

# DepressionOne@LT-EDI-ACL2022: Using Machine Learning with SMOTE and Random UnderSampling to Detect Signs of Depression on Social Media Text.

**Suman Dowlagar**

LTRC

IIIT-Hyderabad

suman.dowlagar

@research.iiit.ac.in

**Radhika Mamidi**

LTRC

IIIT-Hyderabad

radhika.mamidi

@iiit.ac.in

## Abstract

Depression is a common and serious medical illness that negatively affects how you feel, the way you think, and how you act. Detecting depression is essential as it must be treated early to avoid painful consequences. Nowadays, people are broadcasting how they feel via posts and comments. Using social media, we can extract many comments related to depression and use NLP techniques to train and detect depression. This work presents the submission of the DepressionOne team at LT-EDI-2022 for the shared task, detecting signs of depression from social media text. The depression data is small and unbalanced. Thus, we have used oversampling and undersampling methods such as SMOTE and RandomUnderSampler to represent the data. Later, we used machine learning methods to train and detect the signs of depression.

## 1 Introduction

According to Psychiatry, depression is defined as a mental condition characterized by severe despondency and dejection, typically also with feelings of inadequacy and guilt, often accompanied by lack of energy and disturbance of appetite and sleep. Depression remains a significant issue worldwide, and often it progresses to suicidal intention if left undetected (Haque et al., 2021). Thus the diagnosis of depression is an important task. Many existing methods for detecting depression rely on Electronic health records or suicide notes. But such data is limited and challenging to acquire.

In the current generation, online forums on social media act as a means where people vent out how they feel. We can scrape these resources to create datasets. Such data, if annotated, can be helpful to detect depression (Haque et al., 2021). A growing number of studies are using such data for research and diagnostic purposes. A survey on detecting depression using social media data is given in the paper Ji et al. (2020). Detecting depression

represents a significant clinical challenge, both for the advancement of how depression is treated and for implementing interventions (Leonard, 1974).

To encourage work on depression from social media comments/posts, the LT-EDI community has organized a shared task to identify the signs of depression of a person from their social media postings where people share their feelings and emotions ("Sampath et al., 2022). The dataset used for this task has a total of 16,632 train, valid, and test comments in the English language. This task aims to classify the given depression data into three classes, severe, moderate, and not\_depression.

This paper presents a method for detecting/classifying depression text. We have used under sampling and oversampling to represent the data better. Then we used a machine learning classifier to train and classify the given text.

The paper is organized as follows. Section 2 provides related work on depression detection on social media text. Section 3 provides information on the task and datasets. Section 4 describes the our submission. Section 5 presents the experimental setup and the performance of the model. Section 6 concludes our work.

## 2 Literature Survey

This section provides a brief of research done till now on depression detection.

(Salas-Zárate et al., 2022) surveyed on detecting depression using social media data (from 2016 to mid-2021). The survey analyzed and evaluated Thirty-four primary studies. Twitter was the most studied social media. Word embedding was the most prominent linguistic feature extraction method. Support vector machine (SVM) was the most used machine-learning algorithm.

(William and Suhartono, 2021) conducted a for early depression detection in textual data. The review found three concerning issues, i.e., (1) Ethical concerns, (2) Lack of data, (3) Awareness of mental

well-being. The classifiers mostly used were Support Vector Machine and Probabilistic Classifier. The survey observed that the BiLSTM + Attention method yields the best result. The models such as BERT were not suitable for depression detection because of their inability to deal with long sequences. So new methods such as summarizing the text were proposed to deal with long sequences before feeding it into the model.

The given depression dataset has long sequences, and the BERT models could not process long sequences. Also, text summarization techniques are not 100% accurate and will be propagated to BERT models. Thus we opted for the SVM and RNN models for depression detection.

### 3 Task Setup

The goal of this task is to detect depression from social media. The model should classify the signs of depression into three labels, namely “not depressed”, “moderate”, and “severe”. The dataset has 16,632 comments, wherein 8,891 belong to the training set, 4,496 belong to validation, and 3,245 belong to the test set. All the posts are in the English language.

