ML&AI_IIIT Ranchi@LT-EDI-2023: Hybrid Model for Text Classification aimed at Identifying Different Forms of Depression

Kirti Kumari
IIIT Ranchi, India
kirti@iiitranchi.ac.in

Shirish Shekhar Jha
IISER Bhopal, India
shirish20@iiserb.ac.in

Zarikunte Kunal Dayanand
IISER Bhopal, India
zarikunte20@iiserb.ac.in

Praneesh Sharma
KIIT, Bhubaneswar, India
praneeshsharma30@gmail.com

Abstract

DepSign–LT–EDI@RANLP–2023 is a dedicated task that addresses the crucial issue of identifying indications of depression in individuals through their social media posts, which serve as a platform for expressing their emotions and sentiments. The primary objective revolves around accurately classifying the signs of depression into three distinct categories: "not depressed", "moderately depressed", and "severely depressed". Our study entailed the utilization of machine learning algorithms, coupled with features such as sentence embeddings and feature representations like TF-IDF, and Bag-of-Words. Remarkably, the adoption of hybrid models yielded promising outcomes, culminating in a 10^{th} rank achievement out of 31 participants, supported by a macro F1-Score of 0.408. This research underscores the effectiveness and potential of employing advanced text classification methodologies to discern and identify signs of depression within social media data. The findings hold implications for the development of mental health monitoring systems and support mechanisms, contributing to the well-being of individuals in need.

1 Introduction

Depression is a widespread mental health condition that affects a significant number of individuals worldwide. The identification of depression symptoms in individuals is a crucial step in providing timely support and intervention. Given the growing use of social media platforms as channels for emotional expression and experience sharing, there is an increasing interest in leveraging these platforms to detect indications of mental health issues, including depression. Accordingly, the DepSign-LT-EDI@RANLP-2023 task focuses on the detection of depression signs in individuals based on their social media posts.

Past works in this field offer valuable insights into the importance of monitoring mental health and the challenges associated with identifying and addressing mental health problems. The Global Health Data Exchange (GHDx) by the Institute of Health Metrics and Evaluation (Glo) serves as a valuable resource for accessing global health data and understanding the prevalence of mental health disorders. Additionally, the study conducted by (Evans-Lacko et al., 2018) highlights socio-economic disparities in the treatment gap for individuals with anxiety, mood, and substance use disorders, underscoring the need for targeted interventions and support. The eRISK 2017 study (Losada et al., 2017) carried out further emphasizes the importance of early risk prediction and underscores the experimental foundations in the field.

This research aims to employ machine learning algorithms and text classification techniques to identify signs of depression in social media posts. By leveraging features such as sentence embeddings and utilizing feature representations like the TF-IDF, and Bag-oF-Words, we aim to develop an effective classification model capable of categorizing signs of depression into three distinct labels: **"not depressed", "moderately depressed",** and **"severely depressed"**. The hybrid models utilized in our study incorporate a combination of these features, enabling enhanced classification performance.

The ultimate objective of this research is to contribute to the field of mental health monitoring and support systems by providing a reliable and efficient method for detecting signs of depression through social media analysis. By harnessing the abundant data available on social media platforms, we aim to facilitate timely intervention and support for individuals experiencing depressive symptoms.

Through the DepSign-LT-EDI@RANLP-2023 task, we aim to address the urgent need for inno-

LT-EDI 2023 – Third Workshop on Language Technology for Equality, Diversity and Inclusion, pages 223–232, Varna, Bulgaria, Sep 7, 2023.
https://doi.org/10.26615/978-954-452-084-7_034

vative approaches to mental health detection and intervention. By utilizing advanced machine learning algorithms and integrating diverse features, our goal is to offer an easy to use but effective solution for identifying signs of depression in social media data. Although BERTs and other Large Language Models (LLMs) are available for the classification task a major drawback that holds them back is their massive size and training time. Not every individual or organization possesses the capability to leverage the LLM power and supremacy for text classification. Also, there is not much data present to perform the fine-tuning task and get great results. To overcome this and still utilize the abilities of the LLM, we were motivated to get a way around by making use of sentence embeddings that are generated from the LLM but are of smaller size and can be computed in a limited amount of time. The results obtained from our study, as evidenced by our 10th rank and an F1 score of 0.408, demonstrate the potential and promise of our proposed approach.

