's conversational topics would reveal a lot about their mental state.

[2] S. Paul, S. K. Jandhyala, and T. Boss, "Early detection of indications of anorexia and depression over social media utilising effective machine learning frameworks," in CLEF, 2018. To complete the tasks provided, many machine learning frameworks have been offered in this study. The goal is to use the training set to teach a machine learning classifier how to recognise anorexia or depression in the individual documents of the test sets for these activities. A corpus's potential features have a significant impact on how well a text categorization system performs. Hence, utilising both text characteristics and biological features collected from the provided corpus, the effectiveness of various classifiers has been evaluated. The frequency of the various terms is thought to generate the document vectors since, in general, each distinct term of a corpus is considered to be a feature



[3]M. Nadeem, "Detecting Depression on Twitter," CoRR, vol. abs/1607.073842016. Tens of millions of people interact with each other through these internet networks to share ideas, memories, and societal values. Hence, we investigate the potential of social media to identify Major Depressive Disorder (MDD) in online personas even before its manifestation. We use crowdsourcing to create a list of Twitter users who claim to have been diagnosed with depression. We use a Bag of Words method to quantify each tweet using up to a year's worth of previous social media postings. Finally, we use various statistical classifiers to evaluate the risk of depression.

[4] Beyond lda: Exploring supervised topic modelling for depression-related language in twitter, P. Resnik, W. Armstrong, L. Claudino, T. Nguyen, V.-A. Nguyen, and J. Boyd-Graber, Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality, 2015, pp. 99-107. In this research, we investigate the application of supervised topic models in the analysis of language signal for depression detection and present promising results utilising various models. We investigate the usage of the anchor algorithm (Arora et al., 2013, therefore ANCHOR), which offers a quick method for learning topic models and improves interpretability by selecting a single "anchor" word associated with each topic. Multi-task learning for mental health via social media text

[5] A. Benton, M. Mitchell, and D. Hovy, CoRR, vol. abs/1712.03538, 2017. This research analyses the preliminary framework for mental health and suicide risk estimation in a deep learning framework. The algorithm gains the ability to forecast suicide risk and mental health with a low percentage of false positives by simulating a variety of situations. In multitask learning (MTL) paradigm, conditions are modelled as tasks, and gender prediction is added as an auxiliary task. By contrast with a fine-tuned single-task baseline with the same number of parameters, we show the effectiveness of multi-task learning.

## PROPOSED SYSTEM

Several methods are being developed to identify depression in postings. We include a technical discussion of methods used for depression identification utilising NLP and text classification techniques in our study. Data pre-processing, feature extraction, feature analysis for machine learning classifiers, and experimental outcomes make up the framework.

## DATA PRE-PROCESSING

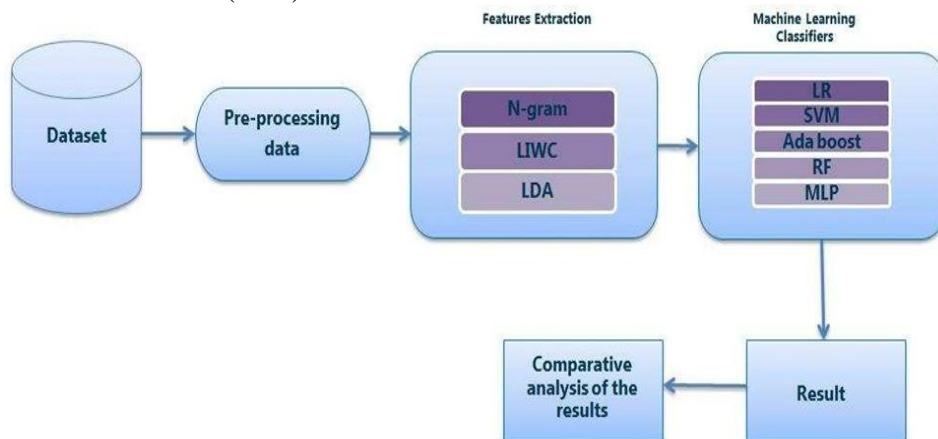
Before moving on to the step of feature selection and training, we pre-process the dataset using NLP tools. To start, we tokenize the posts to separate them into separate tokens. Then, we eliminate all URLs, punctuation, and stop words that, if left in place, could provide unpredictable outcomes. The words are then reduced to their root forms using stemming, which also groups related words together.

