VIT®

Vellore Institute of Technology

(Deemed to be University under section 3 of UGC Act, 1956)

School of Computer Science and Engineering

Persona

A project submitted

in partial fulfilment of the requirements for the degree of

Bachelor of Technology in Computer Science and Engineering

By

Namdev Parth Deendayal	(19BCE0440)
Nikhil Kumar Parashar	(19BCE0076)
Vishakha Kumaresan	(19BCE2678)
Naman Chadha	(20BCE0888)
Yatheendra Nath Reddy	(20BCE2736)

Course Instructor

Dr. Ramesh Babu K

Professor

April 2022

UNDERTAKING

This is to declare that the project entitled “Persona” is an original work done by undersigned, in partial fulfilment of the requirements for the degree “Bachelor of Technology in Computer Science and Engineering” at School of Computer Science and Engineering, Vellore Institute of Technology (VIT), Vellore.

All the analysis, design and system development have been accomplished by the undersigned. Moreover, this project has not been submitted to any other college or University.

Namdev Parth Deendayal (19BCE0440)

Nikhil Kumar Parashar (19BCE0076)

Vishakha Kumaresan (19BCE2678)

Naman Chadha (20BCE0888)

Yatheendra Nath Reddy (20BCE2736)

ABSTRACT

The job recruitment process takes a lot of time and resources as not only one person is involved but a different variety of people are needed for smooth functioning. One of the major tasks of a recruitment process is the first sorting and filtering of candidates. Different organisations and job profiles require different qualities that are best suited for that role. Organisations spend a great deal of time screening potential candidates, in light of their character, that best suits their needs.

With our project Persona, we mean to give a simple yet effective method for the initial screening of candidates by breaking down their scores on psychometric tests and questionnaires. The tests will decide the scores of different character tests dependent on the Big Five Model and MBTI Personality Traits.

We aim to analyse the questionnaire filled by an applicant and then score them based on the responses. These scores are transferred to our AI Model which helps to classify the applicant based on his/her most relevant character trait. Then the recruiter can filter his potential candidate based on the character traits he needs to fulfil his job requirement.

TABLE OF CONTENTS

1. Introduction

1.1 Objective

The fundamental objective of our project is to provide an efficient screening system for the companies to recruit manpower resources. Our model will help the companies to sort their job applicants based on an analysis of scores of psychometric tests and questionnaires. Our main objectives are to help the job recruiters to categorise their candidates based on tests, questionnaires and resumes and make it effortless for an organisation to fasten their enlistment process.

This would help them with picking the most reasonable up-and-comer in less time for which we will provide a software/model for the recruiter to ease this process and also to the job seekers to analyse their resume and build their resume under the requirements.

1.2 Project Scope

There's a very rigorous screening process for pre-interview rounds and still aren't as accurate as it should be. Most of the time this results in vagueness in selections. Our project comes into play in situations like these. Our website provides a set of curated questions to judge the applicant based on Big 5 Factors namely - Openness, Conscientiousness, Extraversion, Agreeableness and Neuroticism. These factors help in accurately determining an applicant's dominant personality. The website also provides an efficient way for recruiters to log in and filter through the applicants. Easing the job for both the applicants and recruiter.

1.3 Definitions

Logistic Regression

When the dependent variable is dichotomous, logistic regression is the appropriate regression analysis to use (binary). The logistic regression, like all regression analyses, is a predictive analysis. Logistic regression is a data analysis technique that is used to describe and explain

the relationship between one dependent binary variable and one or more nominal, ordinal, interval, or ratio-level independent variables.

Big 5 OCEAN

Personality Traits Openness, Conscientiousness, Extroversion, Agreeableness, and Neuroticism are the five major personality traits, also known as OCEAN and CANOE. These five characteristics represent broad domains of human behaviour and account for differences in personality as well as decision making. Today, HR professionals use the model to evaluate potential employees, and marketers use it to understand their product's target audiences. By the 1990s, it was widely accepted that situational and personality factors both influence in-the-moment behaviour. Even as recently as 2016, research and refinement of the OCEAN model were ongoing, demonstrating how influential it is — and continues to be.

The Big Five Personality Traits

Openness to experience: also known as intelligence or imagination, this trait denotes a willingness to try new things and think outside the box. Insightfulness, originality, and curiosity are all characteristics.

Conscientiousness: the desire to be cautious, diligent, and self-disciplined in regulating immediate gratification Characteristics include ambition, discipline, consistency, and dependability.

Extroversion: as opposed to being alone, a state in which an individual draws energy from others and seeks social connections or interaction (introversion). Characteristics include being outgoing, energetic, and self-assured.

Agreeableness: the assessment of how an individual interacts with others, as measured by compassion and cooperation Tactics, kindness, and loyalty are examples of characteristics.

Neuroticism: a proclivity for negative personality traits, emotional instability, and self-destructive thought Pessimism, anxiety, insecurity, and fearfulness are all characteristics.

2. Literature Survey

S. No	Title and Year	Algorithm Used	Dataset being used	Performance Measures	Gap Identified
1	**Personality Predictions Based on User Behaviour on the Facebook Social Media Platform - October 2018**	Logistic Regression, Gradient Boosting, Support Vector Machine and XGBoost	myPersonality dataset (250 users and 9917 status updates)	Accuracy = (correct predictions/total prediction) * 100	Analyses social media behaviour which could be misleading
2	**Personality Prediction System Based on Signatures Using Machine Learning - 2020**	Support Vector Machine	Digital Image	Confusion Matrix	Analyses signatures which is not the best basis for personality prediction
3	**Personality Prediction from social media Text: An Overview - May 2020**	KNN, SVM, LSTM, CNN, TF-IDF Bayes, XGBoost	i.) 2467 essays tagged with their author's traits ii.) Short text of YouTube vloggers	Classification Reports	Analyses social media behaviour which could be misleading
4	**Personality Evaluation and CV Analysis using Machine Learning Algorithm -**	TF-IDF	User Collected Data	Not Specified	Words can have different meanings in different contexts, hence scoring

	May 2019				words could be misleading.
5	**Automated Personality Classification Using Data Mining Techniques - April 2017**	Advanced Naive-Bayes Classification Algorithm, Support Vector Machine	User Collected Data	Not Specified	Does not implement DL algorithms such as XGBoost, etc. And does not implement the Big 5 factor.
6	**Big Five Factor Model, Theory and Structure - 2015**	OCEAN	NA	NA	Being Implemented
7	**A study on the effect of big five personality traits on emotional intelligence - 2014**	OCEAN	NA	NA	Being Implemented
8	**A Study of the Effect of the Myers Briggs Type Indicator on Team Effectiveness - 2003**	ISTJ, ISFJ, INFJ, etc	NA	NA	Does not implement into the program.
9	**Random Forests and Decision Trees - 2012**	J.48, C4.5 algorithm	UCI Machine Learning repository, Dataset Lymphography, Dataset Sonar,	Classification Report	Does not implementi on into the program.

			Dataset Heart-h		
10	**Workshop on Support Vector Machines: Theory and Applications - 2001**	Support Vector Machine (SVM)	set of l training data {(x1, y1) ... (xl, yl)} in Rn × R sampled according to unknown probability distribution P (x, y), and a loss function V (y, f(x))	Not Specified	Being Implemented
11	**KNN Model Based Approach in Classification - 2004**	KNN	UCI machine learning repository	Classification Accuracy and Reduction Rate	Being Implemented
12	**A KNN Research Paper Classification Method Based on Shared Nearest Neighbour - 2010**	KNN	Millions of training sets in about 400 subclasses, 6000 Main groups, 30000 subgroups	Precision	Being Implemented
13	**An Introduction to Logistic Regression Analysis and Reporting - 2002**	Logistic Regression	Sample Data for Gender and Recommendation for Remedial Reading Instruction, 189 Children's Referrals for Remedial Reading Programs by SAS PROC LOGISTIC	Classification Accuracy	Being Implemented

14	**Long Short - Term Memory - 1997**	RNN, LSTM	Not Specified	Classification Accuracy	Not Implemented
15	**Creating Crowd Variation with the OCEAN Personality Model - 2008**	OCEAN	NA	NA	Being Implemented
16	**How the Ocean Personality Model Affects the Perception of Crowds - 2011**	OCEAN	NA	NA	Being Implemented
17	**A logistic regression investigation of the relationship between the learning Assistant model and failure rates in introductory STEM courses - 2018**	Logistic Regression	Admin records at the University of Colorado Boulder. Focuses on 16 cohorts of students who entered the university as full-time freshmen for each fall semester from 2001 to 2016 and took Physics I/II, General Chemistry I/II, Calculus I/II, and/or Calculus I/II for Engineers.	Not Mentioned	Being Implemented

18	**Binary Logistic Regression Analysis in Assessment and Identifying Factors That Influence Students' Academic Achievement - 2016**	Logistic Regression	Questionnaire and checklist were designed to collect data by distributing to students.	Not Mentioned	Being Implemented
19	**A Survey of the Big Five Personality Traits Among Elementary Teachers - 2017**	OCEAN	NA	NA	Being Implemented
20	**XGBoost Based Algorithm Interpretation and Application on Post-Fault Transient Stability Status Prediction of Power System - 2019**	XGBoost	PMU Data	Classification accuracy	Being Implemented

21	**Efficient Way Of Web Developme nt Using Python And Flask. - 2015**	Flask Module	NA	NA	Being Implemente d
25	**Gradient Boosting Machines, a tutorial**	Gradient Boosting GBM	An artificially generated dataset	RMSE	Not implemented

3. Proposed Methodology

4.1 Drawbacks of Existing Systems

The existing work on Personality Prediction is on analysing social media behaviours, analysing signatures or handwriting, etc. These methods tend to be erroneous as people depict a different character on social media altogether. And the handwriting is not an ideal way to judge someone's personality. Also, some systems analyse the applicant's CV (or Resume). These tend to be slightly erroneous too as the context of the word can vary from one CV to another. None of the existing systems uses any verified or highly tested ways of identifying one's personality. Moreover, personality prediction in the recruitment process is not explored much.

4.2 Our Solution

Our system will provide a Questionnaire that would score their five personality tests based on the Big Five Model - OCEAN (Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism). Then based on the score, the model will predict the personality of the person.

In the first stage of the development, we will train our model based on various algorithms like KNN, SVM, Logistic Regression and XGBoost. Then we will compare the precision of

our model for each model using data visualization and some metrics like confusion matrix and classification report.

Considering the results from Phase 1, we finally implement the most appropriate algorithm to our personality prediction algorithm

We will integrate it with a web application, to provide a nice and easy interface for the applicants to take the psychometric test and apply for the job, and for the recruiters to filter through the applicants and pick up the applicants of desired character trait.

The dataset that we will be using is downloaded from Kaggle.

