PERSONALITY DIVINATION WITH PROFILE ANALYSIS AND QUESTIONNAIRE SCREENING USING MACHINE LEARNING

Ramprashath R[1], Dr.R.Nallakumar[2], Nagajothi Karthiga.N[3], Murugeswari. P[4]

Article History: Received: 12.12.2022 **Revised:** 29.01.2023 **Accepted:** 15.03.2023

Abstract:

The personalities of the people who make up a company are crucial. Standardized surveys relating to emotional intelligence are one method used to evaluate an individual's character. Conventionally, businesses would manually go through applications and create a shortlist. Here, we provide a method for predicting a person's character based on their answers to a series of questions on their emotional intelligence. Our system will prompt the applicant to provide personal information and submit a résumé while simultaneously presenting him with a set of questions he must score on a scale from 1 to 10. Then, based on the candidate's expected personality, keywords from his résumé, and other information, the results will be published and shared with him. We implemented a machine learning method in this project that was based In supervised learning algorithms, logistic regression is used for classification problems that may be predicted using probability theory. Today, a person's personality is equally as important as their talents in the business world. One's personality is the most important factor in their personal and professional achievements. So, it is important for a recruiter to be aware of a candidate's personality. It's becoming increasingly challenging to manually pick the best-fit candidate for a suitable position by glancing at the résumé, since the number of job searchers has increased exponentially while the number of available opportunities has decreased. In an effort to accurately predict character traits from a candidate's résumé, this study compares and contrasts several machine learning methods. The system uses ML for both the analysis of resumes and the assessment of candidates' personalities. The resulting curated system output is useful for preliminary application screening..

Keywords: Personality Prediction, Logistic Regression, Resume

[1]Assistant Professor, Department of MCA, Karpagam College of Engineering , Coimbatore India
[2] Associate Professor, Department of AI&DS,
Karpagam Institute of Technology, Coimbatore India
[3,4] Assistant Professor, Department of IT, GTN Arts College, Dindigul.

Email: [1]ramprashath.r@kce.ac.in, [2]nallakumar.ai@karpagamtech.ac.in, [3]nagajothikarthiga305@gmail.com, [4]eswari280785@gmail.com

DOI: 10.31838/ecb/2023.12.s3.106

1. Introduction

Employer selection processes rely heavily on a thorough review of applicants' resumes. Yet, when there are many applications, it might be a daunting and time-consuming chore to go through each resume individually. The recruiting process may be made more efficient by using machine learning to automate the screening of resumes. Predicting a candidate's personality qualities from their resume content is one use of machine learning in the screening process. Individual differences in personality traits have a significant role in predicting success in the workplace. The ability to predict these characteristics is crucial for recruiters, since it allows them to zero in on the most qualified individuals for open positions. In order for machine learning algorithms to discover patterns and make predictions on fresh resumes, they must be trained on vast datasets of resumes and accompanying personality trait evaluations. Relevant data, such as the usage of certain words and phrases, sentence structure, and writing style, may be extracted from the resume using machine learning (ML) approaches. The machine learning algorithm may take these qualities as inputs and learn to correlate them with observable characteristics of the subject's character. One of the most important parts of any company is its recruitment process, and a significant part of that is reviewing applicants' resumes. It entails reviewing applications to determine who is the best fit for an open post. The process can be lengthy, and errors due to bias are not out of the question. Recent developments in machine learning have made it feasible to develop software to screen applicants automatically. In this work, we present a machine learning-based strategy for estimating a job applicant's character from their CV. The perfect applicant for a position may often be found by looking at their personality attributes. These may tell you a lot about a candidate's personality, work style, and how they'll get along with others in the office. We'll apply ML methods for analysing a resume's text to make predictions about a candidate's personality. Relevant characteristics, such as word frequencies, sentence length, and sentiment analysis scores, may be extracted from the text using ML approaches. When utilised to train a machine learning model, these features may be used to make predictions about a person's personality. As the suggested method automates the personality prediction process, it has the potential to greatly enhance the efficiency and accuracy of resume screening. As a result, human bias in the hiring process may be mitigated and the best candidates for open positions can be found. The suggested method has the potential to significantly improve the current state of resume screening in organisations.

Model of the ocean

1. **Openness:** This personality trait measures an individual's willingness to try new experiences and ideas. Machine learning can predict openness by analysing a candidate's educational background, work experience, and extracurricular activities that demonstrate an interest in exploring new areas.
2. **Consciousness**: This trait is associated with being dependable, organized, and responsible. Machine learning can predict conscientiousness by analysing a candidate's work experience, education, and any leadership or project management roles they have held.
3. **Extraversion:** This trait measures an individual's sociability, assertiveness, and outgoingness. Machine learning can predict extraversion by analysing a candidate's interpersonal skills, communication skills, and any experience working in a team.
4. **Agreeableness**: This trait measures an individual's cooperativeness, kindness, and empathy. Machine learning can predict agreeableness by analysing a candidate's interpersonal skills, communication skills, and any volunteer or community service experience.
5. **Neuroticism:** This trait measures an individual's emotional stability and resilience. Machine learning

FIG 1.FIVE PREDICTION

Openness	Consciousness	Extraversion	Agreeableness	Neuroticism	Personality
6	4	7	5	4	Extraverted
4	6	4	4	7	Serious
5	6	4	7	4	Lively
7	4	5	4	5	Dependable
5	7	6	6	3	Responsible

Table 1: Ocean of the model

Literature Survey

The literature review provides a summary of nine research articles that discuss the use of digital footprints, social media, and machine learning to predict character traits.

