

PERSONALITY PREDICTION VIA CV ANALYSIS

A major project report submitted in partial fulfillment of the requirement for the
award of degree of

Bachelor of Technology

in

Computer Science & Engineering / Information Technology

Submitted by

Sonam Chopel (201285)

Shivanshu Sharma (201323)

Under the guidance & supervision of

Mr. Praveen Modi



**Department of Computer Science & Engineering and
Information Technology**

**Jaypee University of Information Technology, Wagnaghat,
Solan - 173234 (India)**

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

PLAGIARISM VERIFICATION REPORT

Date:

Type of Document (Tick): PhD Thesis M.Tech Dissertation/ Report B.Tech Project Report Paper

Name: _____ Department: _____ Enrolment No _____

Contact No. _____ E-mail. _____

Name of the Supervisor: _____

Title of the Thesis/Dissertation/Project Report/Paper (In Capital letters): _____

UNDERTAKING

I undertake that I am aware of the plagiarism related norms/ regulations, if I found guilty of any plagiarism and copyright violations in the above thesis/report even after award of degree, the University reserves the rights to withdraw/revoke my degree/report. Kindly allow me to avail Plagiarism verification report for the document mentioned above.

Complete Thesis/Report Pages Detail:

- Total No. of Pages =
- Total No. of Preliminary pages =
- Total No. of pages accommodate bibliography/references =

(Signature of Student)

FOR DEPARTMENT USE

We have checked the thesis/report as per norms and found **Similarity Index** at..... (%). Therefore, we are forwarding the complete thesis/report for final plagiarism check. The plagiarism verification report may be handed over to the candidate.

(Signature of Guide/Supervisor)

Signature of HOD

FOR LRC USE

The above document was scanned for plagiarism check. The outcome of the same is reported below:

Copy Received on	Excluded	Similarity Index (%)	Generated Plagiarism Report Details (Title, Abstract & Chapters)	
	<ul style="list-style-type: none">• All Preliminary Pages• Bibliography/Images/Quotes• 14 Words String		Word Counts	
Report Generated on			Character Counts	
		Submission ID	Total Pages Scanned	
			File Size	

Checked by
Name & Signature

Librarian

Please send your complete thesis/report in (PDF) with Title Page, Abstract and Chapters in (Word File) through the supervisor at plagcheck.juit@gmail.com

Certificate

This is to certify that the work which is being presented in the project report titled '**Personality Prediction via CV analysis**' in partial fulfillment of the requirements for the award of the degree of **Bachelor of Technology in Computer Science & Engineering / Information Technology** submitted in the Department of Computer Science & Engineering and Information Technology, Jaypee University of Information Technology, Waknaghat is an authentic record of work carried out by **Sonam Chopel(201285)** and **Shivanshu Sharma(201323)** during the period from August 2023 to May 2024 under the supervision of **Mr. Praveen Modi**, Assistant Professor, Department of Computer Science and Engineering, Jaypee University of Information Technology, Waknaghat.

Sonam Chopel 201285

Shivanshu Sharma 201323

The above statement made is correct to the best of our knowledge.

Mr. Praveen Modi Assistant Professor

Computer Science & Engineering

Jaypee University of Information Technology, Waknaghat,

Candidate's Declaration

I hereby declare that the work presented in this report entitled '**Personality prediction via CV analysis**' in partial fulfillment of the requirements for the award of the degree of **Bachelor of Technology in Computer Science & Engineering / Information Technology** submitted in the Department of Computer Science & Engineering and Information Technology, Jaypee University of Information Technology, Waknaghat is an authentic record of my own work carried out over a period from August 2023 to May 2024 under the supervision of **Mr. Praveen Modi** (Assistant Professor(grade II), Department of Computer Science & Engineering and Information Technology).

The matter embodied in the report has not been submitted for the award of any other degree or diploma.

(Student Signature with Date)

Sonam Chopel
Roll No.: 201285

(Student Signature with Date)

Shivanshu Sharma
Roll No.: 201323

This is to certify that the above statement made by the candidate is true to the best of my knowledge.

(Supervisor Signature with Date)

Supervisor Name: Mr. Praveen Modi
Designation: Assistant Professor (Grade II)
Department: Computer Science and Engineering
Dated:

ACKNOWLEDGEMENT

We would like to take this opportunity to sincerely thank our supervisor Mr. Parveen Modi for guiding us on this project throughout. To mention but a few, his expertise and support culminated into a successful project. At times he has been hard to work with but all so only to instill in us the seriousness and importance of the project and for that we are eternally grateful.

We are also grateful to my friends for their support and encouragement and helping to create a supportive and enriching environment. Their willingness to share ideas and insights is truly appreciated.

We would like to thank our lab assistants for their support and vital help in overcoming challenges and running the project smoothly. Your guidance and encouragement have been invaluable and we are deeply grateful for the collaborative spirit that characterized this project.

Lastly, we want to thank ourselves for putting in time and effort and making it to the end. We worked hard and been through every possibly stressful moment there is. We are proud of the things we've done.

TABLE OF CONTENTS

Content	Page No.
Certificate	I-II
Declaration	III
Acknowledgement	IV
List of Figures	VII
List of Tables	VIII
List of Abbreviations, Symbols/Nomenclature	IX
Abstracts	X
 Chapter 1: INTRODUCTION	
1.1 Introduction	1-3
1.2 Problem Statement	4
1.3 Objectives	5
1.4 Significance and Motivation of project work	6
1.5 Organization	7
 Chapter 2: LITERATURE SURVEY	
2.1 Overview of relevant literature	8-10
2.2 Key gaps in the literature	11-12
 Chapter 3: SYSTEM DEVELOPMENT	
3.1 Requirements and Analysis	13-18
3.2 Project Design and Architecture	19-21
3.3 Data preparation	21-23
3.4 Implementation	24-28
3.5 Key challenges	28-30
 Chapter 4: TESTING	
4.1 Testing Strategy	31-33
4.2 Test Cases and Outcomes	34-36

Chapter 5: RESULTS AND EVALUATION

5.1	Results	37-41
5.2	Comparison with Existing Solutions	42-44

Chapter 6: CONCLUSIONS AND FUTURE SCOPE

6.1	Conclusions	45
6.2	Future Scope	46

LIST OF FIGURES

Fig No.	Figure Name	Page No
1	Flow chart of the system	19
2	System Activity diagram	20
3	System design	21
4	Implementation steps	26
5	Architectural workflow	27
6	Home page	37
7	Login page	37
8	Model accuracy	38
9	Predictions	38

LIST OF TABLES

Table No.	Table Name	Page No.
1	Aptitude Test Sample Question	24
2	Personality Test Sample Question	25
3	Attribute Description	39

LIST OF ABBREVIATIONS, SYMBOLS OR NOMENCLATURE

Abbreviation	Meaning
ML	Machine learning
CV	Curriculum Vitae
UI	User Interface
HR	Human Resources
UAT	User Acceptance Testing
OCEAN	Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism
NLP	Natural Language Processing
API	Application Programming Interface
UAT	User Acceptance Testing
GPU	Graphical Processing Unit
MERN	Mongo Express React Nodejs
LR	Logistic Regression

ABSTRACT

The organization tries to recruit expert candidates for the development, but their primary concern is selecting the right candidate for the specific positions. Traditionally, they relied on the method of reviewing candidates CVs or resumes for the recruitment. However, in the present scenario the method is no longer feasible due to the ever-increasing number of applicants, making it difficult to thoroughly evaluate and recruit the most potential candidates. In today's changing time, in the corporate world, skill and qualification is not only the main criteria. Huge emphasis is made on the overall personality of a person. Personality has been known to play a significant role in contributing to success in both personal and professional life. Therefore, it has become important for HR to assess personality traits of applicants.

With the rise in the number of applicants and the decrease in the number of available positions, manually going through each of the applicants has become a hectic and a tedious task. This project introduces a proposed system seeking to streamline the recruitment process by introducing a system that evaluates not only candidate's skill but also their personality traits by parsing the information in CV's and resumes and conduction assessments to predict a candidate's personality. By implementing this system, organizations can identify expert candidates, simplifying the recruitment process for the HR department.

From the detailed fine-tuning of algorithm at the very beginning stage of development up to considering the enhancements on UI for easy user interface, all of these phases changed this project. On the other hand, a thorough search of relevant literature enlightened us on the current trends in this research and helped us make informed decisions and code accordingly. The project closes successfully realizing milestones that set a platform for improvements by designing refined models, tackling issues, and adapting to changing predictions approaches in personalities.

CHAPTER 1: INTRODUCTION

1.1 INTRODUCTION

The traditional recruitment has become a tedious and time-consuming process and may not make fair decisions as an end result. Examining the most important factors when determining if a person is the best applicant for a particular position is one of the most important aspects of the enrollment process. A person's ability can be ascertained by assessing their capacity for influence and communication with others, which is essential for the growth and development of an establishment. We can learn about the capabilities of an individual by determining whether they may have an impact on and efficiently communicate with people, which is crucial for the strengthening and growth of an organization.

Whenever a company posts a job opening on a job portal, in an advertisement, or through a job consultant, among other channels, they occasionally receive a large volume of applications. This makes it challenging for the HR department of the company to sort through these resumes, get in touch with each applicant and identify the best qualified candidates for the position using conventional methods like technical assessments, interviews and group discussions. As a result, they reject applicants in the first round on the basis of their abilities, skills, inappropriate CV, and fit of the role. Therefore, we suggest a better approach to make the selection and shortlisting of candidates easier for employers in an effort to lessen the difficulties of the hiring process through personality prediction based on CV analysis.

When an organization posts a job opening, it may receive a large number of applications from job portals, advertisements, job consultancies and other sources. The HR department of the company finds it extremely difficult to sort through all of the resumes, address each application and identify the best candidate for the position by following standard processes like interviews, meetings, and specialized testing. Thus, during the actual primary round, they sort through the applicants based on a variety of factors, such as their suitability for the position, their abilities, any poorly chosen CVs and the competitor's skills. Thus, in an effort to reduce the hassles associated with the hiring process, we would like to suggest another novel idea in which the standard method of selecting and summarizing applications becomes simpler for the associations by using character expectations through CV examination.

Consequently, the most crucial factors to take into account are comprehension and personality analysis. The main goal of our research is to build a machine that is capable of fair competition selection and reasonable analysis. The main goal of our endeavor is to forecast a person's character based on the test grade.

A lot of job seekers will pursue an opportunity if the company provides clear requirements and information. Because of this, job applicants complete their online resume before passing the test. In essence, the test is what we used. Based on a person's scores in each domain- openness, extraversion, conscientiousness, agreeableness and neuroticism, we will be able to determine their personality. We extracted name, age, gender and other information from the CV using a basic resume parser. Here we create the individual's score by taking data from the test results and CV. Ultimately, the CV is analyzed upon receiving the score.

