Overview of the shared task on Detecting Signs of Depression from Social Media Text

Kayalvizhi Sampath
SSN College of Engineering
kayalvizhis@ssn.edu.in

Durairaj Thenmozhi
SSN College of Engineering
theni_d@ssn.edu.in

Bharathi Raja Chakravarthi
National University of Galway, Ireland
bharathi.raja@insight-centre.org

Jerin Mahibha C
Meenakshi Sundararajan Engineering College
jerinmahibha@gmail.com

Kogilavani Shanmugavadivel
Kongu Engineering College
kogilavani.sv@gmail.com

Pratik Anil Rahood
National University of Galway, Ireland
p.rahood1@nuigalway.ie

Abstract

Social media has become a vital platform for personal communication. Its widespread use as a primary means of public communication offers an exciting opportunity for early detection and management of mental health issues. People often share their emotions on social media, but understanding the true depth of their feelings can be challenging. Depression, a prevalent problem among young people, is of particular concern due to its link with rising suicide rates. Identifying depression levels in social media texts is crucial for timely support and prevention of negative outcomes. However, it's a complex task because human emotions are dynamic and can change significantly over time. The DepSign-LT-EDI@RANLP 2023 shared task aims to classify social media text into three depression levels: "Not Depressed," "Moderately Depressed," and "Severely Depressed." This overview covers task details, dataset, methodologies used, and results analysis. Roberta-based models emerged as top performers, with the best result achieving an impressive macro F1-score of 0.584 among 31 participating teams.

1 Introduction

Depression is considered a common mental disorder that involves mood swings and a lack of interest in any activities [1]. Various aspects of life may be affected by depression. Depression may develop in people who undergo abuse, severe losses, or other stressful events. Depression has emerged as a worldwide concern for public health (Liu et al., 2022). Even though a lot many people suffer from depression, adequate treatment is received by only a fraction of them. Detecting and diagnosing depression is often delayed, imprecise, and missed. Depression is also considered an important cause of suicide. Based on the severity and impact of the symptoms, they can be categorized as mild, moderate, or severe.

Almost 72% of the population is found to be active on social media [2] which provides an opportunity for early detection of depression, particularly in young adults. Social media posts act as important information in identifying people at risk of depression or other mental disorders. There is a lot of research being carried out in the field of detecting depression from social media texts.

To analyse the massive amount of social media text, machine learning has been considered one of the most efficient approaches. The probability of the existence of depression had been predicted using different machine learning algorithms (Aleem et al., 2022). Detection of depression from social media text had also been implemented using Long-Short Term Memory (LSTM) model with Recurrent Neural Network (RNN) (Amanat et al., 2022).

With all these informations in mind, the shared task on detecting the levels of depression from social media posts has been organised which is a continuation of the shared task DepSign-LT-EDI@ACL-2022 (S et al., 2022) conducted during 2022. The shared task DepSign-LT-EDI@RANLP-2023 aims to detect levels of depression in Reddit posts.

2 Task description

The objective of DepSign-LT-EDI@RANLP-2023 is to identify indicators of depression in individuals based on their posts on social media platforms. By analyzing the language, feelings, and emotions expressed in these posts, the system aims to categorize individuals into three levels of depression: "Not Depressed," "Moderately Depressed," and

[1] https://www.who.int/news-room/fact-sheets/detail/depression

[2] https://www.internetlivestats.com/twitter-statistics/

LT-EDI 2023 – Third Workshop on Language Technology for Equality, Diversity and Inclusion, pages 25–30, Varna, Bulgaria, Sep 7, 2023.
https://doi.org/10.26615/978-954-452-084-7_004

”Severely Depressed.” The system works specifically with English social media postings to detect signs of depression.

3 Data description

To determine the level of depression based on social media data, a dataset was created by scraping and labeling posts from Reddit. The dataset is the extension of DepSign-LT-EDI@ACL-2022 (S et al., 2022). It consists of three class labels: ”Not Depressed,” ”Moderately Depressed,” and ”Severely Depressed.” Detailed guidelines and annotation instructions can be found in the publication by (S and D, 2022). The dataset is provided in a ”Tab Separated Value” format and is divided into three sets: the training set, evaluation set, and test set. The distribution of data across these sets is illustrated in Table 1. Examples of sample instances from the dataset are shown in Table 2.