	Gender	Age	openness	neuroticism	conscientiousness	agreeableness	extraversion	Personality (Class label)
2	Male	17	7	4	7	3	2	extraverted
3	Male	19	4	5	4	6	6	serious
4	Female	18	7	6	4	5	5	dependable
5	Female	22	5	6	7	4	3	extraverted
6	Female	19	7	4	6	5	4	lively
7	Male	18	5	7	7	6	4	lively
8	Female	17	5	6	5	7	4	extraverted
9	Female	19	6	6	7	5	4	extraverted
10	Male	18	5	7	5	6	7	dependable
11	Female	19	5	5	7	4	5	lively
12	Male	19	6	7	5	6	3	serious
13	Male	19	7	6	7	7	6	extraverted
14	Male	19	7	6	6	5	6	lively
15	Female	19	6	7	5	5	5	dependable

4. Architecture/Design

4.1 Overview

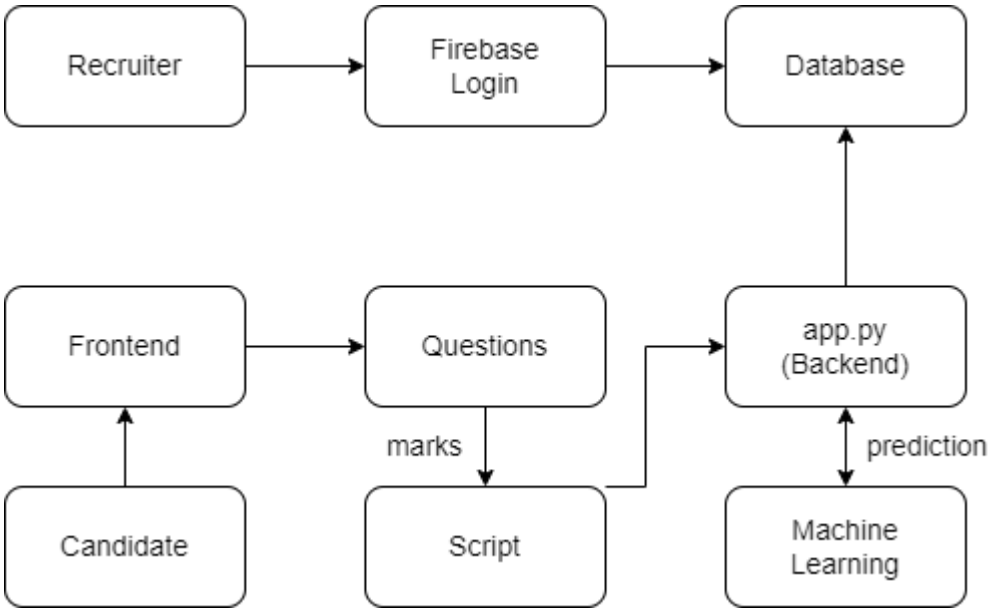
The proposed system or software is divided in three parts

- Candidates - Persona portal
- Employees - Sorting and selecting candidates
- System - Asking questions and predicting marks with help of trained models

An electronic web-based form is created for the candidates on which they are asked questions. As it is web-based, it could be accessed easily on most of the devices.

After candidates submit their answers, the system automatically allot marks and sort them according to criteria set. This makes it easy for the employees to select candidates and makes the screening process smooth.

4.2 System Architecture



- Candidate connects to system frontend portal
- Candidate is displayed with the questions on the portal
- Employee login portal is integrated with firebase login for user authentication
- Marks are calculated automatically by the system using a script
- These marks are sent to backend server
- The server then sends this data to the ML models
- The predicted results are stored in the database

4.3 Components Description

4.3.1 Secured Login for the Recruiters

Identification	Secured login screen
Type	Component/Form
Purpose	The login screen ensures that only employees can access the candidate data.
Function	It enables employees to see the applied candidates data, and make sure any unauthenticated person doesn't have access to it.

Dependencies	The following screen links to this screen: ● Home Screen
Interfaces	The login form is located on the left half of the page, and designed to be easy to view.
Resources	Database Access Requirements: access to the violator information found in the appropriate database tables.
Data	The data for the login Screen is the email and password entered by the user. It is validated using the firebase authentication.

4.3.2 Extensive Questionnaire

Identification	Extensive Questionnaire
Type	Component/Form
Purpose	To ask candidates questions and collect answers
Function	It is used to get answers to the questions asked by the system. These answers are used for ML model prediction.
Dependencies	The following screen links to this screen: ● Register screen
Interfaces	The questions are displayed in an electronic HTML form
Resources	The questions are auto-generated by the system.
Data	The data for this screen are the questions and their respective answers submitted by the candidate.

4.3.3 Candidate Register

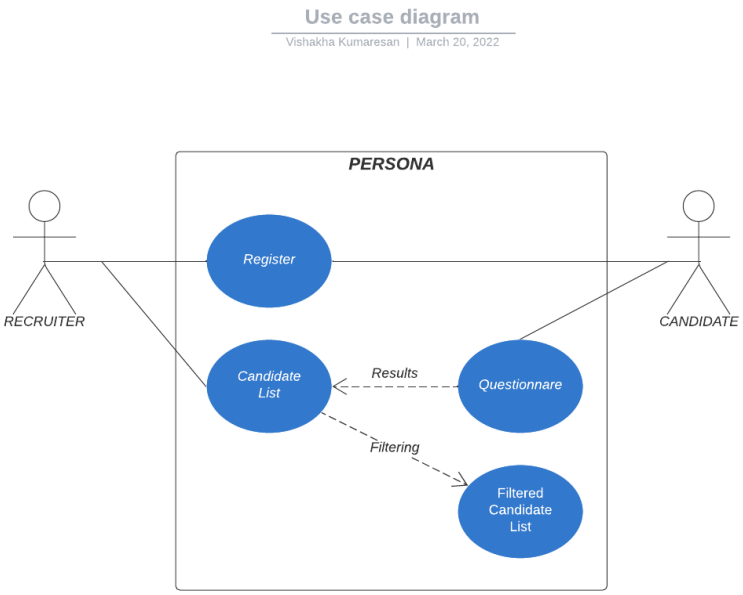
Identification	Candidate register screen
Type	Component/Form

Purpose	To collect personal details of the candidate and register him.
Function	To collect details and store them in the system's database. This database can be accessed by the employees.
Dependencies	The following screen links to this screen: ● Home Screen
Interfaces	The details form is located on the left half of the page, and designed to be easy to view.
Resources	These details are used by employees to recognise and filter candidates.
Data	The data for this screen is the ID, name, age and gender entered by the user.

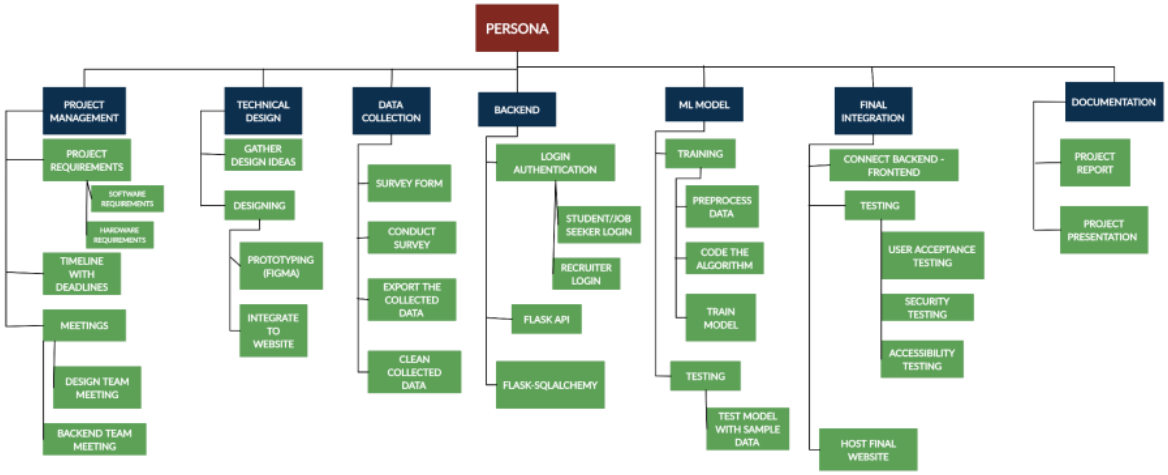
4.3.4 Filter by Category

Identification	Category Filter screen
Type	Component/Page
Purpose	To see candidate data and filter them.
Function	It enables employees to see the applied candidates data and filter them according to different categories.
Dependencies	The following screen links to this screen: ● Recruiter login screen
Interfaces	The login form is located on the left half of the page, and designed to be easy to view.
Resources	Different categories are allotted to candidates based on their marks
Data	The data for this screen is taken from a database.

UML Use Case Diagram



Work Breakdown Structure



5. Description of Various Modules

Module	Description
Backend	The backend server is made using python Flask API. SQLite is used as a database for storing data.
ML	Four algorithms are used including Logistic Regression, K Nearest Neighbours (KNN), Support Vector Classifier (SVC) and XGBoost, for the prediction of different personality traits.
Frontend	It is made using HTML, CSS, Bootstrap and JS.
Data Collection	The dataset is available on kaggle

Product Functions

- **Extensive Questionnaire**: multiple questions have been curated carefully for analysing the candidates' big five traits (OCEAN).
- **Automatic scoring**: Our software calculates the marks of the candidates as soon as they submit their answers.
- **Easy access to candidates**: The candidates can access the system from any laptop or desktop with internet connectivity.
- **Ease the screening process:** The software allows employers to easily filter through a list of candidates and select the ones who fit into the required criteria.
- **Secured login**: A secured login page for employers to login and access the candidate database.

Database Schema

For the proposed system the database used is SQLITE and for employee login firebase authentication is used

The schema used for SQLITE database is

- App_id
- Name
- Gender
- Age

- O, C, E, A, N (Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism)
- Personality

Table name	Field	Datatype	Primary
Applicants	app_id	String	Yes
	name	String	-
	gender	String	-
	age	Integer	-
	o	Integer	-
	c	Integer	-
	e	Integer	-
	a	Integer	-
	n	Integer	-
	personality	String	-

6. Implementation (Sample Code)

Function for Plotting Confusion Matrix

```
def plotCM(y_true, y_pred, name = "Confusion Matrix"):

 labels = unique_labels(y_test)

 table = pd.DataFrame(confusion_matrix(y_true, y_pred), columns = labels,  index = labels)

 sn.heatmap(table, annot = True, fmt = 'd', cmap = 'viridis')
plt.gcf().axes[0].tick_params(colors = 'black')

 plt.gcf().axes[1].tick_params(colors = 'black')

 plt.title(name, color = 'black')
```

```
plt.xlabel('Predicted Label', color = 'black')

plt.ylabel('True Label', color = 'black')

plt.gcf().set_size_inches(10,6)

plt.show()
```

Instantiating the Algorithms

```
lr = LogisticRegression(multi_class='multinomial', solver='newton-cg',max_iter  =1000)

knn = KNN()

svc = SVC()

xgb = XGBClassifier()
```

Testing Logistic Regression

```
lr.fit(X_train, y_train)

lr_pred = lr.predict(X_test)

plotCM(y_test, lr_pred , 'Confusion Matrix for Logistic Regression')
```

Testing KNN

```
knn.fit(X_train, y_train)

knn_pred = knn.predict(X_test)

plotCM(y_test, knn_pred , 'Confusion Matrix for KNN')
```

Testing SVC

```
svc.fit(X_train, y_train)

svc_pred = svc.predict(X_test)

plotCM(y_test, svc_pred , 'Confusion Matrix for SVC')
```

Testing XGBoost

```
xgb.fit(X_train, y_train)

xgb_pred = xgb.predict(X_test)

plotCM(y_test, xgb_pred , 'Confusion Matrix for XGBoost')
```

Calculating F1 Scores and Plotting the Graph

```
algos = ['Logistic Regression', 'KNN', 'SVC', 'XGBoost']

f1_scores = [f1_score(y_test, lr_pred, average='weighted'), f1_score(y_test,  knn_pred, average='weighted'), f1_score(y_test, svc_pred, average='weighted'),  f1_score(y_test, xgb_pred, average='weighted')]

ypos = np.array([0,1,2,3])

plt.xticks(ypos, algos)

plt.bar(ypos, f1_scores)

plt.gcf().axes[0].tick_params(colors = 'black')

plt.title('F1 Scores vs Algos', color = 'black')

plt.xlabel('Algorithms', color = 'black')

plt.ylabel('F1 Scores', color = 'black')

plt.gcf().set_size_inches(10,6)

plt.show()
```