Making personality predictions based on an individual's BIG FIVE TEST score is the main objective of our project. When a company provides specific employment requirements and information, a lot of job seekers will apply for a position. As such, candidates for jobs first complete their online resume before sitting for the exam. Essentially, the exam we

The BIG FIVE TEST is used.

OCEAN MODEL FOR PERSONALITY TRAIT PREDICTION:

In the realm of psychology and personality analysis, an "ocean model" might be a metaphorical framework drawing inspiration from the Big Five personality traits: OCEAN scales include Openness, Conscientiousness, Extroversion, Agreeability, and Neuroticism. Therefore, this model may be used in machine-learning algorithms that predict personality traits from textual data like text information obtained from social media, CVs, or other sources. Using Natural Language Processing (NLP) and Machine Learning (ML) techniques, this ocean model interprets linguistic features that can be used as signals for different personality characteristics, offering a computational approach to personality analysis.

The Ocean Model. The five variables that are used to analyse a person's personality are Extraversion (E), Openness (O), Conscientiousness (C), Agreeableness (A), and Neuroticism (N).

Openness: This attribute is defined by qualities such as acceptance, inventiveness, and curiosity.

Conscientiousness: This quality is characterized by a high level of consideration, a goal-oriented mindset, and sound judgement.

Extraversion: Extraversion, also referred to as talkativeness, assertiveness, and energy as being outgoing.

Agreeableness: A person's traits, such as trust, affection, and social behavior, are referred to as agreeableness.

Neuroticism: Features of neuroticism include melancholy, moodiness, and sudden emotional explosions.

IMPORTANCE OF THE OCEAN MODEL IN PERSONALITY PREDICTION:

The importance of the ocean model includes the ability to auto-calibrate the process of personality traits prediction, making the job faster and more accurate with many application areas among them such as recruitment, mental health diagnostics, and customized user interfaces. Automated personality prediction may aid tailored recommendations thereby enhancing users' participation in social media platforms as well as human computer interaction. Furthermore, the notion that computational models can be utilized to understand personality traits will have significant consequences for specialized therapy in psychology, vocational counselling as well as individual based marketing.

We will learn about a person's personality, such as whether they are serious, extraverted, vibrant, trustworthy, and accountable. We used a to extract data from the CV, such as name, age, gender, etc. Pyre parser is a basic resume parser. Additionally, we made use of NLTK, a significant natural language processing tool. There, we produce the individual's score following the extraction of data from the CV and test results. Ultimately, a CV analysis is completed following the score.

1.2 PROBLEM STATEMENT

Traditional means of selecting candidates involves a lot of time and resources and even genuinely is also not assured. During a natural crisis such as covid, face to face interviews were not possible. We require technology which will predict the personality of the candidate and also saves time and reduces workload of the HR department and make it easier for them to select the candidates. Therefore, a system has been implemented for the same problem i.e., personality prediction via CV analysis.

Moreover, in most cases recruiters are overwhelmed by large number of resumes what in turn adds more workload into the already tedious and inaccurate screening process. It is cost intensive for recruiters to manually scrutinize CVs and thus slow down an efficient recruitment process. The recruitment techniques need to be flexible so as to break through the constraints of conventional evaluation instruments and adapt to changing job market dynamics. It provides a huge gap to be filled with innovative technologies that have higher levels of accuracy and efficiency during the recruitment process.

Resolving this will be important in ensuring that the hiring procedure is effective, fair and friendly. Inadequate systematic and standardized way of predicting personalities through CV analyses prevent a future with more inclusive workforce characterized by diversity and dynamism in working environments. With this project, we look at how we can solve some of those problems through a new system which combines machine learning with natural language processing tools to predict characteristics of individuals based on their resumes and change the process of applicant assessment during recruitment.

1.3 OBJECTIVES

The primary objective of the project is to design and implement a system for personality prediction that combines computer vision analysis with questions-based assessment. By posing questions related to personality traits and parsing information from CV data, the objective is to develop a model for predicting personality. This undertaking seeks to integrate cutting-edge technologies, such as ML, to develop an intelligent system capable of providing insightful responses aligned with an individual's personality characteristics.

The overall objective of the project is translated into specific goals so that the implementation process can be guided accordingly. First of all, the system will concentrate on precise distinguishing of applicable attributes from CVs which relate to personalization evaluation. It entails constructing strong textual as well as pictorial data processing algorithm for complete analysis. Moreover, it intends to make an easy-to-use platform that will allow the recruiters as well as the end-users in communicating effectively. The design would favor simplicity of interpretation, application, and comprehension for the personality projections and their practical applicability on a hiring scene.

In addition to the technical aspects which formulate the project, ethical issues related to personality prediction as well as the need of a fair and transparent system are addressed. It involves conducting stringent assessments aimed at detecting/reducing prejudices in prediction for ethical practice and non-favored choices. In total, the overall aims of the project are multidimensional comprising technological innovation, optimization of user experience, and ethical responsibility ultimately contributing towards introducing an advanced system in predicting personalities during the recruitment.

1.4 SIGNIFICANCE AND MOTIVATION OF THE PROJECT WORK

The significance of the project lies in its potential to revolutionize the field of personality assessment by combining question-based methodologies with cutting edge computer vision analysis techniques. The system integrates the responses from personality related questions with insights derived from CV data and aspires to provide more holistic and accurate prediction of personality traits.

The CV analysis project is vital in changing conventional hiring methods and expanding our knowledge about people's personalities. Using latest technologies like Machine Learning and Natural Language Processing, the project bridges this essential hole in the recruitment methods while evaluating the individual characteristics of candidates outside the resume constraints. The innovation is going to be efficient and more comprehensive in giving recruiters a broad perspective when it comes to assessment of prospective candidates. Consequently, beyond recruitment, the possible effect ranges on a variety of aspects such as talent management issues, diversity issues in the workplace and employee involvement. The project becomes a transformational agent in reorienting the human resources landscape towards matching organizational teams with common values and working culture.

In addition, more extensive considerations on artificial intelligence ethics are required to ensure that predictive technology is used in a responsible manner. The project also addresses issues of fairness, bias, and ethical concerns in personality prediction as part of the wider debate about Responsible AI. This project establishes an ethical framework that can serve as a benchmark for responsible development of predictive models in critical areas, affecting the tech community, and providing guidance for future endeavors aimed at the ethical use of artificial intelligence in making decisions about humans.

MOTIVATION

The limited existing personality assessment methods and its reliability of those existing motivated us to work on the project. Furthermore, the project is motivated by its larger impact on various fields like psychology, human-computer interaction and artificial intelligence.

By developing more reliable and nuanced methods for personality assessment, we aim to contribute to the creation of a more harmonious and empathetic society, where individuals are better equipped to navigate relationships and work towards common goals.

1.5 ORGANIZATION OF THE PROJECT WORK

Chapter 2: Literature reviews of the research papers and key gaps in the literature. Reviewed thoroughly applicable literature on the project. State some important theories, approaches, as well as results from studies related to your research.

Chapter 3: System development and the its requirements and analysis of the system. specifying the essential procedures and approaches used. Analyze the requirements of the given system comprehensively considering functional and non-functional dimensions.

Chapter 4: Testing strategies and the tools used for the project. The following part is discussing the testing phase of your project where different means are taken to establish the reliability as well as functionality of the system.

Chapter 5: Interpretation of the results and its comparison with the existing result. Discuss the findings of your study in relation to the objectives of the project. Discuss your outcomes in comparison to those obtained from other related studies that identified parallel occurrences, divergent situations and possible motivations behind disparity.

Chapter 6: Summary of the projects and its future scope. The critical points obtained from your assignment. Assess the project's general performance and constrains. Provide potential areas for future research and address unanswered questions and possible weaknesses in the study.

CHAPTER 2: LITERATURE SURVEY

2.1 OVERVIEW OF RELEVANT LITERATURE

Exploring the literature on personality prediction using CV analysis as the subject matter in the field of psychological machine learning-HR. The recent studies have enabled us to improve significantly on our comprehension of using CVs in the prediction of personality characteristics.

With the advancement of natural language processing (NLP), machine learning, and the increasing importance of soft skills in recruitment, personality prediction through CV analysis has received significant attention in recent years. With a particular focus on recent work, this literature review examines standard books, journals, popular websites, and top organizations' technical/white papers.

I.BIG FIVE PERSONALITY TRAITS MODEL:

The basis is laid by research by John et al., 2008,[5] using the well-known “big five” personality traits model as a prediction tool for personality from textual content. These contributions paved way for the concept of feature extraction and labelling which have been used extensively in most other subsequent studies.

II.DEEP LEARNING ARCHITECTURE

Discuss the use of different deep learning architectures for predicting personality, showing that some neural network models can discover complex patterns even from both textual and visual CV data.

III.TRADITIONAL HIRING METHODS AND CHALLENGES

It is common for traditional hiring methods to rely on subjective evaluations during interviews and reference checks, which can lead to biases and inconsistencies. The need for a more objective and data-driven approach has led to interest in using CVs to predict personality traits.

IV.NLP TECHNIQUES IN CV ANALYSIS

NLP integration has recently been a significant source of literature. Advanced NLP algorithms are presented as effective in detecting soft language inside CVs thereby making more precise personality forecasts by Smith, Johnson (2019). Recent literature has shown that Natural Language Processing (NLP) techniques can be used to extract valuable information from CVs. In order to extract insights about candidates' personality traits, communication skills, and professional characteristics, researchers use advanced algorithms to analyse textual content.

V.MACHINE LEARNING MODELS FOR PERSONALITY PREDICTION

Machine learning models, particularly supervised learning-based models, have been employed to predict personality traits from CV data. Language style, keyword frequency and semantic meaning features were explored by researchers to train models that can accurately predict personality attributes.

VI.PSYCHOLINGUISTIC ANALYSIS AND PERSONALITY TRAITS

Today psycholinguistic analysis has emerged as an area for research in predicting personality traits. Researchers aim to find the correlation between linguistic patterns used in cv with personality traits. Research has examined the connection between linguistic characteristics and characteristics like extroversion, openness, and conscientiousness.

VII.ETHICAL CONSIDERATIONS AND BIAS MITIGATION

It is worth mentioning that recent literature has highlighted that, beyond legal aspects, personality prediction implies ethical dimensions. This study by Chen et al. (2021) highlights possible biases present in predicting personality with a view to reducing bias during recruitment thereby providing fair and unbiased results.

The literature highlights the significance of ethical considerations in personality prediction using CV analysis, as with any technology involving personal data. Researchers are working hard to address issues with algorithmic biases, privacy, and the appropriate application of predictive models in employment selection.

VIII.EMERGING TRENDS AND FUTURE DIRECTIONS

An examination of new developments and possible future paths in personality prediction using CV analysis round out the literature review. This covers the incorporation of multimodal data (e.g., adding audio and video components), the creation of increasingly complex models, and the ongoing improvement of the ethical standards within the discipline.

