



International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Personality Prediction through Curriculum Vitae Analysis using ML

Ms. Kavana.¹, Ms. Varshitha M K.², Manish K Sharma³ Mr. Vivekananda⁴, Mrs. Aruna M G⁵ Dr. Malatesh S H⁶

^{1,2,3,4} Student, ⁵ Professor and Head of Department, Department of AI & ML, ⁶ Professor and Head of Department, Department of CSE, M. S. Engineering College, Bangalore, India

ABSTRACT—

The development of an organisation and a person's personal growth are both greatly influenced by their personality. Examining the candidate's curriculum vitae or doing a standard review are two common techniques to determine a person's personality. The traditional approach of hiring individuals is manually shortlisting resumes in accordance with the requirements of the business. In this study, a system that automatically separates applicants based on eligibility requirements and personality assessments during the hiring process is proposed. To fulfil this criteria, an online application is created that allows candidates to register their personal information and has personality tests that are multiple-choice questions (MCQs). The machine then compares the uploaded Curriculum Vitae trained datasets to analyse professional eligibility. This system uses the "LogisticRegression" machine learning method, which aids in choosing just decisions when hiring qualified candidates. As a result, both candidates and the admin will receive the results of the personality tests.

Keywords: Personality Prediction, CV Analysis, Machine Learning, Logistic Regression, KNN, Big Five Personality Model, Psychometric Analysis, Hiring, and Selection.

I. INTRODUCTION

The development of the Job Characteristics Model (JCM) has significantly benefited Human Resources Management in designing job roles and understanding employees' capabilities. Effective communication and influencing skills are crucial for organizational development. However, the traditional methods of evaluating candidates, such as technical tests, interviews, and group discussions, become challenging when faced with many applications. Analyzing personality traits manually based on behavior and nature is a laborious task, prone to human bias and inaccurate results. Therefore, there is a need for automated analysis to ensure objective and accurate personality assessment in recruitment processes.

II. EXISTING SYSTEM

Prior to model training, we performed label encoding on the Personality column of our dataset. The final dataset consisted of 708 rows and 6 columns. Employing the sklearn library, we allocated 70% of the data for training and reserved 30% for testing. To predict the personality of potential candidates, we leveraged diverse machine learning algorithms including Logistic Regression and K-Nearest Neighbors (KNN).

Sl. No.	Authors	Methodology Used	Disadvantages
1.	Faliagka et al.	Machine Learning Algorithms	For senior position it required expertise and certain qualifications were screened.
2.	Kalghatgi et al.	Neural Network	Challenging misinformation, conducting automated analysis of tweets, and relying solely on Twitter data are insufficient for accurate personality prediction.
3.	Allan Robey et al	Using NPL and ML	Instead of solely scanning the CVs, the authors suggest the inclusion of an aptitude test and a personality test for the purpose of personality prediction.

Disadvantages of existing system:

Under this model, there are few drawbacks of the application.

- The existing system revealed noise.
- The existing system has the size of the dataset was very small.
- The data may contain Missing values which will cause uncertainty.
- The most common problems among them were the input size, attributes, and validation

III. PROPOSED SYSTEM

In our proposed model, we have utilized machine learning algorithms to improve the precision and efficiency of the personality prediction system through CV analysis. By integrating machine learning techniques, we introduce supplementary attributes that can drive future progress in this domain. Machine learning operates on the principle of centralized data processing, enabling effective assessment and forecasting of personality traits such as Academics, Curriculum, Co-Curriculum, Skills, and Experience. Our model prioritizes harnessing the capabilities of machine learning to extract valuable insights from CV data and deliver dependable predictions.

