Comparative Analysis of Personality Prediction Systems of CV

Ms. Poonam Katyal, Mohit Sharma, Naman Sharma

Department of Computer Science Engineering, Faculty of Engineering
ManavRachna International Institute of Research and Studies,
Faridabad, Haryana, India

ABSTRACT-Organizations recruit expert candidates for their expansion, but the main concern for them is selecting efficient candidates. Every year, they receive a lot of applications and it will be strenuous for them to go through all applications and recruit the best candidates. Traditionally, they go through the candidates' CVs and recruit them. In this paper, the proposed system helps to recruit the right candidates using pyresparser to parse information from a CV. The present research paper is also based on identifying the personality of the individual using machine learning algorithms and Big Five Model (OCEAN). Using this system, organizations can find expert candidates and make the recruitment departments' work easier.

Keywords-Personality prediction, pyresparser, CV, Big Five Model.

I. INTRODUCTION

Human personality plays an indispensable role in an individual's life as well as in the development of an organization. One of the most crucial elements in determining whether a person is a match for the requirements is their personality. We can know the capabilities of a person by whether they can impress and communicate with others effectively. This helps in the development of an organization. A CV is the first stage in the recruitment process. Companies receive thousands of applications for each job. Manually going through CV is quite challenging. So is using traditional techniques like tests, interviews, surveys, etc. that are time consuming and slow in processing. In order to reduce difficulty in the hiring process, we propose a new way, where the process of selecting and shortlisting candidates gets easier using personality prediction. In this research paper, prediction of personality is done using highly valid and reliable Big Five Test (OCEAN) i.e. Openness (O), Conscientiousness (C), Extraversion (E), Agreeability (A), Neuroticism (N). In variation to this, various machine learning methods employed to predict the personality including Logistic Regression, Naïve Bayes, Random Forest, Support Vector Machine (SVM) and KNN.

To extract information from the CV like name, age, gender, etc. we used a simple resume parser, Pyresparser.

Thereafter, extracting information from the CV and score from the test, we generate the score of a person.

Objective: The objective of this research is to make the machine more human by analyzing the applicant in the same way as a human reviewer would.

II. LITERATURE REVIEW

S. No.	Author	Algorithm	Journal/Year	Outcome	Objective
1	Thahira M, Mubeena AK	1. SUM 2. NB 3. RF 4. SLR 5. DT (Decision Tree) 6. LSTR	ICCIDT -2021 Vol-09	Accuracy found using Deep Learning techniques is better than that found using machine learning techniques.	To study personality prediction from social media using machine learning and deep learning methods
2	Warkh Maharani and Veronikha Effendy	1. NB 2. KNN 3. SVM	IJECE Vol-12 2022	SVM outperformed NB and KNN for Indonesians.	To study the Big Five personality prediction based on Indonesian tweets using machine learning
3	PrajwalKaushal et al.	1. SVM 2. KNN 3. NB	IJECE Vol-13 (6) 2021 pp. 48-60	The model succeeds in predicting the personality traits of Twitter users.	To study the Myers-Briggs personality prediction and sentiment analysis for Twitter using machine learning and BERT
4	AvantikaDhar	Random Forest (RF)	Skyfi Labs 2/7/2021	Predicted personality is almost 8-15% close to the original personality of a person.	To study personality prediction using machine learning
5	M. Kalghatgi and Dr. NS Sidral	1. SVM 2. RF 3. NB 4. SLR	IJIRAE Vol-12, Aug. 2015	Twitter alone is not enough to predict personality.	To present Neural Network Approach based on Big Five test to predict personality depending on tweets
6	Allan Robey et al.	1. NB 2. KNN 3. RF	IRJET, Vol-6, Feb. 2019	Proposed system was more effective to shortlist CVs making sure that the ranking is fair and legal.	To propose a system to reduce the load of HRD using two sides of the organization to shortlist CVs from a large pool
7	JunejaZubeda et al.	1. SVM 2. RF	Unpublished	-	To rank CVs using NLP and machine learning
8	Md. Tanzim Reza et al.	1. NB 2. KNN 3. RF	Unpublished	CVs had been classified using multivariate logistic regression.	To analyze CVs using NLP and machine learning
9	Atharva Kulkami et al	1. SVM 2. RF 3. NB 4. KNN 5. LR	IJERT Vol-10 September 2021	RF algorithm achieved better accuracy than other algorithms.	To examine different machine learning approaches for predicting personality through CV analysis using NLP
10	SudhirBagade et al.	TF-IDF	Published on 31 May 2019 in IJCSE	Resulting scores seen in different areas were used to evaluate personality traits that helped in analyzing CVs properly.	To evaluate personality and perform CV analysis using machine learning algorithm