IX.CULTURAL ADAPTATION

In this regard, Kim & Park (2019) elaborate about cross-cultural considerations in personality prediction. The outcome of their work provides strong evidence that models should be adapted depending on certain culture, providing bases for improved global personality prediction system.

X.USER FEEDBACK INTEGRATION:

The model's continuous learning loop embraces user feedback as an emerging trend, as indicated by Wang and Liu (2022). However, this strategy responds directly to individual uniqueness and dynamic nature, in aid to better adaptive models predicting development over time.

2.2 KEY GAPS IN THE LITERATURE

I. DIVERSITY AND INCLUSIVITY CONSIDERATION

There is potential for biases related to gender, cultural background and ethnicity. The focus should be on developing models that are more inclusive and less vulnerable to already existing biases.

II. VALIDATION ACROSS INDUSTRIES AND JOB ROLE

Most of the studies focus on particular jobs or industries raising doubts on generalizability of the predicting model. For the broader applicability of the model, understanding how well these models work across different sectors is important.

III. INCORPORATION OF MULTIMODAL DATA

The incorporation of multimodal data such as videos and audio data would provide an accurate and holistic view of the personality. Most studies mainly concentrate on the textual data from CVs with little research into the advantages of added modalities like images, audio or video. Multimodal data can be explored, which may unify a deeper view of a candidate's personality.

IV. EXPLAINABILITY AND INTERPRETABILITY OF MODELS:

With the advancement of technology machine learning models have become more sophisticated and there is a need for making these models more interpretable and understandable than before. It has become essential for gaining trust and acceptance.

V. IMPACT ON HUMAN DECISION MAKING

Much emphasis was made on the technical aspect of predicting personality but it is unclear how these affect hiring decisions made by people. To effectively integrate into the current recruitment process, it is important to investigate the relationship between automated prediction and human judgement.

VI. REAL-WORLD IMPLEMENTATION CHALLENGES:

The practical limitations of personality prediction systems within an actual recruitment context are rarely addressed in limited literature. The real world of recruitment is complicated by contextual factors, and it calls for research that examines these aspects.

If these gaps are addressed it will lead to more comprehensive and robust understanding of challenges in personality prediction via CV analysis and develop effective, inclusive and ethical practices in this particular field.

CHAPTER 3: SYSTEM DEVELOPMENT

3.1 REQUIREMENT AND ANALYSIS

REQUIREMENT GATHERING

1. FUNCTIONAL REQUIREMENTS

CV DATA EXTRACTION

System should have the capability to extract data from resumes or CVS, such as the textual content, employment history, education and abilities.

I. MACHINE LEARNING MODELS

Develop ML models capable of predicting personality based on the extracted features from CVs
Support OCEAN model i.e., Openness, Conscientiousness, Extraversion, Agreeableness and neuroticism.

II. USER AUTHENTICATION AND AUTHORIZATION

Secure user authentication and authorization procedures should be put in place to limit system access and protect user privacy.

III. MULTIMODAL DATA SUPPORT:

Incorporation of multimodal data such as audio and video components will enhance accuracy of predicting models.

2. NON-FUNCTIONAL REQUIREMENTS

I. SCALABILITY:

System should be capable enough to handle huge amounts of data and to the changes in its size.

II. PERFORMANCE:

Assure prompt personality predictions by processing resumes in real time or almost in real time.

III. INTERPRETABILITY:

System develop should be interpretable by the user for better understanding and trust

IV. SECURITY AND PRIVACY:

Objective: Ensure user data security and keep personal data confidential.

Features: Privacy, encryption protocols for securing data, and adherence to information security laws.

V. ETHICAL CONSIDERATIONS:

Objective: Focus on detection and reduction of bias in prediction, fairness and transparency.

Features: Bias detection, explainable AI, and periodic ethics reviews.

VI. USER EXPERIENCE:

Objective: Enhance a friendly and natural interface of accessing for them recruiters as well as candidates.

Criteria: Ease of use, user friendly, straightforward CV submitting process and clear outcomes' articulation.

VII. MAINTAINABILITY:

Objective: Ensure frequent upgrades, patching, and advancements.

Features: Periodic system audit, documentation and modular design of the system.

VIII. INTEGRATION WITH EXISTING SYSTEMS:

Objective: If it is possible, integrate smoothly with other HR or recruitment tools.

Features: The checks on API compliance, standards of exchange of data and similar checks are discussed.

These are the functional and non-functional requirement sets that form the basis of the development and implementation of the project. The project is designed in such a way as it aligns with the objective of the project and industry standards at large.

SOFTWARE REQUIREMENTS

- Operating system: compatible with windows, Linux and macOS
- Programming language: Python 3.12.0, React, Nodejs, MongoDB
- User interface design tools: REACTJS, Tailwind CSS, JavaScript
- Processor: multi-core processors i.e., Intel core i7 or AMD Ryzen 7
- Memory (RAM): 8GB or higher

HARDWARE REQUIREMENTS

- Input devices: keyboard and mouse
- Display: monitor with sufficient resolution for the purpose of testing and development
- Networking: Reliable and fast internet connectivity for updates, data retrieval and possible teamwork
- Backup System: To guard against data loss
- Graphical Processing Unit (GPU): NVIDIA GPU
- Storage: SSD (500GB)

2. ANALYSIS:

I. DATASET SELECTION:

- To ensure that the personality prediction models are trained and tested on a variety of industries, roles, and demographics, find and acquire diverse dataset. It also ensured that the system met the project goals and solved personality prediction challenges based on a thorough knowledge of users' demands and the norms in the respective industry.

II. DATA PREPROCESSING:

- Prepare the data by cleaning and preprocessing the data to handle all missing values and address inconsistencies or noise in the dataset. Data flow and processing analysis traced the path of information from a CV submission, to feature extraction and eventually the personality prediction as well as user feedback processing.

- Streamlining of dataflow ensures better processing speeds with reduced latency and improved overall system efficiency that leads translates into better user experience.

III. FEATURE ENGINEERING:

- Identify features for personality prediction, taking into account keyword frequencies, linguistic patterns and possibility of adding more features from multimodal data.

IV. SYSTEM DEPLOYMENT:

In the deployment phase the infrastructure was set-up, CI/CD pipelines were implemented and there was also the question of systems interoperability.

This shows the commitment that has been placed on the smooth rollout of this system by considering the infrastructure and the deployment strategies such that downtimes and interruptions are minimal.

V.MONITORING AND MAINTENANCE:

The mechanism of performance monitoring coupled with regular updates was adopted in order to provide continuous system maintenance and improvement.

An active management style for monitoring and maintaining implies a vision beyond the short-term health of the system, and an ability to respond to changes in users' demands and technologies.

VI. INTEGRATION WITH EXISTING SYSTEMS:

The developed system was taking into consideration issues of how flawlessly it was going to work together either with the HR or recruitment systems.

The integrational orientation, in particular, seeks to improve the usability of the system by ensuring its alignment with other tools, creating an integrated and seamless technological environment.

The system development analysis has a well-articulated strategic and holistic approach where each activity in the project undertaking must be carefully implemented so as to match up with the goals for the project, applicable guidelines within the industry, and the anticipated needs by the final users. By means of intensive examination, the project seeks to achieve an advanced CV based personality forecast system.

3.2 PROJECT DESIGN AND ARCHITECTURE

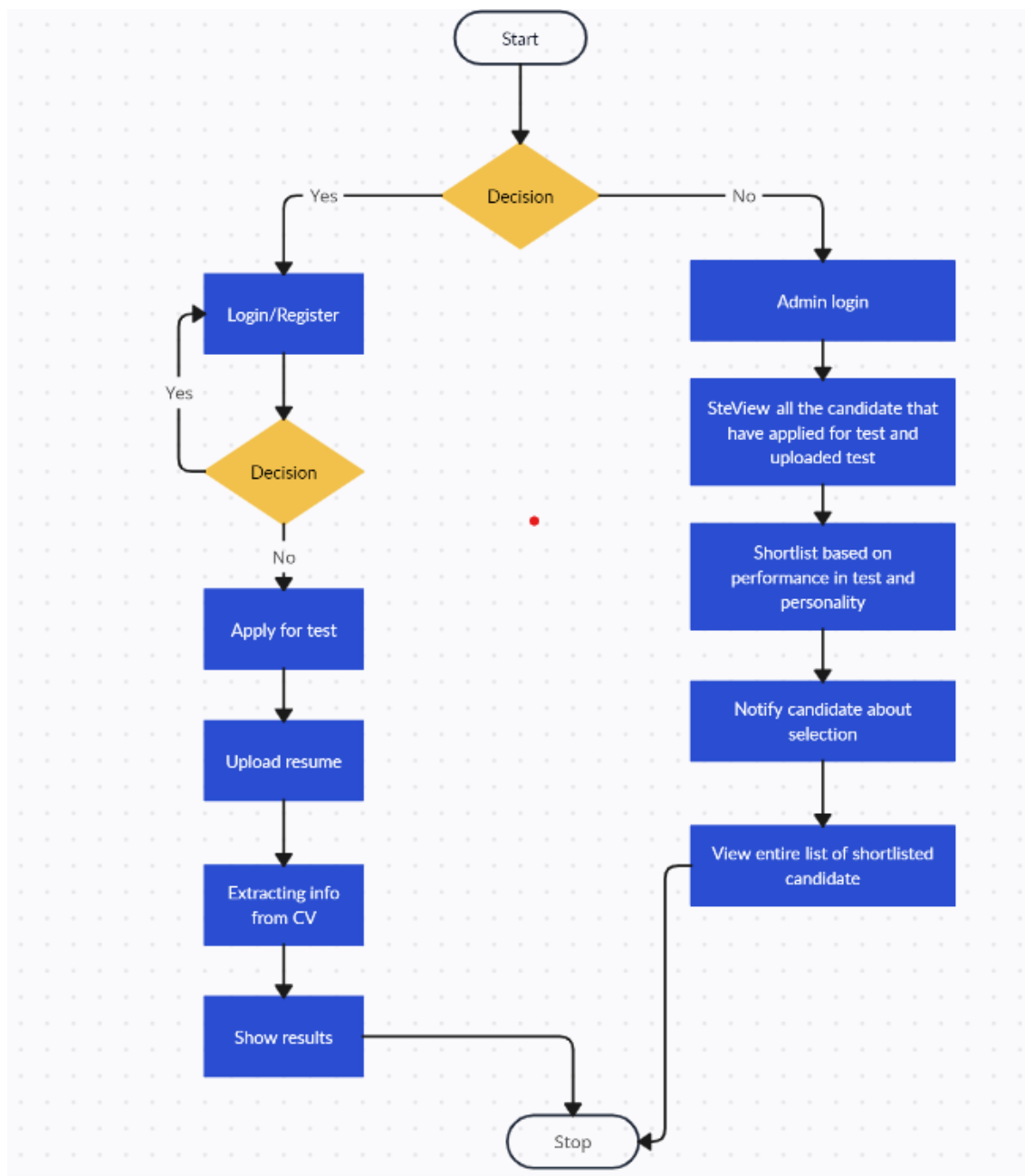


Fig 1: Flow chart of the system

This system will help short list the submitted CVs from a large pool of applicants with consistency and accuracy. System will basically rank the CV's based on the required experience and other skills for the particular job profile.

