International Journal of Innovative Research in Engineering
Volume 4, Issue 3(May-June 2023), PP: 421-423.
www.theijire.com ***ISSN No: 2582-8746***

Personality Prediction and Resume Screening using Machine Learning

Dr. A.S. Shanthi[1], P. Leander Felix[2], K.A. Raghul[3], M. Utchimakali[4], S. Yogaraj[5]
[1,2,3,4,5]*Department of Computer Science and Engineering, Tamilnadu College of Engineering, Tamilnadu, India*

How to cite this paper:
Dr. A.S. Shanthi[1], P. Leander Felix[2], K.A. Raghul[3], M. Utchimakali[4], S. Yogaraj[5], "Personality Prediction and Resume Screening using Machine Learning", IJIRE-V4I03-421-423.

Abstract: *The personality of a human plays a major role in his personal and professional life. Many organizations have also started shortlisting the candidates based on their personality as this increase efficiency of work because the person is working in what he is good at than what he is forced to do The project is based on identifying the personality of an individual using machine learning algorithms and Model classify the personality The prediction of the personality of an individual is a critical problem in both areas whether it is considered in the context of organizations or in the case of our daily lives. Prediction of personality depends on many factors and these factors may vary from one individual to another. Personality prediction is identifying the personalities of individuals through their actions in different situations and observing their behaviour in various circumstances Five characteristics of different individuals commonly known as big five characteristics namely, openness, neuroticism, conscientiousness, agreeableness and extraversion are stored in a dataset along with gender and age of individual and used for training. Before training the model, data is preprocessed like handling missing values, data discretization, standardization etc. This pre-processed data is then used to train the model. User rates himself for different behavioural characteristics and based upon the information provided by the user his/her personality is predicted using trained ML model. Personality traits show the different characteristics of different people based on their thoughts The accuracy of personality prediction achieved by using Logistic regression Classifier is 75.25%. We refer to them as personality types. In this Project people with similar personalities are grouped together based on identifying personality models.*

Key Word: *Classify, dataset, accuracy, Logistic Regression Classifier*

I.INTRODUCTION

The word personality comes from the Latin word persona, which means mask. Worn by actors to perform. But individuality may now be more than just a mask. We may determine whether a person is suitable for a particular job profile. Humans have the full potential to guide, influence, and communicate effectively with others. The first step in recruitment is the application, which consists of personal data. Experience and most important resume. Companies typically receive thousands of. There is an application for each job opening and we have a team of professional judges who select qualified individuals candidates. It's very difficult to manually check everyone's resume. Many candidates have already been eliminated in the first round for the following reasons. Qualifications to apply, inadequate resume, lack of qualifications. Recruiting the right candidates is a huge challenge. It's a challenge because no candidate is perfect and some candidates may not be competent enough, or even competent enough. doesn't have the right personality Therefore, we propose the following method. Personality prediction simplifies and speeds candidate lists. Resumes can be thought-provoking Reflects an individual's professional qualifications, but not their personality Man. Personality is one of the most important indicators of a person's ability. Analyzing and understanding personality is important to working in a particular role. Our the purpose of this project is to make machines more human and analyze them. Evaluate candidates just like a real human evaluator. This paper attempts to explore this. Implement different machine learning algorithms and analyse which ones are. It offers the highest accuracy with the widest range of data provided. we will also try.Visualize data and establish relationships between different elements

II.RELATED WORK

Kunte et al., described while Machine learning has recently received a great deal of attention among researchers. It can be widely applied to tasks such as prediction and classification. Current research focuses on effectively recognizing the personality of social network users. Personality is the combination of your own thoughts and actions. Knowledge of an individual's personality has many real-world applications, such as in various recommender systems and in human resources departments. A person's personality can be better understood through interaction with people. Personality prediction using social media is a new approach that eliminates direct human interaction and enables accurate prediction. Researchers have recently used various machine learning techniques for prediction tasks, but the use of ensembles has not yet been explored. Current research focuses on advanced classifiers such as XGBoost and Ensemble for prediction. Experiments using the real-time Twitter dataset showed a high accuracy of 82.59% for the ensemble.

Later Ramsha Ali et al, This article describes the use of TF-IDF (term frequency-inverse document frequency) in investigating document-keyword relationships in a corpus. This research focuses on how this algorithm can be applied to different documents. First, we detail the working principles and steps to follow to implement TF-IDF. The results are then displayed, comparing the strengths and weaknesses of the TF-IDF algorithms and validating the insights gained from running the algorithms. This document also explained how to address such weaknesses. Finally, we summarise the research contents and discuss future research directions.