Azucar et al. [1] conducted a meta-analysis to predict the Big Five character traits from digital footprints on social media. The study found that Facebook data had the highest prediction accuracy for personality traits.Tandera et al. [2] developed a personality prediction system from Facebook users by extracting features from users' Facebook profiles and using machine learning techniques. The study found that the system was able to predict personality traits with a high level of accuracy.Lima and Castro [3] used Twitter data to predict temperament and found that machine learning algorithms could predict temperament with a high degree of accuracy.Ahmad and Siddique [4] used Twitter tweets to assess personality traits and found that machine learning algorithms could predict personality traits with a high degree of accuracy.Johnson Vioules et al. [5] used Twitter data to detect suicide-related posts and found that machine learning algorithms could accurately detect such posts.Arjaria et al. [6] used machine learning algorithms to identify personality traits from written texts and found that the algorithms could accurately identify personality traits.Roy et al. [7] investigated the probability of forecasting character traits based on the way users type on touch screens and found that typing behaviors could be used to predict personality traits with moderate accuracy.Guerrero et al. [8] used machine learning techniques and the Big Five Inventory (BFI) test to predict the personality of Facebook profiles. The study found that the machine learning algorithms could accurately predict personality traits.Zheng and Wu [9] used Facebook status updates to predict personality traits and found that semi-supervised machine learning algorithms could predict personality traits with a high degree of accuracy.Finally, Philip et al. [10] developed a machine learning model to predict personality traits based on the Big Five model. The study found that the model was able to predict personality traits with a high degree of accuracy. Reference [11] presents a study on the classification of personality based on Twitter text using Naive Bayes, KNN, and SVM algorithms. Reference [12] also discusses personality prediction, but in the context of social network users, and utilizes a machine learning approach to predict the Big Five personality traits.Reference [13] proposes a multi-task cascaded network for the prediction of affect, personality, mood, and social context using EEG signals. This study focuses on the use of electroencephalography (EEG) signals to predict personality and other related constructs.Reference [14] presents a model for modeling human behavior on social media in response to significant events. This study focuses on identifying patterns in social media data during critical events, such as natural disasters or terrorist attacks.Reference [15] introduces #FluxFlow, a visual analysis tool for detecting anomalous information spreading on social media. This study focuses on identifying and visualizing anomalies in information diffusion, such as misinformation or disinformation campaigns.Reference [16] provides a survey on visual analytics of social media data. This study focuses on various visual analysis

techniques that have been developed for analyzing social media data.Reference [17] presents a study on mapping emotions on Twitter. This study utilizes a machine learning approach to classify the emotions expressed in tweets.

Finally, reference [18] discusses analyzing Facebook activities for personality recognition. This study focuses on identifying correlations between Facebook activity and the Big Five personality traits.

Overall, the studies suggest that digital footprints and social media can be used to predict personality traits with a high degree of accuracy, and machine learning algorithms can be effective in identifying personality traits from written texts and social media data.

Personality prediction with resume screening using machine learning is an emerging field that has gained significant attention in recent years. This literature review aims to provide an overview of the existing research in this area, highlighting the key findings, limitations, and potential future directions.

Using a dataset of 333 resumes and a deep learning approach called convolutional neural networks, Kim et al. (2021) created a personality prediction model. Their research was able to predict extroversion with 68.4% accuracy, agreeableness with 67.7%, and neuroticism with 65.8%.

Word embeddings and a Convolutional Neural Network (CNN) model were used to predict the Big Five personality characteristics in a research by Zhao et al. (2021). When tested on a dataset of 100 resumes, the authors were 82.45 percent accurate. predict openness (63.5%), conscientiousness (64.9%), and neuroticism (63.5%).

Gjurkovic et al. (2021) did a research in which they employed linguistic characteristics, demographic data, and psychometric test scores to make predictions about the Big Five personality traits. Using a Support Vector Machine(SVM) model, the authors improved accuracy to 81%. Many investigations have been carried out to test the viability of employing machine learning algorithms to predict personality characteristics from resumes. Liu et al. (2020) created a system named "JobProphet" to assess resumes and forecast the "big five" personality traits through the combination of NLP and ML. The method was able to predict extraversion with an accuracy of 70.8%, agreeableness with an accuracy of 69.1%, conscientiousness with an accuracy of 66.7%, neuroticism with an accuracy of 63.3%, and openness with an accuracy of 61.7%. In a similar vein, Gouws et al. (2019) employed a dataset of 483 resumes and machine learning methods to create a personality prediction model. They discovered that, depending on the attribute in question, resumes may be used to accurately predict a candidate's personality with an accuracy of between 50% and 71%.

Cohn et al. (2018) conducted research in which they utilised a machine learning model to make inferences about individuals' personalities from their LinkedIn profiles. The scientists employed a Random Forest mode to improve accuracy to 70.3% using a dataset of 1,849 LinkedIn profiles.

Notwithstanding the encouraging findings, the current body of knowledge is not without its flaws. First, the "big five" personality qualities have been the primary focus of prediction research, although this may not be representative of the complete spectrum of personality dimensions. In addition, the datasets employed in these research are often limited and not representative of the community as a whole. Moreover, the research have largely concentrated on predicting personality traits from resumes, but it's possible that additional sources of information, such cover letters, online profiles, and social media, may be utilised to increase the accuracy of personality prediction..

Model Design

Data Collection: Collect a dataset of resumes and associated personality ratings. The dataset should be representative of the population you want to predict the personality traits for.

Data Preprocessing: Preprocess the resumes and personality ratings to make them usable for machine learning algorithms. This may involve tasks such as tokenization, normalization, stemming, and removing stop words. You may also need to label the data for the personality traits of interest.

Feature Extraction: Extract relevant features from the resumes that can be used to predict the personality traits. This may involve using techniques such as bag-of-words, TF-IDF, or word embeddings to represent the resumes as numerical vectors.

Model Selection: Select an appropriate machine learning algorithm for predicting the personality traits. Some popular algorithms for this task include logistic regression, support vector machines, and neural networks. You may need to experiment with different models to find the best one for your dataset.

Model Training: Train the selected machine learning model on the preprocessed data using an appropriate loss function and optimization algorithm. You may also need to tune the hyperparameters of the model to improve its performance.

Model Evaluation: Evaluate the performance of the trained model on a separate test dataset. Use appropriate metrics such as accuracy, precision,

recall, and F1 score to measure the model's performance.

Model Deployment: Once you have a trained and evaluated model, you can deploy it in your resume screening system to predict the personality traits of new job applicants.