IV. SYSTEM ARCHITECTURE

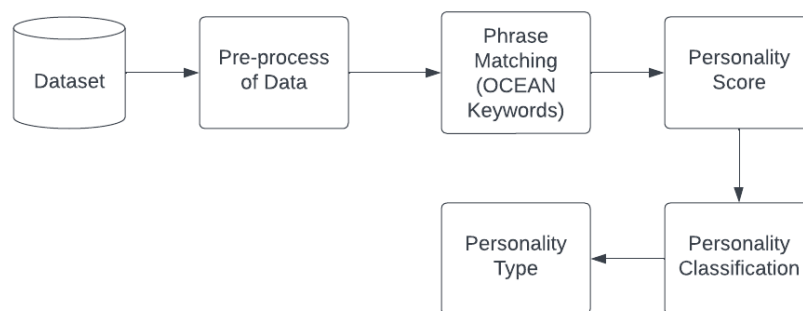


Fig 4.1: System Architecture

Figure 4.1 illustrates the System Architecture of the Personality Prediction Structure employed in the system. The resume of the job seeker is uploaded, and they are required to complete a personality test. This test encompasses questions related to each characteristic described in the previous section [5]. Upon completion, the test results are evaluated, stored in the database, and utilized for personality prediction [13]. The user's personality is determined based on training, and further analysis is performed using multinomial logistic regression on individual personalities. Prior to testing the data, appropriate preprocessing techniques are applied.

V. METHODOLOGY

Before training our model, we performed label encoding on the Personality column of our dataset. The resulting dataset contained 700 rows and 6 columns. For model training, 70% of the data was used, while the remaining 30% was allocated for testing. The sklearn library was utilized for these tasks. In order to predict the personality of potential candidates, we employed diverse machine learning algorithms including Logistic Regression and KNN.

- Logistic Regression is a machine learning algorithm commonly used for classification tasks, particularly in predicting binary outcomes such as True or False. It finds applications in various domains, including plagiarism detection, fraud detection, and sentiment analysis.

- KNN (k-nearest neighbors) is a well-known supervised machine learning algorithm suitable for regression and classification tasks. It operates on the assumption that similar data points are usually located close to each other, reflecting the idea that "Birds of a feather flock together." This algorithm is widely employed in diverse applications to identify similarities and make predictions based on neighboring data points.

PERSONALITY PREDICTION MODULE

Pseudo code:

Procedure Test ()

// Input: Trained model and user values

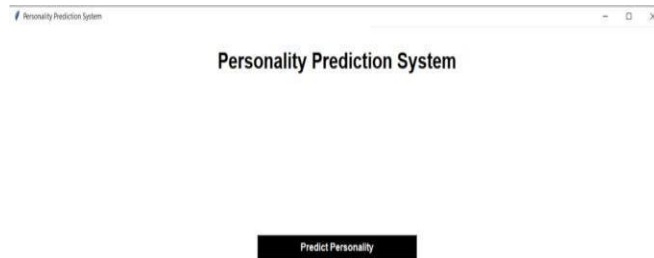
// Output: Personality prediction Begin

Step 1: Read user values Step 2: Load Trained model

Step 3: Prediction Personality using predict () Return predicted Personality

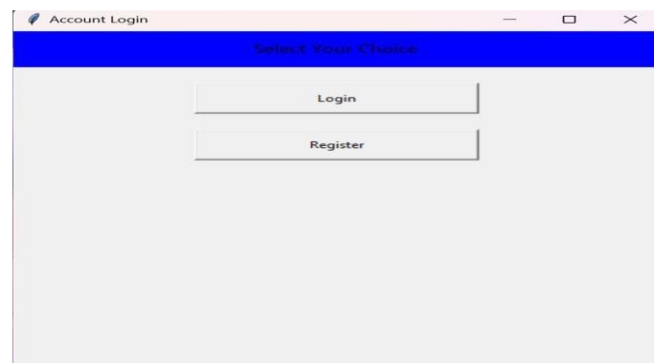
VII.SCREENSHOTS

Screenshot 1 Shows the landing page



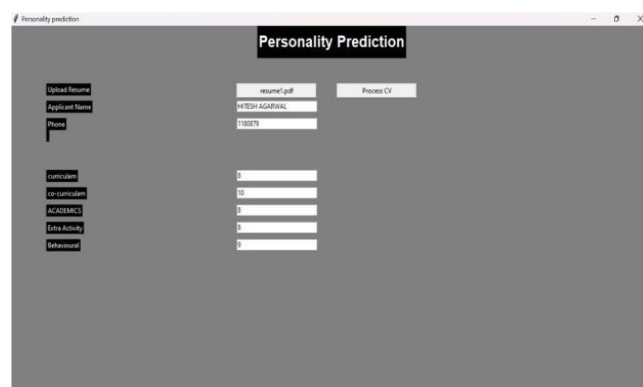
In the figure above Personality Prediction system homepage is shown.