This paper presents the related work Section 2, dataset description Section 3, task description Section 4, methodology and validation results Section 5, result Section 6, and conclusions Section 7 of our research, shedding light on the effectiveness of text classification techniques in detecting signs of depression. Furthermore, the findings contribute to the broader field of mental health research and pave the way for the development of scalable and efficient solutions for mental health monitoring and support systems.

2 Related Work

In the past many efforts have been made in the field of depression identification from social media texts. The authors of (De Choudhury et al., 2013) were among the pioneers in researching the detection of depression through social media posts. Their study focused on Twitter users who had been diagnosed with depression, and they collected one year's worth of posts from this group. Using this dataset, they developed a statistical classifier that aimed to estimate the risk of depression. The classifier utilized various linguistic and behavioral features extracted from the users' posts to predict the likelihood of depression. This research marked an important milestone in the field, providing insights into the potential of social media data for identifying mental health conditions. By employing a rigorous statistical approach, it also laid the groundwork for subsequent studies on detecting depression through social media, contributing to the advancement of research in this domain.

The authors of (William and Suhartono, 2021) conducted a study that explores the use of machine learning techniques to detect depression from social media posts. By employing linguistic features and classifiers, the authors achieve promising results, demonstrating the potential of utilizing social media data for identifying signs of depression. They compare various machine learning algorithms and feature selection methods to determine the most effective combination. The study highlights the importance of pre-processing techniques and feature engineering in improving classification performance. The findings contribute to the development of reliable and efficient methods for detecting signs of depression through social media analysis, supporting timely intervention and support for individuals experiencing depressive symptoms. The authors of (Dessai and Usgaonkar, 2022) carried research that focuses on the scientific aspect of detecting depression from social media data using machine learning techniques and text mining. The authors propose a novel approach that combines text and image information for improved depression detection accuracy. They extract textual features from posts and visual features from associated images, and then employ a multi-modal fusion model to integrate these features. The study evaluates the proposed approach on a large-scale dataset and compares it with existing methods, demonstrating its superior performance. The findings highlight the importance of considering both textual and visual cues for accurate depression detection, offering valuable insights for developing effective mental health monitoring systems. Observing the multi modality of the task, advancements have been made to make the text based identification more robust. Work presented by The authors of (Wolohan et al., 2018) have led emphasis on detecting the linguistic features from the text corpora generated from Reddit, so as to classify them into the categories of depression based on lexical and predictive analysis. The introduction of deep learning architectures like transformer models have also significantly improvised the results (Devlin et al., 2018) presents work that leverages the power of transfer learning to classify text into depression.

The authors of (Salas-Zárate et al., 2022) pre-

sented a thorough search of the literature, the authors found 34 primary papers that satisfied their inclusion requirements. The studies employed several techniques, such as language feature extraction, machine learning, and statistical analysis, to find symptoms of depression on social media. The research findings were conflicting, but taken together, they imply that social media can be used to identify depression symptoms with some degree of accuracy.
Our work has been inspired by past related works and motivated us to develop a simple system to test and identify depression. Past methods have inculcated various experimentation's like statistical analysis, pre-trained model-based research, and employing classical machine learning algorithms on various features; considering them, we have devised a solution that involves sentence embeddings as features and further classified them using multiple machine learning classifiers.

3 Dataset

The English-language postings in the dataset for the competition were taken from the Reddit platform. "Not depression", "moderate", or "severe" are the three labels that have been manually added to each post in the dataset (Kayalvizhi et al., 2022; Sampath and Durairaj, 2022; S et al., 2022; Sampath et al., 2023). When there were no signs of depression found in the post, the label "not depression" is used to denote those cases. Alternatively, the terms "moderate" and "severe" denote increasing degrees of depression symptoms in the text.

