**An Experiment of Discourse and Sentiment Analysis for the Prediction of Empathy, Distress and Emotion**

# Introduction

The paper outlines the task of predicting empathy, distress, and emotion, and discusses various learning strategies and language models used in the study. The authors' main objective was to analyze and extract emotion-intensive features from essays in the training data and news articles to improve the prediction of empathy and distress scores using discourse and sentiment analysis. To overcome the issue of limited training examples for fine-tuning pre-trained language models, the authors propose several text feature extraction schemes, including methods based on Rhetorical Structure Theory (RST) parsing, cosine similarity, and sentiment score.

# Related Work

In order for early intervention to be viable, depression detection models need to be highly accurate and fast. (Shen et al., 2017) proposed the extraction of six feature groups, which were then used to train a multi-modal depression dictionary learning model for identifying depressed Twitter users. (Burdisso et al., 2019) introduced the SS3 text classification system, which is both easy to use and effective for early diagnosis of depression in social media streams. (Lin et al., 2020) presented SenseMood, a system that utilizes a BERT classifier and a CNN to classify social media messages and images as either depressed or not depressed.

BERT models are commonly utilized in the field of natural language processing. To gain a deeper understanding of how these models operate, (Van Aken et al., 2019) conducted a Layer-Wise Analysis of Transformer Representations. (Devlin et al., 2018) demonstrated how pre-trained models can be leveraged for natural language interpretation. (Acheampong et al., 2021) provide an overview of BERT-based models for text-based emotion recognition. (Xin et al., 2020) proposed an early stopping modification of BERT to improve inference speed.

# Experiments

## Dataset

The dataset utilized in the shared task was originally collected for conducting experiments on the prediction of empathy using Batson's Empathic Concern and Personal Distress Scale (Batson et al., 1987) as the basis. The participants were provided with news articles to read and then instructed to write a brief essay describing their emotional response to the news. Subsequently, they were given questions to answer, and their responses were evaluated to determine their levels of empathy and distress, which were graded on a scale ranging from level 1 to 7. The data collection process also involved gathering demographic and personality information from the participants to examine how these variables might impact their empathy and distress levels. The emotion labels assigned to the data were generated through a semi-automated process that involved manual corrections of the automatic predictions made by deep learning models. The class distribution in the dataset is presented in Figure 1.

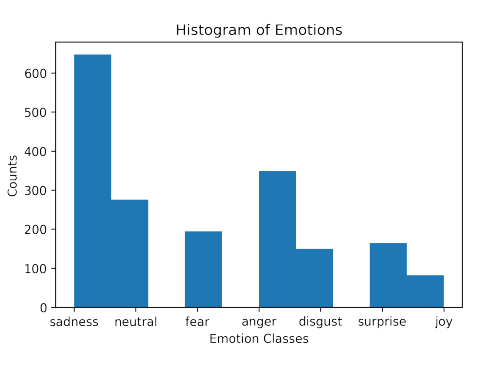


Figure 1: Distribution of Emotion Classes in the dataset.

## Implementation

The implementation of the paper # was performed using the GitHub repo provided by the author of the paper at this link: <https://github.com/shenbinqian/WASSA_SharedTask>

## Results

By using the provided code, we calculated the f1 score for the emotion prediction. The fine-tuning the model, we calculated the highest 0.597 f1 score. The comparison of the classification results is presented in the below table. However, the regression results are difficult to reproduce using the provided code.

|  |  |  |
| --- | --- | --- |
| Metric | Reported Results | Reproduced Results |
| Accuracy | 0.646 | 0.651 |
| Precision | 0.595 | 0.604 |
| Recall | 0.559 | 0.587 |
| F1 Score | 0.559 | 0.597 |

## Discussion

The code of the selected paper is available via the GitHub repository that was used here for the reproduction of the results. The Roberta classifier showed a maximum 0.597% f1 score after hyperparameter tuning. The results of the tuned model are slightly higher than the reported results. Numerous factors can play a role in this regard including the different values of parameters and different split ratios and seed values. If the exact split ratio and seed values are used in future studies, then the exact result can be reproduced.

## Error Analysis

For the error analysis, a trained model must be required. As we are unable to reproduce the regression results that resultantly didn’t provide any error analysis. Moreover, the paper didn’t discuss any error analysis related to classification and regression models.

# Conclusion

By following the model parameter tuning and model training, we conclude that the results of the paper are partially reproducible. The classification results can be reproduced from the development of the code from scratch or by using the GitHub repository of the paper provided by the author. But the regression results were not reproduced using the provided code. The reprodiction of reported results is very difficult currently as it references different models for feature extraction that require extensive attention. Collectively, the methodology of the paper is partially implementable, and results are reproducible but similar results may require extensive attention.

# References

Acheampong, F. A., Nunoo-Mensah, H., & Chen, W. (2021). Transformer models for text-based emotion detection: a review of BERT-based approaches. *Artificial Intelligence Review*, *54*(8), 5789–5829. https://doi.org/10.1007/S10462-021-09958-2/TABLES/18

Anantharaman, K., Rajalakshmi, S., Angel Deborah, S., Saritha, M., & Sakaya Milton, R. (2022). SSN\_MLRG1@LT-EDI-ACL2022: Multi-Class Classification using BERT models for Detecting Depression Signs from Social Media Text. *LTEDI 2022 - 2nd Workshop on Language Technology for Equality, Diversity and Inclusion, Proceedings of the Workshop*, 296–300. https://doi.org/10.18653/V1/2022.LTEDI-1.44

Burdisso, S. G., Errecalde, M., & Montes-y-Gómez, M. (2019). A text classification framework for simple and effective early depression detection over social media streams. *Expert Systems with Applications*, *133*, 182–197. https://doi.org/10.1016/J.ESWA.2019.05.023

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference*, *1*, 4171–4186. https://arxiv.org/abs/1810.04805v2

Lin, C., Hu, P., Su, H., Li, S., Mei, J., Zhou, J., & Leung, H. (2020). SenseMood: Depression detection on social media. *ICMR 2020 - Proceedings of the 2020 International Conference on Multimedia Retrieval*, *5*(20), 407–411. https://doi.org/10.1145/3372278.3391932

Shen, G., Jia, J., Nie, L., Feng, F., Zhang, C., Hu, T., Chua, T. S., & Zhu, W. (2017). Depression detection via harvesting social media: A multimodal dictionary learning solution. *IJCAI International Joint Conference on Artificial Intelligence*, *0*, 3838–3844. https://doi.org/10.24963/IJCAI.2017/536

Van Aken, B., Löser, A., Winter, B., & Gers, F. A. (2019). How does BERT answer questions? A layer-wise analysis of transformer representations. *International Conference on Information and Knowledge Management, Proceedings*, 1823–1832. https://doi.org/10.1145/3357384.3358028

Xin, J., Tang, R., Lee, J., Yu, Y., & Lin, J. (2020). DeeBERT: Dynamic Early Exiting for Accelerating BERT Inference. *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, 2246–2251. https://doi.org/10.18653/v1/2020.acl-main.204