And , Chris et al, Social media websites are among the most popular destinations for Internet users today and offer social scientists an excellent way to understand online behaviour. There is a growing body of research on social media, a few of which focus on personality prediction. While previous research has usually focused on the "Big 5" personality traits, one relatively under-researched area is the antisocial traits of narcissism, Machiavellianism, and psychopathy, which are commonly called the "Dark Triad". This study investigated the extent to which antisocial personality traits can be determined based on Twitter usage. For this purpose, the Dark Triad and Big Five personality traits of 2,927 of his Twitter users were compared with profile attributes and language usage. Analysis revealed a statistically significant relationship between these variables. Using crowdsourced machine learning algorithms, we show that, although machine learning provides useful predictive rates, it is imperfect in predicting an individual's Dirk-His triad traits from Twitter activity. Predictive models may be poor at predicting individual personality traits, but when applying models to large groups of people, for example, whether there is an increase in antisocial traits within the population, or Gaining the ability to detect whether weight is being lost is still of practical importance. Our results raise important questions regarding the unregulated use of social media analytics for screening purposes. It is important to better understand the practical and ethical implications of making inferences about personal information embedded on social media sites.

And Hassanein et al Online social networks are becoming a very rich source of user-generated content. This content launches different types of applications based on personalization. Such as recommendation systems and online marketing. Personality recognition by mining publicly available social data represents an important and relevant problem that Web-based systems can support. Several approaches have been introduced to infer a user's personality using publicly available social data. This article presents an approach to personality trait inference based on semantic analysis of texts. Various representations of user text combining multiple semantic measures have been proposed to predict user personality based on Facebook status updates. The proposed approach was tested and validated against data published by the my Personality project of the Computer-Based Personality Recognition Workshop. The results show that information content-based measures achieve the best average predictions of personality traits with 64% accuracy.

III.METHODOLOGY

Machine learning is a branch of artificial intelligence (AI) and computer science that focuses on using data and algorithms to mimic the way humans learn with incremental improvements in accuracy. Machine learning is an important part of the growing data science field. Using statistical techniques, algorithms are trained to classify, predict, and gain key insights in data mining projects. These insights drive decisions within applications and organizations, ideally influencing key growth indicators. As big data continues to expand and grow, we believe the market demand for data scientists will increase. They should help identify the most relevant business questions and the data to answer them.

Supervised learning (also known as supervised machine learning) is defined by using labelled datasets to train algorithms to classify data and accurately predict outcomes. As input data is fed into the model, the model adjusts the weights until it fits well. This is done as part of the cross-validation process to ensure that the model is not over- or under-fitted. Supervised learning helps organisations solve a variety of real-world problems at scale, such as sorting spam into a folder separate from the inbox. Techniques used in supervised learning include neural networks, naive bayes, linear regression, logistic regression, random forests, and support vector machines (SVMs).

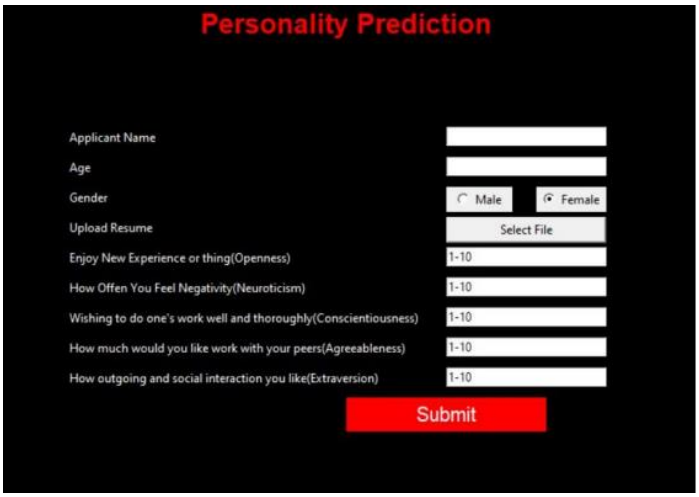
Unsupervised learning, also known as unsupervised machine learning, uses machine learning algorithms to analyse and group unlabeled datasets. These algorithms discover hidden patterns and groupings of data without the need for human intervention. This method can discover similarities and differences in information, making it ideal for exploratory data analysis, cross-selling strategies, customer segmentation, image and pattern recognition. It is also used to reduce the number of features in the model through the process of dimensionality reduction. Two common approaches to this are principal component analysis (PCA) and singular value decomposition (SVD). Other algorithms used in unsupervised learning include neural networks, k-means clustering, and probabilistic clustering techniques.

