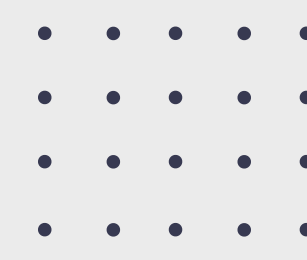




Mini Project - I


Predicting Big Five Personality Traits from Facial Features using ML

Under the guidance of
-Dr. Chinmayananda A





Index

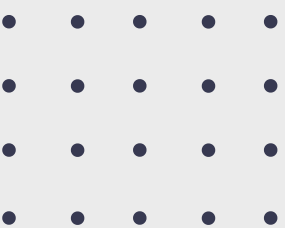
- Introduction
 - Literature study
 - Enhanced Personality Contribution Assessment
 - Problem Statement
 - Progress
 - Tools used
 - Data set
 - Working
 - Preparation of training dataset
 - Demo
 - Final Outcome
 - Challenges Faced
- 



01

Introduction

- The human face is a powerful tool that conveys a wide range of emotions and nonverbal cues. Advances in machine learning algorithms and image processing techniques have made it possible to analyze facial features with greater accuracy and speed than ever before. In this project, we aim to leverage these techniques to predict personality traits based on facial features.
- We will use a dataset of facial images along with self-reported personality assessments to train our machine learning algorithm. Our goal is to develop an accurate and reliable system that can predict personality traits based solely on facial features.





02


Literature Study

"2.5D Facial Personality Prediction Based on Deep Learning" by

Jia Xu, Weijian Tian, Guoyun Lv, Shiya Liu, and Yangyu Fan(2020)

Dataset : The personality dataset included 5,560 male and 8,547 female college students aged 18 to 25, measuring the Big Five traits with scores ranging from 0 to 60 for each dimension, where higher scores indicate a more easygoing and pleasant personality [21, 22].

Details : This paper utilized deep learning algorithms, specifically CNN, to analyze and extract features from 2.5D facial images. They used MobileNetV2 and residual network version 50 (ResNet50), two deep learning networks that are popular in academia. To verify the experimental results, 5-fold cross-validation method was used.






03

Enhanced Personality Contribution Assessment


Our project sets itself apart as a groundbreaking endeavor in the field of personality prediction from facial features. While drawing inspiration from existing research, we have pushed the boundaries by introducing an innovative machine learning model. What truly makes our project exceptional is its ability to not only accurately predict personality traits but also provide an unparalleled level of transparency and interpretability. By offering the percentage or weighted contribution of each facial feature type for different personalities, our project unveils the intricate nuances of how facial cues influence individual traits.





04

Problem Statement

- The ability to predict someone's personality can have numerous applications, ranging from personalized recommendations to identifying mental health issues.
 - Research has focused on various approaches, including self-assessments, behavioral observations, and physiological measurements. However, these approaches have their limitations, and their accuracy is often questioned
 - In this project, we aim to explore the relationship between facial features and personality traits and develop a machine learning algorithm to predict personality based on facial features.
- 