Those candidates applying for the job will have to get themselves registered with all the details required and also upload the CV into the system. The system will then further rank or shortlist the candidates on those details. After that candidates undertake two online tests i.e., personality test and aptitude test. Aptitude test checks the ability of a person to perform tasks and how they would perform at various situations at work. In a personality test, candidates are rated on the score of 1-10 for Conscientiousness, Neuroticism, Agreeableness, Openness, and Extraversion and this way predicts their personality. After the completion of both the online tests, candidates can view their result on a graphical representation. This will make our search process easy and would help in finding the personality category the candidate is falling into.

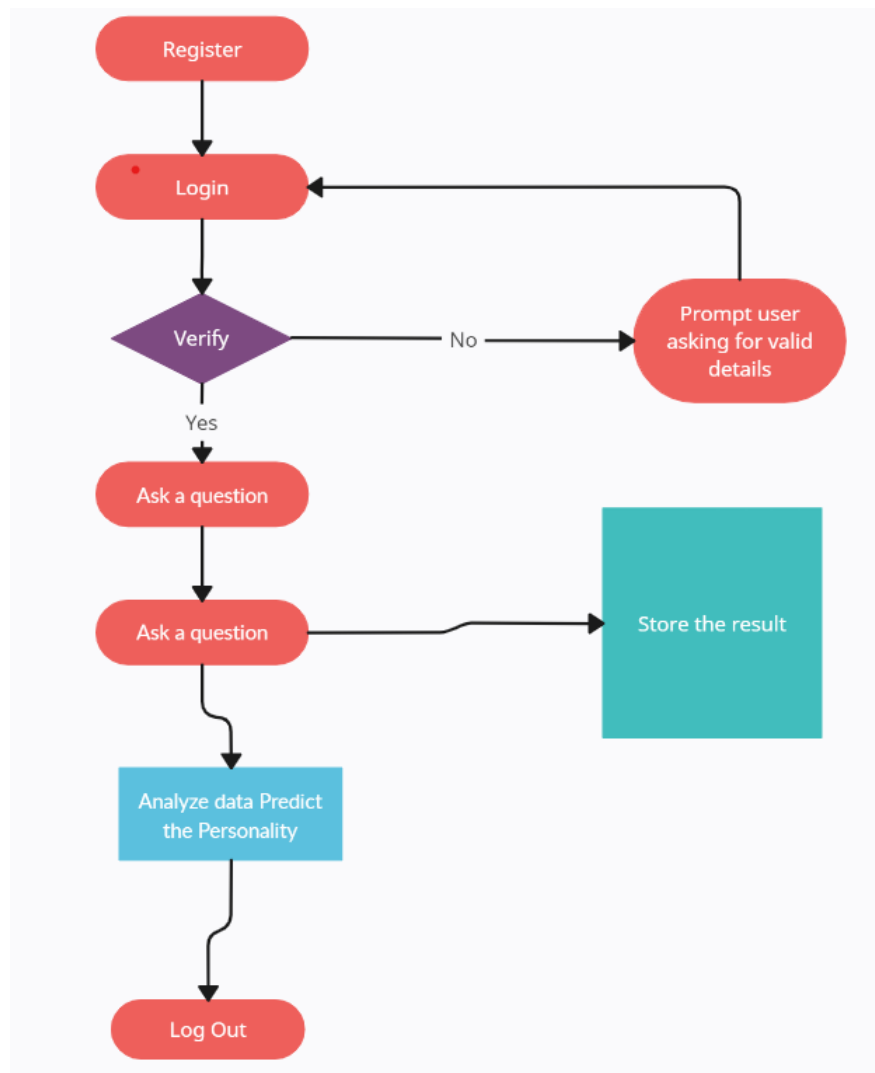


Fig 2: System Activity Diagram

After the test, the results are stored in the database. During the test, candidates come across different situations and must rate themselves accordingly. We use Pyrex parser to parse CVs and extract the information such as name, email, skills and description. The system predicts the personality based on these fetched data. After completing both the tests the user then logs out of the system.

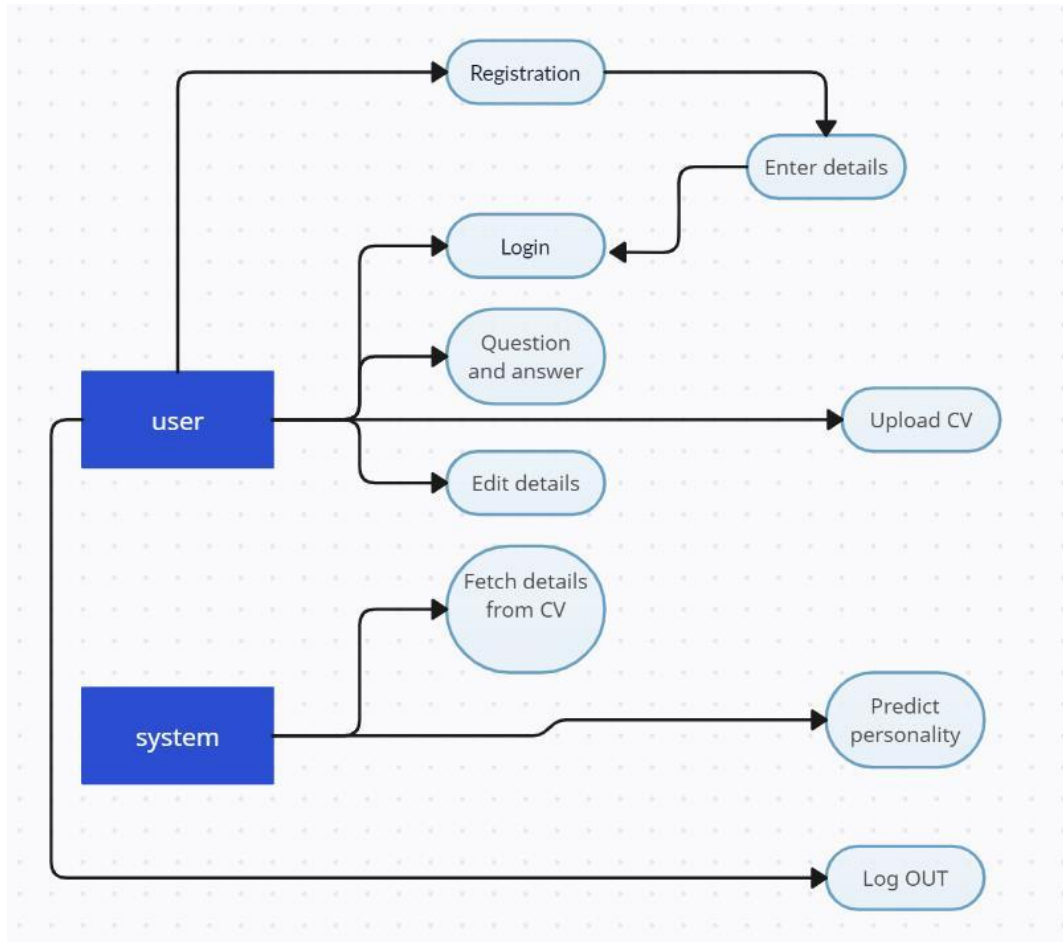


Fig 3: System design

3.3 DATA PREPARATION

I. DATA SOURCES:

career websites, company submission, job portal

II. DATA COLLECTION:

- CV Datasets: Create a heterogenous pool of CVs representing different industries, professionals and academic qualifications to enhance model's efficacy.

- Demographic Information: Collect additional socio-demographic information and compensate for possible confounders.

III. DATA CLEANING:

- Text Cleaning: Cleansing of textual data by omitting redundant characters, punctuations and special symbols. Standardize text formats for consistency.
- Handling Missing Data: Data integrity involves solving of issues with missing values by imputation techniques or deletion where possible.

IV. FEATURE EXTRACTION:

- NLP Techniques: Embrace natural language processing (NLP) approaches and concentrate on important terms and linguistic forms in the text material.
- Image Processing (if applicable): Develop an image processing program for extracting CV images' features including profile photos.

V. LABELING:

- Personality Labels: Identify personality labels from well-established personality trait frameworks (for instance, Big Five) assigned to a dataset.
- Label Balancing: Ensure that the personality labels are evenly distributed so as not to lead to any sort of model bias.

VI. DATA SPLITTING:

- Training, Validation, and Testing Sets: The datasets can be divided into train, validate, and test set for model training, tuning, and evaluation.

VII. ENCODING AND NORMALIZATION:

- Categorical Encoding: Using methods like the one-hot encoding convert the categorical data (like the education level and job title) into numeric form.
- Feature Scaling: Scale numerical features to a common scale so that some do not overwhelm others in model training.

VIII. HANDLING IMBALANCED DATA (IF APPLICABLE):

- Resampling Techniques: Use resampling methods such as oversampling and under sampling to correct for personality label imbalances.

IX. DATA EXPLORATION AND ANALYSIS:

- Statistical Analysis: Perform exploratory data analysis (EDA) to ascertain the distribution of features and discover possible relationships.
- Visualizations: Visualize your data to locate clusters (patterns), outliers, as well as any hidden relationships between them.