Table 1 displays examples of text excerpts together with their matching labels to help the reader comprehend the dataset. These illustrations from the dataset highlight the variety of postings that each label can be applied to.

The training set, the development set, and the test set were the three separate parts of the dataset that were divided up for the competition. It is important to note that the labels for the test set were withheld by the competition's administrators because this section was only used to assess the competitors' solutions. Table 2 provides an overview of the label distribution in the dataset and lists the number of instances for each label category.

Notably, the training set contains a larger number of instances compared to the development set. This discrepancy in size enables more effective fine-tuning of hyperparameters for both machine learning algorithms and deep learning neural networks. The availability of a larger training set facilitates more robust model optimization, ultimately leading to improved performance and generalization capabilities.

4 Task Description

The objective of this task is to develop a system that can effectively classify signs of depression within social media postings written in English. The system should be able to categorize these signs into one of three distinct labels: "not depressed", "moderately depressed", and "severely depressed". By automating this classification process, the system can provide valuable insights into individuals' mental well-being and potentially facilitate timely intervention and support.

The dataset provided for this task consists of a collection of social media posts expressed in the English language. Each post has been carefully annotated with one of the three aforementioned labels, indicating the level of depression detected within the content. This labeling scheme enables the development of a comprehensive classification system that can effectively gauge the severity of depression symptoms expressed in social media postings.

The authors of (Lin et al., 2020), have shown depression identification with the use of deep visual-textual multimodal learning approach which embarks a great development in this field and simultaneously introduces the domain where machine learning approaches can be applied where the features can be multimodal, meaning it can be in the form of text, video or both. The authors of (Poświata and Perełkiewicz, 2022; AlSagri and Ykhlef, 2020) have performed on a similar task where they have used large pre-trained models to make accurate predictions. They finally used the features and greatly performed using an average ensemble approach. The text cited acted as our motivation for this task, and we leveraged various machine-learning techniques and methodologies. These include traditional machine learning algorithms, deep learning models, or a combination of both. By training and fine-tuning such models on the provided dataset, we aim to create a robust classifier capable of accurately predicting the level of depression exhibited in unseen social media posts.

Table 1: Text Excerpt From Dataset

S.No.	Text	Label	Dataset
train_pid_7197	arent tired ive depressed month lost trust peo...	severe	Train
dev_pid_1	im scared lie every day say ill make think mig...	moderate	Dev
test_id_3	But here I am, 24 years old man and doing exac...	moderate	Test

Table 2: Dataset Distribution

Class	Training	Development	Test	Total
moderate	3678	2169	275	6122
severe	768	228	89	1085
Not depression	2755	848	135	3738
Total	7201	3245	499	10945

5 Methodology and Validation Results

The methodology employed in the classification task comprises several sequential steps: data pre-processing, TF-IDF features based classification, bag of words features based classification, and the utilization of sentence embeddings for classification.

In the initial data pre-processing step, the raw English social media postings underwent a series of text cleansing procedures. These include the removal of punctuation, stop words, and special characters, as well as tokenization and stemming operations to standardize the textual data. This pre-processing stage ensures the text is appropriately formatted for subsequent analysis.

Subsequently, TF-IDF based classification is implemented. The TF-IDF (Term Frequency-Inverse Document Frequency) technique is employed to represent each document, i.e., social media post, as a numerical feature vector. Following TF-IDF based classification, the methodology incorporates bag of words classification. In this approach, the text data is represented using a bag of words model, which establishes a vocabulary consisting of unique words derived from the corpus. Finally, sentence embeddings are leveraged for classification. Sentence embeddings aim to capture the semantic meaning of a sentence through a compact vector representation.

By adhering to this methodology, which encompasses data pre-processing, TF-IDF and bag of words classification, and the utilization of sentence embeddings, a classical and effective approach was formulated for accurately categorizing signs of depression within English social media postings.