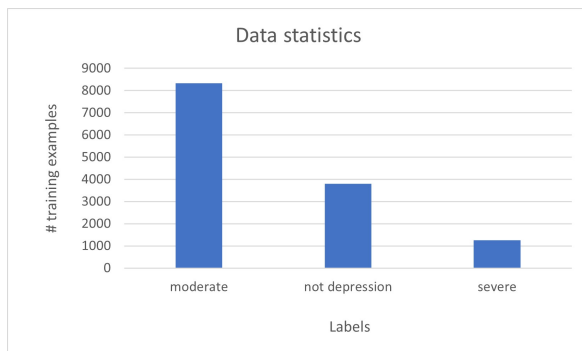


Figure 1: Data Statistics w.r.t the three labels

As given in Figure 1, we can see that there are more instances of moderate classes when compared to the not-depressed and severe classes in the given data. Also, the data has a wide range of sentence lengths as given in Figure 2. We have also observed that more than 6% of the sentences are long<sup>1</sup>. The long sentences are not suitable for the BERT model as it only works if the tokens are less than 512. So we chose robust classification algorithms such as SVM to classify the data and detect depression in the dataset.

<sup>1</sup>where long means tokens in the sentence are greater than 512

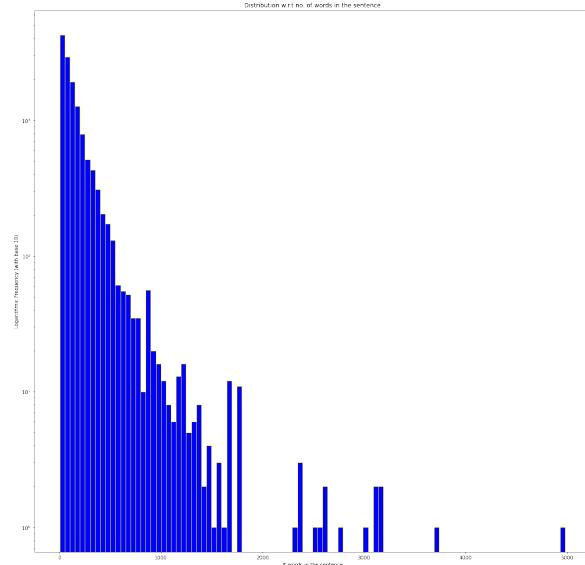


Figure 2: Logarithmic Distribution of data with respect to sentence length

### 4 Our Submission

As mentioned above, there are more instances of moderate class labels when compared to not-depressed and severe. It leads to an imbalance in data. So we chose the resampling method. Resampling involves creating a new altered version of the training dataset, in which the selected examples have a similar class distribution. The simple way is to choose instances for the transformed dataset randomly. Thus it is called random resampling. It is a simple and effective strategy to handle imbalanced classification problems.

The two main methods of random resampling are oversampling and undersampling.

**Random Oversampling** Random oversampling involves randomly selecting examples from the minority class, with replacement, and adding them to the training dataset.

**Random Undersampling** Random undersampling involves randomly selecting examples from the majority class and deleting them from the training dataset until a more balanced distribution is reached.

This technique is practical where the skewed distribution affects the classification models, and multiple examples for a given class can overfit the model. It makes the model to be biased towards the class that has the majority of instances.

If we only use random undersampling for the given majority class, i.e., moderate, then the data

Model	Labels (F1-score)				Accuracy
	moderate	not-depression	severe	macro-avg	
<b>SVM</b>	0.65	0.47	0.37	0.50	0.56
<b>RNN</b>	0.57	0.56	0.28	0.48	0.54
<b>CNN</b>	0.58	0.57	0.27	0.47	0.53
<b>BERT</b>	0.58	0.57	0.29	0.48	0.54
<b>Our Submission</b>	0.72	0.47	0.37	0.60	0.65

Table 1: The performance of the models on the depression dataset (On validation data).

might lose some specific points, which might degrade the model’s performance. Also, If only use random oversampling for the minority classes, i.e., not depression and severe. This oversampling method can balance the class distribution but does not provide additional information to the model.