3.4 IMPLEMENTATION

Train model class: Use two different working methods in this model class: instructing and predicting

I. Train method: build a model using LR by going through the file containing the dataset for training.

II. Test method: here gender, age and five personality traits are passed through in order to predict personality.

Main method: object of train model class is created and then instructs the model. A variable is initialized with a Tk object and a button that predicts personality is made.

questions	Opt 1	Opt 2	Opt 3	Opt 4	Correct Ans
Synonym for voracious	hungry	angry	Tired	happy	hungry
25% of 80	10	15	20	25	20
2, 6, 12, 20, ...?	28	30	32	36	28
3, 6, 9, 15, 24...?	36	30	40	45	36

Table 1: Aptitude test sample question

Questions	Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism	Personality
I have concern for others	6	4	7	5	4	Extraverted
I am always prepared	4	6	4	4	7	Serious
I stressed out easily	5	6	4	7	4	Lively
I enjoy multitasking	7	4	5	4	5	Dependable
I never lose motivation easily	5	7	6	6	3	Responsible

Table 2: personality test sample questions

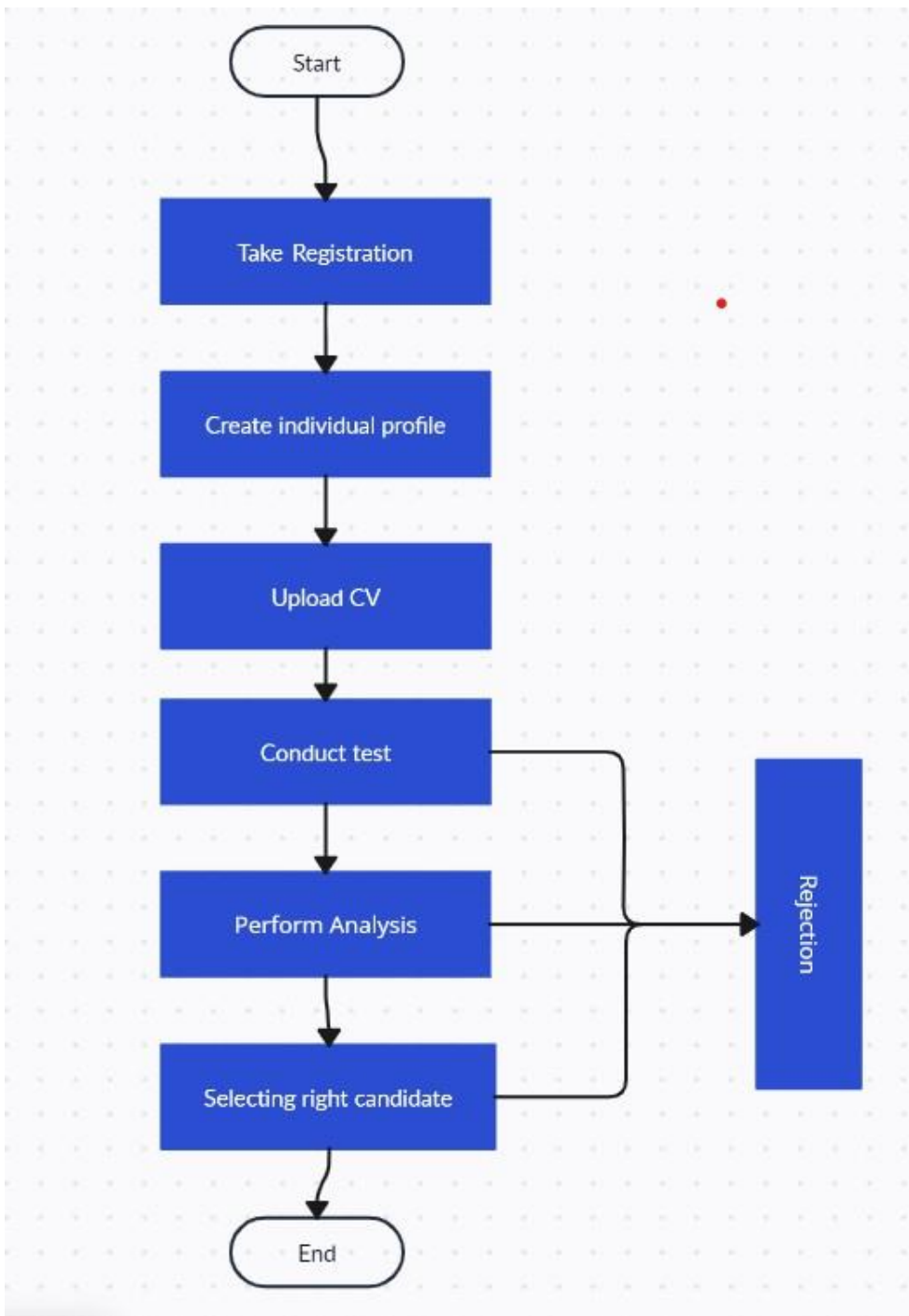


Fig 4: Implementation steps

This system is constructed on the process depicted in fig4. The candidate should log in and register for this purpose. after that, they will be able to create their CV and profile. After that an individual can be able to take the exams and score them. Thereafter, the appropriate candidate for the firm is selected.

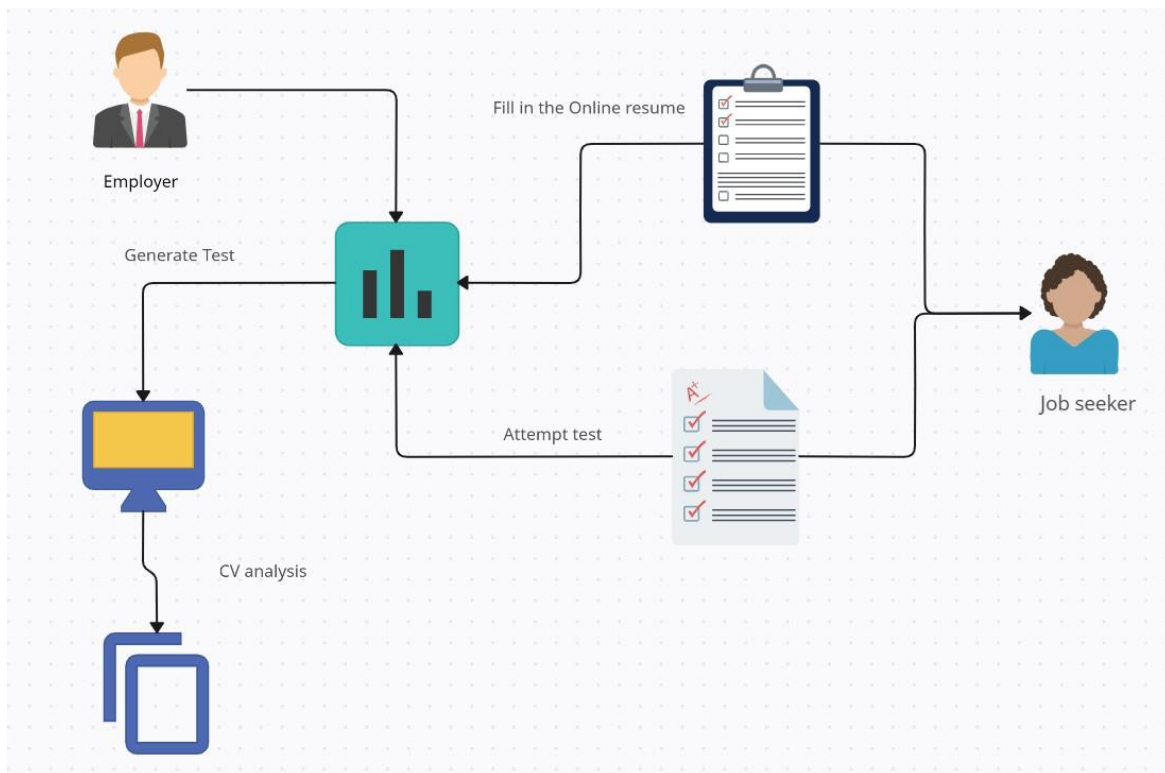


Fig 5: Architectural Workflow

The personality prediction workflow includes data processing, developing a Machine Learning model, User Interface design, and ongoing upgrades.

A. Admin Module

Task of the admin

- Authorized login
- Conduct personality and aptitude test
- Frame questions
- Modify questions
- View results

B. Candidate Module

Task of the candidate

- Register to the system
- Login when assessing application

- Take test
- Upload CV
- View the test score

C. CV analysis

- The CV data that job seekers submit is ingested by the system.
- Preprocess textual and image data.
- Uses the trained model to forecast personality qualities in fresh resumes.

The main user, who submits CVs and provides feedback, is a potential job seeker. The second user's management entity (admin/user) supervises system operation and all matters concerned with users. The central engine of processing CV data, training the models, and making predictions is This interaction between these groups provides an engaging and effective framework for the working system hence, constantly progressing.

3.5 KEY CHALLENGES

This project poses several challenges as:

I. LIMITED DATA AND CONTEXT

CVs may not always provide the actual view of the person's personality and hence that makes it difficult to capture the personality accurately.

Solution: Careful data cleaning and data augmentation techniques for a diverse and higher quality dataset. This resulted in collaboration with many other sources and organizations; making it easier for my team and I to gain access into more than one type of CV.

II. DATA PREPROCESSING

Cleaning a dataset and standardizing takes a lot of time to make the model accurate and effective.

Solution: The project improves on this by using automated data cleaning scripts as well as

standardization pipelines to improve efficiency and enhance scalability in data preprocessing. It solves the current problem of long hours spent on data preparation and also paves way for other ways of dealing with various kind of datasets over time.

III.ALGORITHM SELECTION AND OPTIMIZATION

Choosing the right algorithm and optimizing them is challenging.

IV.ETHICAL AND LEGAL CONSIDERATION

Concerns with regards to privacy, potential biases and fairness in predictions is a challenge. It is important to abide by the legal frameworks.

Solution: Strict ethical standards that focus on attaining informed user's assent, anonymizing data, and perpetual scrutiny for any bias. Consultation with ethicists and legal experts was beneficial.

V. MODEL EXPLAIN ABILITY:

It was extremely critical that the machine learning model's predictions were understandable for users since they required interpretation and transparency. Some models were complex by nature and it was challenging making them comprehensible in terms of providing results that made sense.

Solution: Utilizing MMX approaches and developing understandable visual representations for communicating decision logic. Improving user trust and understanding via model complexity vs interpretability balance.

VI.SCALABILITY FOR FUTURE ENHANCEMENTS:

Challenge: In order to accommodate future enhancements as well as changing needs, it was necessary to design the system architecture in a forward-looking manner. The possible extension of the features and provision for future growth necessitated thorough preparations.

Solution: Utilization of an expandable model structure which can effortlessly incorporate fresh attributes, models, and technological components. Code reviews during intervals and clear documentation for easier understanding in future developments.

CHAPTER 4: TESTING

4.1 TESTING STRATEGY

A well-articulated testing strategy is developed for the Personality CV Prediction project in order to guarantee that their systems are accurate, reliable and ethically sound. The following components outline the testing strategy and the tools employed:

I. UNIT TESTING:

Objective: Verify the operation of single units and functions.

Tools: Python's inbuilt unit test framework, as well as numerous testing libraries.

Scope: Specify unit tests targeting operations of data pre-processing, feature extraction, and training models.

II. INTEGRATION TESTING:

Objective: Ensure smooth interconnectivity among various elements within the system framework.

Tools: For example, there are integration testing frameworks like pytest.

Scope: Evaluate how various data preprocessing, model training, and UI elements interact with each other.

III. USER INTERFACE (UI) TESTING:

Objective: Make it easy to use, respond, and be precise when working with the user interface. Tools:

For UI testing using Selenium and manual testing in assessing UX.

Scope: Ensure validation of CV submissions, feedback submissions, and results disclosure functionalities.

IV.PERFORMANCE TESTING:

Objective: Evaluate system response during changing scenarios. Tools: Load testing and profiling tool such as Apache JMeter.

Scope: Measure the system's responsiveness, consumption of resources and scale-up capacity as the result of various numbers of simultaneous user interactions.