DepSign-LT-EDI@RANLP-2023	Train	Dev	Test
Not depressed	2755	848	136
Moderate	3678	2,169	275
Severe	768	228	88
Total instances	**7201**	**3245**	**499**

Table 1: Data set distribution

4 Methodology

A total number of 62 submissions were submitted by 31 teams.

- **DeepLearningBrasil**(Garcia et al., 2023) The submission of the team used domain-specific RoBERTa and DeBERTa models by further pre-training them on a scraped Reddit comments corpus extracted from subreddits with a mental health theme. The process was evaluated using different techniques, such as different truncation mechanisms, ordinal classification specific losses, and sample weights on the loss function, to deal with the unbalanced data. The three submissions consisted of different ensemble approaches for nine models with various techniques applied.

- **DeepBlueAI** The submission of the team had implemented the process of detecting levels of depression by fine-tuning with XLM-RoBERTa as the base model.

- **Cordyceps**(Ninalga, 2023) The team had implemented the process of self-distillation using unlabeled data from the ”Reddit Mental Health Dataset”. The process of prediction used a modified RoBERTa model named ”MentalRoBERTa”.

- **iicteam**(Vajrobol et al., 2023) The team has applied MentalRoBERTa, which is a model initialized with RoBERTa-Base (cased_L-12_H-768_A-12) and trained with mental health-related posts collected from Reddit.

- **CIMAT-NLP**(María de Jesús García Santiago and Monroy, 2023) The first two submissions used a transformer-based approach with differences in the dataset training. The dataset provided by the organizers was used but was structured differently. The third submission used an ensemble of BOW.

- **IJS**(Caporusso and Hanh Tran, 2023) The team applied language models, such as monolingual and multilingual BERT and XLNet, to predict depression levels based on given sentences. These models leverage their ability to understand contextual relationships within the text to capture nuanced linguistic features associated with depression.

- **NLP_CHRISTINE**(Christodoulou, 2023) The team had used the majority ensemble learning of three DeBERTa-V3-Large model architectures.

- **Biasbusters**(Nedilko, 2023) Three runs utilizing baseline BoW XGBoost with numeric engineered features, ChatGPT zero-shot, and the GPT-3 DaVinci model had been submitted.

- **NLP_JA**(Kumari and Kumar, 2023) The team had fine-tuned RoBERTa for detecting the depression level associated with the given text.

- **ML&AI_IIITRanchi**(Kumari et al., 2023) The submission made use of features like Tf-idf and BOW, along with sentence embeddings that were trained with deep neural networks and ensemble machine learning techniques.

- **TechSSN1**(Sivanaiah et al., 2023) The first run is based on a pre-trained BERT model

PID	Text Data	Class label
train_pid_1	My life gets worse every year : That's what it feels like anyway....	moderate
train_pid_2	Words can't describe how bad I feel right now : I just want to fall asleep forever.	severe
train_pid_3	Is anybody else hoping the Coronavirus shuts everybody down?	not depressed

Table 2: Sample instances of data set

that has been fine-tuned for depression analysis by training it on a specific dataset. To convert the target labels into numerical values, a label encoder was employed. The second run used Word2Vec to generate word embeddings that capture the contextual meaning of words in our vocabulary. The Support Vector Classifier (SVC) had been employed to train and predict based on the word vectors. The third run employed the TfidfVectorizer, which converts text into numerical feature vectors. Subsequently, the vectors were fitted to the LinearSVC model.

- **Interns**(L et al., 2023) The team had submitted three runs, of which the first run used linear SVM, the second run used textblob, and the third run was based on bi-LSTM.

- **Team-KEC** The training data had been preprocessed and balanced using SMOTE. Word embedding techniques such as N-Gram (Trigram) and Fasttext were used for feature extraction. Models like SVM, CNN, and BERT were used for prediction. Various combinations of word embedding and ML and DL models have been tried to achieve the best outcome.

- **BLP Navigator** Depression had been detected using transformer-based models.

- **Deepalaksmi** ALBERT model was used for detecting the signs of depression provided in the test dataset. Due to the large number of words in each instance of the training dataset, the text summarization method is used to extract the core words without losing the originality of the text. Also, text preprocessing methods were used to enhance the performance.