Screenshot 2 Shows the Account login page



In the above figure, registration and login can be done.

Screenshot 3 Shows the Upload Resume and Details pa



VI. IMPLEMENTATION

KNN ALGORITHM

Pseudo code

Step - 1: Select the number K of the neighbors.

Step - 2: Calculate the Euclidean distance of K number of neighbors.

Step - 3: Take the K nearest neighbors as per the calculated Euclidean distance.

Step - 4: Among these k neighbors, count the number of the data points in each category.

Step - 5: Assign the new data points to that category for which the number of the neighbor is maximum.

Step - 6: Our model is ready

In the above figure, here the CV is uploaded from which we extract required information through NLP and predict the personality based on the designed model.

Screenshot 4 shows the Command Line Outputs

```
C:\Users\hites\Desktop\Personality Prediction through CV\python main.py
##### Candidate Entered Data #####
Candidate Name: 'hitesh Agwal', 'CV Location': 'C:\Users\hites\Desktop\Personality Prediction through CV\Test Cases\resume.pdf' [0, 'M', 'T', 'T', 'T', 'T', 'T']
##### Predicted Personality #####
['responsible']
##### Resume Parsed Data #####
email : hitesh.1108179@ou.in
mobile number : 991530719
skills : ['C', 'Php', 'Mobile', 'C++', 'Python', 'Cis', 'Training', 'Java', 'Electrical', 'Technical', 'Research', 'Javascript', 'Html', 'Email', 'Programming', 'Automation', 'System', 'Github', 'Rocket']
degree : ['B. Tech (ISE)']
designation : ['NONE AUTOMATION USING GOOGLE ASSISTANT', 'Medical Assistant', 'Technical Head']
company names : ['appliance using Google Assistant on mobile']
no. of pages : 3
total experience : 0
```

Figure above show the command line outputs.

REFERENCES

1. Faliagka, E., Kozanidis, L., Stamou, S., Tsakalidis, A. and Tzimas, G. (2011a), "A personality mining system for automated applicant ranking in online recruitment systems", Proceedings of International Conference on Web Engineering (ICWE 2011), pp. 379-82.
2. In their publication titled "Neural Network-based Personality Prediction using the Big-Five Model," M. Kalghatgi, M Ramannavar, and Dr. N. S. Sidnal (2015) proposed an approach to predict personality traits using artificial neural networks. The research was conducted in the International Journal of Innovative Research in Advanced Engineering.
3. In their research article titled "Personality Prediction System through CV Analysis" published in the International Research Journal of Engineering and Technology (IRJET) in February 2019, Allan Robey, K. Shukla, K. Agarwal, K. Joshi, and Professor S. Joshi proposed a system to predict personality traits using CV analysis.
4. In their unpublished work titled "Analyzing CV/Resume using natural language processing and machine learning," Md Tanzim Reza and Md. Sakib Zaman conducted a study on analyzing CVs using NLP and machine learning techniques.
5. Pennebaker, J.W., King, L.: Linguistic Styles: Language Use as an Individual Difference. In: Journal of Personality and Social Psychology. vol. 77, pp. 1296 – 1312. (1999).
6. Barrick, Murray R, The Personality Dimensions, and job performance. Gaytri Vaidya, Pratima Yadav, Reena Yadav, Prof. Chandana Nighut, Personality Prediction by Discrete Methodology, IOSR Journal of Engineering (IOSRJEN) ISSN (e): 2250-3021, ISSN (p):22788719 Volume 14, PP 10-13 International Conference on Innovative and Advanced Technologies in Engineering.
7. Barrick, M. R., Mount, M. K. (1991). Big Five Personality dimensions and job performance: Ameta analysis Personnel Psychology, Vol. 44, pp. 1-26.
8. F. Harrell, Regression modelling strategies: with applications to linear models, logistic regression, and survival analysis Springer, New York (2001)
9. Diseth A, (2003). The Personality and approaches to learning as predictors of academic achievement. European Journal of Personality, 17, 143-155.
10. Michael M Tedasse; Lin Boxchu; Liang Yang, Personality Predictions Based on User Behavior on the Facebook Social Media Platform, 2018.