iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** IV **Month of publication:** April 2023

DOI: https://doi.org/10.22214/ijraset.2023.50832

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue IV Apr 2023- Available at www.ijraset.com

Psychometric Test and Personality Assessment by using Machine Learning

Dr. V. K. Taksande[1], Lakhan Shende[2], Gunjan Bhure[3], Nitish Kumar[4], Vaibhav Waghule[5]
[1] *HOD, Electronics & Telecommunication Department, Priyadarshini College of Engineering, Nagpur*
[2, 3, 4, 5] *Student, Electronics & Telecommunication Department, Priyadarshini College of Engineering, Nagpur*

Abstract: ***Psychometric tests and personality assessments are widely used in a variety of settings, from academic research to employment screening. Traditional methods of administering and scoring these tests can be time-consuming and labor-intensive, making them difficult to scale or automate. However, recent advances in machine learning have made it possible to develop more efficient and accurate psychometric tests and personality assessments. This paper provides an overview of the use of machine learning in the development and administration of psychometric tests and personality assessments. We discuss the benefits and limitations of machine learning-based approaches, including the ability to process large amounts of data quickly and identify complex patterns that may not be immediately apparent to human observers. We also review some of the key challenges associated with using machine learning in this context, including the need for high-quality data and the potential for algorithmic bias. Overall, we argue that machine learning has the potential to revolutionize the field of psychometric testing and personality assessment, enabling researchers and practitioners to develop more efficient, accurate, and equitable assessment tools. However, we caution that careful consideration must be given to issues of data quality, bias, and fairness in order to ensure that machine learning-based assessments are truly effective and equitable***

I. INTRODUCTION

The project's goal is to create a working prototype of a platform that does away with the manual screening of countless applications and resumes to determine which candidates best match the requirements of the job being offered.

We are looking for a more effective method of ranking submitted candidate CVs and twitter account from a big pool of applicants in a way that is both consistent and legal.

Based on the CV rating policy, this method will enable the HR department to quickly shortlist candidates. The Big Five, Rorschach test, and MBTI test are just a few of the tests that can help identify personality types.

In this project, the MBTI test is taken into account when predicting personality. As a systematization of archetypal personality types utilized in therapeutic practice, the MBTI personality classification system emerged from Jungian psychoanalytic psychology. The system is organized into 16 different individuals along four binary orthogonal personality traits.

The dimensions are as follows

Extraversion (E) and introversion (I) are indicators of how much a person enjoys their inner or outward environment, respectively.

Sensing (S) vs. Intuition (N): A comparison of how much information is processed by a person using their five senses vs their perception of patterns.

Thinking (T) vs. Feeling (F) is a test of one's preference for objective principles and facts over consideration of other people's emotive viewpoints.

A gauge of a person's preference for a structured and ordered life vs a flexible and spontaneous one is the Judging (J) vs. Perceiving (P) scale.

II. BODY OF PAPER

The aptitude test aids in identifying the underlying trends in a candidate's interests and forecasting the stream in which the candidate will be most interested.

For an organization, it is extremely important to understand a candidate's natural aptitude. Applicants can test their aptitude, and a report is then prepared to gauge their level of interest. Also candidate can be screened by using his/her Twitter account by scanning his/her posts on twitter. Based on this, the human resource manager can choose the best candidate for a position and assign them to the appropriate team.

©IJRASET: All Rights are Reserved

A. Assessment

The registered applicants may take the aptitude test by completing this section of the application. Through a portal, the applicant registers, completes the required profile information, and logs into the application. The student looks at the evaluation report. The tool enables the preparation of a series of questions by the human resource staff. Despite the fact that a human resources employee developed the exam paper, the questions are automatically assigned to the candidates. Additionally, it enables staff to oversee students and change their own profiles. The human resource manager registers the employees.

B. Question Format

Every question is an MCQ. This programme can support four different types of inquiries.

1) Typical MCQ tests

Question with a picture and an answer

Image-based question and answer pair.

There are four primary categories of questions: science, business, humanities, and aptitude. The human resources staff selects roughly fifteen questions from each component to create the test. Sixty questions are on each exam paper. The weighting of the questions is determined by the category under which each one is placed. The weighting is put to good use in the candidate assessment evaluation. The pupils must tackle each of the sixty questions.