## FEATURE EXTRACTION

Following data pre-processing, we feed our models with attributes that correspond to users' Reddit forum language preferences. We use the LIWC dictionary, LDA topics, and N-gram features to examine the language used by the users in the posts. These text encoding techniques are used to encode the words that various classifiers will come after. The postings' features are examined using N-gram modelling. The chance of each input sentence occurring as both a unigram and a bigram is a feature for depression identification that is frequently used in text mining and NLP. The term frequency-inverse document frequency (TF-IDF), a numerical statistic that highlights the significance of a word in relation to each document in a corpus, is used for n-gram modelling. The basic objective of its use is to reduce the impact of empirically less informative tokens that appear frequently in order to make room for more informative terms that occur less frequently.

If a term appears in one post but not another, its TF-IDF value is ranked higher in that post. In our project, we extract 194,613 unigrams and bigrams using the TF-IDF vectorizer from the scikit-learn

Python library. We eliminate all stop words from the dataset and limit the term-document matrix to the top 3000 bigrams and unigrams in frequency. Additionally, to filter sporadic bigrams, we used Pointwise Mutual Information (PMI).



**Figure: Proposed System Architecture**

**TEXT CLASSIFICATION TECHNIQUES**

We apply classification algorithms to assess the likelihood of depression among the users in order to estimate the presence of depression. Logistic Regression, Support Vector Machine, Random Forest, Adaptive Boosting, and Multilayer Perceptron Classifier are used to create the suggested framework [4]. The linear classification method known as logistic regression (LR) is used to calculate the likelihood that a binary response will occur given one or more characteristics and predictors. The Support Vector Machine (SVM) model represents the instances as points in a highly dimensional space that is used for classification and in which the points of the various categories are distinctly separated. Then, new examples are projected to fit into a category and mapped into the same space.

By doing so, we are able to calculate the scores for three categories at a higher level while taking into account linguistic norms, psychological processes, and individual preferences. One of the biggest components of the LIWC psycholinguistic vocabulary set is the common linguistic processes. It was designed to measure the words' usage in psychologically meaningful classifications and to recognise the relationships that people have with one another in social cooperation. In order to categorise the users' content in our study, we first selected 9 linguistic elements (articles, auxiliary verbs, adverbs, conjunctions, impersonal and personal pronouns, negations, prepositions, and verbs). In computational linguistics, topic modelling is a useful approach for condensing the feature space of the incoming textual data to a predetermined set of subjects. Unsupervised text mining can be used to extract subjects from the chosen documents that are hidden, such as those related to anxiety and sadness. It is not produced by a LIWC, as opposed to LIWC a predetermined, rigid list of terms. A unique type of artificial neural network utilised frequently for simulating intricate interactions between input and output layers is the multilayer perceptron (MLP) [11]. It can identify the data that is not only non-linearly separable due to its numerous layers and non-linear activation[12]. To achieve a consistent comparison in our study, we used the MLP approach and two hidden layers with 4 and 16 perceptrons to correct for all the features.

**Table:Text Classification Techniques**

Type Feature	Methods	Number of selected features
N-Grams	Unigram	3000
	Bigram	2736



Linguistic dimensions psychological processes	LIWC	68
Topic modelling	LDA	70

## SYSTEM METHODOLOGY

The paper makes four distinct contributions: first, it investigates the connection between language usage and depression; second, it develops three LIWC features specifically for our research problem; third, it assesses the performance accuracy of N-gram probabilities, LIWC, and LDA as single features; and fourth, it demonstrates the predictive power of single and combined features with proposed classification approaches to improve depression identification performance. With correct feature selection and different feature combinations, the suggested system seeks to find a performance enhancement solution. In order to characterise the content of the posts, we first select the best language features used for depression diagnosis. The correlation significance, hidden topics, and word frequency that were retrieved from the text are all examined next. We concentrate on the LIWC dictionary and its three feature kinds with regard to the correlation (linguistic dimensions, psychological processes, and personal concerns). For the topic examination, we chose the LDA method as one of the successful characteristics. The Multilayer Perceptron (MLP) classifier achieves the best performance for depression, with 91% accuracy, demonstrating the power and efficacy of the combined features (LIWC+LDA+bigram).