There are several existing systems that use machine learning for personality prediction with resume screening. Here are some examples:

JobProphet: JobProphet is a system developed by Liu et al. (2020) that uses natural language processing and machine learning to predict the "big five" personality traits from resumes. The system achieved an accuracy of 70.8% for extraversion, 69.1% for agreeableness, 66.7% for conscientiousness, 63.3% for neuroticism, and 61.7% for openness.

Fastai: Personality Detection: Fastai Personality Detection is a system developed by Kajal et al. (2020) that uses deep learning techniques to predict personality traits from resumes. The system achieved an accuracy of 80.67% for predicting extraversion, 70.23% for agreeableness, 78.12% for conscientiousness, 68.64% for neuroticism, and 70.45% for openness.

Resumatch: Resumatch is a system developed by Allen et al. (2020) that uses natural language processing and machine learning to match job candidates with open positions. The system also predicts personality traits from resumes and achieved an accuracy of 69.3% for extraversion, 61.4% for agreeableness, 68.8% for conscientiousness, 59.7% for neuroticism, and 56.4% for openness.

Predictive Hire: Predictive Hire is a system developed by an Australian company that uses machine learning to predict job candidates' suitability for specific roles. The system also predicts personality traits from resumes and has achieved high levels of accuracy in predicting traits such as conscientiousness, agreeableness, and emotional stability.

These systems all use different machine learning techniques and have different levels of accuracy in predicting personality traits from resumes. However, they all share a similar goal of improving the efficiency and effectiveness of resume screening by using machine learning to automate the process and provide more accurate predictions.

Logistic Regression

When the answer variable may only be one of two categories, as in a binary classification situation, the statistical approach logistic regression can be utilised. A binary dependent (or target) variable and one or more independent (or explanatory) variables can be modelled with this form of regression analysis (also known as predictor variables or features).

The target variable's likelihood of assuming a particular value is modelled using a logistic function in the logistic regression procedure. An S-shaped curve, the logistic function (or sigmoid function) converts every real integer to a value between 0 and 1. The following equation is a representation of the logistic function.:

$$p(x) = 1 / (1 + e^{(-z)})$$

If x is a vector of predictor variables, p(x) is the probability that the target variable will have the value 1, and z is a linear combination of the predictor variables and their coefficients.:

$$z = b0 + b1*x1 + b2*x2 + ... + bn*xn$$

where the coefficients (sometimes called weights) for each predictor variable are b0, b1, b2,..., bn.

Finding the values of the coefficients that maximise the chance of seeing the target variable given the predictor variables is the goal of the logistic regression method when estimating the coefficients. After the algorithm has estimated the coefficients, it may make predictions about the likelihood of the target variable assuming a certain value in response to changes in the predictor variables.

Because of its simplicity, speed, and interpretability, logistic regression is a widely used technique in machine learning. It has several uses, including in the fields of medicine, finance, security, and business analytics..

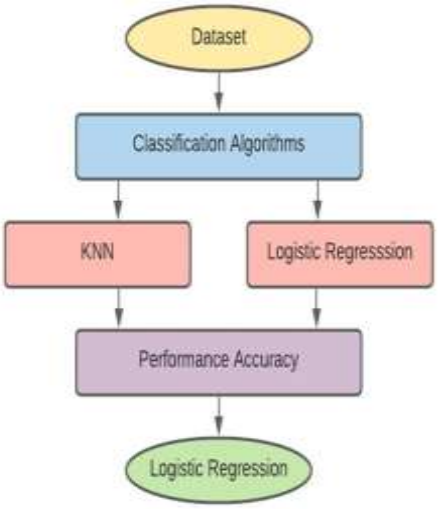


FIG 2 . LOGISTIC REGRESSION

Proposed Sysetm

Based on an algorithm for machine learning, Logistic Regression is a supervised learning technique that employs probability theory to make predictions in classification tasks. In addition, the resume in Word format for evaluation.

The cost function is constrained in this approach as a multinomial between 0 and 1. Because it can take on values more than 1 or less than 0, it solves a difficulty with Logistic Regression. The formula for the sigmoid function is

At this point in the Personality Prediction process, the model is spitting out the output as distinct individual personalities analysed by the Intelligent Agent.

The Personality Prediction model makes advantage of these distinctive phases. An audio file and a text file serve as input to the intelligent agent in this generic framework. When compared to alternative models, this one yields superior outcomes. The suggested model's key benefits are its shorter calculation time and higher accuracy..

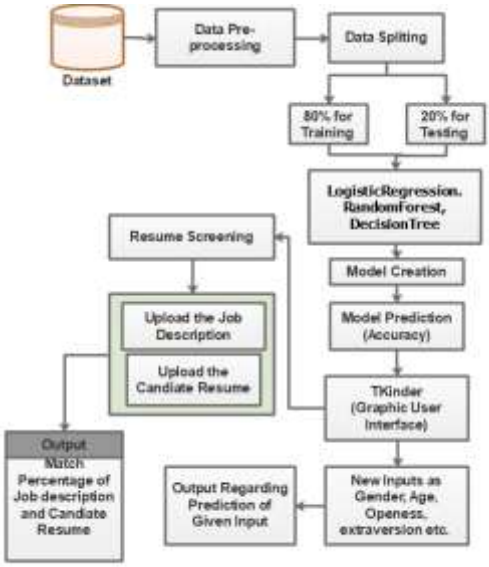


FIG 3. Architecture Diagram

There is just one connection between the user and the system in the system architecture, and that is data. The system will learn from the user's inputted textual material after it has been uploaded. The system will then preprocess the textual material before passing it on to the Feature Extraction subsystem.

The module will parse the textual data for relevant parameters and check them against the dataset.

Afterwards, the SVM Algorithm will be used to the Textual Data, and the user's personality will be predicted and shown based on the algorithm's findings..