- **SENTIZEN** The team had submitted runs that implemented the process of classification using logistic regression, random forest, and the K nearest neighbors algorithm.

- **Ramya Sivakumar** Machine learning-based passive classifier was used to evaluate and predict values for the given dataset.

- **KEC_NL_DEP**(Shanmugavadivel et al., 2023) Different machine learning algorithms like logistic regression, decision tree, multinomial Naive Bayes, Gaussian Naive Bayes, and random forest have been used for detecting depression from social media text.

- **mucs**(Coelho et al., 2023) The submitted run used TF-IDF vectorizers for feature extraction and BERT models for the process of classification.

- **SIS**(B K et al., 2023) The model used bagging, which is an ensemble learning technique that helps improve the performance and accuracy of machine learning algorithms. The prediction has been done using Multi-Layer Perceptron (MLP), Recurrent Neural Network (RNN), and then Linear SVM.

- **meghaAarthi** In this system, multiple models from simple transformers have been used, and the final output is predicted on the basis of voting classification.

- **codemonkeys** Basic-bert-base-uncased had been incorporated, and the model had been fine-tuned for the specific dataset. A custom-made stopword list without using the nltk libraries has been used, and data augmentation has been done while predicting the model values.

- **the_mavericks**(Sathvika et al., 2023) The process of feature extraction from the preprocessed text data had been implemented using bag-of-words representation and count vectorization weighting. The Naive Bayes model was trained using the labeled data, learning the probabilities associated with each feature for the depressive and non-depressive classes.

- **Flamingos_python**(P S et al., 2023) An ensemble model combining three machine learning algorithms, namely Random Forest, SVM, and Naive Bayes classifier, was used to train the model for detecting the level of depression in a text.

- **KEEMS** Machine learning algorithms such as Random Forest, Support Vector Machine, and Ensembled Model with both Random Forest and Support Vector Machine were used for the three submissions. Google Translator had been used for up-sampling the dataset.

- **Tercet** The method that had been employed for the task of identifying the level of depression was Support Vector Machines. A Tf-idf vectorizer was used to extract the features based on which the SVM model had been applied.

- **Techwhiz**(M et al., 2023) Transformer-based models, namely ALBERT and RoBERTa, was used to implement the process of classification.

- **spr** Three runs were submitted by the team, which used different machine learning models. The first run was based on Logistic Regression, the second run used the Random Forest classifier, and the third run applied voting to Logistic Regression and Random Forest classifier.

- **Supernova**(Reddy et al., 2023) The team used the TF-IDF feature extraction mechanism and a Support Vector Machine to implement the process of classification.

5 Evaluation

The evaluation encompassed the utilization of all performance metrics available in sklearn. To account for the dataset's inherent imbalance, the submitted runs were assessed and ranked primarily based on the macro F1 score. The teams' rankings are presented in Table 3.

A notable observation from the table is that the system developed by the DeepLearningBrasil team excelled, achieving the highest macro F1 score of 0.584 and the top accuracy score of 0.565.

6 Analysis and discussion

Early detection of depression is crucial as it is a prevalent mental illness that profoundly affects an individual's mood and emotions. Failure to recognize and address depression at its early stages can have severe consequences. The DepSign-LT-EDI@RANLP-2023 shared task focused on utilizing Reddit postings to detect various levels of depression. The detection process involved assigning three labels to individuals: "Not Depressed," "Moderately Depressed," and "Severely Depressed." By analyzing the content of these posts, the aim was to identify and categorize individuals based on their level of depression.

7 Conclusion

Depression is a widespread mental health issue that profoundly affects a person's emotions and well-being. Detecting it early is crucial to prevent potential harm. The DepSign-LT-EDI@RANLP-2023 challenge aimed to identify depression levels from Reddit posts, categorizing them as "Not Depressed," "Moderately Depressed," or "Severely Depressed."

In this shared task, 31 teams participated, employing various models, with a strong focus on transformer-based methods and machine learning. The systems were assessed using the macro-averaged F1-score. Remarkably, Team DeeplearningBrasil excelled by combining Roberta in an ensemble approach, achieving an impressive F1 score of 0.584. This underscores the promise of advanced natural language processing techniques in early depression detection.