5.1 Data Pre Processing

To commence our preparations for the task, we initiated the process by conducting data pre-processing and visualization. Initially, we conducted an inspection of the data to identify any instances of null or missing values. Upon confirming the absence of such values, we proceeded to perform an analysis of the text statistics. This involved examining the word count, character count, and word density per sentence. The statistical insights derived from this analysis proved instrumental in the subsequent generation of sentence embeddings.

$$\text{WordCount(T)} = |\text{words(T)}|$$
$$\text{CharacterCount(T)} = |\text{characters(T)}|$$
$$\text{WordDensity(T)} = \frac{WordCount(T)}{SentenceCount(T)}$$

Given the nature of the data sourced from social media platforms, we made the assumption that certain text elements required cleaning. Consequently, we focused on the removal of character encodings deemed improper, contractions, and special characters. Furthermore, we undertook the task of eliminating hyperlinks and social media hashtags, as well as alphanumeric characters. In order to enhance the cleanliness and quality of the text, we also implemented the removal of stop words. Additionally, we performed tokenization to segment the text into individual units and applied lemmatization to obtain a refined version of the text suitable for subsequent feature representation.

Through these systematic and formalized data pre-processing steps, we obtained a cleaner and more refined version of the text suitable for further feature representation. These operations laid the foundation for subsequent stages of feature extraction, classification, and analysis in our task.

5.2 Classification using TF-IDF Features

In this research study, we started experimentation with the application of the TF-IDF (Term Frequency-Inverse Document Frequency) technique for machine learning-based classification tasks. The primary aim of our investigation was to assess the efficacy of TF-IDF-based approaches for text classification.

TF-IDF is a widely employed method in the field of natural language processing, which assigns weights to individual terms based on their occurrence frequency within a specific document and their rarity across the entire corpus. By considering both the local significance within a document and the global distinctiveness across the corpus, TF-IDF enables the identification of discriminative features crucial for classification purposes.

$$\text{TF-IDF: TF-IDF}(t,d) = \text{TF}(t,d) \times \text{IDF}(t)$$
$$\text{Max Document Frequency (max\_df) = 0.9}$$
$$\text{Min Document Frequency (min\_df) = 5}$$

To implement the TF-IDF-based classification, we utilized a range of machine learning algorithms, such as the random forest classifier and logistic regression wrapped in OneVsRest Classifier to perform the task. These algorithms are renowned for their effectiveness in handling text classification tasks. Additionally, we explored other similar algorithms to evaluate their performance and compare the obtained results.

For the configuration of the TF-IDF approach, we selected a value of 0.9 for max_df, indicating that we disregarded terms appearing in more than 90% of the documents. Furthermore, min_df was set to 5, implying the exclusion of words appearing in fewer than five documents. This parameter selection aimed to strike a balance between capturing diverse vocabulary while avoiding computational complexity. By limiting the dictionary to the most prevalent and informative terms, we sought to ensure robust classification performance while managing the dimensionality of the feature space. Table 3 provides a summary of the outcomes obtained from employing various machine learning algorithms using the training and development datasets.

5.3 Bag-Of-Words Features based Classification

In the domain of natural language processing, Bag-of-Words (BoW) features based text classification has emerged as a prevalent methodology for representing text documents as numerical feature vectors. Within the context of this research study, we also employed BoW-based text classification in conjunction with machine learning algorithms to effectively analyze and classify textual data.

The initial step in the BoW-based text classification process involved constructing a dictionary or vocabulary comprising unique words or terms. This dictionary served as the foundation for representing the documents. To ensure comprehensive coverage, we curated a dictionary consisting of 10,000 terms, encompassing the most frequent and informative terms derived from the training dataset.

Once the dictionary was established, each document was transformed into a sparse vector representation. This representation captured the presence or absence of terms from the dictionary within the document, along with their respective frequencies or weighted values, using techniques such as term frequency-inverse document frequency (TF-IDF).

To facilitate the training and classification of the BoW representations, we employed a diverse range of machine learning algorithms, including established models such as logistic regression, support vector machines, random forests, and decision trees, among others. These algorithms underwent training using a labeled dataset comprising documents and their respective class labels.

