International Journal of Innovations in Engineering and Science, www.ijies.net

Review on Text-Based Personality Prediction Using Social Media Data

Prof. Dr. Minal V. Gade[1], Akshata Kharat[2], Snehal Thoke[3], Krutika Khare[4]

[1] Professor, [2,3,4]UG, student
Sandip Institute of Technology and Research centre, Nashik, India,422012

akshatakh2001@gmail.com

Abstract *–One factor that influences how people interact with others is personality. One could consider a person's personality to be a crucial aspect of their behavior. People's interpersonal interactions shape their personalities. The numerous personality prediction studies that have been carried out by different researchers are covered in this paper. This system will be useful for businesses and other groups that hire candidates more for their personality than for their technical expertise.*

Keywords- Natural language processing; Myers Briggs type indicator; Machine learning; Support Vector Machine

I. INTRODUCTION

Human behavior and significant elements like happiness, emotions, and melancholy are reflected in personality. People's choices in life are influenced by their personalities. In humans, personality is a crucial psychological trait. Personality is a reflection of who a person is, what they like, and how they act in different situations. The idea is to develop a system which fetch data from the social media account and predict the personality of the user. The usage of social media has increased extremely. Social media is used extensively by many people in daily life. They share content and their thoughts by posting status updates, images, and videos. This allows for a prediction of the user's personality based on an evaluation of their online persona.

II. LITERATURE REVIEW

[1] Christian, Suhartono, D., Chowanda, A., and others with Christian, H. Text-based personality prediction using pre-trained language models and model averaging from various social media data sources. 68, J. Big Data (2021).

Using the personality prediction technique, features from the digital content are extracted and mapped to a personality model. Due to its simplicity and established efficacy, the big five personality traits, a well-known personality model, have frequently been embraced in the literature as the de facto benchmark for personality evaluation. For personality prediction systems, there are many algorithms that can be used to extract embedded contextualized words from textual input; some of them are based on ensembled models and deep learning.[1]

International Journal of Innovations in Engineering and Science, www.ijies.net

[2]25 Tweets to Know You: A New Model to Predict Personality with social media Pierre-HadrienArnoux, Anbang Xu, Neil Boyette, Jalal Mahmud, Rama Akkiraju, Vibha Sinha 2017."

Personality prediction is crucial for social applications that promote human-centered activities, but existing modelling techniques using users' written language require too much input information to be applied in a social media setting. With this study, we hope to significantly minimize the amount of data needed for personality modeling and create a model that can be used by the majority of Twitter users. Our model includes features from Word Embedding and Gaussian Processes Regression. Our model achieves equivalent or superior accuracy with 8 times less data than state-of-the-art methods, according to the examination of more than 1.3K Twitter users. [2]

[3]Srilakshmi Bharadwaj, Srinidhi Sridhar, Persona Traits Identification based on Myers-Briggs Type Indicator (MBTI) - A Text Classification"

In this project, a personality profile of a person is developed by the study of material that person has produced, such as an essay, tweet, or blog post. The primary focuses of the research are the type of data collected, text pre-processing techniques, and machine learning algorithms employed to estimate personality scores. It has been explored how different feature vector combinations and machine learning models compare, as well as how to put answers into practice. The techniques described in this study have been used to reach accuracy levels of up to 88%.[3]

[4]NosheenSabahat, Sagar Ashraf, and Zeeshan Mushtaq Gradient boosting and K-means Clustering for MBTI Personality Type Prediction"

In this research, we present a method for integrating two current machine learning methods, such as K-Means Clustering and Gradient Boosting, to evaluate user data posted on social media in order to predict user personality type. Additionally, this study aids in the analysis of the empirical link between a user's personality and the data they share on social media. The Myers-Briggs Type Indicator (MBTI), developed by Swiss psychotherapist Carl Jung, was used in this essay. The sixteen personality types that make up the MBTI can be used to understand a person's unique personality.Compared to the widely used naive Bayes classification technique and other algorithms, the method

of combining these two machine learning methods yielded more accurate results. The results of this study can assist bloggers and social media users in understanding the personality types they are portraying through their online sharing.[4]

[5]SandhyaKatiyar, PersonalityClassificationSystemusingDataMining
In this study, automated personality classification—a system that analyses a user's personality based on particular traits—is covered. This paper offers a way for assessing an applicant's behaviors. This system will be useful for businesses and other organizations that hire candidates more for their personalities than for their technical expertise. The Big Five Personality qualities serve as the foundation for the personality prediction results, which are classified using the Naive Bayes Algorithm and Support Vector Machine.