2) Team of Human Resources

The only individual who can add a member to the Staffing team is the Human Resource Manager.

The questions a candidate should answer are chosen by human resources staff.

The weight that each question should have is determined by the human resources staff.

The Archive

pupil profile

responses provided by pupils

instructor information

question specifics

Answer specifics, including the correct response and all possible answers

3) Test of Personality

Social media creates constant contact between its users and the outside world by allowing them to share intimate details and opinions about every facet of life. Examining how Twitter (a dataset) may be used to predict personality is the main goal of this study. We demonstrate how to create machine learning models using textual data. Since Twitter is the traditional entrance point for practicing machine learning, candidates can utilize their account names to view their anticipated Myers-Briggs Personality Types utilizing this tool. Using Twitter data, you could conduct research thanks to the intriguing combination of data (tweet contents) and meta-data (location, hashtags, users, re-tweets, etc.). Some subjects are covered in this training:

4) Managing Imbalanced Datasets in Exploratory Data Analysis

Model Development Model Training Model Evaluation: Vectorization of Text Data

Collection Description

In this study, 8675 rows of the Myers-Briggs personality type dataset from Kaggle, which is available to the general public, were used. Each row in this dataset has two columns. A person's MBTI personality type is listed in the first column, and fifty posts from their social media accounts are listed in the second column. Three pipe characters have been used to separate each post. This information was gathered from participants in an online forum where participants first completed a quiz to determine their MBTI type and then engaged in conversation with other participants.

C. Four-dimensional Classification of Type Indicators

The type indicators were divided into four groups in order to comprehend how they were distributed throughout the dataset.

The first category was for introversion (I) and extraversion (E), the second for intuition (N) and sensing (S), the third for thinking (T) and feeling (F), and the fourth for judging (J) and perceiving (P) (P).

As a result, one letter will return for each category, resulting in a string of four letters that each stand for one of the MBTI's 16 personality types. For instance, the applicable personality type would be INTJ if the first category returned I, the second category returned N, the third category returned T, and the fourth category returned J.



Fig-1: 16 Types of personality in MBTI technique

D. Common Terms in a Word Cloud

The magnitude of each word in a word cloud, a data visualization technique for expressing text data, shows its frequency or relevance. With a word cloud, significant textual data points can be highlighted. Word clouds are frequently employed for social network data analysis. The most commonly occurring words for each of the characters were examined using word clouds.

Model Development, Model Training, and Model Preservation Predictions were made using the classification techniques listed below:

1) *Random Forest:* Based on various sets of instances from the dataset, Random Forest is a classification technique that mixes numerous decision trees. Each tree in the forest is dependent on the values of a random vector that was sampled individually, using the same distribution for every tree. The most popular class is chosen using the Random Forest technique, which combines the outcomes of decision trees.
2) *Naive Bayes:* This classification method relies on the Bayes Theorem and makes the assumption that predictors are independent. Simply put, a Naive Bayes classifier believes that the existence of one feature in a class has no bearing on the existence of any other features.
3) *Support Vector Machine (SVM):* For two-group classification issues, a support vector machine (SVM) is a supervised machine learning model that employs classification techniques. A non-probabilistic binary linear classifier, SVM training algorithm creates a model that categorizes fresh instances into one of two groups (although methods such as Platt scaling exist to use SVM in a probabilistic classification setting). SVM assigns training samples to spatial coordinates in order to maximize the distance between the two categories.

With a decision tree, categorization criteria are represented by the paths from root to leaf. Each internal node of a decision tree represents a test on a feature, each leaf node a class label (decision made after computing all features), and the branches the conjunctions of features that result in those class labels.

Extreme Gradient Boosting, or XGBoost, is a well-known supervised-learning technique for classification on huge datasets. In order to produce reliable results, it employs shallow decision trees that are formed sequentially and a highly scalable training technique that prevents overfitting.