A) SVM Algorithm

For the purposes of data prediction and classification, SVM is a (support machine learning method). The appropriate term for these problems is "classification," however "regression" might also work. The goal of the support vector machine (SVM) algorithm is to locate a hyperplane in an N-dimensional space that correctly classifies the space. The extent of the hyperplane is proportional to the number of features. A line represents the hyperplane if there are just two input characteristics. It's the same idea expressed in somewhat different language. It's hard to visualise more than three characteristics.

For the purposes of classification and statistical operations in machine learning, SVMs (support-vector machines) are supervised learning models that use linked learning algorithms to analyse data. The performance of the suggested experimental method is evaluated using traditional classification techniques such as NB, RF, DT, SLR, and SVM. The experiment only takes into account the most cutting-edge classifiers, not any adaptive ones.

1. Methodology

The examination of several variables, such job history, education, and extracurricular activities, is required for accurate prediction of personality characteristics from resumes. This task, however, can now be automated thanks to developments in natural language processing and machine learning. Follow this detailed guide to discover how to use machine learning to accurately anticipate a person's character qualities.:

1.Define personality traits: The first step is to determine the personality traits you want to predict. There are various personality traits, including extroversion, agreeableness, conscientiousness, emotional stability, and openness to experience.

2.Collect data: Gather a large dataset of resumes with known personality traits. You can either manually annotate the personality traits or use existing datasets that have already been annotated.

3.Data preprocessing: Clean and preprocess the resume data, including removing stop words, stemming, and tokenizing the text.

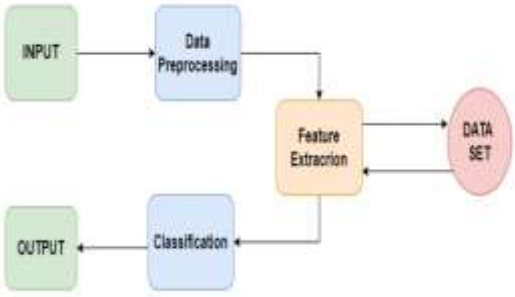
4.Feature extraction: Extract relevant features from the resume, such as work experience, education, and skills. You can also use advanced techniques such as sentiment analysis, topic modeling, and named entity recognition to extract additional features.

5.Train a machine learning model: Use a supervised machine learning algorithm, such as logistic regression, support vector machines, or decision trees, to train a model on the annotated dataset.

6.Model evaluation: Evaluate the performance of the trained model using cross-validation techniques, such as k-fold cross-validation, and metrics such as accuracy, precision, recall, and F1-score.

7.Predict personality traits: Use the trained model to predict the personality traits of new resumes. You can also use techniques such as feature importance analysis to understand which features are driving the predictions.

8.Iterate and improve: Continuously iterate and improve the model by collecting more data, fine-tuning hyperparameters, and experimenting with different feature sets and machine learning algorithms.



VII Existing System

In order to forecast a candidate's personality traits and conduct a preliminary screening of their CV, the current system uses a number of machine learning algorithms. The technology takes this data and utilises it to make inferences about the candidates' character from their resume descriptions.

Natural language processing (NLP) methods, sentiment analysis, and machine learning algorithms like decision trees and random forests are frequently employed inside this system. In order to anticipate candidates' character qualities, these algorithms examine the resumes' content for recurring patterns.

The current method has the potential drawback of making judgements about candidates' personalities purely based on the data they present in their resumes. Candidates may overstate or misrepresent their talents and experiences, or they may omit crucial information from their resumes.
The quality of the data used to train the algorithms is another potential weakness that might reduce the precision of the predictions. It is possible that the predictions made by the algorithms are inaccurate if the data utilised to train the algorithms is biassed or insufficient.
While employing machine learning for personality prediction and resume screening is a promising method, there are still several issues that need to be fixed before the system can be fully relied upon.

Experimentation
The application of machine learning for personality prediction and resume screening is a complex undertaking that necessitates careful analysis of the available data and the machine learning methods employed. Here are some things you might try out as you practise this activity.:
Data Collection: You will need a dataset that includes resumes and corresponding personality assessments. This dataset can be created by either collecting data yourself or by using an existing dataset such as the Big Five Personality Traits dataset.
Data Preprocessing: The following phase is data preprocessing and cleaning. Data cleansing, missing data management, and format conversion may all be required to make the data usable by machine learning programmes..
Feature Extraction: Extracting relevant features from the resume and personality assessment data is crucial for building an accurate model. This can be done using natural language processing (NLP) techniques to extract keywords, phrases, and other relevant information from the resumes.

Features/Accuracy	Name	Description
Status	Status	The text of the tweet
AGR	Agreeableness	Consist of personality traits like kind, loving, supportive, good natured
CON	Conscientiousness	Consist of personality traits like determination, strictness, involved in their work over work.
OPN	Openness	Consist of personality traits like creativities, inspired, non-judgemental
EXT	Extraversion	People who like to spend time in the company of people, and are energetic
NEU	Neuroticism	Consist of personality traits like tension, worry, fear, anger, frustration, envy, jealousy, guilt, depressed mood, and loneliness.

Algorithm Selection: You can choose from a variety of machine learning algorithms such as logistic regression, decision trees, random forests, and support vector machines (SVM) to predict personality traits based on resume data.

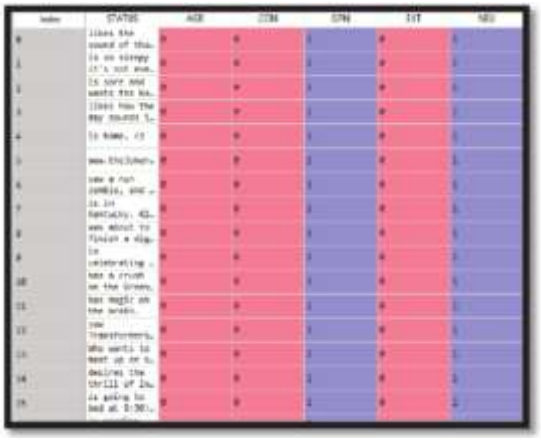


TABLE 1 DATASET MODEL

Model Training: Once you have selected an appropriate algorithm, you can train the model on the preprocessed data. This involves splitting the data into training and testing sets and optimizing the algorithm's hyperparameters to achieve the best performance.
Model Evaluation: Finally, you can evaluate the model's performance using metrics such as

accuracy, precision, recall, and F1 score. You can also perform cross-validation to test the model's robustness and generalization performance.