Final Implementation of Model into Web App

```
from os import name

from flask import Flask, render_template, url_for, flash, redirect, request from flask_sqlalchemy import SQLAlchemy

from marks import calculate

import pickle
```

```
import numpy as np

with open ('personality_prediction_lr', 'rb') as f:

 model = pickle.load(f)

app = Flask(__name__)

app.config['SQLALCHEMY_DATABASE_URI'] = 'sqlite:///Applicants.db'
app.config['SQLALCHEMY_TRACK_MODIFICATIONS'] = False db = SQLAlchemy(app)

15

class Applicants(db.Model):

 app_id = db.Column(db.String(100), primary_key = True)  name =
db.Column(db.String(100), nullable = False)

 gender = db.Column(db.String(10), nullable = False)

 age = db.Column(db.Integer, nullable = False)

 o = db.Column(db.Integer, nullable = False)

 n = db.Column(db.Integer, nullable = False)

 c = db.Column(db.Integer, nullable = False)

 a = db.Column(db.Integer, nullable = False)

 e = db.Column(db.Integer, nullable = False)

 Personality = db.Column(db.String(20), nullable = False)

def __repr__(self) -> str:

 return f"{self.app_id}-{self.name} - {self.gender} - {self.age} - {self.o} - {self.n} - {self.c} -
{self.a} - {self.e} - {self.Personality}"

 def __init__(self, app_id, name, gender, age, o, n, c, a, e, Personality):  self.app_id = app_id

 self.name = name
```

```
 self.gender = gender
 self.age = age
 self.o = o
 self.n = n
 self.c = c
 self.a = a
 self.e = e
 self.Personality = Personality


@app.route('/')
def index():
 return render_template("index.html")
16
@app.route('/register', methods = ['GET', 'POST'])
def register():
 if request.method == 'POST':
 app_id = request.form['id1']
 name = request.form['name']
 gender = request.form['gender']
 age = request.form['age']
 detail = app_id + "-" + name + "-" + gender + "-" + age
 return redirect(url_for("quiz", details = detail))
```

```
 return render_template("register.html")


@app.route('/quiz/<details>', methods = ['GET', 'POST'])
def quiz(details):
 if request.method == 'POST':
 ans = []
 ans.append(request.form.getlist('mycheckbox')[0])
 ans.append(request.form.getlist('mycheckbox1')[0])
 ans.append(request.form.getlist('mycheckbox2')[0])
 ans.append(request.form.getlist('mycheckbox3')[0])
 ans.append(request.form.getlist('mycheckbox4')[0])
 ans.append(request.form.getlist('mycheckbox5')[0])
 ans.append(request.form.getlist('mycheckbox6')[0])
 ans.append(request.form.getlist('mycheckbox7')[0])
 ans.append(request.form.getlist('mycheckbox8')[0])
 ans.append(request.form.getlist('mycheckbox9')[0])
 essentials = details.split('-')
 o1, c1, e1, a1, n1 = calculate(ans)
 data = []
 if(essentials[2].lower()=='male'):
 data.append(1)
 else:
```

```
 data.append(0)

 data.append(int(essentials[3]))

 data.append(o1)

 data.append(n1)

 data.append(c1)

 data.append(a1)

 data.append(e1)

 data = np.array(data).reshape(1,7)

 p = model.predict(data)

apply = Applicants(app_id = essentials[0], name = essentials[1], gender =  essentials[2], age = int(essentials[3]),o = o1,n= n1,c= c1,a= a1,e = e1,  Personality = p[0])

 db.session.add(apply)

 db.session.commit()

 return redirect(url_for("submit"))

 return render_template("quiz.html")

@app.route('/submit')

def submit():

 return render_template("submit.html")

@app.route('/recruiter-login', methods = ['GET', 'POST'])

def recruiter():

 if request.method == 'POST':

 return render_template("recruiter.html")
```

```
 return render_template("recruiter.html")

@app.route('/data')

18

def data():

 apply = Applicants.query.all()

 return render_template("data.html", applicants = apply)

if __name__=="__main__":

 app.run(debug = True, port = 4000)
```

Function to calculate the marks from the questionnaire

```
def calculate(answers):

 o = 0

 c = 0

 e = 0

 a = 0

 n = 0

 #for ques 1

 temp0 = int(answers[0])

 n+=temp0

 #for ques 2

 temp1 = int(answers[1])

 c+=temp1

 #for ques 3
```

```
temp2 = int(answers[2])

o+=(6-temp2)

#for ques 4

temp3 = int(answers[3])

e+=temp3

#for ques 5

temp4 = int(answers[4])

c+=(6-temp4)

#for ques 6
```

19

```
temp5 = int(answers[5])

a+=(6-temp5)

#for ques 7

temp6 = int(answers[6])

n+=(6-temp6)

#for ques 8

temp7 = int(answers[7])

e+=temp7

#for ques 9

temp8 = int(answers[8])

o+=temp8

#for ques 10
```

```
temp9 = int(answers[9])

a+=temp9

return o,c,e,a,n
```

7. Testing (ALL Screen Snapshots)

7.1 Phase 1 Comparison of Various Models

Four algorithms were used in the first stage of development including Logistic Regression, K Nearest Neighbours (KNN), Support Vector Classifier (SVC) and XGBoost. To compare these various models, Confusion Matrix and F1 Score measures were employed. In both the metric cases, Logistic Regression turned out to perform the best for our dataset. F1 score for Logistic Regression turned out to be approximately 0.8 which was quite good. Depending on the aforementioned results, we decided to move forward and use logistic regression as the building model for our project. Here are the confusion matrices for each type and the graph for final comparison.