An improvement in duplicating instances from the minority class is synthesizing new instances from the minority class. The most widely used approach to synthesize new instances is called the Synthetic Minority Oversampling Technique (abbreviated as SMOTE) (Chawla et al., 2002). SMOTE selects instances that are close in the feature space by fitting a line between the instances in the feature space and selecting a new sample at a point along that line.

Chawla et al. (2002) suggests that, using random undersampling to trim the number of in the majority class, then use SMOTE to oversample the minority class to balance the class distribution. The combination of SMOTE and under-sampling performs better than plain under-sampling.

After resampling the classes, we have used an SVM classifier to train the transformed data and applied the model to the test set.

## 5 Experiments

The section presents the baselines, hyper-parameter settings, and analysis of observed results.

### 5.1 Baselines

The baselines used are:

**SVM with TF-IDF** Term frequency and inverse document frequency-based vectorization is used to represent the text data, and the support vector machine is used to classify the data.

**CNN** (Kim, 2014) This convolutional neural network-based text classifier is trained by considering pre-trained FastText word vectors.

**Bi-LSTM** (Hochreiter and Schmidhuber, 1997)

A two-layer, bi-directional LSTM text classifier with pre-trained FastText word embeddings as input was considered for the task of text classification.

**Pre-trained BERT** (Devlin et al., 2018) A pre-trained BERT model with a feed-forward network for classification

### 5.2 Hyperparameters and Libraries used

The SMOTE is obtained from the imblearn library<sup>2</sup>. The Random oversampling, SVM with TF-IDF vectorization, is obtained from the scikit-learn library<sup>3</sup>. The default parameters are used to train the SVM for multiclass classification. The pre-trained BERT with sentence classification is obtained from the huggingface transformers library<sup>4</sup>. The optimizer used is weighted Adam with the learning rate of 1e-5 and epsilon value equal to 1e-8. The loss function used is the BERT’s inbuilt cross-entropy loss. The number of epochs used for training the model is 30. We have used PyTorch<sup>5</sup> for implementing Bi-LSTM and CNN models. The number of Bi-LSTM layers is given as 2. For CNN, we took three kernels of sizes 2,3,4. We have used the adam optimizer with cross-entropy loss for the given models. The batch size is 64. The models were run on GPU notebooks.

### 5.3 Results

From the results, we can see that the performance of the SVM and our approach are better when compared to the neural network(NN) models. The NN models didn’t perform better on all the labels. The models didn’t distinguish between the “moderate” and “not depression” labels and “severe”

<sup>2</sup>[https://imbalanced-learn.org/stable/references/generated/imblearn.over\\_sampling.SMOTE.html](https://imbalanced-learn.org/stable/references/generated/imblearn.over_sampling.SMOTE.html)