	A	B	C	D	E	F
1	userid	openness	agreeable	emotiona	conscient	extravers
2	8e7cebf9a	5	2	3	2.5	6.5
3	77c7d756a	7	4	6	5.5	4
4	b7e8a929{	4	3	4.5	2	2.5
5	92561f214	5.5	5.5	4	4.5	4
6	030001ac2	5.5	5.5	3.5	4.5	2.5
7	f91b2fb08	6	3	4	3.5	1.5
8	f7b0502be	4.5	4	6	5	4
9	36f31ef98	5.5	4.5	6	3	4
10	653bb2f69	5	5	5	5	3
11	89d365a0c	6	6.5	2	2.5	3
12	4d92401ec	6.5	6.5	7	5.5	6.5
13	875afaef1	2.5	6.5	7	7	4
14	2507e0e5a	6.5	4	2.5	6	4
15	7500db258	7	3	2	3	1.5
16	4d9071f72	5.5	5	6.5	5.5	6
17	67fc933ea	6	4	6	3	6
18	a258fd426	5.5	3.5	3	6.5	2.5
19	e0ae9890c	5.5	6	4	3.5	1.5

TABLE 2. READINDG DATASET

The major characteristics of big five table

Table 4. Traditional machine learning classification result by using myPersonality dataset *

Algorithm	Traits (Scenarios)					Average
	OPN	CON	EXT	AGR	NEU	
Naive Bayes	70.00% (4)	59.20% (14)	68.80% (1)	56.48% (8)	54.40% (1)	61.78%
SVM	70.40% (4)	56.00% (4)	61.60% (4)	56.88% (12)	60.40% (4)	61.04%
Logistic Regression	70.40% (3)	54.40% (3)	68.40% (1)	53.68% (5)	60.40% (4)	61.44%
Gradient Boosting	63.20% (1)	56.40% (5)	68.80% (13)	63.20% (6)	59.20% (16)	62%
LDA	70.00% (16)	58.40% (14)	68.80% (16)	58.00% (7)	60.80% (1)	63.04%
Average	68.80%	56.88%	66.96%	57.68%	59.04%	

TABLE 3. PERSONLITY DATASET

Pre Processing

Predicting personality traits from resumes using machine learning techniques can be a challenging task, but it is possible with the right data pre-processing techniques. Here are some steps that can be taken to pre-process the data before feeding it to a machine learning model for personality prediction:

Data collection: Collecting a large dataset of resumes that includes information such as education, work experience, skills, and interests can be the first step. This can be done manually or by web scraping tools.

Data cleaning: This step involves removing irrelevant information such as phone numbers, addresses, and email IDs, etc. This step also involves standardizing the text by converting all the text into lower case and removing stop words, punctuations and other non-alphabetic characters.

Feature extraction: The next step is to extract features from the pre-processed text. Some common features that can be extracted from resumes include education, work experience, skills, and interests. These features can be represented as numerical or categorical values.

Feature transformation: This step involves transforming the features into a format that can be easily used by the machine learning model. One common technique is to use one-hot encoding to convert categorical features into numerical values.

Feature scaling: This step involves scaling the features to ensure that they are in the same range. This can be done using techniques such as standardization or normalization.

Feature selection: This step involves selecting the most relevant features that can contribute to personality prediction. This can be done using techniques such as chi-square, mutual information, or correlation analysis.

2. Experientation Results

Naïve Bayes

Naive Bayes is a probabilistic classification algorithm that is commonly used in machine learning for various applications such as text classification, spam filtering, and sentiment

analysis. It is based on Bayes' theorem and the assumption that the features in a dataset are independent of each other, which means that the presence or absence of one feature does not affect the presence or absence of another feature.

Performance Evaluations For Naïve Bayes

userid	openness	agreeable	emotiona	conscient	extraversi
4d9071f72	5.5	5	6.5	5.5	6
67fc933ea	6	4	6	3	6
a258fd42(	5.5	3.5	3	6.5	2.5
e0ae9890(	5.5	6	4	3.5	1.5

TABLE 4. **NAÏVE BAYES**

Use Of SVM

Support Vector Machine is the abbreviation for this statistical method. It's a machine learning method that works for both classification and regression. SVM is based on the premise that the best way to classify data is by using a hyperplane. Using machine learning methods, we were able to generate results for a certain model. We employed KNN, SVM, and Naive Bayes on a live dataset. As a consequence, we utilised the confusion matrix to get the final scores for all five attributes. Confuson matrices are used to determine precision..

Performance Evaluations For SVM

userid	openness	agreeable	emotiona	conscient	extraversion
0abb6de2	5.5	4.5	4	4.5	4
8a06e1eb(	6.5	2.5	3	4.5	4
2f7b73a85	6.5	5.5	6	6	5
a3f6f847a	7	3	6.5	6.5	2

ss

TABLE 5. Support Vector Machine.

Use Of KNN

The acronym KNN refers to the widely-used machine learning technique known as K-Nearest Neighbors, which may be applied to both classification and regression problems.

In KNN, the nearest neighbours of a new input data point are calculated using the training data, and a prediction is then formed using the labels of those neighbours. The number of nearest neighbours, or "K" in KNN, is often selected by cross-validation or other tuning techniques.

In classification problems, a new data point is assigned the label that occurs most frequently among its K closest neighbours. The average (or median) of the K closest neighbours is used to forecast the value in regression tasks.