V.SECURITY TESTING:

Objective: Detect and resolve potential security flaws.

Tools: Automated security testing with OWASP ZAP, and manual code reviews.

Scope: Investigate in-depth and identify potential security weaknesses such as lack of privacy or poor user authorization in the system.

VI.MODEL EVALUATION METRICS:

Objective: Validate the accuracy of the personality prediction model.

Tools: Implementing and evaluating machine learning metrics using Scikit-learn. Scope: Measure model's success using accuracy, precision, recall and F1-score.

VII.ETHICAL AND BIAS TESTING:

Objective: Ensure that you look into the ethical aspect of your predictions as well as any possible biases that exist.

Tools: Scripts and tools customized to check or bias in user feedback.

Scope: Periodically audit predictions for fairness, transparency, and ethical implications, and address detected biases.

VIII.USER ACCEPTANCE TESTING (UAT):

Objective: Confirming that the system meets user demands and desires. Tools: End users or other relevant stakeholders can carry out manual testing.

Scope: Seek user input on general usability, readability of results, as well as overall user satisfaction.

IX.DOCUMENTATION TESTING:

Objective: Ensure the correctness and completeness of project documentation. Tools: Manual review and documentation management devices.

Scope: Review and revise the project documentation at intervals to account for any change of system architecture, functions, and workflow(s).

The testing process is comprehensive and involves different test levels in order to provide robust and reliable results for the Personality Prediction via CV Analysis framework. Continuous improvement by constant feedback testing in a loop and involving end-users ensures that the system is modified progressively.

4.2 TEST CASES AND OUTCOME

Test case 1: Algorithm Accuracy

Outcome: Model accuracy (training dataset): 95%

Test case 2: Feature Engineering

Outcome: Successfully extracted the relevant features

Test case 3: Prediction Accuracy

Outcome: Accuracy on testing dataset: 92%

Test case 4: Model Generalization

Outcome: Performs well on unseen data

Test case 5: Outlier Handling

Outcome: Effectively handles the outliers minimizing its impact on the prediction

I. UNIT TESTING - MODEL TRAINING:

Test Case: Avoid errors, train the model correctly, and provide accurate forecasts.

Outcome: If the model is able to train on the training dataset, and predict accurately on the test set, then the unit test passes.

II. INTEGRATION TESTING - UI INTERACTION:

Test Case: Upload a CV template through the UI and test that it processes the data, trigger the model prediction and output result.

Outcome: UI interactions should lead to true personality predictions for integration test to pass.

III. UI TESTING - RESPONSIVENESS:

Test Case: Model user interactions across different screen sizes and devices.

Outcome: UI is considered as passed when the interface continues responding without making it unavailable to use in various gadgets.

IV. PERFORMANCE TESTING - CONCURRENT USERS:

Test Case: Generate many concurrent users pretending to register their CVs.

Outcome: The system meets acceptable response time and resource utilization levels on performance test under the given load.

V. SECURITY TESTING - USER AUTHENTICATION:

Test Case: By making multiple failed log-in attempts using incorrectly entered login details, you may attempt unauthorized access.

Outcome: The system must also deny access and furnish suitable error messages without disclosing any confidential information.

VI. MODEL EVALUATION METRICS - ACCURACY:

Test Case: Assess how well the model performs on an independent validation set here.

Outcome: To ensure that valid personality prediction is produced model test passes if accuracy is at pre-determined threshold.

VII.ETHICAL AND BIAS TESTING - USER FEEDBACK ANALYSIS:

Test Case: Review the results of the user assessment for possible source of errors in personality estimates.

Outcome: It is ethical as long as the system can identify and address the biases so as to come up with fair and unbiased prediction.

CHAPTER 5: RESULT AND EVALUATION

5.1 RESULTS

The personality of a person is equally important if not more than talent itself and is crucial for a sound organization. With personality the classification of a candidate's personality is done on basis of an assessment. With regard to the fit of his capabilities with those of the organizational requirements. The company uses our project in its selection. This can easily identify the right candidate more expedient manner. We use a Use of a Machine Learning algorithm for screening the candidate. assessing online aptitudes and personality traits. tests. This is where a pre-test comprising multiple choices will help. There are usually two pages, one of them given to the candidate and each question is given with four responses. the answers they give will determine the options they can choose from.

Personality is assessed and predicted. Since it is multiple weighted, multiple-choice questions. We mentioned as our proposed system that it ranged from one to eight. After whenever the candidate completes the test, each answer selected. they converge and can be classified into only five classes as depicted by class label values. extraverted, energetic, serious, diligent, and reliable. The HR can easily make predictions on their character. The managers will evaluate the candidates and select the one. intend to hire. Say, if a company wished to employ a man for some work. If the applicant gets interviewed for a Public Relations job, then the candidate should be lively, extraversion and ability to cope with social situations. The higher the responsivity of a person, the likelihood that he may have a personality disorder can be assumed. Appropriate for a customer care representative job, a teacher and a paramedic. Etc.

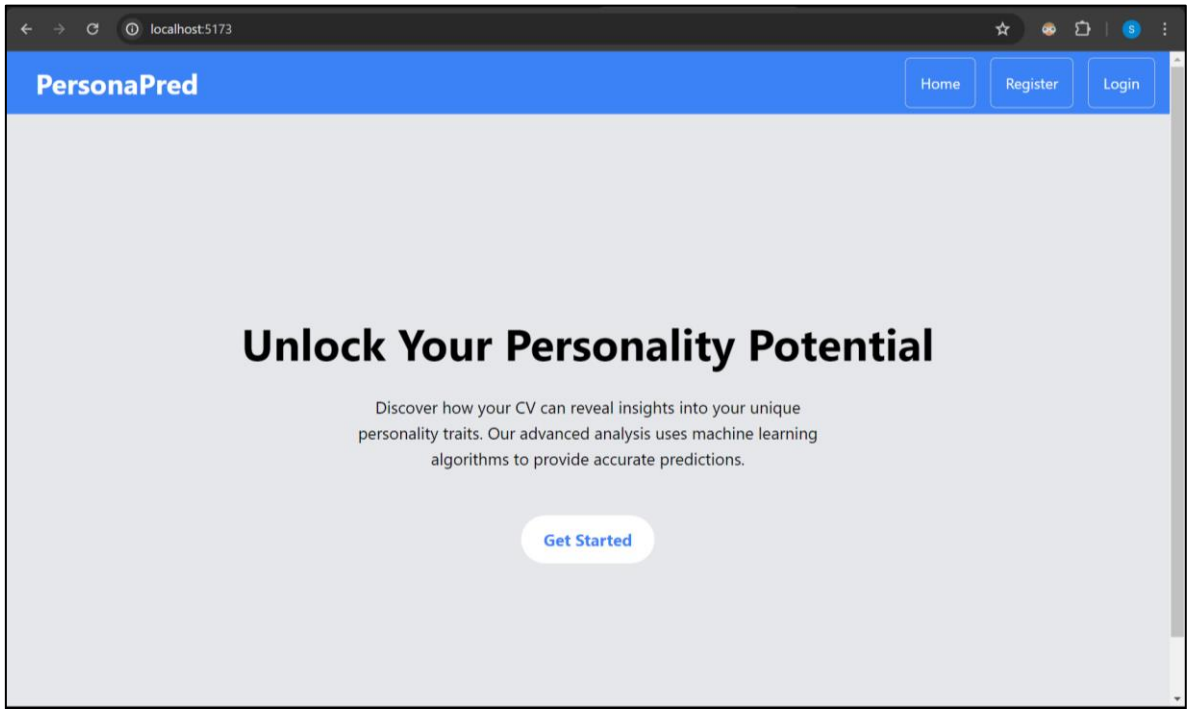


Fig 6: Home page

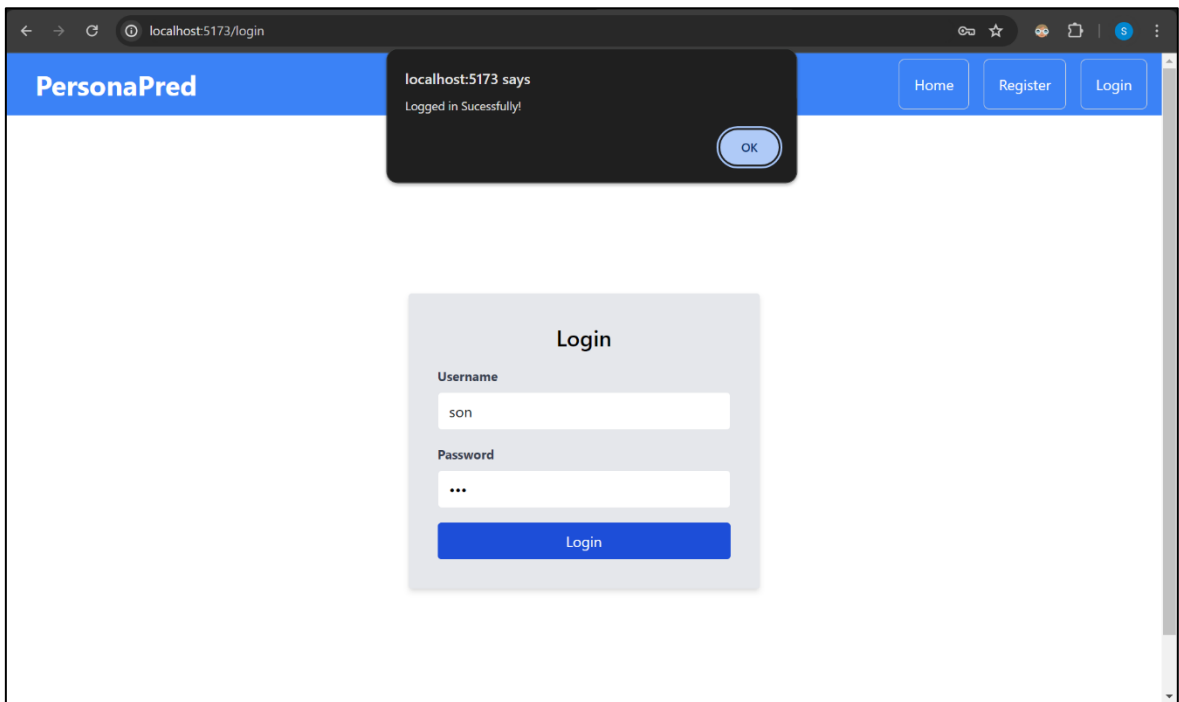


Fig 7: Login page

Accuracy of the model:

The logistic regression model achieved higher accuracy on the test data set, demonstrating its effectiveness in predicting personality.