E. Working of system

1) *Case Study:* We will examine numerous research publications in this module and use the information we learn there to create our project.
2) *Project Plan:* Establishing a project plan where we will choose the technology to be used, the methodology, and the required vs. actual results.
3) *Technology Stack Selection:* Python programming language serves as our backend language for the web portal, with HTML, CSS, and JavaScript serving as the frontend languages. MySQL will be used for data storing and manipulation. Python is used to code the machine learning model in the Jupyter notebook.

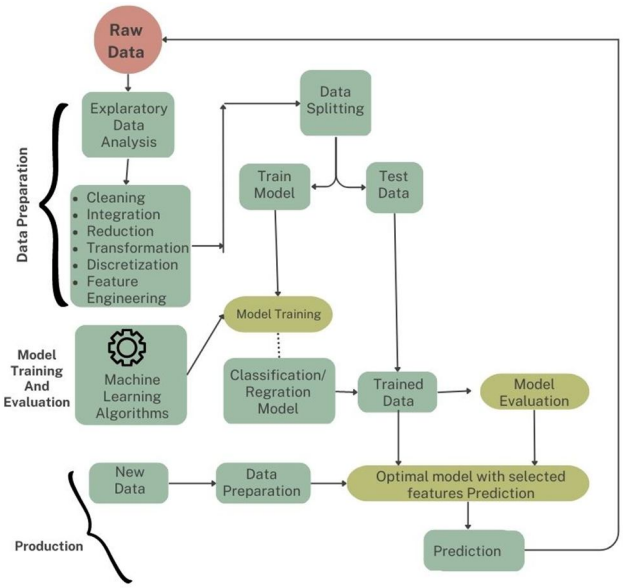


Fig-2: Block Diagram of Proposed System

4) *Building Web Portal:* We have created a website that will serve as the testing environment using all available web technologies and computer languages.
5) *Collecting Data By Conducting Tests:* We will administer tests to a number of candidates after the portal is ready, and we will save the results in our database.

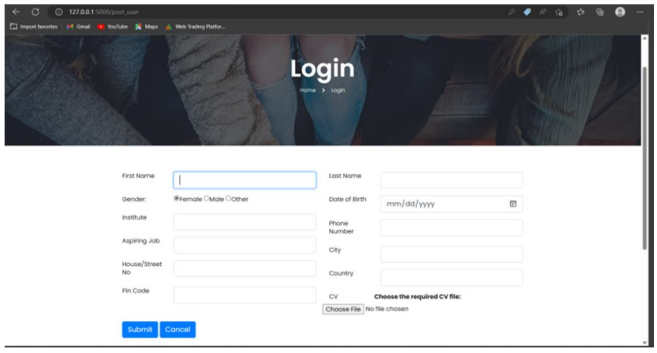


Fig-3: User login interface

6) *Building, Testing And Running Machine Learning Model:* We'll employ machine learning algorithms in this module. By obtaining the test data from the database and loading it into our machine learning model, we may train the model first and then obtain the outcome by adding more data. Foreseeing Results and Forecasting Output. We shall compare the needed and actual results at the conclusion.