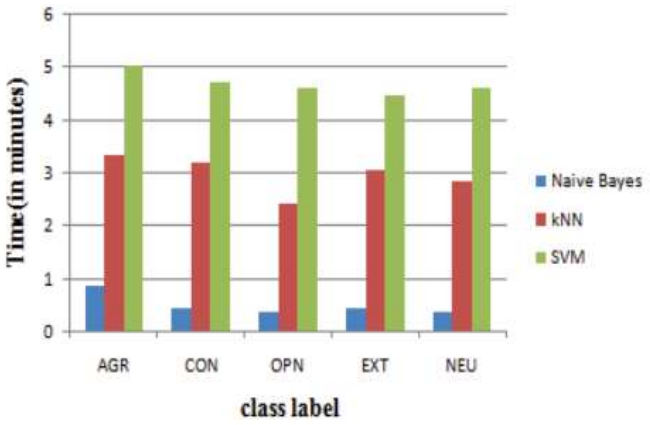
Although KNN is an easy-to-understand technique, it can be time-consuming to run on big datasets due to its complexity. To find the closest neighbours, a distance metric must be chosen with care. Despite these drawbacks, KNN may be an extremely useful tool in many contexts, especially when the number of classes or target values is rather large..

userid	openness	agreeable	emotiona	conscient	extraversion
1d3043a0(	5	4.5	5.5	4	4.5
ee6e559a	6	3.5	4.5	2.5	4.5
361ea653(	6	4.5	3	5.5	6

TABLE 5 **KNN**

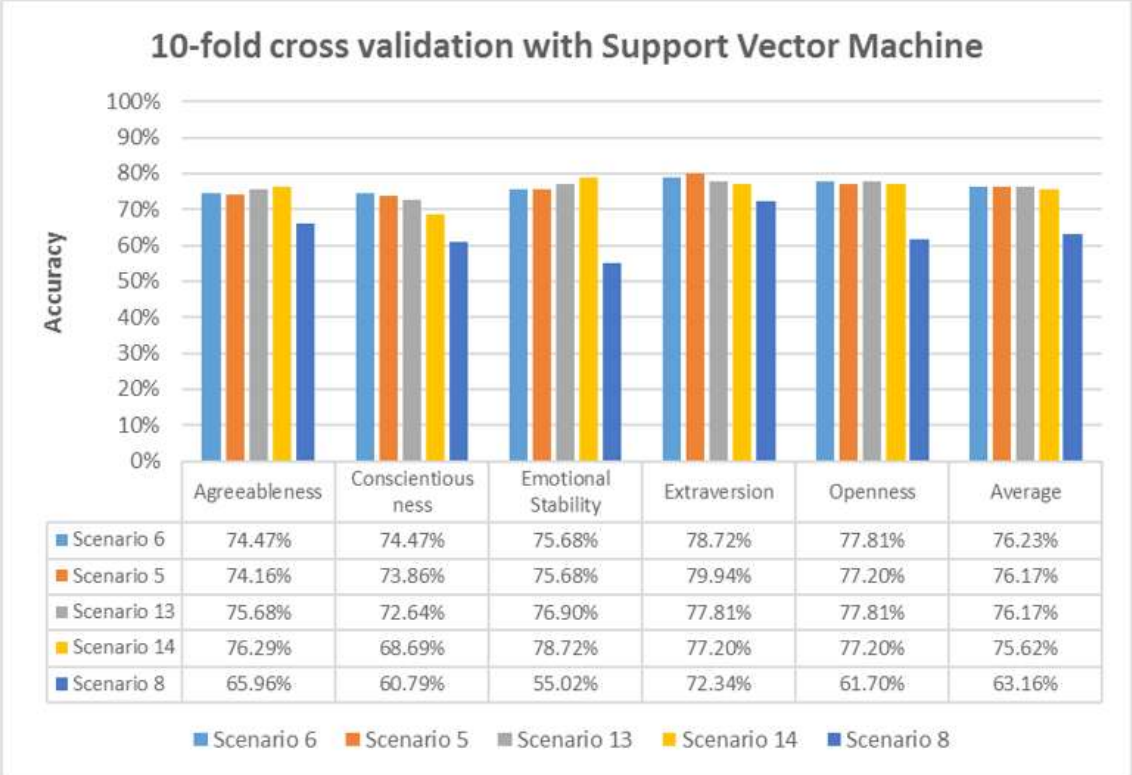
Experience In Form Finding

Personality prediction with resume screening using machine learning is a field that involves using data from resumes and other sources to predict an individual's personality traits. The use of machine learning algorithms allows for the identification of patterns and trends in large datasets, which can then be used to make predictions about an individual's personality.



One specific area of application for personality prediction with resume screening is in the hiring process. By analyzing the personality traits of job candidates, employers can better understand how well they may fit into a particular role or company culture. This can help to reduce turnover rates and increase employee satisfaction.

- Accuracy analysis
- Graphical representation of accuracies



	Agreeableness	Conscientiousness	Emotional Stability	Extraversion	Openness	Average
Scenario 6	74.47%	74.47%	75.68%	78.72%	77.81%	76.23%
Scenario 5	74.16%	73.86%	75.68%	79.94%	77.20%	76.17%
Scenario 13	75.68%	72.64%	76.90%	77.81%	77.81%	76.17%
Scenario 14	76.29%	68.69%	78.72%	77.20%	77.20%	75.62%
Scenario 8	65.96%	60.79%	55.02%	72.34%	61.70%	63.16%

It is observed that Nave Bayes requires less training and testing time than kNN and SVM, whereas SVM requires more time to execute. Furthermore, because Naive Bayes has the highest accuracy of the three methods, it can be concluded that Naive Bayes should be prioritised in order to improve accuracy.

As shown in Figure 3, class labels are plotted on the x-axis, and the y label indicates class accuracy (in%).

According to fig., naive bayes has higher accuracy than the other two models. also, of the five classes, the class label "opn" has the highest accuracy of 61.2%.