### LIWC (Linguistic Inquiry Word Count)

LIWC, or the Linguistic Inquiry and Word Count dictionary, is frequently used in computational linguistics as a source of features for psychological and psycholinguistic research. With a set of words and a behavioural relationship, it functions as a baseline measurement [2]. It frequently appears in a number of mental health undertakings. In order to complete our experiment, we chose 68 out of 95 features based on psycholinguistic measurements and converted each depressive and non-depressive post into a numerical number. By doing so, we are able to calculate the scores for three categories at a higher level while taking into account linguistic norms, psychological processes, and individual preferences. One of the biggest components of the LIWC psycholinguistic vocabulary set is the common linguistic processes. It was designed to measure the words' use in psychologically meaningful classifications and to recognise the relationships that exist between people in social cooperation. In order to categorise the users' content in our study, we first selected 9 linguistic elements (articles, auxiliary verbs, adverbs, conjunctions, impersonal and personal pronouns, negations, prepositions, and verbs). The psychological processes are then divided into subcategories from which we used effective processes (anxiety, sadness, positive or negative emotion), biological processes (sexual, body, ingestion, and health), social processes (family, friend, male, female), cognitive processes (cause, always, never, because), personal concerns (job, cook, cash, bury, kill), and time orientations (present, past, season).

### LDA (Latent Dirichlet Allocation)

The validation set's LDA model performs best when it is restricted to 70 topics. We only take into account words for subject selection if they appear in at least 10 or more postings. Every post is included as one document that needs to be further tokenized and stemmed. This technique allows us to compute the themes over the collection of documents and annotate them according to discovered subjects. All stop words are eliminated prior to the topic modelling procedure. The Mallet toolbox offers LDA implementation.

### N-gram



To find postings in Reddit social media forums that are relevant to depression, feature extraction can be done using N-gram analysis. An n-gram is a group of n words that appear in a text, and studying the frequency of various n-grams might provide information about the linguistic traditions of a culture. For instance, by looking at the frequency of various n-grams in posts about depression, we can spot some linguistic patterns that are suggestive of depression. Among the n grams that could be useful for identifying posts connected to depression are the following:

- Unigrams: single words like "lonely," "sad," "depressed," and "anxious."
- Bigrams: a group of words that are paired together, such as "mental health," "feeling depressed," "hopeless thoughts," etc.
- Trigrams are three-word groups that include phrases like "can't stop sobbing," "continuous bad thoughts," "lack of energy," etc.

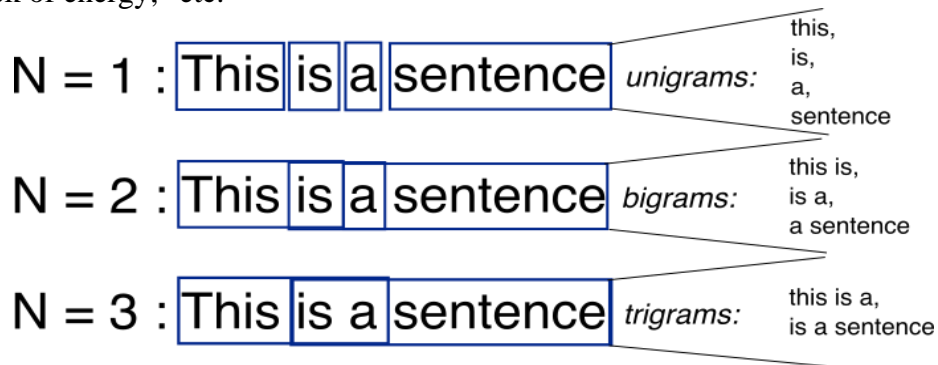


Figure : Natural Language Processing

### PREDICTIVE POWER OF N GRAM

We categorise the full tagged corpus of Reddit posts to compare the lexical variations. We compute the frequencies of all the unigrams and bigrams in both depression-indicative posts and regular posts to examine the existence of depression. For each category, the top 100 bigrams and unigrams are chosen.