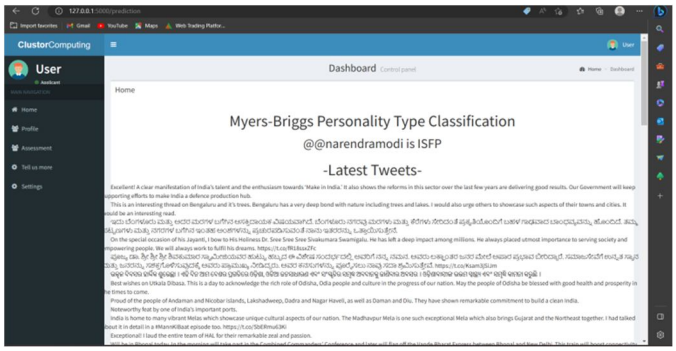


Fig-4: User's personality result using Twitter ID

III. CONCLUSIONS

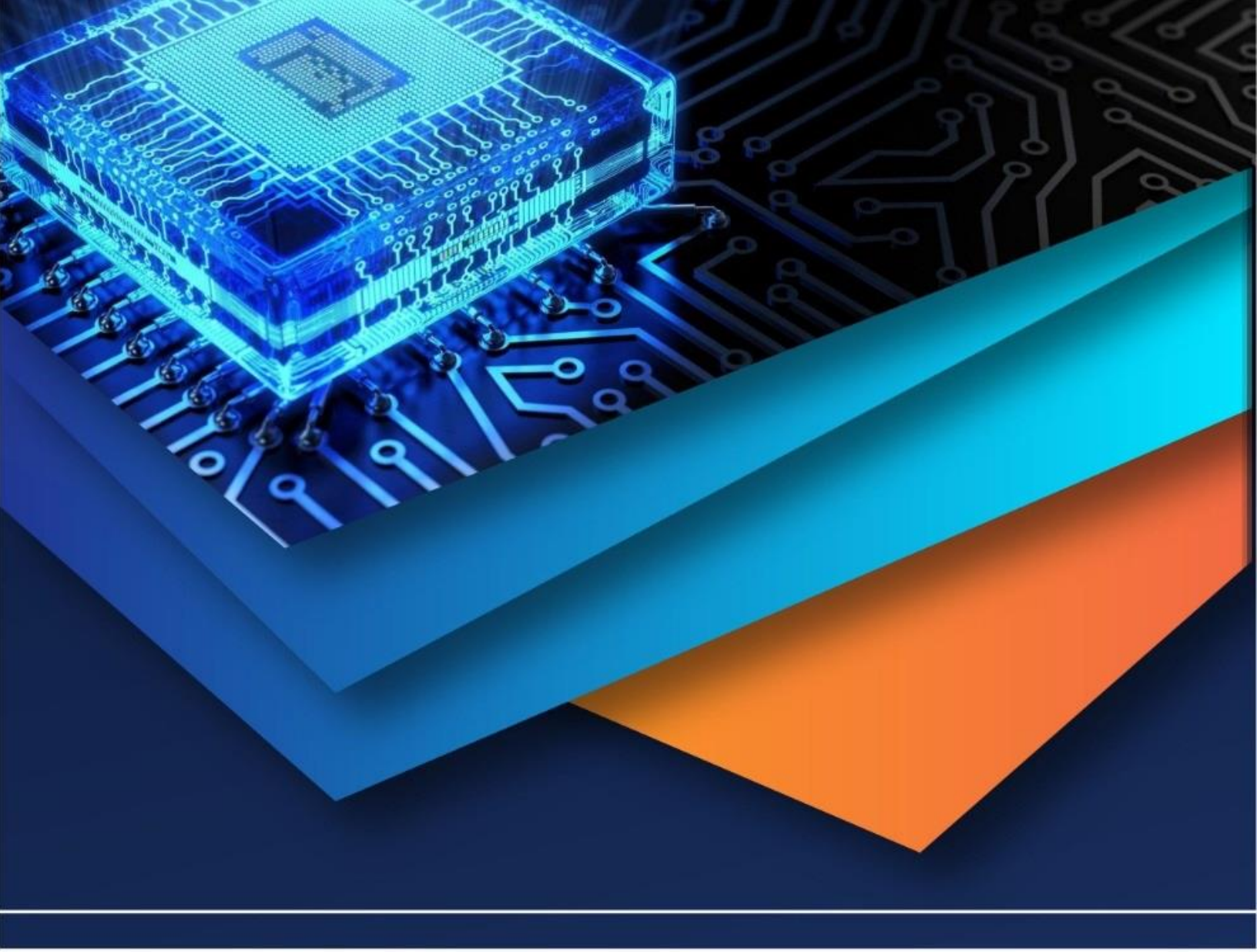
The AI Personality Prediction System using CV Analysis predicts a person's personality traits based on facial expressions and other visual cues using state-of-the-art computer vision and machine learning algorithms. The approach has the potential to totally change a number of industries; including employment, customer service, and targeted advertising, by providing insights into a person's personality that were previously difficult to gain without protracted interviews and exams. Yet, the development and application of such a system also raises important moral and legal issues, including those of privacy, bias, and discrimination. The system needs to be developed and put into use in a responsible and moral manner, with a focus on justice, accuracy, and privacy. This means obtaining their informed consent, ensuring that the system does not discriminate on the basis of factors like gender, race, or ethnicity, and protecting the data from misuse and unauthorized access before collecting and analyzing the data of any individual. Other factors that affect the accuracy of the system include the performance of the facial recognition and analysis algorithms, the performance of the machine learning model, and the model's capacity to generalize to new data. The effectiveness of the system must be monitored and improved by ongoing testing. Based on respondents' preferences in the four dichotomies of extraversion against introversion, sensing versus intuition, thinking versus emotion, and perceiving versus judging, the widely used MBTI personality test categorizes individuals into one of 16 personality types. The test has been utilized for many years across a number of academic fields, including psychology, education, and career counselling. The MBTI test has benefits including usability and simplicity, but it also has some disadvantages. The test has drawn a lot of flak for lacking scientific validity and dependability. According to some, the test's results are inconsistent and it is not a reliable predictor of behavior or job success. Academics have also criticized the dichotomies' binary nature and the absence of a continuous range of personality traits. Despite these concerns, the MBTI test is still a popular way to evaluate personality, and many people find the results to be insightful and helpful. It can provide people a common language and act as a springboard for reflection and personal growth. As a result, the AI Personality Prediction System using CV Analysis has the potential to revolutionize a variety of industries and provide valuable knowledge about a person's personality traits. To ensure fairness, accuracy, and privacy, however, such a system must be developed and implemented with care and accountability. Only then will we be able to fully benefit from this ground-breaking technology while guaranteeing that it has no adverse effects on people or society. Despite the fact that it might not be the best tool for personality analysis, the MBTI exam can be a great and helpful resource for those who want to better understand themselves and others. It is essential to be aware of its limitations, utilize the results with caution, and incorporate them with other assessments and recommendations.