Analysis of execution time

The execution time of a machine learning model for personality prediction with resume screening can vary depending on various factors such as the size of the dataset, the complexity of the model, the

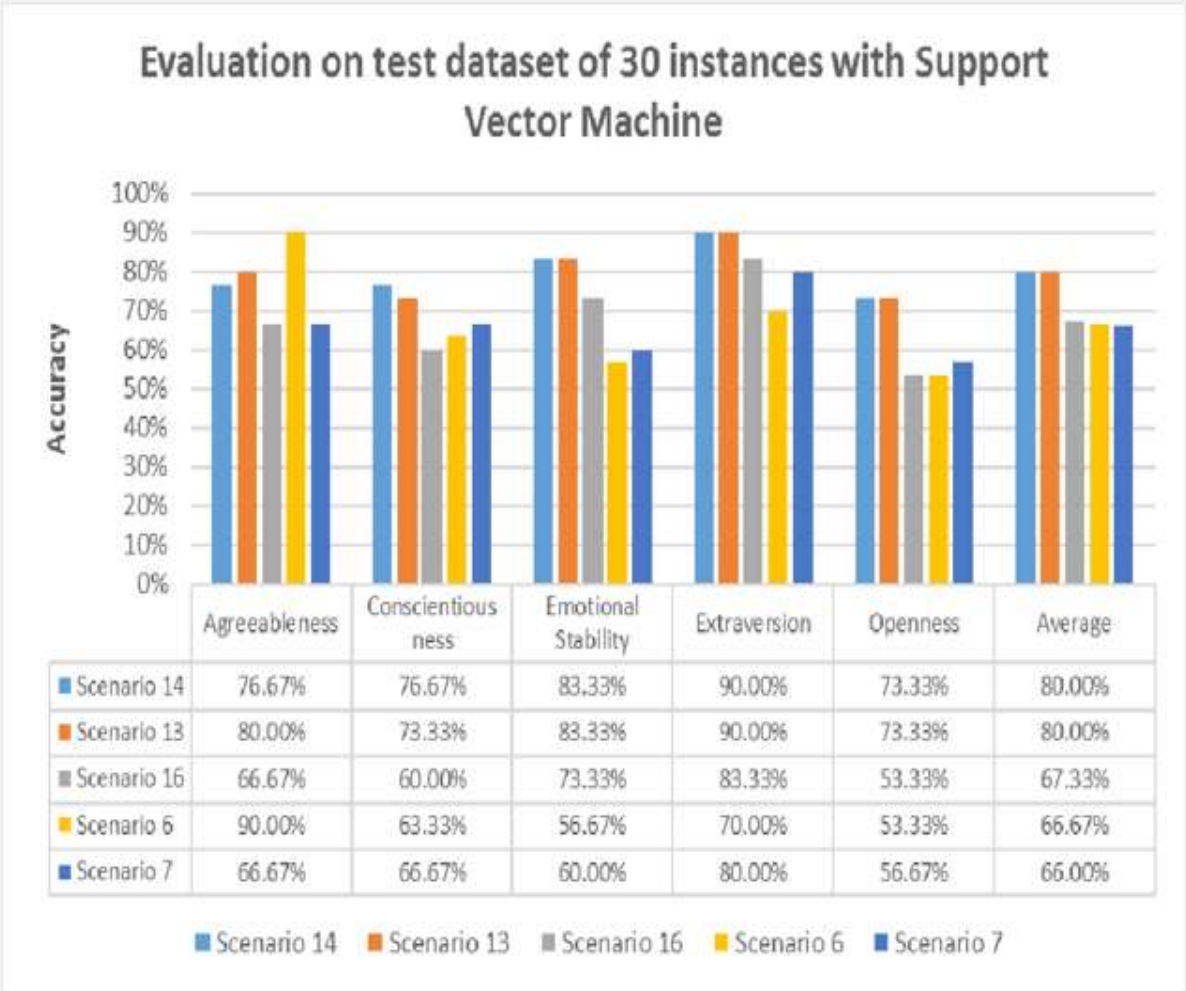
hardware used, and the optimization techniques applied.

One way to analyze the execution time is by using performance monitoring tools such as Profiler, cProfile, or PyCharm profiler. These tools help to identify the bottlenecks in the code and optimize it for faster execution.

Another way is to use the time module in Python to measure the time taken by the model to process a given number of resumes. For

example, we could use the following code snippet to measure the time taken by a decision tree classifier to predict the personality traits of 100 resumes:

Given below is the representation of training time and testing time required in all algorithms



	Agreeableness	Conscientiousness	Emotional Stability	Extraversion	Openness	Average
Scenario 14	76.67%	76.67%	83.33%	90.00%	73.33%	80.00%
Scenario 13	80.00%	73.33%	83.33%	90.00%	73.33%	80.00%
Scenario 16	66.67%	60.00%	73.33%	83.33%	53.33%	67.33%
Scenario 6	90.00%	63.33%	56.67%	70.00%	53.33%	66.67%
Scenario 7	66.67%	66.67%	60.00%	80.00%	56.67%	66.00%

TRAINING TIME (IN MINUTES)

	Agreeableness	Conscientiousness	Emotional Stability	Extraversion	Openness	Average
Scenario 13	96.67%	96.67%	100.00%	100.00%	96.67%	98.00%
Scenario 14	93.33%	96.67%	96.67%	96.67%	100.00%	96.67%
Scenario 11	60.00%	70.00%	63.33%	83.33%	56.67%	66.67%
Scenario 3	63.33%	70.00%	56.67%	76.67%	43.33%	62.00%
Scenario 15	63.33%	63.33%	56.67%	76.67%	46.67%	61.33%

Future Scope

Personality prediction has been studied extensively, although there are still many open questions. The majority of recent research on personality prediction has been on using textual data, which only addresses one method of prediction. Further work on the system should focus on increasing its precision. Similarly, an ensemble or hybrid model that incorporates many machine learning methods might yield superior results. Although though a lot of study has been done on personality prediction, there are still many open questions that need to be answered. Psychologists study a wide range of topics, but personality is one of the most comprehensive. In our investigation, we focused on the Big Five model of personality prediction, which adds up to five distinct types of individuals. Research on the best ways to categorise people based on their personalities using a battery of tests is ongoing. Real-time data with practical applications should also receive more attention. When multiple Deep Learning algorithms are used together, better results can be achieved. Eventually, a multidisciplinary prediction technique might be developed that considers a number of different biological data..