During the training phase, the machine learning algorithms learned the underlying patterns and relationships between the BoW features and their associated classes. Subsequently, we evaluated the trained models on a separate development dataset, employing performance metrics such as macro precision, macro recall, and macro F1-score, to assess their effectiveness and generalization capabilities.

In selecting the size of the dictionary, we consci-

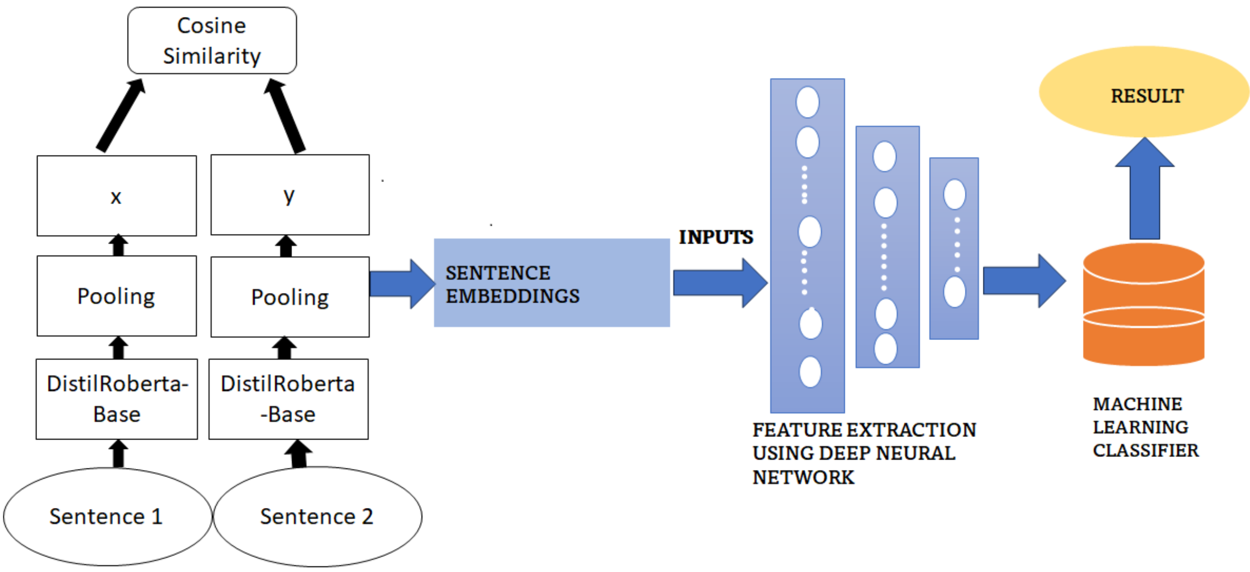

Figure 1: Overview of Proposed System

Table 3: TF-IDF features based Classification Results Summary on Training and Validation Datasets

Classifier	**Macro Precision**	**Macro Recall**	**Macro F1**
Ridge-Classifier	0.653	0.465	0.488
Perceptron-Classifier	0.482	0.471	0.474
SGD-classifier	0.583	0.487	0.511
Passive-Aggressive Classifier	0.498	0.465	0.477
Decision-Tree-Classifier	0.417	0.455	0.393
Random-Forest-Classifier	0.625	0.430	0.432
AdaBoost-Classifier	0.536	0.496	0.511
Gradient Boosting -Classifier	0.601	0.477	0.499
SVM Classifier	0.560	0.489	0.510

entiously considered the trade-off between capturing an adequate diversity of vocabulary and managing computational complexity. By opting for a dictionary size of 10,000, we aimed to strike an optimal balance between these considerations. Table 4 provides a summary of the experimental results obtained from training and evaluating the models using the training and validation datasets. Notably, it was observed that the Bag-of-Words features exhibited enhanced efficacy in classifying the text data.