IV. ACKNOWLEDGEMENT

The well-known and established Priyadarshini College of Engineering upholds the ideal of producing technocrats who will take on difficult tasks in the technical area for the benefit of people. We would like to take this opportunity to extend our sincere gratitude to our guide Dr. V. K. Taksande, HOD of the Electronics and Telecommunication Department at PCE Nagpur) and co-guide Dr. O.G. Hastak who consistently oversaw our work with the utmost care and devotion. They have consistently provided us with guidance as we have worked to present our project on "Psychometric Test And Personality Assessment By Using Machine Learning". It gives us great pleasure to extend our sincere gratitude and appreciation to Dr. S.A. Dhale, the principal of Priyadarshini College of Engineering, Nagpur. We extend our sincere gratitude to the lab assistants from PCE's Electronics and Telecommunication Department for their invaluable assistance in carrying out the project work. The Institute's teaching and non-teaching staff, as well as all of our friends and family members, who directly and indirectly assisted us in completing and realizing our project, deserve our sincere gratitude.

REFERENCES

[1] Ethem Alpaydin. 2004. Introduction to Machine Learning. Cambridge, MA.
[2] Elaf Abu Amrieh, Thair Hamtini, and Ibrahim Aljarah. 2016. Mining educational data to predict student's academic performance using ensemble methods. International Journal of Database Theory and Application 9(8), 119-136.
[3] S. K. Card and J. Mackinlay. 1997. The structure of the information visualization design space. In: Proceedings of the 1997 IEEE Symposium on Information Visualization IEEE, 92-99.
[4] G. Dekker, M. Pechenizkiy, and J. Vleeshouwers. 2009. Predicting students drop out: A case study. In: Educational Data Mining 2009, 41-50.
[5] Pedro Domingos. 2012. A few useful things to know about machine learning. Communications of the ACM 55(10), 78-87.
[6] Wayne W. Eckerson. 2007. Predictive analytics. Extending the Value of Your Data Warehousing Investment. TDWI Best Practices Report 1, 1-36.
[7] Tom Fawcett. 2005. An introduction to ROC analysis. Pattern Recognition Letters 27(8), 861-874.
[8] J. Friedman, T. Hastie, and R. Tibshirani. 2001. The Elements of Statistical Learning. Springer, Berlin: Springer Series in Statistics.
[9] Isabelle Guyon and André Elisseeff. 2003. An introduction to variable and feature selection. Journal of Machine Learning Research 3, 1157-1182.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)