III. RESULT & DISCUSSION

The goal of this survey is to have an idea about Personality and prediction solutions offered on these platforms and how they meet the requirements as well as to identify personality prediction solutions. Since the survey does not include quantifiable information and in-depth research, we evaluated it tolerantly rather than strictly.

During this survey, we conclude that there are several models for prediction likethe big 5 model, MBTI model, Rorschach Inkblot Test, Thematic Apperception Test, etc. MBTI model has excellent accuracy but not as much exploration as the Big Five model, despite having a lot of research. The other two models Rorschach Inkblot Test and Thematic Apperception Test were also utilized in a number of studies, but they did not provide many accurate or reliable predictions.

From this study, we found that MBTI is an effective model for prediction.

And for classification, some algorithm which are used in previous projects are SVM algorithm, naïve bayes algorithm etc.

The basic classification algorithm results from this survey are used to forecast personalities
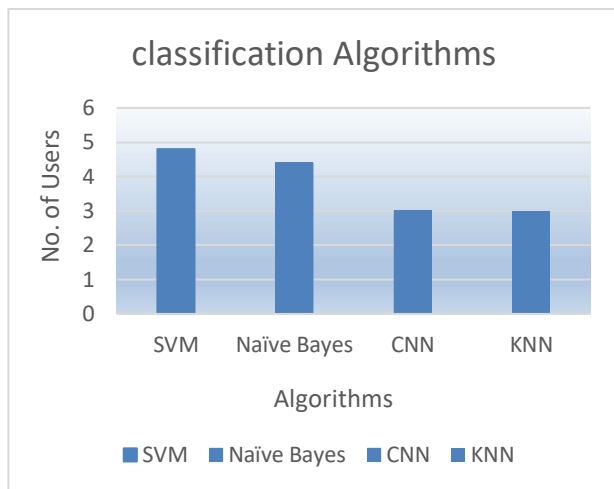
International Journal of Innovations in Engineering and Science, www.ijies.net

Fig 1. Classification Table

This figure shows that the SVM algorithm was most utilized and investigated. Additionally, the SVM algorithm is extensively used for classification.

According to Accuracy, SVM algorithm accuracy is higher than that of other classification algorithms.

Consequently, using SVM will result in greater accuracy.

VI. CONCLUSION

Different personality models claim that a user's actions on social media platforms might provide insight into their personality. The use of questionnaires in the past could be expensive and time-consuming. This essay seeks to give a summary of the research on social media text personality prediction and a prediction of future trends.

Personality prediction is a developing study topic that predicts automatically a user's personality qualities from publicly available information on social media platform. Social media platforms are expanding quickly, which results in an increase in user engagement. Users publish their thoughts and ideas on various subjects, and this information can be used to forecast a user's personality. Assisting with job or product recommendations, political assessments, mental health diagnoses, etc., personality characteristic prediction is useful in many different ways. This study presents an insight into the previous personality prediction based on text from twitter.

From this paper, we conclude that MBTI, Big 5 model and for classification SVM and Naïve bayes are reliable models.

REFERENCES

[1] *H. Christian, D. Suhartono, A. Chowanda, et al. Text-based personality forecast using a pre-trained language model and model averaging from various social media data sources. J Huge Data.*

[2] *media Rama Akkiraju, Anbang Xu, Neil Boyette, Jalal Mahmud, Pierre-HadrienArnoux, A New Model to Predict Personality Using Social Media: 25 Tweets to Get to Know You*

[3] *Srilakshmi Bharadwaj, Srinidhi Sridhar, Persona Traits Identification based on Myers-Briggs Type Indicator(MBTI) - A Text Classification*

[4] *NosheenSabahat, Sagar Ashraf, and Zeeshan Mushtaq Gradient boosting and K-means Clustering for MBTI Personality Type Prediction*

[5] *"Psychometric properties of the hexaco-100," Assessment, K. Lee and M. C. Ashton, 5.*

[6] *The big five personality traits and job performance, by S. Rothmann and E. P. Coetzer. Industrial Psychology SA Journal.*

[7] *The big five personality traits and job performance: A meta-analysis, Personnel Psychology, M. R. Barrick and M. K. Mount.*

[8] *"Personality classification based on Twitter text using Naive Bayes, KNN, and SVM," B. Y. Pratama and R. Sarno.*

[9] *Bernard Theodore L. Pollo P. Edward Tighe "Personality Trait Classification of Essays with the Application of Feature Reduction," by Jennifer C. Ureta"*

[10] *Based on user behaviour on the Facebook social media platform, Michael Mesfin Tadesse, Hongfei Lin, and Xu made predictions about people's personalities in October 2018.*