### Output screens

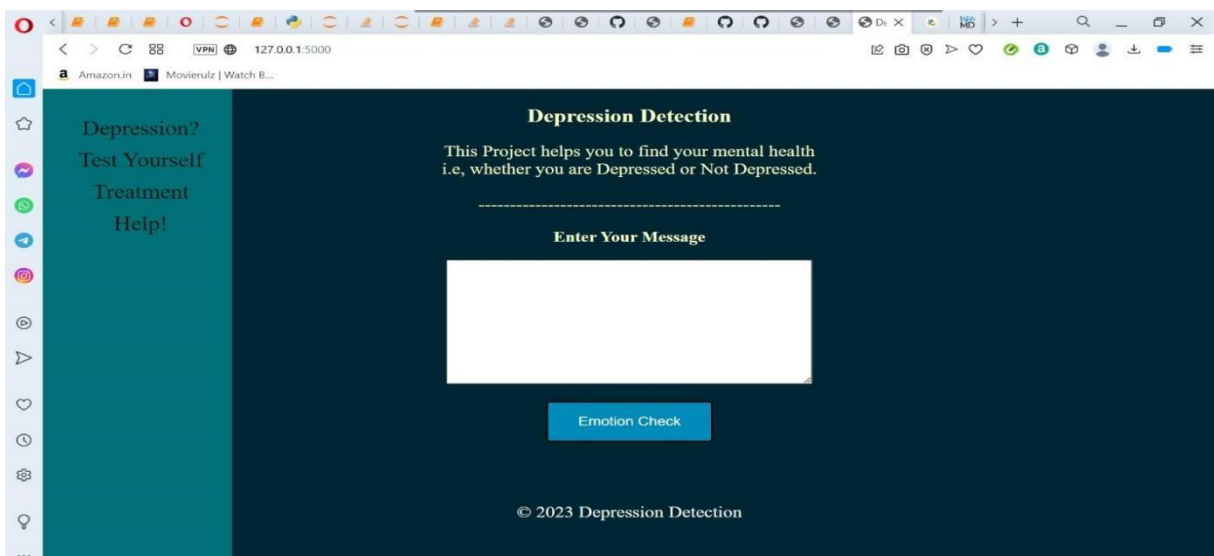
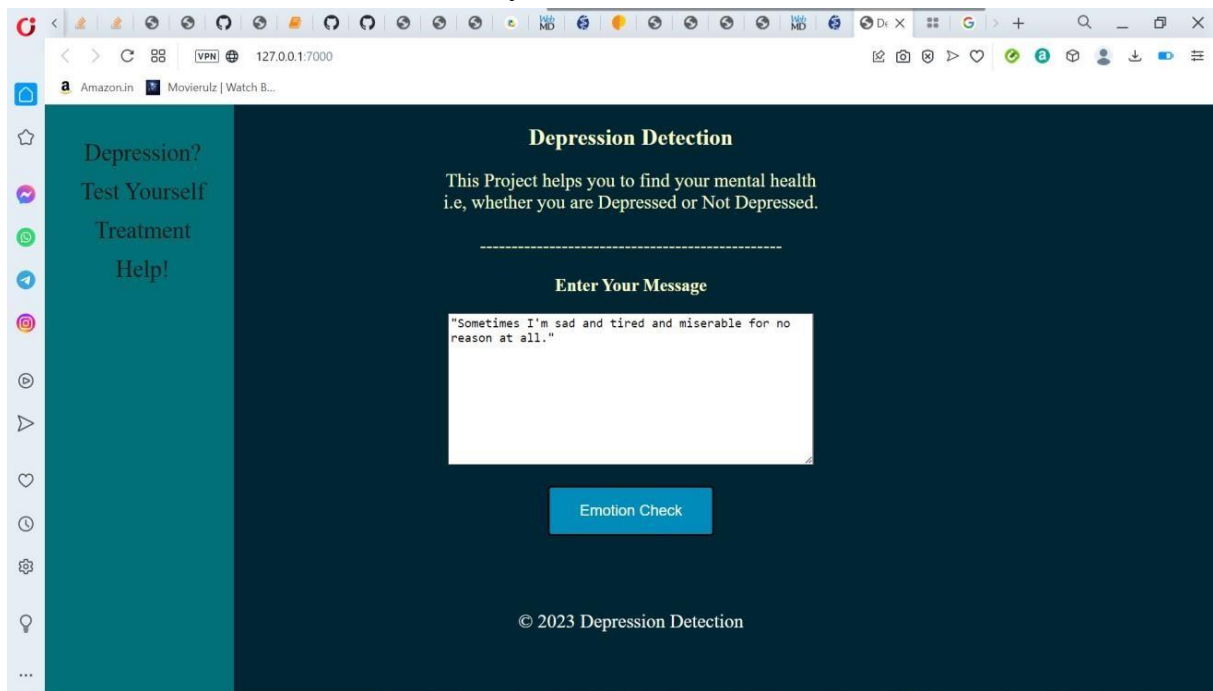
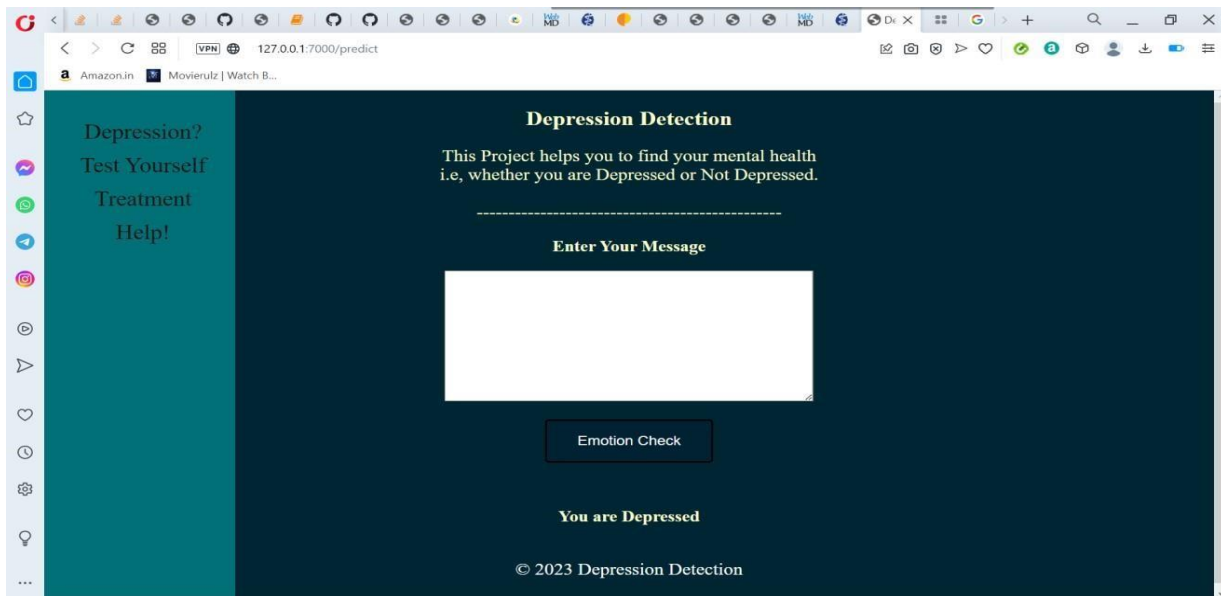