<sup>3</sup><https://scikit-learn.org/stable/>

<sup>4</sup><https://huggingface.co/>

<sup>5</sup><https://pytorch.org/>

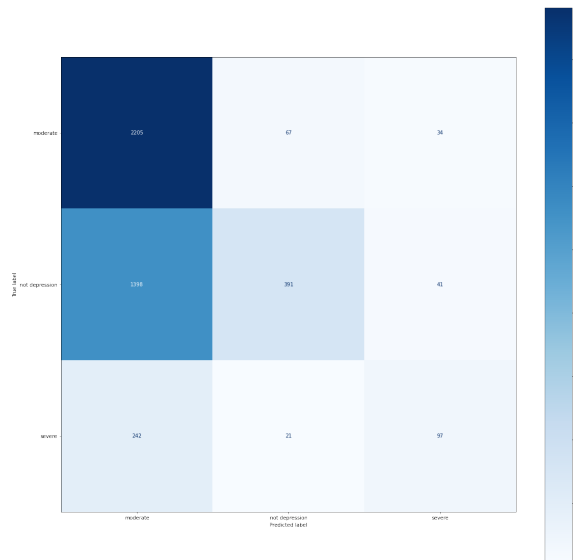


Figure 3: Confusion matrix of our submission

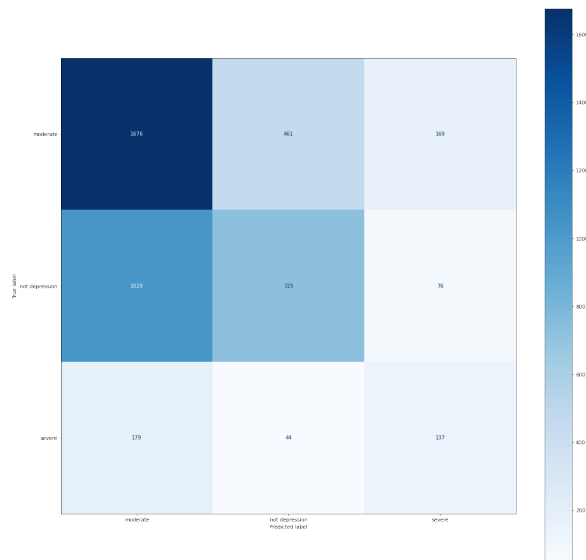


Figure 4: Confusion matrix of SVM + TF-IDF baseline

and “not depression” labels, resulting in decreased performance. In contrast, SMOTE and Random undersampling helped the model generate synthetic points that helped the model tune better, thus leading to improved performance. The SVM model didn’t distinguish between the “moderate” and “not depression” labels. Whereas it relatively showed improved performance on “severe” and “not depression” labels compared to the NN models. We also compared the confusion matrices of our model with the top-performing baseline (SVM with TF-IDF). The confusion matrices are given in the Figures 3 and 4. We have observed that our submission showed better performance on “moderate” labels when compared to SVM with TF-IDF baseline.

But our model showed a decreased performance on the “not-depressed” model. But the number of instances of correctly classified “moderate” instances was more, resulting in increased accuracy.

## 5.4 Conclusion

We used SMOTE and random undersampling with an SVM classifier to detect signs of depression. The dataset is in English and has a wide range of sentence lengths, and it is imbalanced. In the dataset, 6% of sentences have more than 500 words. We used SMOTE and random undersampling methods to balance the dataset. We tested the method on other neural network baselines. The results showed that using the oversampling and undersampling methods handled the problem of imbalanced data. It, in turn, helped the machine learning classifier, i.e., SVM, to perform better on the transformed dataset. Due to the presence of long sentences, the BERT model didn’t perform better on the given dataset. We hope to test the meta embedding models on the given dataset in the future.

## References

- Nitesh V Chawla, Kevin W Bowyer, Lawrence O Hall, and W Philip Kegelmeyer. 2002. Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16:321–357.
- 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*.
- Ayaan Haque, Viraj Reddi, and Tyler Giallanza. 2021. Deep learning for suicide and depression identification with unsupervised label correction. In *International Conference on Artificial Neural Networks*, pages 436–447. Springer.
- Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation*, 9(8):1735–1780.
- Shaoxiong Ji, Shirui Pan, Xue Li, Erik Cambria, Guodong Long, and Zi Huang. 2020. Suicidal ideation detection: A review of machine learning methods and applications. *IEEE Transactions on Computational Social Systems*, 8(1):214–226.
- Yoon Kim. 2014. [Convolutional neural networks for sentence classification](#). In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1746–1751, Doha, Qatar. Association for Computational Linguistics.
- CV Leonard. 1974. Depression and suicidality. *Journal of consulting and clinical psychology*, 42(1):98.

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. Multidisciplinary Digital Publishing Institute.

Kayalvizhi "Sampath, 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*. Association for Computational Linguistics.

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