05 Progress

- **PHASE - I [22 DEC - 1 JAN]**

Topic Research

- **PHASE - II [2 JAN - 10 FEB]**

Learning ML & Deep Learning

- **PHASE - III [11 FEB - 30 APR]**

Data collection and implementation



06 Tools used



Google Colaboratory

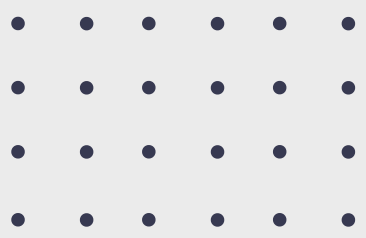


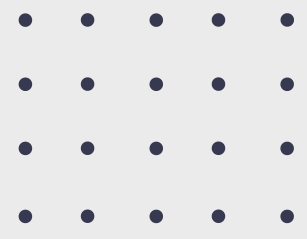


07 Dataset

1	Eyebrows				Forehead		Face Shape		Cheek		Nose				PERSONALITY
2	Long	Short	Thick	Thin	Broad	Tiny	Long	wide	Fleshy	Normal	Long	Short	Sharp	Blunt	
3	1	0	1	0	1	0	1	1	0	1	1	0	0	1	1
4	1	0	1	0	0	1	1	0	0	1	1	0	1	0	2
5	0	1	0	1	1	0	0	1	1	0	1	0	0	1	2
6	0	1	0	1	1	0	0	1	0	1	0	1	0	1	0
7															
8	1	0	1	0	0	1	1	1	0	1	1	0	0	1	0
9	1	0	0	1	1	0	0	1	1	0	1	0	1	0	2
10	0	1	1	0	1	0	0	1	1	0	0	1	0	1	4
11	1	0	1	0	0	1	0	1	1	0	0	1	0	1	0
12															
13	1	0	1	0	0	1	1	0	1	0	0	1	0	1	2
14	0	1	0	1	1	0	1	0	0	1	1	0	0	1	0
15	1	0	0	1	1	0	0	0	1	0	0	1	0	1	0
16	1	0	1	0	1	0	0	1	1	0	0	1	0	1	0
17	1	0	1	0	0	1	0	0	0	1	1	0	0	1	2
18	0	1	1	0	1	0	1	0	0	1	1	0	1	0	3

We collected primary data from college students using Google Forms. A total of 138 volunteers participated, aged between 18 to 25 years, with 7 females and the rest males. The form had 38 questions related to five personality traits and an image section where volunteers uploaded their frontal facial image without glasses. Each question was rated on a scale of 0 to 5. The data is stored on a Google Drive accessible only to team members to ensure privacy. However, since there was no invigilator present during data collection, there may be potential biases in the data.





08 How personalities are Assigned :

Steps :

- After collecting the response for the questions we manually divide all the questions into five categories(which question belongs to which category) and calculate overall score for each person using the shown formulas and assign him/her a personality on the basis of the score obtained.
- After that, we analyze every image and extract its five features: eyebrows, nose, forehead, face and cheeks.

Formula For Calculating Score

A	B
Personality Type	Formula to get personality score
Openness (F3)	$(Q3+Q10+Q14+Q16+Q22+Q24+Q38)/7$
Extraversion (F1)	$(Q5+Q7+Q12+Q18+Q23+Q26+(6-Q31)+Q35)/8$
Agreeableness (F4)	$(Q2+Q9+(6-Q11)+(6-Q17)+Q21+Q29+Q32+Q34+Q36)/9$
Neuroticism (F2)	$(Q4+Q6+Q13+Q20+Q25+(6-Q28)+Q30+Q37)/8$
Conscientiousness (F5)	$(Q1+Q8+Q15+Q19+(6-Q27)+Q33)/6$
	Qi is the score that you get for ith question from the questionnaire

Features	Categories
Eyebrows	Long, Short, Thick, Thin
Forehead	Broad, Tiny
Face	Long, Wide
Cheeks	Fleshy, Normal
Nose	<u>Long</u> , Short, Sharp, and Blunt



09

How the Algorithm works :

- For training purpose, we created two NumPy array , one contains all the features(2D array) and other contains different personalities (1D array) and then split whole data set into 3:2 ratio, 3/5th portion for training purpose and remaining 2/5th portion for testing purposes using 5 different algorithms .
- Then we compared the accuracies of all the training models and took KNN classifier as best fit model with accuracy of approx. 46%.

10





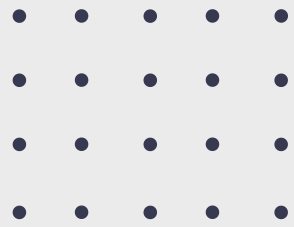
11

Final Outcome

- The model was able to predict the personality with moderate accuracy, which is essential for self-awareness and self-development, psychological assessment and therapy, etc.
- The accuracy of our personality prediction model stands at 43%, reflecting its effectiveness in capturing and predicting personality traits with a moderate level of precision.
- Additionally, our model provides the percentage contribution of each face part to the overall prediction of personality traits.