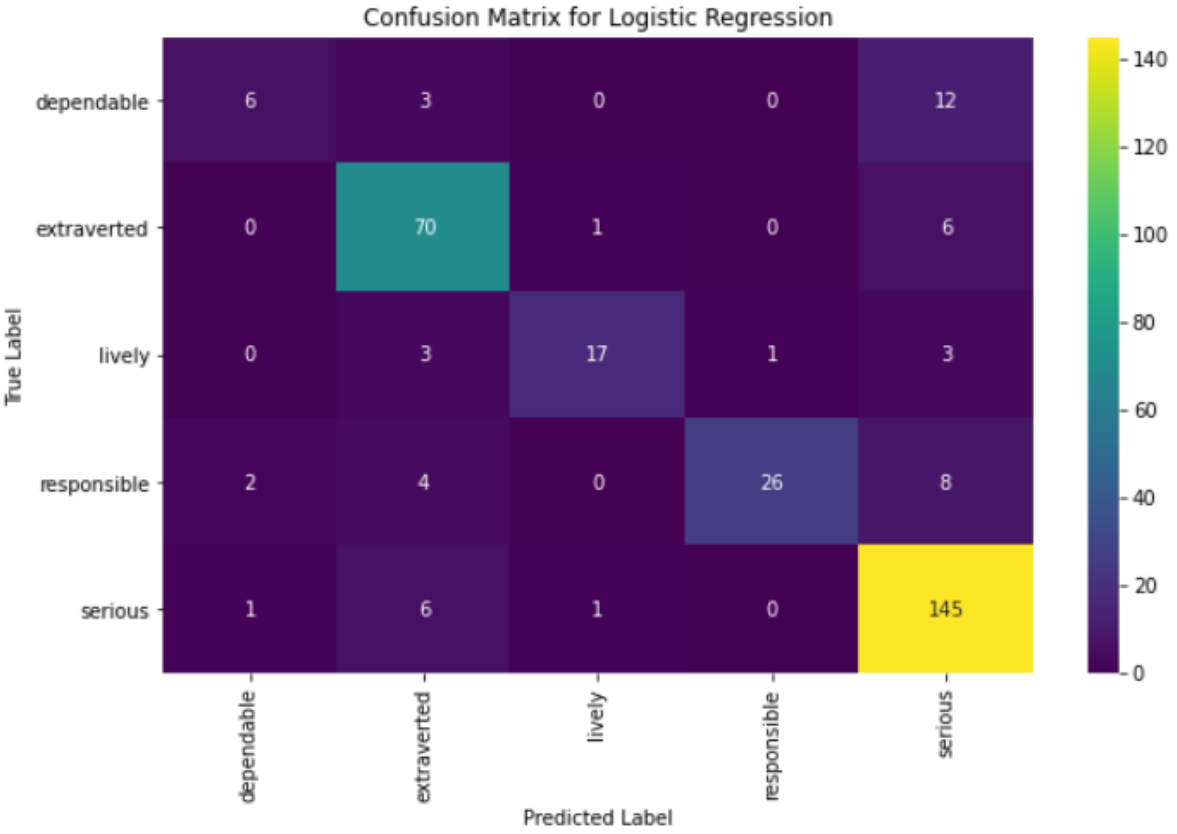


Figure 1 Confusion Matrix for Logistic Regression

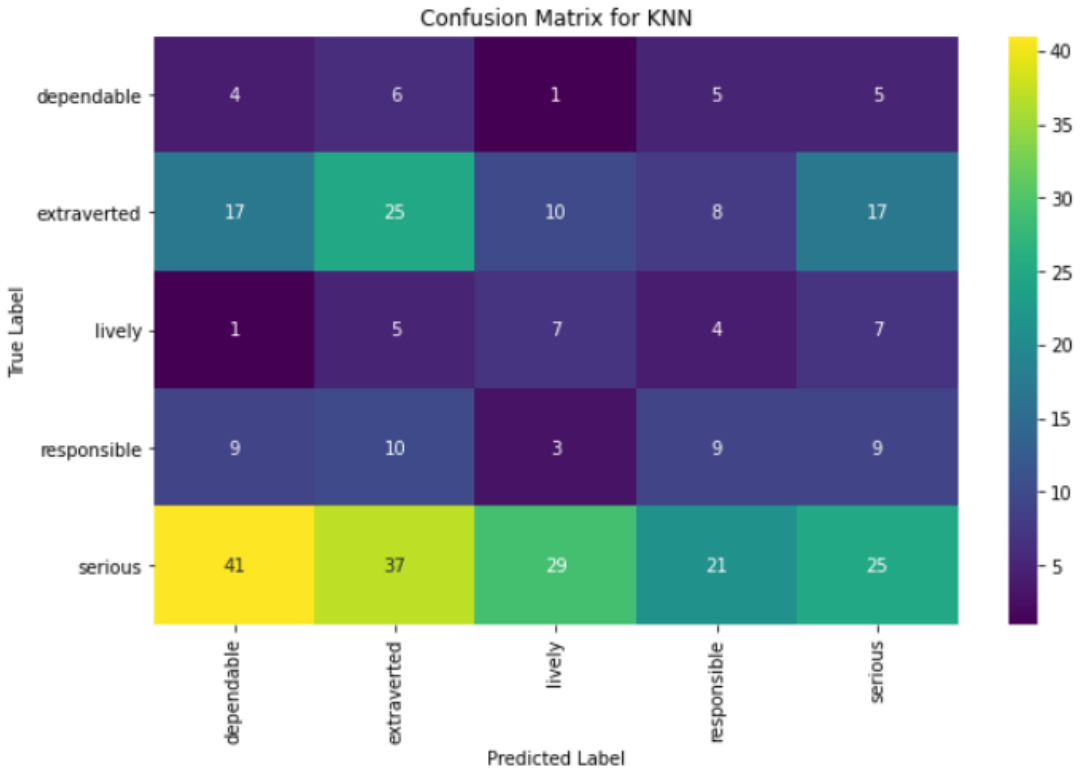


Figure 2 Confusion Matrix for KNN

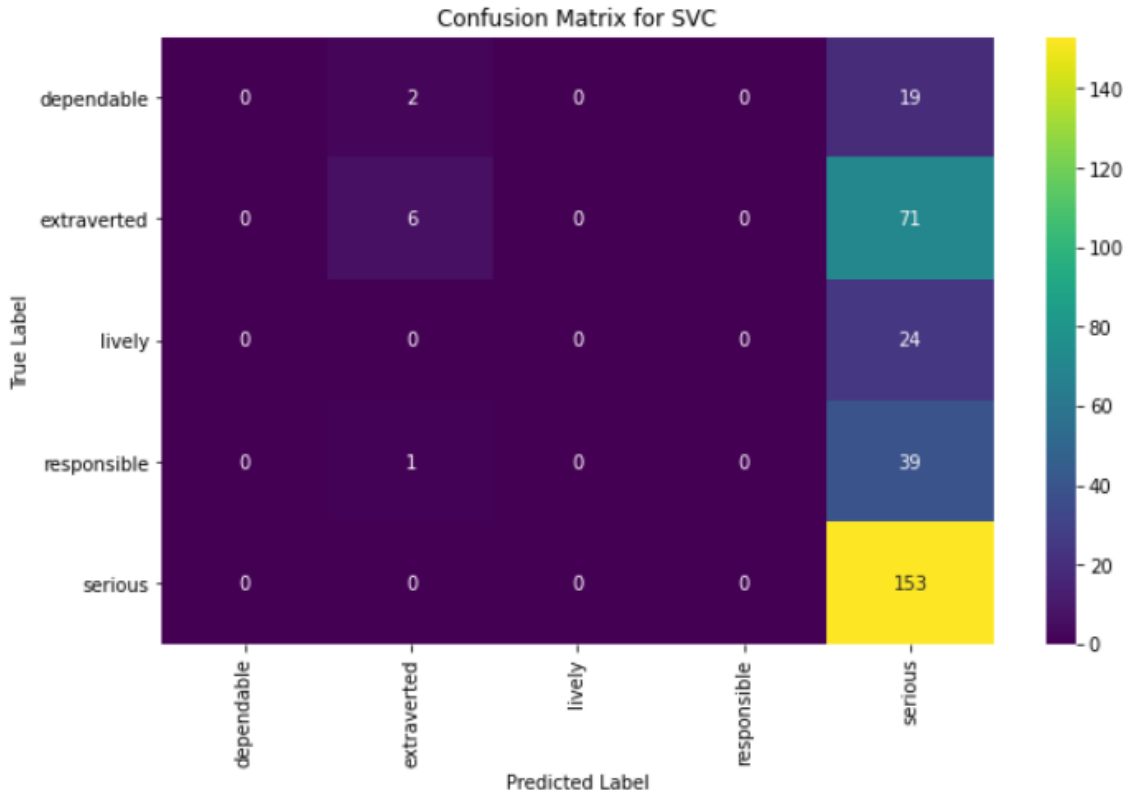


Figure 3 Confusion Matrix for SVC

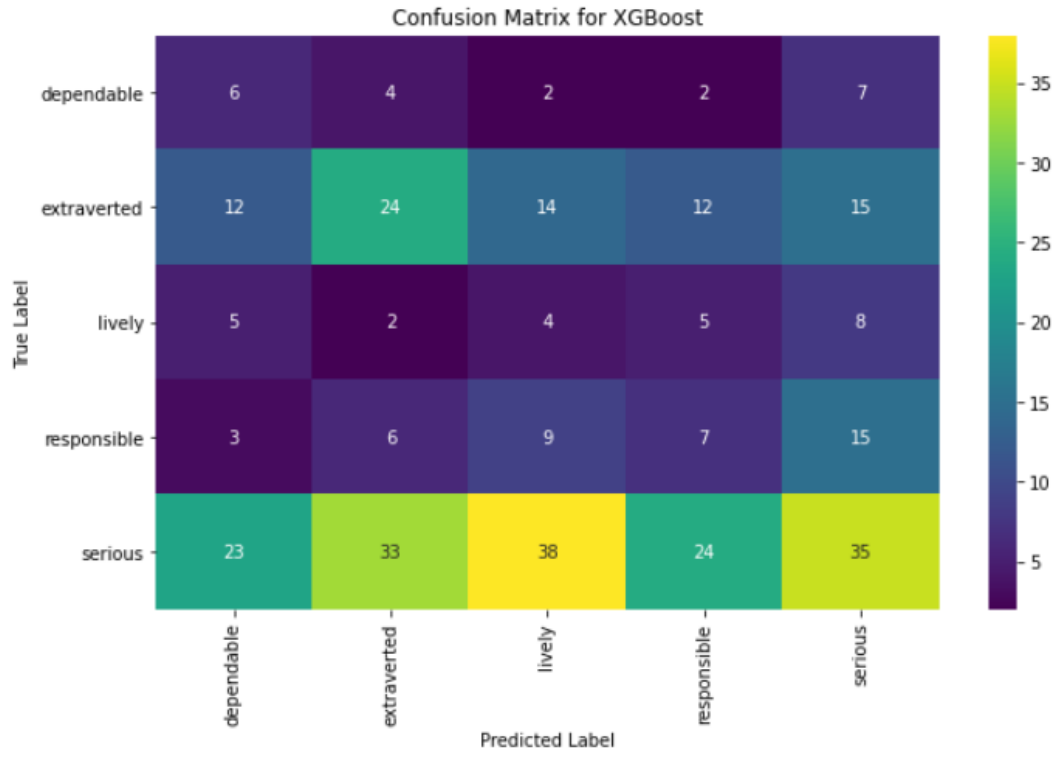


Figure 4 Confusion Matrix for XGBoost

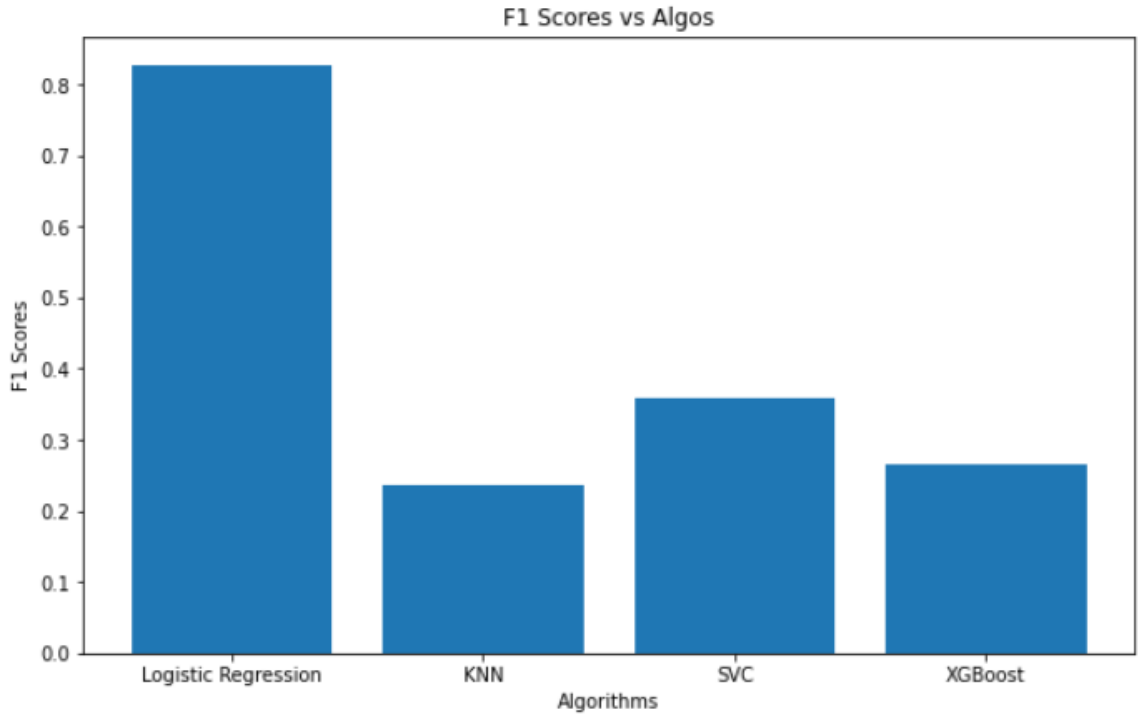


Figure 5 Comparison of F1 Scores

7.2 Phase 2 Implementation of the Chosen Model

In the second stage, we implemented our Logistic Regression model and integrated the same into our Web Application using Flask. We also created a database in SQLite using Flask SQLAlchemy.

Dataset Description:

Attribute Description: No. of attributes are 7 as listed below.

S. No	Attribute	Type	Range
1	Gender	Nominal	Male/Female
2	Age	Numeric	17-28
3	Openness	Numeric	1-8
4	Neuroticism	Numeric	1-8
5	Conscientiousness	Numeric	1-8
6	Agreeableness	Numeric	1-8
7	Extraversion	Numeric	1-8

Class label description:

No. of class labels: 5

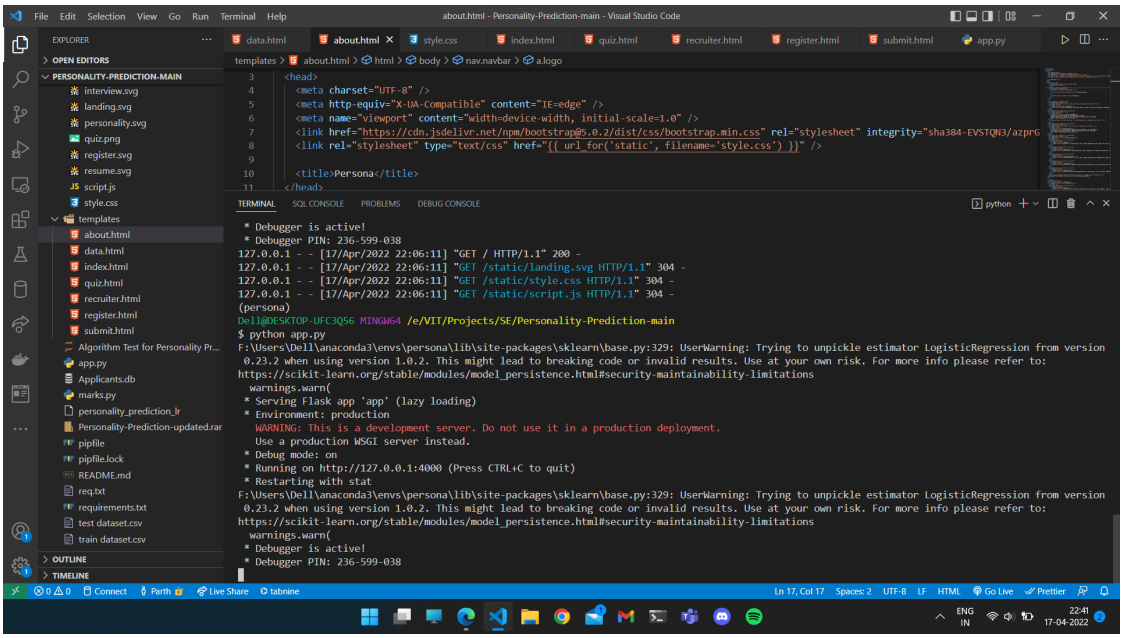
Type: Nominal

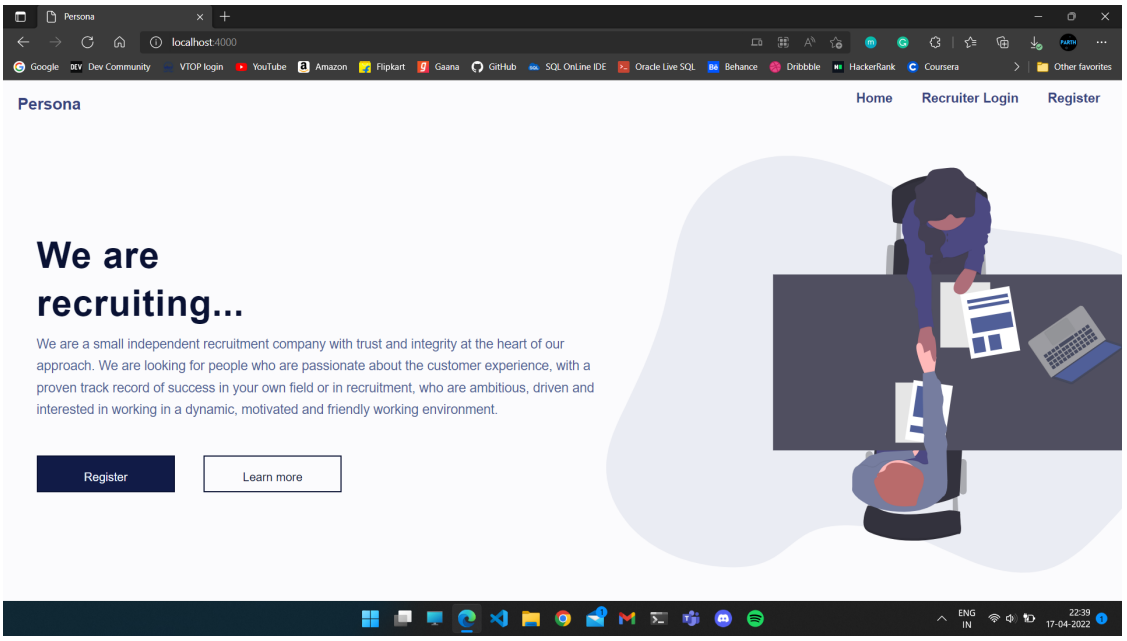
Class Labels:

- Extraverted
- Serious
- Responsible
- Lively
- Dependable

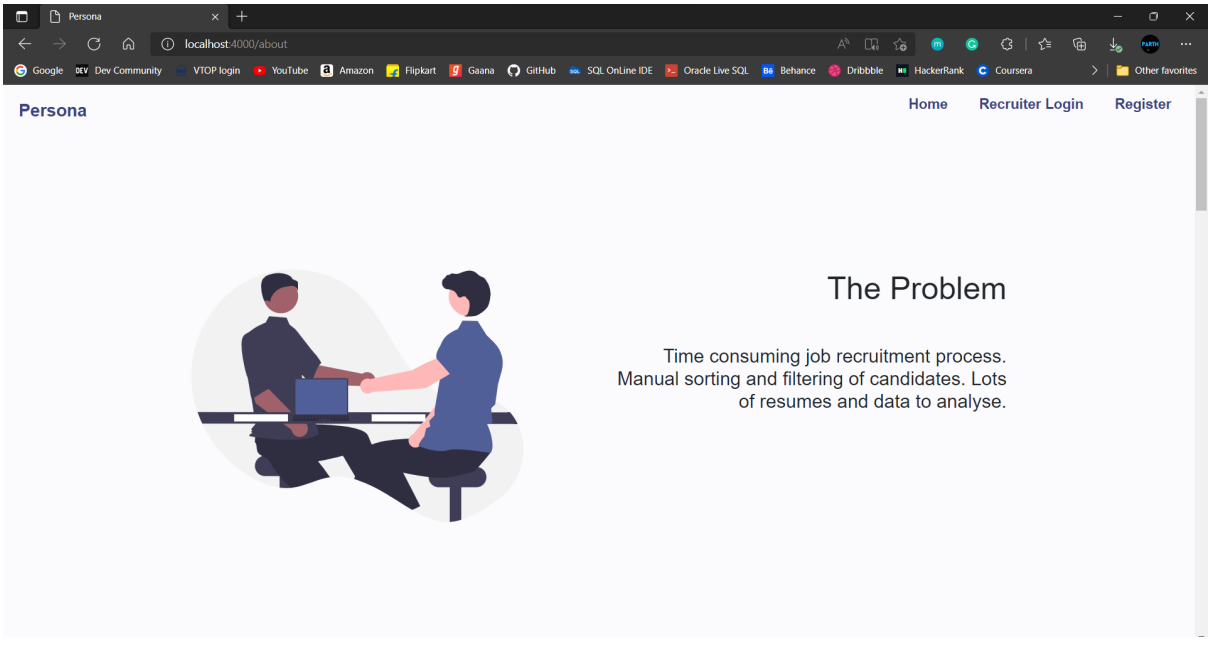
7.3 Test cases

Test case	Objective	Pre-requisite	Steps	Output	Status
TC1	Index/Home page of the software/webapp	Flask server must be started	Open browser and fetch “/” in the url	Home page is displayed	PASS





Test case	Objective	Pre-requisite	Steps	Output	Status
TC2	Opening Learn More page	Learn more button should be present	Open home page and click on learn more button in the left side pan.	Learn more page is displayed It shows about page, scoring criteria and personalities description	PASS



What is PERSONA?

Efficient screening system for the companies to recruit manpower resources. It helps the job recruiters to categorise their candidates based on tests, questionnaires and resume.

Here you go

A questionnaire that would score candidates' five personality tests based on the Big Five Model - OCEAN. Then based on the score, the model will predict the personality of the person.

Scoring Criteria

Openness

Openness is a characteristic that includes imagination and insight. The world, other people and an eagerness to learn and experience new things is particularly high for this personality trait. It leads to having a broad range of interests and being more adventurous when it comes to decision making.

Neuroticism

Neuroticism is characterised by sadness, moodiness, and emotional instability. Often mistaken for anti-social behaviour, or worse a greater psychological issue, neuroticism is a physical and emotional response to stress and perceived threats in someone's daily life.

Conscientiousness

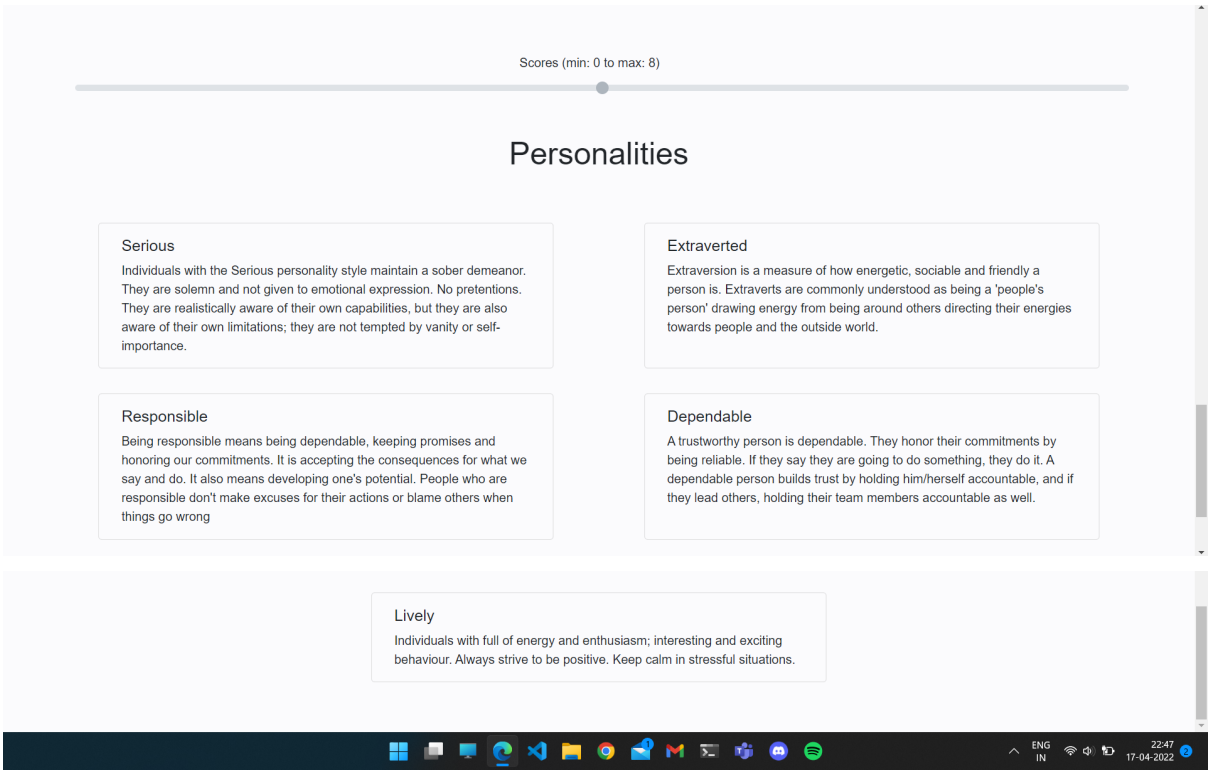
Conscientiousness is a trait that includes high levels of thoughtfulness, good impulse control, and goal-directed behaviours. This organised and structured approach is often found within people who work in science and even high-retail finance where detail orientation and organisation are required as a skill set.

Extraversion

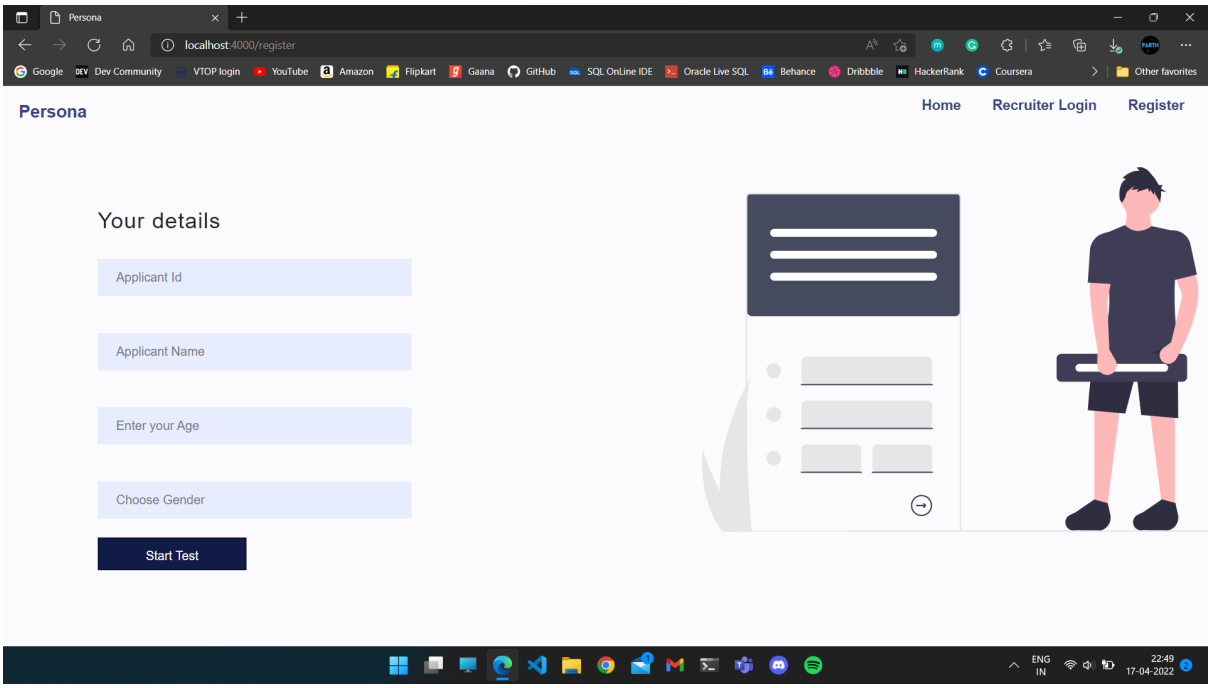
Extraversion (sometimes referred to as Extroversion) is a trait that many will have come across in their own lives. It's easily identifiable and widely recognisable as "someone who gets energised in the company of others."