Acknowledgments

We extend our gratitude for the financial support received from the Department of Science and Technology - Science and Engineering Research Board (DST-SERB).

One of the authors Bharathi Raja Chakravarthi was supported in part by a research grant from Science Foundation Ireland (SFI) under Grant Number SFI/12/RC/2289_P2 (Insight_2)

References

Shumaila Aleem, Noor ul Huda, Rashid Amin, Samina Khalid, Sultan S Alshamrani, and Abdullah Alshehri. 2022. Machine learning algorithms for depression:

Team Name	Accuracy	Weighted F1-score	Weighted Recall	Weighted Precision	Macro Recall	Macro Precision	Macro F1-score	Rank (based on Macro F1 score)
DeepLearningBrasil	**0.565**	0.470	0.473	0.588	0.540	0.565	**0.584**	1
DeepBlueAI	0.525	0.446	0.457	0.554	0.506	0.525	0.554	2
Cordyceps	0.521	0.441	0.451	0.517	0.503	0.521	0.535	3
iicteam	0.525	0.439	0.439	0.513	0.506	0.525	0.528	4
CIMAT-NLP	0.505	0.439	0.447	0.467	0.493	0.505	0.503	4
TechSSN4_English	0.479	0.437	0.436	0.501	0.475	0.479	0.509	6
IJS	0.487	0.425	0.438	0.487	0.476	0.487	0.513	7
NLP_CHRISTINE	0.543	0.420	0.474	0.459	0.491	0.543	0.513	8
Biasbusters	0.499	0.419	0.442	0.489	0.481	0.499	0.518	9
NLP_JA	0.523	0.413	0.426	0.510	0.494	0.523	0.527	10
ML&AI_IIITRanchi	0.517	0.408	0.411	0.517	0.488	0.517	0.524	11
TechSSN1	0.549	0.407	0.416	0.537	0.504	0.549	0.545	12
Interns	0.445	0.402	0.408	0.400	0.449	0.445	0.458	13
Team-KEC	0.451	0.401	0.414	0.458	0.445	0.451	0.486	14
BLP Navigator	0.453	0.387	0.415	0.474	0.439	0.453	0.500	15
Deepalaksmi	0.463	0.382	0.395	0.439	0.447	0.463	0.473	16
SENTIZEN	0.425	0.371	0.378	0.392	0.420	0.425	0.434	17
Ramya Sivakumar	0.451	0.365	0.385	0.399	0.436	0.451	0.455	18
KEC_NL_DEP	0.433	0.362	0.378	0.408	0.422	0.433	0.451	19
mucs	0.475	0.361	0.397	0.435	0.441	0.475	0.471	20
testresult	0.409	0.359	0.378	0.376	0.407	0.409	0.439	21
SIS	0.501	0.345	0.371	0.527	0.452	0.501	0.522	22
meghaAarthi	0.511	0.328	0.363	0.305	0.449	0.511	0.405	23
codemonkeys	0.409	0.323	0.398	0.432	0.368	0.409	0.470	24
the_mavericks	0.551	0.263	0.341	0.338	0.412	0.551	0.430	25
Flamingos_python	0.477	0.262	0.316	0.226	0.388	0.477	0.328	26
KEEMS	0.477	0.262	0.316	0.226	0.388	0.477	0.328	26
Tercet	0.533	0.259	0.331	0.591	0.403	0.533	0.541	28
Techwhiz	0.421	0.258	0.303	0.227	0.370	0.421	0.331	29
spr	0.407	0.245	0.287	0.216	0.356	0.407	0.318	30
Supernova	0.341	0.155	0.158	0.385	0.324	0.341	0.477	31

Table 3: Team Wise results

diagnosis, insights, and research directions. *Electronics*, 11(7):1111.

Amna Amanat, Muhammad Rizwan, Abdul Rehman Javed, Maha Abdelhaq, Raed Alsaqour, Sharnil Pandya, and Mueen Uddin. 2022. Deep learning for depression detection from textual data. *Electronics*, 11(5):676.

Sulaksha B K, Shruti Krishnaveni S, Ivana Steeve, and Monica Jenefer B. 2023. Detecting signs of depression from social media text. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, Varna, Bulgaria. Recent Advances in Natural Language Processing.