5.4 Embeddings based Classification

In our research experiment, we employed classical Deep Learning algorithms, specifically the Multilayer Perceptron (MLP), to classify the sentence embeddings generated. To generate these embeddings, we utilized the widely used SBERT all-distilroberta-v1 model (Reimers and Gurevych, 2019). For this a pre-trained model distilroberta-base was used and then it has been trained on an extensive dataset consisting of 1 billion training pairs. It maps input sentences to a 768-dimensional vector space, enabling a comprehensive representation of the semantic information.To create embeddings(x,y) (Reimers and Gurevych, 2020) with a high level of semantic similarity, the all-distilroberta-v1 model uses a contrastive learning strategy. It measures language similarity using the

Table 4: BOW features based Classification Summary on Training and Validation Datasets

Classifier	Macro Precision	Macro Recall	Macro F1
Ridge-Classifier	0.505	0.473	0.484
Perceptron-Classifier	0.429	0.422	0.425
SGD-classifier	0.493	0.478	0.482
Passive-Aggressive Classifier	0.445	0.459	0.450
Decision-Tree-Classifier	0.401	0.447	0.384
Random-Forest-Classifier	0.609	0.426	0.430
AdaBoost-Classifier	0.53	0.48	0.50
Gradient Boosting -Classifier	0.593	0.472	0.493
SVM Classifier	0.476	0.476	0.475

Table 5: Embedding Based Classification Summary on Training and Validation Datasets

Classifier	Macro Precision	Macro Recall	Macro F1
SBERT + Decision Tree Classifier	0.449	0.452	0.450
SBERT + Random Forest Classifier	0.717	0.376	0.428
SBERT + DNN	0.635	0.511	0.540
SBERT+DNN+Random Forest	0.626	0.510	0.552
SBERT+DNN+AdaBoost	0.578	0.547	0.560
SBERT+DNN+Gaussian NB	0.498	0.598	0.505
SBERT+DNN+Decision Tree Classifier	0.506	0.515	0.510
SBERT+DNN+Ensemble	0.601	0.558	0.573

cosine similarity, dot product, and Euclidean distance. The model is appropriate for our classification assignment because it was created primarily as a sentence and brief paragraph encoder.

To generate the sentence embedding we set the maximum sequence length of the embedding model as 512. Following the generation of the sentence embeddings using SBERT all-distilroberta-v1, we proceeded to deploy various machine learning classifiers to make predictions based on these embeddings. Additionally, we constructed a Deep Neural Network (DNN) tailored for the classification of these embeddings. For training and evaluation purposes, we utilized the complete set of training and development embeddings, with a validation split of 0.1 to ensure reliable performance assessment.

To enhance the classification performance further, we extracted the last layer features of the DNN, which served as input for the machine learning classifiers. These last layer features encapsulated the learned representation of the embeddings and captured their discriminative properties. To exploit the complementary strengths of classical machine learning algorithms, we constructed a custom ensemble model that employed these extracted features for classification. This ensemble model combined the predictions from multiple classifiers, aiming to leverage their collective intelligence and improve overall classification performance.

The incorporation of the DNN in our methodology played a crucial role in learning the generalized distribution of the data. By employing deep neural networks, we enabled the model to capture complex patterns and relationships within the embeddings, enhancing its ability to generalize to unseen instances. The DNN's architecture, consisting of multiple layers with interconnected nodes, facilitated the extraction of hierarchical representations, enabling the model to uncover intricate features and capture underlying dependencies. To complement the textual description (Reimers and Gurevych, 2019), we have included Figure 1 showcasing the architecture of the proposed model. This visual representation provides a comprehensive overview of the connections and flow of information within the current model, enhancing the clarity and understanding of the methodology employed.

To provide comprehensive insights about this experimentation, we have documented all the obtained results in Table 5. This table presents a detailed overview of the performance achieved by the different classification models utilized in our study. Furthermore, we have submitted the architecture of the DNN in Table 6, illustrating the configuration and arrangement of the model's layers and nodes. This architectural representation facilitates a clear understanding of the underlying structure and organization of the DNN. The comprehensive results and model details provided offer valuable insights into the effectiveness and performance of our approach.