Agreeableness

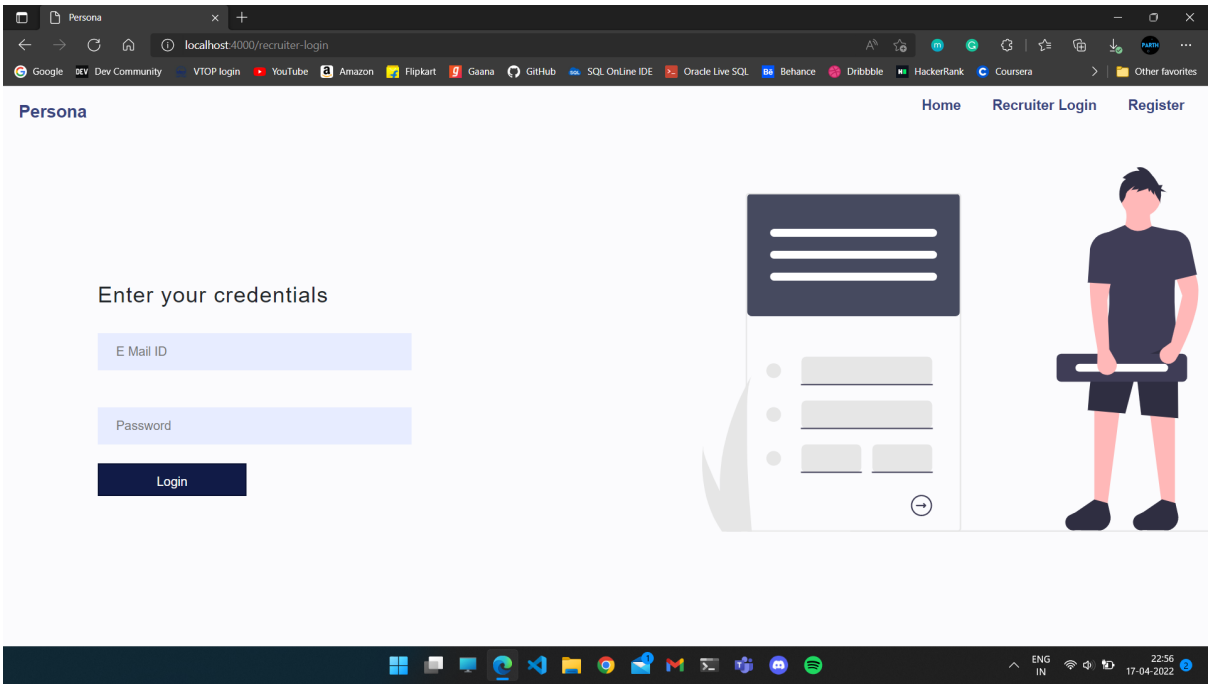
People who exhibit high agreeableness will show signs of trust, altruism, kindness, and affection. Highly agreeable people tend to have high prosocial behaviours which means that they're more inclined to be helping other people.

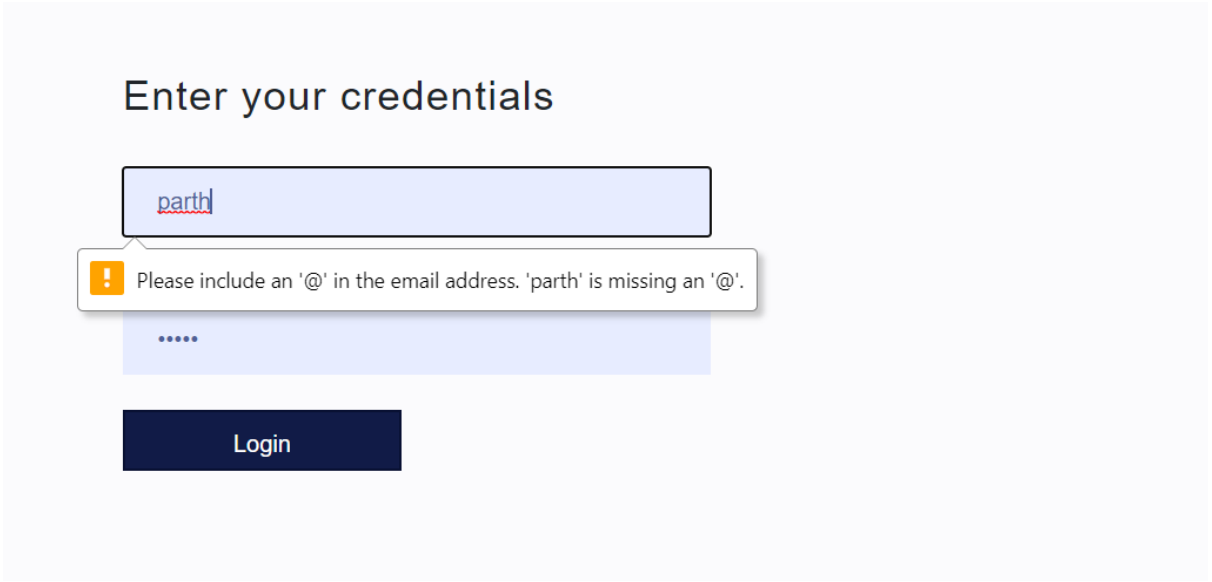


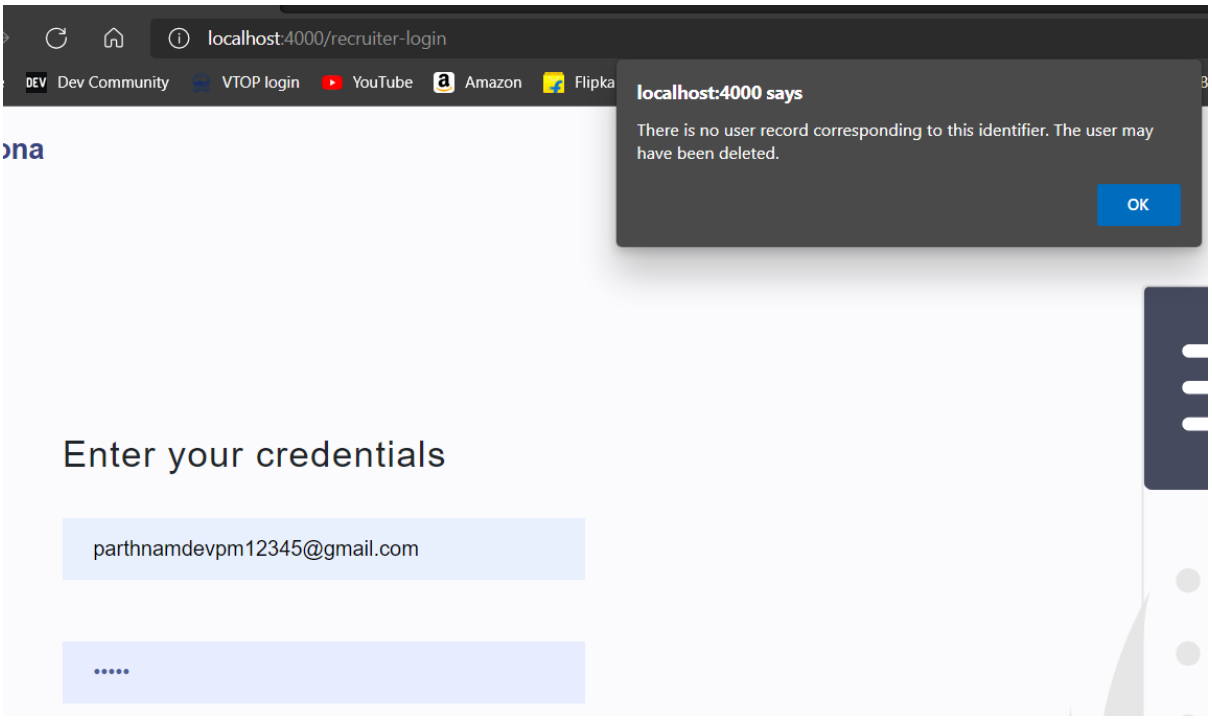
Test case	Objective	Pre-requisite	Steps	Output	Status
TC3	Candidate register page	Register button should be present	Open home page Click on register button in the left side pan or on the top right corner of home page.	Candidate register page is displayed It shows a form in the left pan and an SVG illustration on the right. Form includes four input field and a submit button	PASS



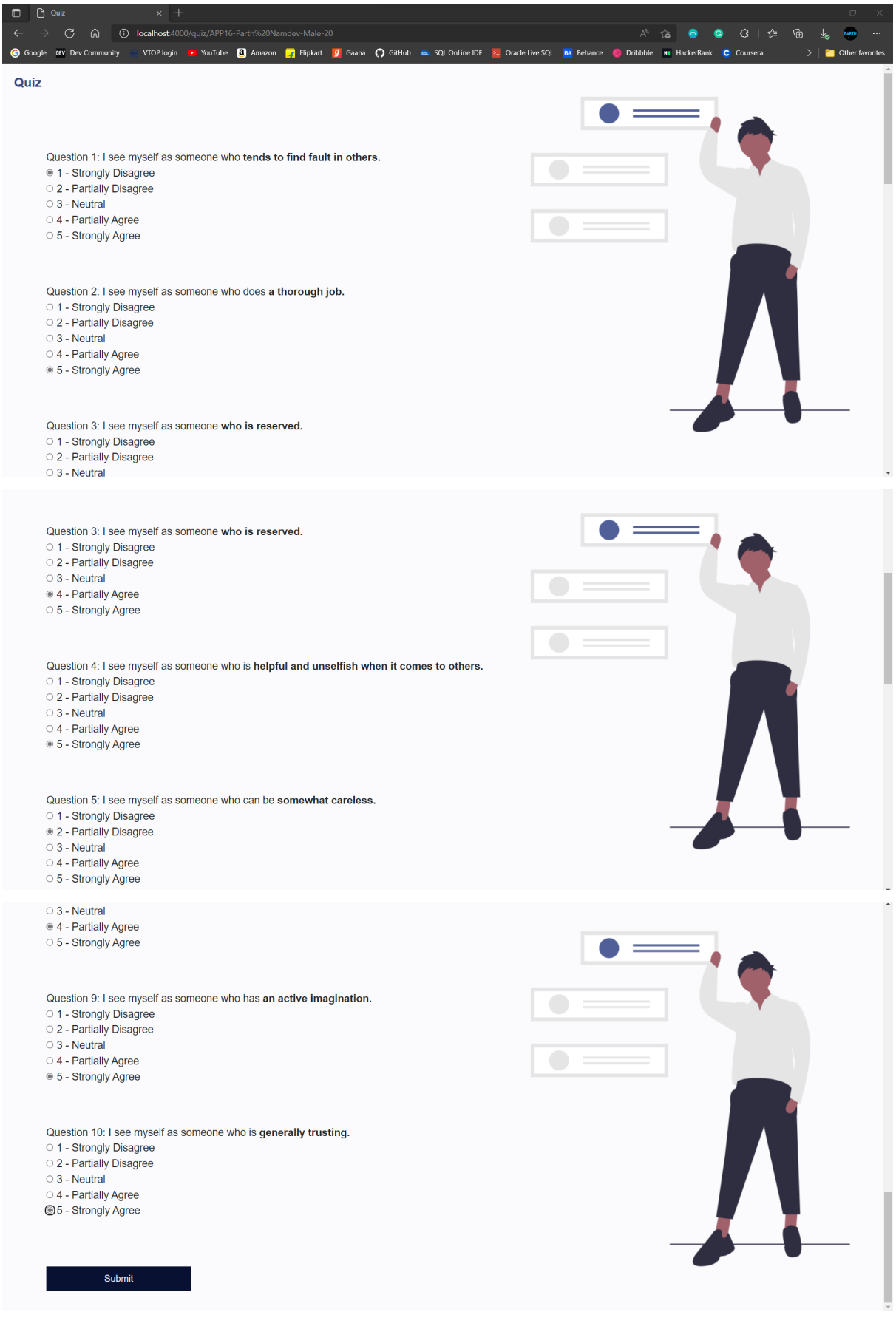
Test case	Objective	Pre-requisite	Steps	Output	Status
TC4	Recruiter login page Valid email and password	Register button should be present	Open home page Click on register button in the left side pan or on the top right corner of home page. Check with both valid and invalid emails Check with valid and invalid passwords	Candidate register page is displayed It shows a form in the left pan and an SVG illustration on the right. Form includes four input field and a submit button Shows alert for invalid email and password	PASS