Jaya Caporusso and Thi Hong Hanh Tran. 2023. Ensemble approaches to detect signs of depression from social media text. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, Varna, Bulgaria. Recent Advances in Natural Language Processing.

Christina Christodoulou. 2023. Roberta deberta finetuning for detecting signs of depression from social media text. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, Varna, Bulgaria. Recent Advances in Natural Language Processing.

Sharal Coelho, Asha Hegde, Kavya G, and Hosahalli Lakshmaiah Shashirekha. 2023. Detecting signs of depression in social media text. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, Varna, Bulgaria. Recent Advances in Natural Language Processing.

Eduardo Garcia, Juliana Gomes, Adalberto Ferreira Barbosa Junior, Cardeque Henrique Bittes de Alvarenga Borges, and Nadia Félix Felipe da Silva. 2023. Detecting signs of depression from social media text. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, Varna, Bulgaria. Recent Advances in Natural Language Processing.

Fernando Sánchez Vega María de Jesús García Santiago and Adrián Pastor López Monroy. 2023. Finegrain depression detection by multiple binary problems approach. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, Varna, Bulgaria. Recent Advances in Natural Language Processing.

Jyoti Kumari and Abhinav Kumar. 2023. Empowering mental health assessment: A roberta-based approach for depression detection. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, Varna, Bulgaria. Recent Advances in Natural Language Processing.

Kirti Kumari, Shirish Shekhar Jha, Zarikunte Kunal Dayanand, and Praneesh Sharma. 2023. Hybrid model for text classification for identification of various types of depression. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, Varna, Bulgaria. Recent Advances in Natural Language Processing.

Koushik L, Hariharan R. L, and Anand Kumar M. 2023. Detecting signs of depression from social media text. In *Proceedings of the Third Workshop on Speech and*

Language Technologies for Dravidian Languages, Varna, Bulgaria. Recent Advances in Natural Language Processing.

Danxia Liu, Xing Lin Feng, Farooq Ahmed, Muhammad Shahid, Jing Guo, et al. 2022. Detecting and measuring depression on social media using a machine learning approach: systematic review. *JMIR Mental Health*, 9(3):e27244.

Madhumitha M, Jerin Mahibha C, and Thenmozhi Durairaj. 2023. Transformer models to detect levels of depression from social media text. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, Varna, Bulgaria. Recent Advances in Natural Language Processing.

Andrew Nedilko. 2023. Detecting signs of depression with generative pretrained transformers. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, Varna, Bulgaria. Recent Advances in Natural Language Processing.

Dean Ninalga. 2023. Depression detection with reddit and self-training. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, Varna, Bulgaria. Recent Advances in Natural Language Processing.

Abirami P S, Amritha S, Pavithra Meganathan, and Jerin Mahibha C. 2023. An ensemble model to detect severity of depression. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, Varna, Bulgaria. Recent Advances in Natural Language Processing.

Ankitha Reddy, Pranav Moorthi, and Ann Maria Thomas. 2023. Detecting signs of depression from social media text. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, Varna, Bulgaria. Recent Advances in Natural Language Processing.

Kayalvizhi S and Thenmozhi D. 2022. Data set creation and empirical analysis for detecting signs of depression from social media postings. *CoRR*, abs/2202.03047.

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.

V S Sathvika, S Vaishnavi, S Angel Deborah, S Rajalakshmi, and T T Mirnalinee. 2023. Detection of signs of depression from social media text quotes using navie bayse approach. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, Varna, Bulgaria. Recent Advances in Natural Language Processing.

Kogilavani Shanmugavadivel, Malliga Subramanian, Vasatharan K, Prethish GA, Sankar S, and Sabari S. 2023. Detecting signs of depression from social media texts. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, Varna, Bulgaria. Recent Advances in Natural Language Processing.

Rajalakshmi Sivanaiah, Venkatasai Ojus Yenumulapalli, Vijai Aravindh R, and Angel Deborah S. 2023. Depression detection and classification using bert model for social media texts. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, Varna, Bulgaria. Recent Advances in Natural Language Processing.

Vajratiya Vajrobol, Nitisha Aggarwal, and Karanpreet Singh. 2023. Leveraging pre-trained transformers for fine-grained depression level detection in social media. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, Varna, Bulgaria. Recent Advances in Natural Language Processing.