Table 6: Deep Neural Network Architecture and Hyperparameters

Hyperparameters	Values
Number of Layers	4
Activation Function(s)	Tanh and ReLU
Dropout Rate	0.2
Optimizer	Adam
Number of Epochs	2

6 Results

Table 7: Final Model results on the test dataset

Method	Macro F1-Score
SBERT+DNN+Ensemble	0.408

We demonstrate the outcomes of the task we submitted in this part. We utilised the configuration SBERT+DNN+Ensemble Model for prediction because we could see that it produced a significantly superior overall result. We were assessed using the Macro F1-Score, Macro Precision, and Macro Recall. On the test dataset shown in Table 7, we received an macro F1-Score of 0.408. The confusion matrix, which details the classification of numerous classes as well as classes that were incorrectly classified, is shown below as Figure 2. It is an essential tool for assessing the effectiveness and performance of our model. From the confusion matrix it is made clear that the classifier is more biased towards the label *Not depression*; as for the training data, it was more in number as compared to other labels, which are less in number Table 2. The Figure 3 shows that the classifier is biased and has not very well adapted towards the minority class. The problem can be overcome in future if the minority class is assigned a class weight and utilization of data augmentation methods for addressing class balance issues.

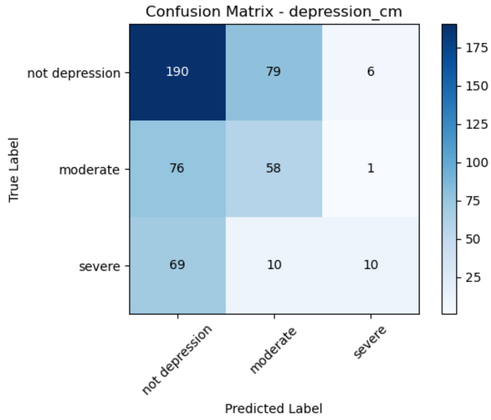

Figure 2: Confusion Matrix of Test Predictions

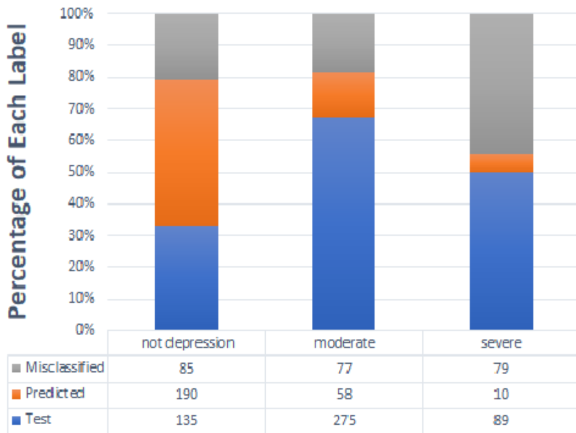

Figure 3: Comparison Results on Different Depression Classes

7 Conclusion

The successful completion of this task holds great potential for various applications and benefits. By accurately classifying the signs of depression within social media postings, the developed system can aid in the identification of individuals who may be at risk or in need of mental health support. This classification system could be integrated into social media platforms or utilized as a standalone tool to provide real-time insights into users' mental well-being. Early identification and intervention can play a crucial role in promoting mental health and well-being, and the system developed through this task has the potential to contribute significantly in this regard.

8 Future Work

The study can be further expanded for the related domains. One can empower one's studies with newer techniques like active learning. Semisupervised learning can also propose some findings as we might be able to generate synthetic data that act as an element for our training process.

Acknowledgements

We are thankful to Indian Institute of Information Technology Ranchi for all the support during our research.

References

Hatoon S AlSagri and Mourad Ykhlef. 2020. Machine learning-based approach for depression detection in twitter using content and activity features. *IEICE Transactions on Information and Systems*, 103(8):1825–1832.

Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. 2013. Predicting depression via social media. In *Proceedings of the international AAAI conference on web and social media*, volume 7, pages 128–137.

