LITERATURE REVIEW

There is a significantly increasing interest in automatic personality prediction using social media among researchers in both Natural Language Processing and Social Sciences [1]. Most of the studies on personality prediction have focused on the Big Five or MBTI personality models, which are the two most used personality models. The Big Five personality model can be defined as a set of five broad trait dimensions: (1) extroversion, (2) agreeableness, (3) conscientiousness, (4) neuroticism, and (5) openness [2]. The Myers–Briggs Type Indicator (MBTI) classifies 16 personality types into four dimensions: (1) introversion/extraversion, (2) perceiving/intuition, (3) thinking/feeling, and (4) judging/perceiving [3]. Considering the controversy about the reliability and validity of these two models, studies suggest that the MBTI model has applications in more fields [4].

Golbeck et al. [5] stated that the personality type of users could be accurately predicted based on the Big Five personality type indicator and taking into account the information presented on their Twitter. In this study, two different regression algorithms (ZeroR and Gaussian Processes) were used, and similar performances were obtained from both algorithms on personality traits.

Komisin and Guinn [6] used Naive Bayes (NB) and Support Vector Machine (SVM) techniques to predict an individual's personality type based on word choices. The databases were created based on classroom writing samples from 40 graduate students with MBTI personality types. When the performance of these two methods is compared, it is seen that NB performs better than SVM in small datasets.

Wan et al. [7] estimated the Big Five personality types of users through their texts on Weibo. In this study, where they used Logistic Regression (LR) and NB algorithms, they used personality test results to validate the predicted results. In this study, it has been shown that the Big Five personality type can be predicted by user-generated information on social media.

Cui and Qi [8] tested Baseline, LR, NB and SVM algorithms for personality type prediction using the Myers–Briggs Personality Type Data Set from Kaggle. In the study in which the results of all these methods were compared, it was found that SVM performed better.

Hernandez and Knight [9] studied with the Myers–Briggs Personality Type Data Set from Kaggle. They used machine learning to classify people into MBTI personality types based on text samples. In this study, where they produced word clouds for data visualization, instead of training a multiclass classifier, they trained four different binary classifiers, each to be used in

one of the personality dimensions. It was observed that the model they created achieved significant success in accurately predicting the binary class for each MBTI personality dimension.

Another study by Ontoum and Chan [10] used the Myers–Briggs Personality Type Data Set from Kaggle. In the study, NB, SVM and Recurrent Neural Networks (RNN) algorithms were used to predict people's personalities from MBTI-based text. The overall accuracy performances, it was seen that a higher performance was obtained from RNN.

REFERENCES

1) Gjurković, M., & Šnajder, J. (2018, June). Reddit: A gold mine for personality prediction. In Proceedings of the second workshop on computational modeling of people's opinions, personality, and emotions in social media (pp. 87-97).
2) Soto, C. J. (2018). Big Five personality traits. The SAGE encyclopedia of lifespan human development, 240-241.
3) Myers, I. B., McCaulley, M. H., & Most, R. (1985). Manual, a guide to the development and use of the Myers-Briggs type indicator. consulting psychologists press.
4) Barbuto Jr, J. E. (1997). A critique of the Myers-Briggs Type Indicator and its operationalization of Carl Jung's psychological types. Psychological Reports, 80(2), 611-625.
5) Golbeck, J., Robles, C., Edmondson, M., & Turner, K. (2011). Predicting personality from twitter. In 2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing (pp. 149-156). IEEE.
6) Komisin, M. C., & Guinn, C. (2011). Identifying personality types using document classification methods (Doctoral dissertation, University of North Carolina Wilmington).
7) Wan, D., Zhang, C., Wu, M., & An, Z. (2014). Personality prediction based on all characters of user social media information. In Chinese national conference on social media processing (pp. 220-230). Springer, Berlin, Heidelberg.
8) Cui, B., & Qi, C. (2017). Survey analysis of machine learning methods for natural language processing for MBTI Personality Type Prediction.
9) Hernandez, R., & Knight, I. S. (2017). Predicting Myres-Briggs Type Indicator with text classification. Curran Associates Inc., New York.

10) Ontoum, S., & Chan, J. H. (2022). Personality Type Based on Myers-Briggs Type Indicator with Text Posting Style by using Traditional and Deep Learning. arXiv preprint arXiv:2201.08717.