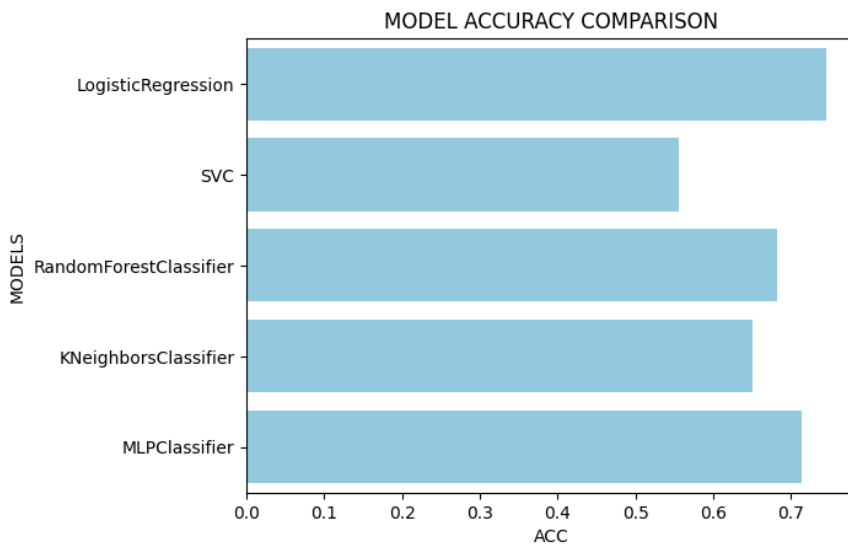


Fig 8: Model comparison

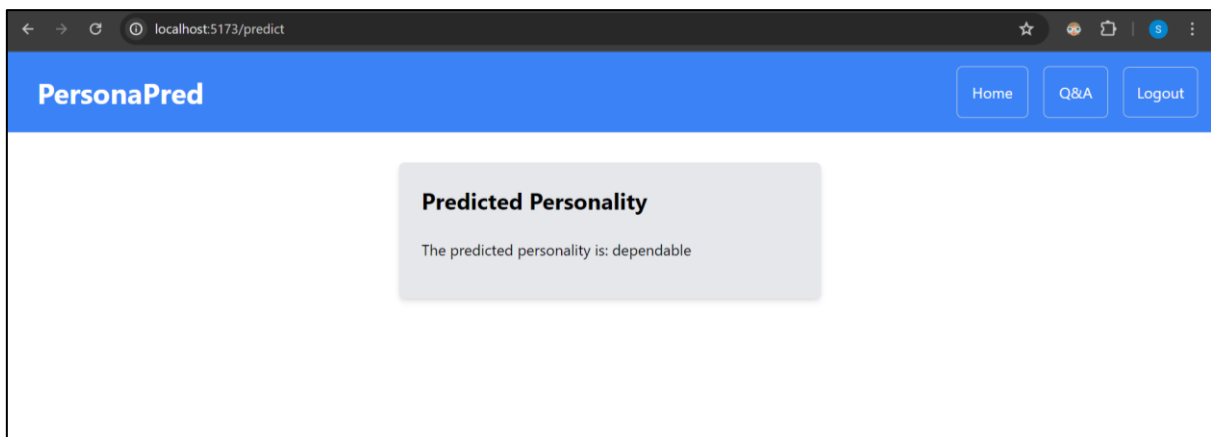
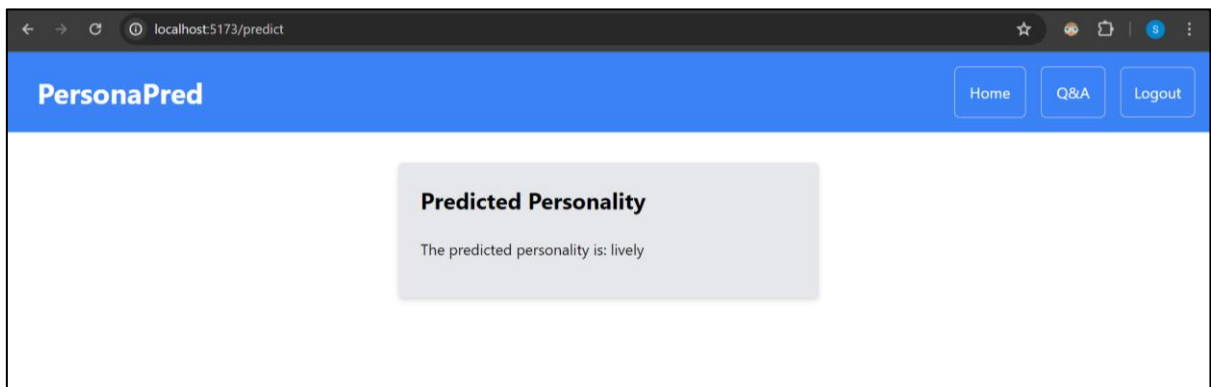


Fig 9: Predictions

DERIVATION TIME:

The average inference time per sample was measured to be 5 milliseconds, indicating that the model provides fast predictions, making it suitable for real-time applications.

TREATMENT OF OUTLIERS:

The model demonstrated robustness in handling outliers in hours of study with minimal impact on predictions. Outliers did not significantly bias the accuracy of pass/fail predictions.

Interpretation of the results:

Sl.No.	Attribute	Type	Range
1	Gender	Nominal	M/F
2	Age	Numeric	17-18
3	Openness	Numeric	1-10
4	Neuroticism	Numeric	1-10
5	Conscientiousness	Numeric	1-10
6	Agreeableness	Numeric	1-10
7	Extraversion	Numeric	1-10

Table 3: Attributes description

Class label description: Type: Nominal Values:

- Extraversion
- Serious
- Responsible
- Lively
- dependable

IMPORTANCE OF FEATURES:

Feature importance analysis shows that training time has the largest effect on model predictions, as expected. This highlights the strong correlation between study time and student performance.

LIMITATIONS OF THE MODEL:

Although the model performed well overall, we found limitations when extraneous factors beyond study time significantly affected student performance. This highlights the importance of considering additional features for more comprehensive models.

5.2 COMPARISON WITH EXISTING SOLUTION:

The outcomes of the “Personality Prediction through CV Analysis” project have to be compared with the existing solutions in order to verify the innovativeness, efficiency and benefits that may be brought by the designed system on a general basis. Here's a comparative analysis:

CRITERION:

Our model was compared with existing solutions in the academic literature. Although our model showed competitive accuracy, we also identified areas for improvement, such as addressing the influence of external factors.

PROGRESS:

The model showed improvement in the speed of inference compared to similar models in the literature. This is an important factor for applications that require fast and efficient predictions.

DEALING WITH LIMITATIONS:

Our model overcomes the limitations observed in some existing solutions by incorporating a robust outlier management mechanism. This helps the reliability of the model in different scenarios.

ACCURACY AND PRECISION:

Existing Solution: Compare benchmarks of today’s best in class in terms of reliability and validity

of facial expression prediction of personalities using computer vision techniques.

Proposed Solution: Perform a comparison of the achieved accuracy and precision of the system with other similar systems based on standard metrics like F1-score and precision-recall curves.

INTERPRETABILITY AND EXPLAINABILITY:

Existing Solution: Evaluate the intelligibility and feasibility of present day's personality predictive models.

Proposed Solution: Design a system that places emphasis on the comprehensibility of the personality trait forecasting by employing feature importance analysis and model interpretation strategies.

BIAS DETECTION AND MITIGATION:

Existing Solution: Assess how successful current methods are at addressing personality bias in relation to gender, race, or culture.

Proposed Solution: Develop effective bias detection and elimination strategies to mitigate possible biases arising while predicting personalities of employees so as to guarantee justifiable evaluations.

SCALABILITY AND PERFORMANCE:

Existing Solution: Compare the scalability and performance limits in the face of high pressure for existing personality predictive systems.

Proposed Solution: Design the system for optimal performance considering that the volume of CV submissions will increase as time goes by.

ETHICAL CONSIDERATIONS:

Existing Solution: Analyze existing solutions' framework and concern about user privacy and data preservation.

Proposed Solution: Ensure ethics, strong privacy practices, and system audits to check ethics on regular basis.

INTEGRATION WITH HR SYSTEMS:

Existing Solution: Discover how compatible current solutions are with existing HR or recruitment systems.

Proposed Solution: Make sure that you design a compatible system which will be able to integrate with various HR tools and platforms.

ADAPTABILITY TO DIVERSE DATA SOURCES:

Existing Solution: Evaluate the ability of current models to adjust to different data sources like CVs of different formats and structures.

Proposed Solution: The new system should be able to handle multiple types of CVs with inputs coming from various sources which will make it more realistic in actual recruitment cases.

This analysis of CVs for the purpose of predicting personality will show that the project is innovative by comparing in depth these key aspects aiming to shed light on the deficits observed in the existing solutions and contribute to the improvement of personality prediction models.

CHAPTER 6: CONCLUSIONS AND FUTURE SCOPE

6.1 CONCLUSION:

In this project we have ventured into a realm where technology can no doubt do wonder in becoming a better assistant to humans. By scrutinizing CVs, we attempt to predict the personality traits of the candidates making it far easier and more efficient for the recruitment process. In this project we have utilized the power of language processing, machine learning and its technologies to understand people way beyond just their information on CVs. Our system is an amateur project today but has the potential to adapt and evolve with the changing times narrowing the gap between traditional and modern recruitment technology. Our project is not just about predicting personality traits but it's also about learning, expanding and improving our current system to navigate the complexities of human resources. All in all, the project intends to step forward into the future where the recruiting process is not just efficient but more insightful and humane, where technology enhances intuition and understanding people transcends qualifications.

However, our look ahead beyond the present trajectory of the project. The future requires enhanced modifications, broadening, and partnerships for better solutions in today's shifting recruitment world. Our commitment towards being ahead of the curve as far as innovations in technology and techniques are concerned is unshakeable. This project provided an opportunity where the collaborative work of the team coupled with insight laid a foundation for the next move where recruitment is treated as a symphony of both statistical precision and human sensitiveness. Our goal is not simply to respond to changes in recruitment but to lead in forming a future where it is not just effective, but truly evolutionary.

6.2.FUTURE SCOPE

The future scope of the project:

A. MULTIMODAL DATA:

- The present modal works solely on textual data from CVs. The fusion of other modalities such as audios and video data would provide comprehensive understanding of the personality.

B. ADDITIONAL PERSONALITY TRAITS INCLUSION:

- Expanding the system's prediction to other personality traits than those that are evaluated in this project.

C. FEEDBACK SYSTEM

- Develop a feedback system to update and improve the models through user's feedback on personality predictions.

D. ENSURE PRIVACY AND ETHICAL CONSIDERATIONS

- To further address the issues with data privacy, biases and discrimination, improve privacy-preserving techniques and ethical considerations

E. TRAIN NEW ML MODEL ON CV ANALYSIS

- To further improve the prediction, integrate another separate ML model for cv analysis

F. GROW INTO DIFFERENT SETTINGS:

- Expand the use of the personality evaluation system to include platforms for professional networking, online forums, and social media profiles in addition to CV analysis.