Figure: Depression Detection Home Page



**Figure: Text given to check Depressed or not.**



**Figure: Output Depressed**

## CONCLUSION & FUTURE SCOPE

In this study, we sought to determine whether melancholy was present in Reddit social media solutions to depression detection that work better. We used NLP and text classification methods to identify a closer link between language use and depression. A list of terms with higher frequency was found by us among the gloomy reports. Our research revealed that the words associated with self-centeredness, sadness, anxiety, anger, hostility, or thoughts of suicide were more prevalent in the language predictors of depression than other words. These words also placed a stronger focus on the present and future. Using a variety of text classifying techniques; we evaluated the performance of both single feature and combined feature sets in order to assess the symptoms of depression. Our findings demonstrate that a better predictive performance is concealed in appropriate. Moreover, among the single feature sets, the



best characteristic is bigram; using an SVM classifier, it is possible to identify melancholy 80 percent precision and a 0.79 F1 rating. Taking into account LIWC and LDA features, LIWC fared better than topic models created by LDA. Despite the fact that our experiment demonstrates that the performances of the employed methodologies are fairly good, the absolute values of the metrics show that this is a difficult job deserving of further investigation. We think that this experiment may strengthen the foundation for novel approaches to estimating depression and related variables that could be used in various healthcare settings. The ability to take more initiative to hasten their recovery can be helpful for those with mental health issues.

#### REFERENCES

- [1] D. Maupomé & M.-J. Meurs, "Using topic extraction on social media information for the early identification of depression,".
- [2] S. Paul, S. K. Jandhyala, and T. Boss, "Early detection of indications of anorexia and depression over social media utilising effective machine learning frameworks," in CLEF, 2018.
- [3] M. Nadeem, "Identifying Depression on Twitter", CoRR, vol. abs/1607.073842016.
- [4] P. Resnik, W. Armstrong, L. Claudino, T. Nguyen, V.-A. Nguyen, and J. Boyd-Graber, "Beyond lda: Exploring supervised topic modeling for depression-related language in twitter," in Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality, 2015, pp. 99–107.
- [5] A. Benton, M. Mitchell, and D. Hovy, "Multi-task learning for mental health using social media text," CoRR, vol. abs/1712.03538, 2017.