Persona
localhost:4000/recruiter-login
Persona
Home
Recruiter Login
Register
Enter your credentials
E Mail ID
Password
Login

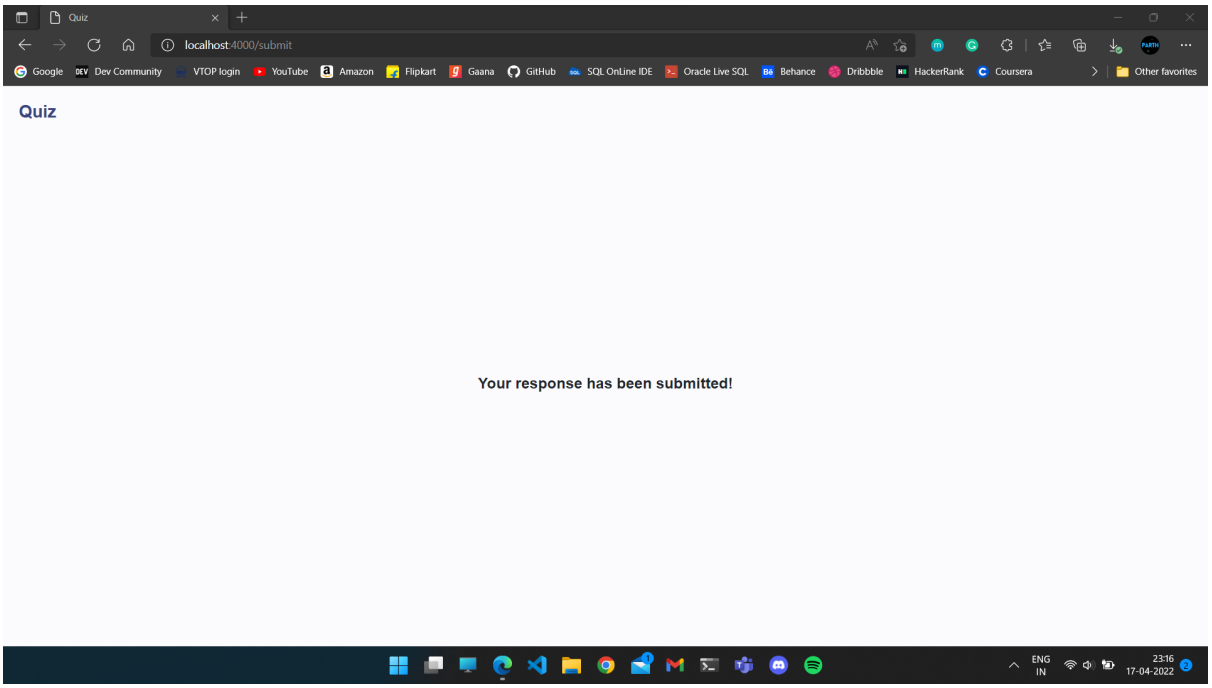

Enter your credentials
parth
Please include an '@' in the email address. 'parth' is missing an '@'.
Login



Test case	Objective	Pre-requisite	Steps	Output	Status
TC5	Quiz/Questionnaire page	User authentication via candidate register page	Register as a candidate Submit the form	Quiz/questionnaire page is displayed Left side includes 10 questions Right pan shows an SVG illustration All questions have 5 options and could be answered using radio buttons.	PASS

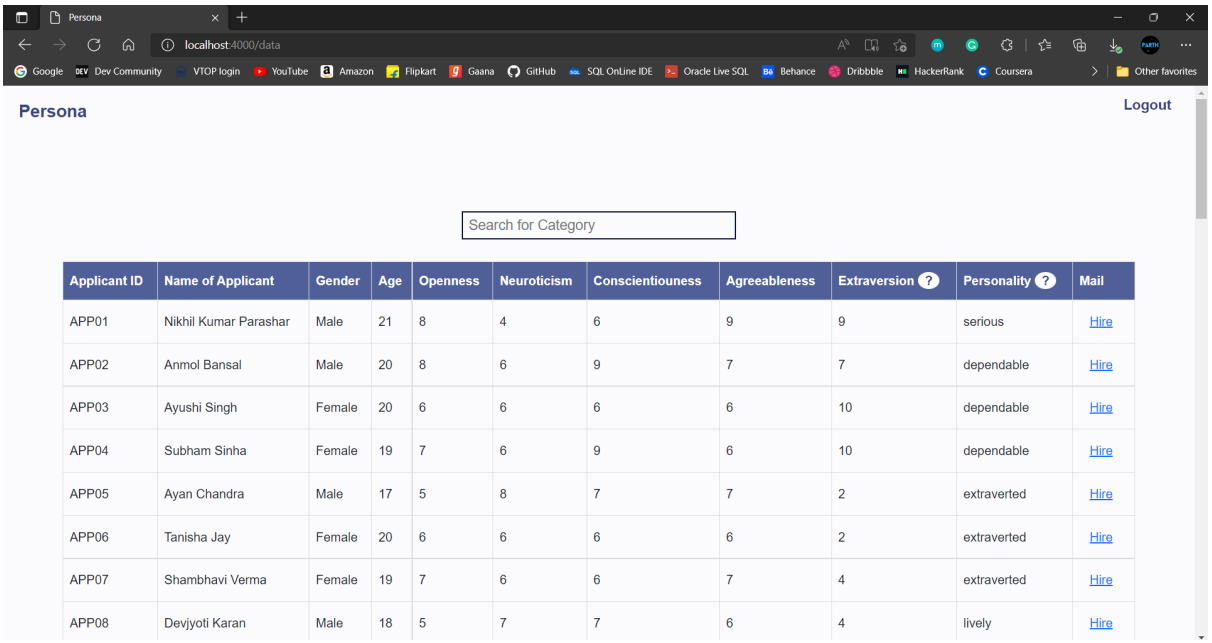

Quiz
localhost:4000/quiz/APP16-Parth%20Namdev-Male-20
Google
Dev Community
VTOP login
YouTube
Amazon
Flipkart
Gaana
GitHub
SQL OnLine IDE
Oracle Live SQL
Behance
Dribbble
HackerRank
Coursera
Other favorites
Quiz
Question 1: I see myself as someone who tends to find fault in others.
1 - Strongly Disagree
2 - Partially Disagree
3 - Neutral
4 - Partially Agree
5 - Strongly Agree
Question 2: I see myself as someone who does a thorough job.
1 - Strongly Disagree
2 - Partially Disagree
3 - Neutral
4 - Partially Agree
5 - Strongly Agree
Question 3: I see myself as someone who is reserved.
1 - Strongly Disagree
2 - Partially Disagree
3 - Neutral
Question 3: I see myself as someone who is reserved.
1 - Strongly Disagree
2 - Partially Disagree
3 - Neutral
4 - Partially Agree
5 - Strongly Agree
Question 4: I see myself as someone who is helpful and unselfish when it comes to others.
1 - Strongly Disagree
2 - Partially Disagree
3 - Neutral
4 - Partially Agree
5 - Strongly Agree
Question 5: I see myself as someone who can be somewhat careless.
1 - Strongly Disagree
2 - Partially Disagree
3 - Neutral
4 - Partially Agree
5 - Strongly Agree
3 - Neutral
4 - Partially Agree
5 - Strongly Agree
Question 9: I see myself as someone who has an active imagination.
1 - Strongly Disagree
2 - Partially Disagree
3 - Neutral
4 - Partially Agree
5 - Strongly Agree
Question 10: I see myself as someone who is generally trusting.
1 - Strongly Disagree
2 - Partially Disagree
3 - Neutral
4 - Partially Agree
5 - Strongly Agree
Submit

Test case	Objective	Pre-requisite	Steps	Output	Status
TC6	Response submitted page	Quiz to be submitted	Regiter and open quiz page Answer all questions Submit the form	Success/submitted message displayed in center of the page	PASS



Test case	Objective	Pre-requisite	Steps	Output	Status
TC7	Candidates' data page	Recruiter login/authentication	Open recruiter login page Input valid credentials Submit form To filter candidates, a search category in the field given	A data table is displayed that includes or shows all candidates' data It has all five big traits' scores and personality category for each	PASS

			To send mail for hiring, press corresponding button	candidate Buttons for mail/hire is given in the last column of each row Know more button given for scoring criteria and personality categorisation/classification	



Applicant ID	Name of Applicant	Gender	Age	Openness	Neuroticism	Conscientiouness	Agreeableness	Extraversion	Personality	Mail
APP01	Nikhil Kumar Parashar	Male	21	8	4	6	9	9	serious	Hire
APP02	Anmol Bansal	Male	20	8	6	9	7	7	dependable	Hire
APP03	Ayushi Singh	Female	20	6	6	6	6	10	dependable	Hire
APP04	Subham Sinha	Female	19	7	6	9	6	10	dependable	Hire
APP05	Ayan Chandra	Male	17	5	8	7	7	2	extraverted	Hire
APP06	Tanisha Jay	Female	20	6	6	6	6	2	extraverted	Hire
APP07	Shambhavi Verma	Female	19	7	6	6	7	4	extraverted	Hire
APP08	Devjyoti Karan	Male	18	5	7	7	6	4	lively	Hire
APP09	Vidushi Gupta	Female	19	6	4	6	3	4	responsible	Hire
APP10	Komal Kayastha	Female	19	7	7	2	6	5	serious	Hire
APP11	Madan Lal	Male	19	3	7	2	6	7	responsible	Hire
APP12	Aryan Pandey	Male	19	5	4	6	1	3	lively	Hire
APP13	Manisha Pandey	Female	47	5	3	6	8	9	serious	Hire
APP14	Ranjan Kumar	Male	50	6	4	9	10	10	serious	Hire
APP15	Nikhil Ranjan	Male	26	6	6	8	4	8	serious	Hire
APP16	Parth Namdev	Male	20	7	4	9	10	9	serious	Hire

Scoring Criteria

Openness

Openness is a characteristic that includes imagination and insight. The world, other people and an eagerness to learn and experience new things is particularly high for this personality trait. It leads to having a broad range of interests and being more adventurous when it comes to decision making.