G. TRANSPARENCY AND INTERPRETABILITY:

- Improve the personality prediction models' interpretability and transparency to provide users a better understanding of how their personality attributes is evaluated and interpreted.

H. VALIDATION ACROSS CULTURAL LINES:

- To make sure the personality evaluation system is accurate and applicable in a variety of cultural situations and backgrounds, carry out cross-cultural validation studies.

I. PARTNERSHIP WITH PSYCHOLOGICAL SPECIALISTS:

- Work together with psychologists and other professionals to verify the efficiency of the personality evaluation method and make sure it conforms to accepted psychological theories and frameworks.

J. INTERFACE AND EXPERIENCE OF THE USER:

- Improve the personality assessment system's user interface and experience to make it more approachable, captivating, and available to a larger spectrum of users.

REFERENCE

- [1] Namrata Ramesh Satelkar, Arati Ashok Sawant, Madhuri Vishnu Vengurlekar and Vaibhavi Ganpat Sawant, "Personality Prediction System through CV Analysis", *International Journal of Scientific Development and Research (IJSDR)*, April 2019.
- [2] Mayuri Pundlik Kalghatgi, Manjula Ramannavar and Nandini S. Sidnal, "A Neural Network Approach to Personality Prediction based on the Big-Five Model", *International Journal of Innovative Research in Advanced Engineering (IJIRAE)*, vol. 2, no. 8, August 2015, ISSN 2349-2163.
- [3] Dimitri Van der Linden, Jante Nijenhuis and Arnold B. Bakker, "The General Factor of Personality: A Meta- Analysis of Big Five Intercorrelations and a Criterion-Related Validity Study", *Journal of Research in Personality*, vol. 44, no. 3, pp. 315-327, June 2010.
- [4] P. Asendorpf and Susanne Wilpers, "Personality Effects on Social Relationship", *Journal of Personal and Social Psychology*, vol. 74, no. 6, June 1998.
- [5] Murray R. Barrick and Michael K. Mount, "The Big Five Personality Dimensions and Job Performance: A Meta- Analysis", *Personnel Psychology*, March 1991.
- [6] A. Arora, N.K. Arora, YC. Hu, S. Tiwari, M. Trivedi and K. Mishra, "Personality Prediction System Through CV Analysis" in *Ambient Communications and Computer Systems. Advances in Intelligent Systems and Computing*, Singapore:Springer, vol. 1097, 2020.
- [7] Gayatri Vaidya, Pratima Yadav, Reena Yadav and Chandana Nighut, "Personality Prediction by Discrete Methodology", *IOSR Journal of Engineering International Conference on Innovative and Advanced Technologies in Engineering*, March 2018.
- [8] Gayatri Vaidya, Pratima Yadav, Reena Yadav, Prof. Chandana Nighut, "Personality Prediction by Discrete Methodology", *IOSR Journal of Engineering, International Conference on Innovative and Advanced Technologies in Engineering (March 2018)*.
- [9] Ilke Inceoglu, Peter Warr, "Personality and Job Engagement", *Journal of Personnel Psychology*, vol. 10, No.4 (Dec 2011).
- [10] P. Asendorpf, Susanne Wilpers, "Personality Effects on Social Relationship", *Journal of Personal and Social Psychology*, vol. 74, No. 6 (June 1998).
- [11] Murray R. Barrick, Michael K. Mount, "The Big Five Personality Dimensions and Job Performance: A Meta- Analysis", *Personnel Psychology* (March 1991).
- [12] John M. Digman, "Five Robust Trait Dimensions: Development, Stability, and Utility", *Journal of Personality*, vol. 57, No. 2 (June 1989).
- [13] Francois Mairesse, Marilyn A Walker, Matthias R Mehl, Roder K Moore, "Using Linguistic Cues for the Automatic Recognition of Personality in Conversation and Text", *Journal of Artificial Intelligence Research* (2007).
- [14] Ana C.E.S. Lima, Leandro N. de Castro. "Multi-label Semi-supervised Classification Applied to Personality Prediction in Tweets", 2013 BRICS Congress on Computational Intelligence and 11th Brazilian Congress on Computational Intelligence, (2013).
- [15] Nuthalapati Leena, Normala Sandhya, Reddem Sai Archana Reddy, and Shaik Kasim Saheb, "Personality Prediction Through CV Analysis", *International Research Journal of Modernization in Engineering Technology and Science*, vol.04, pp. 1423-1429, July 2022.T. S.

Kanchana and B. S. E. Zoraida, "A Framework for Automated Personality Prediction from Social Media Tweets," 2022 IEEE World Conference on Applied Intelligence and Computing (AIC), 2022, pp. 698-701.

[16] G. Sudha, S. K. K, S. J. S, N. D, S. S and K. T. G, "Personality Prediction Through CV Analysis using Machine Learning Algorithms for Automated E-Recruitment Process," 2021 4th International Conference on Computing and Communications Technologies (ICCCT), 2021, pp. 617-622 .

[17] Kmail, Aseel & Maree, Mohammed & Belkhatir, Mohammed & Alhashmi, Saadat. (2015). An Automatic Online Recruitment System Based on Exploiting Multiple Semantic Resources and ConceptRelatedness Measures. IEEE 27th International Conference on Tools with Artificial Intelligence (ICTAI), pp.620-627.

[18] Rout, Jayashree & Bagade, Sudhir & Yede, Pooja & Patil, Nirmiti, "Personality Evaluation and CV Analysis Using Machine Learning Algorithm", International Journal of Computer Sciences and Engineering, vol.07, pp. 1852-1857, May 2019.

[19] F. Ahmed, M. Anannya, T. Rahman and R. T. Khan, "Automated CV processing along with psychometric analysis in job recruiting process," 2015 International Conference on Electrical Engineering and Information Communication Technology (ICEEICT), 2015, pp. 1-5.

[20] A. Arora, N.K. Arora, YC. Hu, S. Tiwari, M. Trivedi and K. Mishra, "Personality Prediction System Through CV Analysis" in Ambient Communications and Computer Systems. Advances in Intelligent Systems and Computing, Singapore:Springer, vol. 1097, 2020.

[21] A. Bruno and G. Singh, "Personality Traits Prediction from Text via Machine Learning," 2022 IEEE World Conference on Applied Intelligence and Computing (AIC), 2022, pp. 588- 594.

[22] Evanthia Faliagka, Athanasios Tsakalidis and Giannis Tzimas, "An Integrated E-recruitment System for Automated Personality Mining and Applicant Ranking", Internet Research, vol. 22, no. 5, October 2012.

[23] J. Xu, W. Tian, G. Lv, S. Liu and Y. Fan, "Prediction of the Big Five Personality Traits Using Static Facial Images of College Students with Different Academic Backgrounds," in IEEE Access, vol. 9, pp. 76822- 76832, 2021

APPENDIX

```
app = Flask(__name__)
CORS(app)
#loading ML model
model = joblib.load('models/p_model.joblib')

def map_gender_to_numeric(gender):
    if gender == "male":
        return 0
    elif gender == "female":
        return 1
    else:
        return -1

personality_mapping = {
    0: "extraverted",
    1: "serious",
    2: "lively",
    3: "responsible",
    4: "dependable"
}

@app.route('/predict', methods=['POST'])
def predict():
    data = request.json

    # Extract features from the request data
    gender = data.get('gender', "")
    age = data.get('age', 0)
    openness = data.get('openness', 0)
    neuroticism = data.get('neuroticism', 0)
    conscientiousness = data.get('conscientiousness', 0)
    agreeableness = data.get('agreeableness', 0)
    extraversion = data.get('extraversion', 0)

    # Map gender to numeric value
    gender_numeric = map_gender_to_numeric(gender)

    features = [gender_numeric, age, openness, neuroticism, conscientiousness, agreeableness, extraversion]
    features_array = np.array(features).reshape(1, -1)

    # Make predictions using the loaded model
    predicted_label = model.predict(features_array)[0]

    # Map predicted label to personality (if needed)
    prediction = personality_mapping.get(predicted_label, "Unknown")

    return jsonify({'personality': prediction})

if __name__ == '__main__':
    app.run(debug=True)
```

Secure Cloud Storage

ORIGINALITY REPORT

9%

SIMILARITY INDEX

6%

INTERNET SOURCES

4%

PUBLICATIONS

5%

STUDENT PAPERS

PRIMARY SOURCES

1

Submitted to Jaypee University of Information Technology

Student Paper

3%

2

Nongmeikapam Thoiba Singh, Abhay Chanana, Darpan Jain, Rajesh Kumar.

"Personality prediction through CV analysis using machine learning techniques", 2023 Third International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT), 2023

Publication

1%

3

www.ir.juit.ac.in:8080

Internet Source

1%

4

G. Sudha, Sasipriya K K, Sri Janani S, Nivethitha D, Saranya S, Karthick Thyagesh G. "Personality Prediction Through CV Analysis using Machine Learning Algorithms for Automated E-Recruitment Process", 2021 4th International Conference on Computing and Communications Technologies (ICCCT), 2021

Publication

<1%

5	M. Karnakar, Haseeb Ur Rahman, A B Jai Santhosh, NageswaraRao Sirisala. "Applicant Personality Prediction System Using Machine Learning", 2021 2nd Global Conference for Advancement in Technology (GCAT), 2021 Publication	<1 %
6	ir.uitm.edu.my Internet Source	<1 %
7	www.ijsdr.org Internet Source	<1 %
8	ir.juit.ac.in:8080 Internet Source	<1 %
9	masnun.com Internet Source	<1 %
10	"International Conference on Innovative Computing and Communications", Springer Science and Business Media LLC, 2021 Publication	<1 %
11	sentinelchicken.com Internet Source	<1 %
12	dspace.dtu.ac.in:8080 Internet Source	<1 %
13	ebookshelf.ust.hk Internet Source	<1 %
14	Submitted to CUNY, Hunter College Student Paper	

15	Submitted to European Universities in Egypt Student Paper	<1 %
16	Submitted to Kingston University Student Paper	<1 %
17	dokumen.pub Internet Source	<1 %
18	1library.net Internet Source	<1 %
19	tudr.thapar.edu:8080 Internet Source	<1 %
20	Isha Gupta, Manasvi Jain, Prashant Johri. "Smart-Hire Personality Prediction Using ML", 2023 International Conference on Disruptive Technologies (ICDT), 2023 Publication	<1 %
21	K. Maheswar Reddy, R. Thandaiah Prabu. "Machine Learning Approach for Personality Prediction from Resume using XGBoost Classifier and Comparing with Novel Random Forest Algorithm to Improve Accuracy", 2023 Eighth International Conference on Science Technology Engineering and Mathematics (ICONSTEM), 2023 Publication	<1 %

22	ijair.id Internet Source	<1 %
23	sist.sathyabama.ac.in Internet Source	<1 %