Neural Networks simulate the workings of the human brain using a large number of interconnected processing nodes. Neural networks are good at recognizing patterns and play an important role in applications such as natural language translation, image recognition, speech recognition, and image creation. This algorithm is used to predict numerical values based on linear relationships between different values. For example, this technology can be used to predict property prices based on historical data for the area. This supervised learning algorithm makes predictions for categorical response variables such as: B. Yes/No answers to questions. It can be used for purposes such as spam classification and quality control on production lines. Using unsupervised learning, clustering algorithms can identify patterns in the data and group the data. Computers can help data scientists by identifying differences between data items that humans miss. The Decision Tree can be used both for numerical prediction (regression) and for classifying data into categories. A decision tree uses a branching sequence of linked decisions that can be represented by a tree diagram. One advantage of decision trees is that they are easy to

validate and test, unlike neural network black boxes. In a random forest, a machine learning algorithm combines the results of a set of decision trees to predict a value or category

IV.RESULTS AND DISCUSSION

So when a user tries to classify, the system automatically classifies their personality. Investigate using the dataset provided by the backend. Personality analysis and predictions are much more. These days, more personality traits may be added in the future. Using datasets and algorithms, improvements can be made to improve accuracy. If the user has good storytelling and persuasion skills, it will be useful for the Career Counseling module. This project is about personality analysis and prediction. We withdraw the root tkinter window and create a new top level window and configure its size and attributes. We label the heading of the window followed by various labels and their entries. For selecting a resume file, the user needs to press the choose file button which then calls the Open file method that takes an argument of the button. In the predict person method, various entries are taken for predicting the personality. Submit button passes all the values to prediction result. First, collect data from a set of messages and a data set containing those answer. Then remove punctuation, extra symbols, etc. Extract what you don't need from this dataset. Then use the bag of words technique to identify the Input message class. Pass the processed input query/message to the neural network. A sequential model used to train the chat bot. then predict and return the class



V.CONCLUSION

The goal of this project is to investigate different methods for extracting and analyzing spoken language. Non-verbal data for predicting a person's personality. We also discussed how Predict personality using machine learning models. Studies show they were using different personality prediction methods, some less accurate, others less accurate. Gives high accuracy. Offered high accuracy but had limited use person/record, or used orally or non-verbally. Emotional trait performance Very low. In this case, multimodal measures are effective to make the task automatic and intelligent. Multimodal agents are good at identifying personality traits based on both verbal and nonverbal information features. In conclusion, although several techniques were used, Much research has been done on personality prediction, but nothing on systems. Totally perfect, but still room for improvement.

References

[1]. *M.F. Arani, A.A. Jahromi, D. Kundur and M. Kassouf, "Modelling and simulation of the aurora attack on micro grid point of common coupling," in 2019 7th workshop on Modelling and Simulation of Cyber-Physical Energy Systems (MSCPES). IEEE, 2019, pp. 1-6.*

[2]. *R. Tan, V. Badrinath Krishna, D.K. Yau and Z. Kalbarczyk, "Impact of integrity attacks on real-time pricing in smart grids," 2013, pp. 439-450.*

[3]. *Y. Dafalla, B. Liu, A.A Hahn, H.Wu, R. Ahmadi and A.G. Bardas, "Prosumer nanogrids: A cybersecurity assessment," IEEE Access, vol. 8, pp. 131 150 – 131 164, 2020, event IEEE Access.*

[4]. *B. Liu and H. Wu, "Optimal d-facts placement in moving target defense against false data injection attacks," IEEE Transactions on Smart Grid, pp. 1-1, 2020, event: IEEE Transactions on smart grid.*

[5]. *Z. Zhang, R. Deng, D.K Yau, P. Cheng and J. Chen, "Analysis of moving target defense against false data injection attacks on power grid," IEEE Transactions on Information Forensics and Security, vol. 15, pp.2320 – 2335, 2019.*

[6]. *S. Lakshminarayana and D.K. Yau, "Cost-benefit analysis of moving target defense in power grids," IEEE Transactions on Power Systems, 2020.*

[7]. *J. Tian, R. Tan, X. Guan and T. Liu, "Enhanced hidden moving target defense in smart grids," IEEE Transactions on Smart Grid, vol. 10, no. 2, pp. 2208 – 2223, 3 2019 44.*

[8]. *X. Niu, J. Li , J. Sun and K. Tomsovic "Dynamic detection of false data injection attack in smart grid using deep learning," in 2019 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT), 2019, pp. 1-6.*

[9]. *Y. Li, Y. Wang and S. Hu, "Online generative adversary network based measurement recovery in false data injection attacks: A cyber-physical approach," IEEE Transactions on Industrial Informatics, vol. 16, no. 3, pp. 2031-2043, 2020.*

[10]. *G. Ding, Q. Wu, Y.D. Yao, J. Wang and Y.Chen, "Kernel-based learning for statistical signal processing in cognitive radio networks: Theoretical foundations, example applications, and future directions," IEEE Signal Processing Magazine, vol. 30, no. 4, pp. 126-136, 2019.*