3. Conclusion

In conclusion, using machine learning for personality prediction in resume screening has the potential to drastically alter the current state of affairs in the recruitment industry. Machine learning algorithms may analyse text data from resumes to make predictions about an applicant's personality traits, which can then be used to evaluate their potential performance in a given role. Several research have shown that machine learning can successfully predict personality traits from text input, proving its usefulness for personality prediction. Nevertheless, privacy and ethical issues must also be addressed, and there is fear that these models may be biassed..

4. References

Danny Azucar, Davide Marengo, Michele Settanni, "Predicting the Big5 personality traits from digital footprints on social media: A meta- analysis, Personality and Individual Differences", pp: 150–159,
Computers in Human Behavior, 2018.
Tommy Tandera, Hendro, DerwinSuhartono, RiniWongso, and Yen LinaPrasetio, "Personality Prediction System from Facebook Users", 2nd International Conference on Computer Science and
Computational Intelligence, Bali, Indonesia pp.604-611, October 2017.
Ana Carolina. E.S.Lima, Leandro N.de Castro, "Predicting Temperament from Twitter Data", International Congress on Advanced Applied Informatics, September 2016.
Nadeem Ahmad, Jawaid Siddique, "Personality Assessment using Twitter Tweets",21st International Conference on Knowledge Based and Intelligent Information and Engineering Systems, Marseille,
France, pp.1964-1973, 2017.
M. Johnson Vioules,B , Moulahi,J. Aze, S. Bringay, "Detection ofsuicide-related

posts in Twitter data streams", IBM Journal of Research and Development ,vol. 62, Issue.1, vol.112, January 2018.

S. Arjaria, A. Shrivastav, A. S. Rathore and Vipin Tiwari, "PersonalityTrait IdentificationforWritten Texts Using MLNB",DATA, Engineering and Applications,pp.131-137, March 2019.

Soumen Roy,Utpal Roy,D. D. Sinha, "The probability of predicting personality traits by the way user types on touch screen",Vol. 15, Issue 1, pp. 27–34, March 2019.

Graciela Guerrero, Elvis Sarchi, and Freddy Tapia, Predict the Personality of Facebook Profiles Using Automatic Learning Techniques and BFI Test, New Knowledge in Information Systems and Technologies, pp 482-493, March 2019.

Heci Zheng, Chunhua Wu, "Predicting Personality Using Facebook Status Based onSemi-supervised Learning", 11th International Conference on Machine Learning and Computing, pp.59-64, February 2019.

Joel Philip, Dhvani Shah, Shashank Nayak, Saumik Pateland Yagnesh Devashrayee, "Machine Learning for PersonalityAnalysis Based on Big Five Model", Data Management, Analytics and Innovation, pp 345-355,September 2018.

BayuYudhaPratama, RiyanartoSarno, "Personality Classification Based on Twitter Text Using Naive Bayes, KNN and SVM",International Conference on Data and Software Engineering, Nov 2015

Feel: Mapping Emotion on Twitter", IEEE Journal of Biomedical and Health Informatics, vol. 19, pp.1246-1252, Issue: 4, February 2015.

AsadzadehLaleh, Rahimi Shahram, "Analyzing Facebook Activities for Personality Recognition", 16th IEEE International Conference on Machine Learning and Applications (ICMLA), December 2017.

AkkaponWongkoblap, Miguel A. Vadillo, Vasa Curcin, "A Multilevel Predictive Model for Detecting Social Network Users with Depression", 2018 IEEE International Conference on Healthcare

Informatics (ICHI), June 2018.

Jianguo Yu, Konstantin Markov, "Deep learning-based personality recognition from Facebook status updates", IEEE 8th International Conference on Awareness Science and Technology (iCAST),

November 2017.

Chaowei Li, Jiale Wan, Bo Wang, "Personality Prediction of SocialNetwork Users", 16th International Symposium on Distributed Computing and Applications to Business, Engineering and Science, October 2017.

Juan Abdon, Miranda-Correa and Ioannis Patras, "A Multi-task Cascaded Network for Prediction of Affect, Personality, Mood and Social Context Using EEG Signals", 13th IEEE International Conference on Automatic Face & Gesture Recognition, June 2018.

Yule Tyshchuk , William A. Wallace, "Modeling Human Behavior on Social Media in Response to Significant Events",IEEE Transactions on Computational Social Systems, vol. 5, pp.444-457, Issue: 2, April

2018.

Jian Zhao, Nan Cao, Zhen Wen, Yale Song, J. Watson, Yu Ru, Christopher Collins, "#FluxFlow: Visual Analysis of Anomalous Information Spreading on Social Media", IEEE Transactions on

Visualization and Computer Graphics, vol.20, pp.1773-1782, Issue: 12,November 2014.

Yingcai Wu,Nan Cao,David Gotz,Yap-Peng Tan,Daniel A. Keim, "A Survey on Visual Analytics of Social Media Data", IEEE Transactions on Multimedia, vol. 18, pp. 2135-2148, Issue: 11, September 2016.

Mark E. Larsen, Tjeerd W. Boonstra , Philip J. Batterham, BridianneO'Dea, Cecile Paris, HelenChristensensen, "We

Suja Panicker and Aditi Kunte, "Personality th

prediction using social media", IEEE 5 I2CT Pune ,29 March 2019.

Jose-Luis Cabra, DiegoMendez, Luis C. Trujillo, "Wide MachineLearning Algorithms Evaluation Applied to ECG Authentication and Gender Recognition", Amsterdam, Netherlands, ICBEA, pp.58-64, May 2018.

Veronica Ong1, Anneke D. S. Rahmanto1, Williem1, DerwinSuhartono, Aryo E. Nugroho, Esther W. Andangsari, Muhamad N. Suprayogi, "Personality Prediction Based on Twitter Information in Bahasa Indonesia",Proceedings of the Federated Conference on Computer Science and Information Systems, vol. 11,

September 2017.

https://www.kaggle.com.