Neuroticism

Neuroticism is characterised by sadness, moodiness, and emotional instability. Often mistaken for anti-social behaviour, or worse a greater psychological issue, neuroticism is a physical and emotional response to stress and perceived threats in someone's daily life.

Conscientiousness

Conscientiousness is a trait that includes high levels of thoughtfulness, good impulse control, and goal-directed behaviours. This organised and structured approach is often found within people who work in science and even high-retail finance where detail orientation and organisation are required as a skill set.

Extraversion

Extraversion (sometimes referred to as Extroversion) is a trait that many will have come across in their own lives. It's easily identifiable and widely recognisable as "someone who gets energised in the company of others."

Agreeableness

People who exhibit high agreeableness will show signs of trust, altruism, kindness, and affection. Highly agreeable people tend to have high prosocial behaviours which means that they're more inclined to be helping other people.

Scores (min: 0 to max: 10)

Personalities

Serious

Individuals with the Serious personality style maintain a sober demeanor. They are solemn and not given to emotional expression. No pretentions. They are realistically aware of their own capabilities, but they are also aware of their own limitations; they are not tempted by vanity or self-importance.

Extraverted

Extraversion is a measure of how energetic, sociable and friendly a person is. Extraverts are commonly understood as being a 'people's person' drawing energy from being around others directing their energies towards people and the outside world.

Responsible

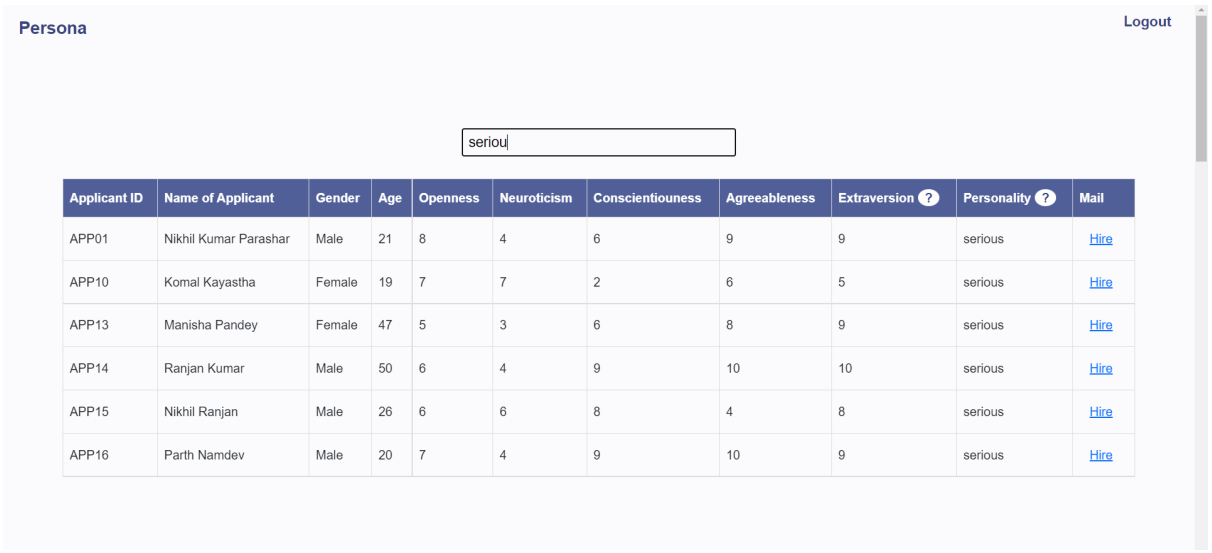
Being responsible means being dependable, keeping promises and honoring our commitments. It is accepting the consequences for what we say and do. It also means developing one's potential. People who are responsible don't make excuses for their actions or blame others when things go wrong

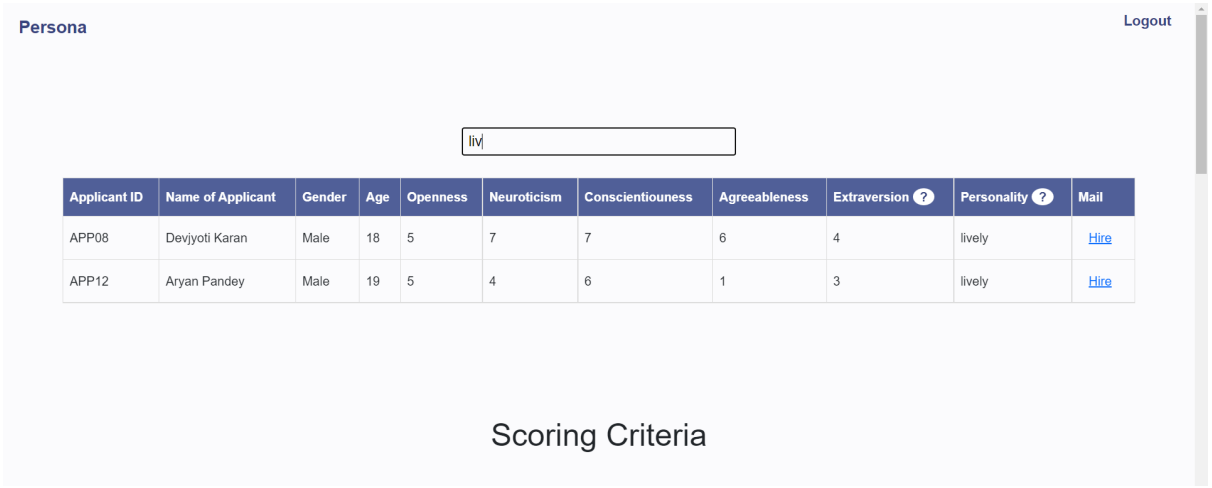
Dependable

A trustworthy person is dependable. They honor their commitments by being reliable. If they say they are going to do something, they do it. A dependable person builds trust by holding him/herself accountable, and if they lead others, holding their team members accountable as well.

Lively

Individuals with full of energy and enthusiasm; interesting and exciting behaviour. Always strive to be positive. Keep calm in stressful situations.





Test case	Objective	Pre-requisite	Steps	Output	Status
TC8	Answering all questions as neutral	Quiz page open	Submit all answers as neutral	Submit page is displayed The personality category shows “serious”. The person thinks deeply for all different cases, hence marked as “serious”.	PASS

8. Conclusion and Future Enhancements

We believe that our model can be helpful for recruiters around the globe to identify the perfect candidate for their job in a quick manner. Logistic Regression and the Big 5 Model can be useful in identifying the personality of an individual. Further, we would like to enhance the model by incorporating a portal for job seekers to identify their strengths and weaknesses and also find other applications of our Personality Prediction Model - Persona.

Team Members Contribution

Registration Number	**Name**	**Contribution / Role in this Project**
19BCE0440	Namdev Parth Deendayal	Backend
19BCE0076	Nikhil Kumar Parashar	ML
19BCE2678	Vishakha Kumaresan	Frontend
20BCE0888	Naman Chadha	Frontend
20BCE2736	Yatheendra Nath Reddy	Data Collection

References

[1] M. M. Tadesse, H. Lin, B. Xu and L. Yang, "Personality Predictions Based on User Behavior on the Facebook Social Media Platform," in IEEE Access, vol. 6, pp. 61959-61969, 2018, doi: 10.1109/ACCESS.2018.2876502.

[2] Maliki, I., & Sidik, M. A. (2020). Personality prediction system based on signatures using machine learning. *IOP Conference Series: Materials Science and Engineering*, *879*(1), 012068. https://doi.org/10.1088/1757-899x/879/1/012068

[3] Hetal Vora , Mamta Bhamare , Dr. K. Ashok Kumar, 2020, Personality Prediction from Social Media Text: An Overview, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 09, Issue 05 (May 2020)

[4] Rout, Jayashree & Bagade, Sudhir & Yede, Pooja & Patil, Nirmiti. (2019). Personality Evaluation and CV Analysis using Machine Learning Algorithm. International Journal of Computer Sciences and Engineering. 7. 1852-1857. 10.26438/ijcse/v7i5.18521857.

[5] Ombhase, Manasi & Gogate, Prajakata & Patil, Tejas & Nair, Karan & Hegde, Prof. (2017). Automated Personality Classification Using Data Mining Techniques. 10.13140/RG.2.2.35949.59363.

[6] Dehghanan, Hamed & Abdollahi, Hadi & Rezaei, Mohammad. (2014). A study on effect of big five personality traits on emotional intelligence. Management Science Letters. 4. 10.5267/j.msl.2014.4.016.

[7] Varvel, Tricia & Adams, Stephanie & Pridie, Shelby. (2003). A Study of the Effect of the Myers-Briggs Type Indicator on Team Effectiveness. ASEE Annual Conference Proceedings.

[8] Ali, Jehad & Khan, Rehanullah & Ahmad, Nasir & Maqsood, Imran. (2012). Random Forests and Decision Trees. International Journal of Computer Science Issues(IJCSI). 9.

[9] Evgeniou, T., Pontil, M. (2001). Support Vector Machines: Theory and Applications. In: Paliouras, G., Karkaletsis, V., Spyropoulos, C.D. (eds) Machine Learning and Its Applications. ACAI 1999. Lecture Notes in Computer Science(), vol 2049. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-44673-7_12

[10] Durupinar, F., Pelechano, N., Allbeck, J. M., Gudukbay, U., & Badler, N. I. (2011). How the ocean personality model affects the perception of crowds. IEEE Computer Graphics and Applications, 31(3), 22–31. https://doi.org/10.1109/mcg.2009.105

[11] Natekin, Alexey & Knoll, Alois. (2013). Gradient Boosting Machines, A Tutorial. Frontiers in neurorobotics. 7. 21. 10.3389/fnbot.2013.00021.

[12] Rohani, Afrina R., "A Survey of the Big Five Personality Traits Among Elementary Teachers" (2017). Honors Undergraduate Theses. 257

[13] Aslam, Fankar & Mohammed, Hawa & Lokhande, Prashant. (2015). Efficient Way Of Web Development Using Python And Flask.. International Journal of Advanced Research in Computer Science. 6.

[14] Randles, B. M., Pasquetto, I. V., Golshan, M. S., & Borgman, C. L. (2017). Using the Jupyter notebook as a tool for Open science: An empirical study. *2017 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*. https://doi.org/10.1109/jcdl.2017.7991618

[15] Rish, Irina. (2001). An Empirical Study of the Naïve Bayes Classifier. IJCAI 2001 Work Empir Methods Artif Intell. 3.

[16] Kaviani, Pouria & Dhotre, Sunita. (2017). Short Survey on Naive Bayes Algorithm. International Journal of Advance Research in Computer Science and Management. 04.

[17] Song, Shuang & Chaudhuri, Kamalika & Sarwate, Anand. (2013). Stochastic gradient descent with differentially private updates. 2013 IEEE Global Conference on Signal and Information Processing, GlobalSIP 2013 - Proceedings. 245-248. 10.1109/GlobalSIP.2013.6736861.