Sukanya Dessai and Soniya Shakil Usgaonkar. 2022. Depression detection on social media using text mining. In *2022 3rd International Conference for Emerging Technology (INCET)*, pages 1–4.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

SAGS Evans-Lacko, Sergio Aguilar-Gaxiola, Ali Al-Hamzawi, Jordi Alonso, Corina Benjet, Ronny Bruffaerts, WT Chiu, Silvia Florescu, Giovanni de Girolamo, Oye Gureje, et al. 2018. Socio-economic variations in the mental health treatment gap for people with anxiety, mood, and substance use disorders: results from the who world mental health (wmh) surveys. *Psychological medicine*, 48(9):1560–1571.

S Kayalvizhi, Thenmozhi Durairaj, Bharathi Raja Chakravarthi, et al. 2022. Findings of the shared task on detecting signs of depression from social media. In *Proceedings of the Second Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 331–338.

Chenhao Lin, Pengwei Hu, Hui Su, Shaochun Li, Jing Mei, Jie Zhou, and Henry Leung. 2020. Sensemood: depression detection on social media. In *Proceedings of the 2020 international conference on multimedia retrieval*, pages 407–411.

David E Losada, Fabio Crestani, and Javier Parapar. 2017. erisk 2017: Clef lab on early risk prediction on the internet: experimental foundations. In *Experimental IR Meets Multilinguality, Multimodality, and Interaction: 8th International Conference of the CLEF Association, CLEF 2017, Dublin, Ireland, September 11–14, 2017, Proceedings 8*, pages 346–360. Springer.

Rafał Poświata and Michał Perełkiewicz. 2022. Opi@ lt-edi-acl2022: Detecting signs of depression from social media text using roberta pre-trained language models. In *Proceedings of the Second Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 276–282.

Nils Reimers and Iryna Gurevych. 2019. Sentence-bert: Sentence embeddings using siamese bert-networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics.

Nils Reimers and Iryna Gurevych. 2020. Making monolingual sentence embeddings multilingual using knowledge distillation. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics.

Kayalvizhi S, Thenmozhi Durairaj, Bharathi Raja Chakravarthi, and Jerin Mahibha C. 2022. Findings of the shared task on detecting signs of depression from social media. In *Proceedings of the Second Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 331–338, Dublin, Ireland. Association for Computational Linguistics.

Rafael Salas-Zárate, Giner Alor-Hernández, María del Pilar Salas-Zárate, Mario Andrés Paredes-Valverde, Maritza Bustos-López, and José Luis Sánchez-Cervantes. 2022. Detecting depression signs on social media: a systematic literature review. In *Healthcare*, volume 10, page 291. MDPI.

Kayalvizhi Sampath and Thenmozhi Durairaj. 2022. Data set creation and empirical analysis for detecting signs of depression from social media postings. In *Computational Intelligence in Data Science: 5th IFIP TC 12 International Conference, ICCIDS 2022, Virtual Event, March 24–26, 2022, Revised Selected Papers*, pages 136–151. Springer.

Kayalvizhi Sampath, Thenmozhi Durairaj, Bharathi Raja Chakravarthi, Jerin Mahibha C, Kogilavani Shanmugavadivel, and Pratik Anil Rahood. 2023. Overview of the Second Shared Task on Detecting Signs of Depression from Social Media Text. In *Proceedings of the Third Workshop on Language Technology for Equality, Diversity and Inclusion*, Varna, Bulgaria. Recent Advances in Natural Language Processing.

David William and Derwin Suhartono. 2021. Text-based depression detection on social media posts: A systematic literature review. *Procedia Computer Science*, 179:582–589.

JT Wolohan, Misato Hiraga, Atreyee Mukherjee, Zeeshan Ali Sayyed, and Matthew Millard. 2018. Detecting linguistic traces of depression in topic-restricted text: Attending to self-stigmatized depression with nlp. In *Proceedings of the first international workshop on language cognition and computational models*, pages 11–21.