III. EXISTING SYSTEM

We label encoded the Personality column of our dataset before training our model. There were 708 rows and 6 rows in our final dataset. We used 70 percent of our data for training and 30 percent for testing the findings using the sklearn library. We employed a variety of machine learning methods to predict the personality of a potential candidate, including Logistic Regression, Naive Bayes, Random Forest, Support Vector Machine (SVM), and KNN.

IV. PROPOSED SYSTEM

In this paper, we suggest employing machine learning algorithms for personality assessment and CV analysis. This gives the company access to a skilled workforce which will make it easier to choose the best applicant for a certain job profile. Apart from this, machine learning algorithms i.e. Logistic Regression, Naïve Bayes, Random Forest, Support Vector Machine (SVM) and kNN are used to build the model that will parse the information from the CV and display it as the output.

V. RESULT

We were able to increase the accuracy to around 0.71 after feeding the data to the models. As shown in Table, the Random Forest algorithm has the highest accuracy, followed by Bayes, kNN, SVM, and Logistic Regression. Random Forest, as expected, has the smallest mean squared error, which is calculated as the average of the square of the difference between real and estimated values.

Table: - Accuracy and MSE values

Model	Accuracy	Mean square error (MSE)
Logistic regression	0.62	0.37
Naïve Bayes	0.62	0.37
SVM	0.63	0.36
Random Forest	0.71	0.36
kNN	0.64	0.29

VI. CONCLUSION AND FUTURE SCOPE

For Personality prediction using CV Analysis, we employed several Machine Learning Algorithms such as Logistic Regression, Naive Bayes, Random Forest, SVM, and KNN. We were able to predict the personalities of numerous candidates using pyresparser, spaCy, and PhraseMatcher. The results show that Random Forest has a maximum accuracy of 0.71, however this is substantially lower than expected due to a lack of relevant data. Various firms can utilize the proposed technique to streamline the recruitment process by taking into account the personality of possible recruits. Future work could be done to improve the efficiency and performance of the suggested system in order to improve the accuracy of personality prediction using CV analysis. This technique can be applied to a variety of industries that require expert candidates. The human resource department's workload will be reduced as a result of this technology. This method will assist the human resource department in selecting the best candidate for a specific job description, resulting in an expert workforce for the company. Admin or the concerned person can quickly shortlist candidates based on their online test results and choose the best candidate for the job.

REFERENCES

[1]. Md. Tanzim Reza and Md. Sakib Zaman, analyzing CV/resume using natural language processing and machine learning (unpublished)

[2]. Atharva Kulkarni, TanujShankarwar et al., published a research paper on "Personality Prediction via CV Analysis Using Machine Learning" in September 2021, IJERT, Vol-10

[3]. ManasiOmbhase, PrajaktaCogate et al., "Automated Personality Classification Using Data Mining Techniques" in 2016

[4]. Automated CV Processing along with Psychometric Analysis in Job Recruitment Process", Firoz Ahmed, MehrinAnannya, Tanvir Rahman et al., Institute of Information Technology, May 2015

[5]. Warth Maharani, Veronikha Effendy, "Big Five Personality Prediction Based in Indonesian tweets Using Machine Learning Methods" published in International Journal of Electrical and Computer Engineering (IJECE) in January 2022, Vol-12.

[6]. PrajwalKaushal, Nitin Bharadwaj et al., "Myers-Briggs Personality Prediction and Sentiment Analysis of Twitter Using

Machine Learning classifiers and BERT", published in International Journal of Information Technology and Computer Science in 2021, Vol 13 (6) pp. 48-60

[7].   Tahira, M., Mubeena AK et al., "Comparative Study of Personality Prediction from Social Media Using Machine Learning and Deep Learning" published in IJERT, Vol-09 in May 2021.