Oleksandr Mazurets
Ph.D in Engineering Science, Associate Professor
Oleksandr Ovcharuk
Postgraduate student
Khmelnytskyi National University, Ukraine

EFFICIENCY RESEARCH OF METHOD FOR DETECTING MENTAL DISORDERS BY ANALYSIS OF USER CONTENT

Abstract. The paper proposes the efficiency research of method for detecting mental disorders based on analysis of user content, which allows performing multiclass classification of mental disorders using the neural network model-transformer DistilBERT and knowledge distillation. This approach ensures fast execution and increases Accuracy to 0.854, and Precision to 0.867.

Over the years, the scientific community has explored approaches to monitor and detect certain mental disorders and risk behaviors, such as depression, eating disorders, gambling, and suicidal ideation, among others, to activate prevention or mitigation strategies and, in severe cases, clinical treatment [1].

Analysis of user content, such as messages in social networks, blogs, forums and comments, allows to obtain unique insights into the emotional and mental state of victims. The application of machine learning and natural language processing methods to detect signs of PTSD in user texts is a modern and promising direction of scientific research, because it can not only identify individual cases of PTSD, but also help in the development of preventive measures and support programs for persons who are at risk [2].

Accordingly, linguistic indicators of mental health covering several aspects were investigated. The authors used BERT for multiclassification, but they achieved an average accuracy of about 64% and an average F1 score of about 0.645, which is not a high result for a mental health problem.

At the same time, the proposed method of detecting mental disorders based on user content analysis automatically classifies textual user content regarding the presence of mental disorders and their symptoms. As part of the work, the following types of mental disorders were considered: «Anger/ Intermittent Explosive Disorder», «Anxiety Disorder», «Depression», «Narcissistic Disorder», «Panic Disorder».

DistilBERT was chosen as the classifier, which has a knowledge distillation approach to learning and allows maintaining high classification results when reducing the model. The following metric values were obtained: Accuracy: 0.856, Precision: 0.867, Recall: 0.856, F1-score: 0.854. Figure 1 shows the confusion matrix on the validation sample. As can be seen from Figure 1, the highest indicators for all metrics are in the class «Narcissistic Disorder», and the worst in classes according to somemetrics «Anger/ Intermittent Explosive Disorder» and «Depression». This is due to the fact that these classes have similar signs of mental disorders. The low Precision of the Anger/Intermittent Explosive Disorder class (0.68) indicates that many examples that the model classifies as this class actually belong to other classes

at Recall values of 0.93. The model has problems separating this class from others due to the similarity of features.

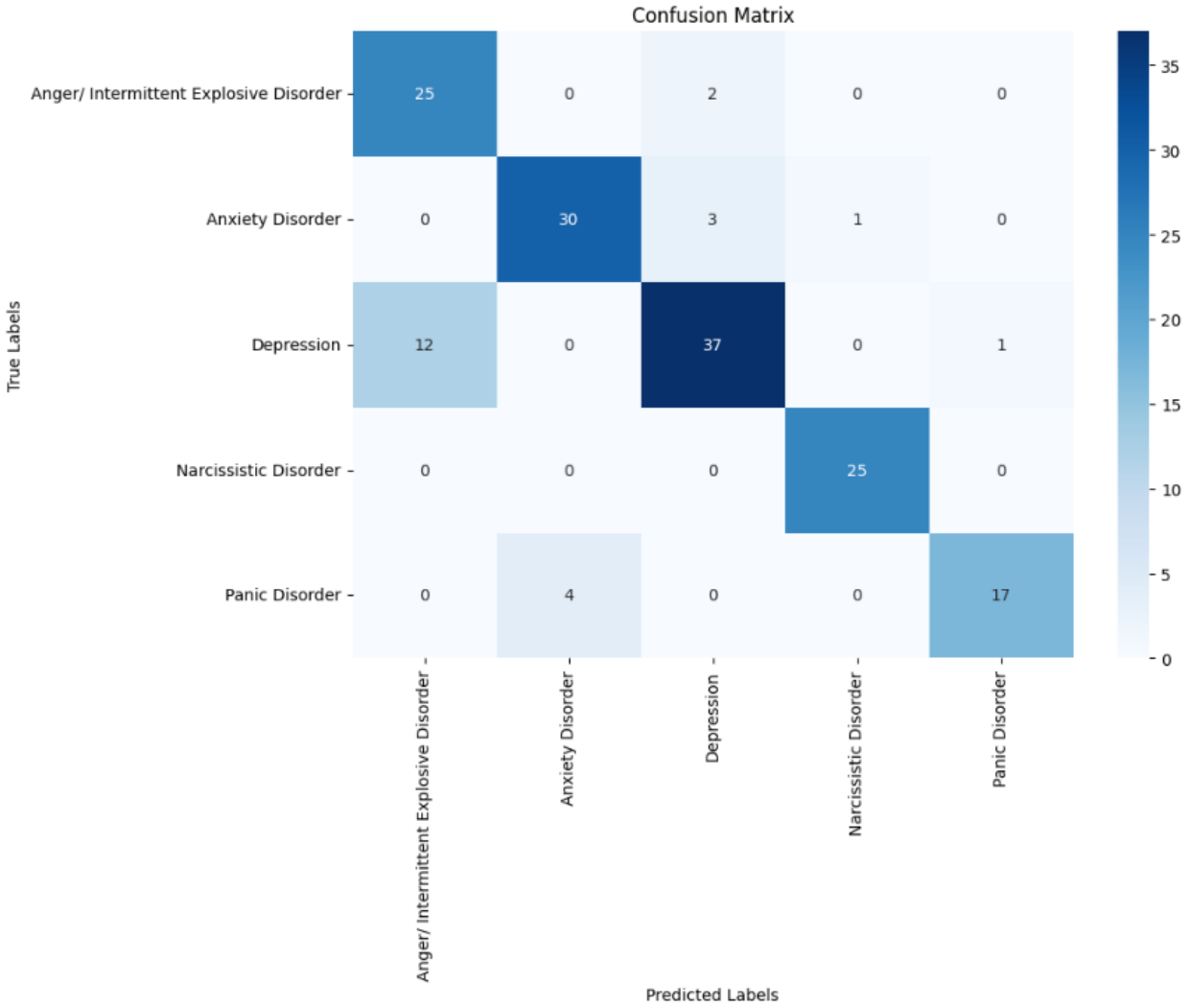

Figure 1 – Confusion matrix for multiclassification

So, the effectiveness of the method of detecting mental disorders based on the analysis of user content was investigated, which allows to increase the effectiveness of the multiclass classification of mental disorders to 0.856 according to the Accuracy metric (in similar studies 0.64), and to 0.854 according to the F1-score metric (in similar studies 0.645).

References:

1. O. V. Mazurets, O. M. Ovcharuk, O. O. Tyschenko, O. O. Zalutska, Effectiveness Research of Method for Values Forecasting of Epidemiological Danger Indicators by Means of Neural Network Modeling, in: Science and society: modern trends in a changing world, Proceedings of the 6th International scientific and practical conference, MDPC Publishing. Vienna, Austria, 2024, pp. 136-142.

2. Krak, O. Zalutska, M. Molchanova, O. Mazurets, R. Bahrii, O. Sobko, O. Barmak, Abusive Speech Detection Method for Ukrainian Language Used Recurrent Neural Network, in: CEUR Workshop Proceedings, 3688, 2024, pp. 16–28.