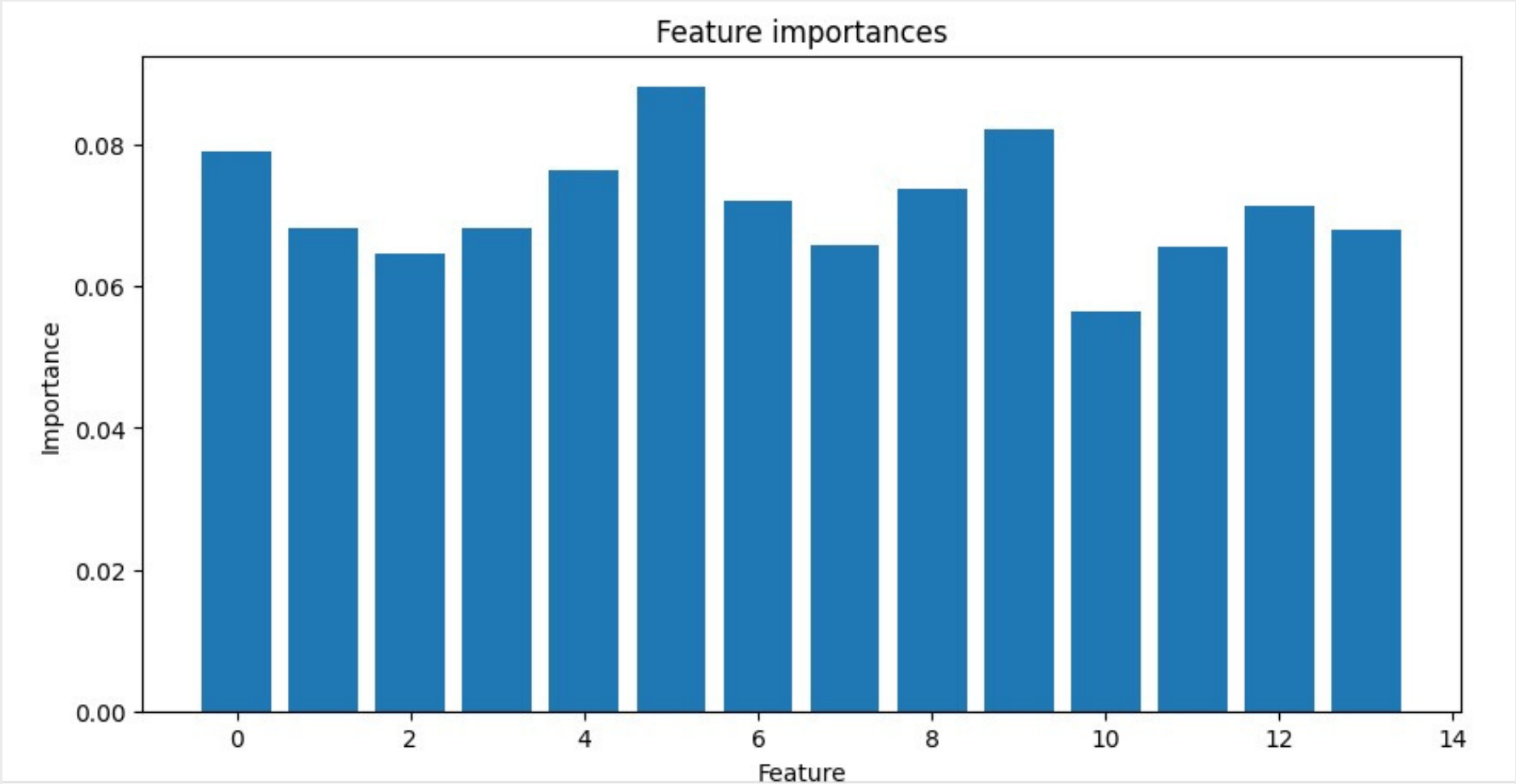
Contribution of each facial features to personality

	long(Eyebrows)	thick(Eyebrows)	broad(Forehead)	Long(Face_shape)	fleshy(Cheek)	Long(Nose)	Sharp(Nose)
openness	0.543478	0.565217	0.630435	0.500000	0.543478	0.608696	0.391304
conscientiousness	0.461538	0.653846	0.615385	0.538462	0.461538	0.730769	0.230769
extraversion	0.658537	0.512195	0.634146	0.487805	0.585366	0.609756	0.536585
agreeableness	0.363636	0.454545	0.545455	0.363636	0.545455	0.545455	0.363636
neuroticism	0.375000	0.875000	0.875000	0.500000	0.500000	0.375000	0.375000

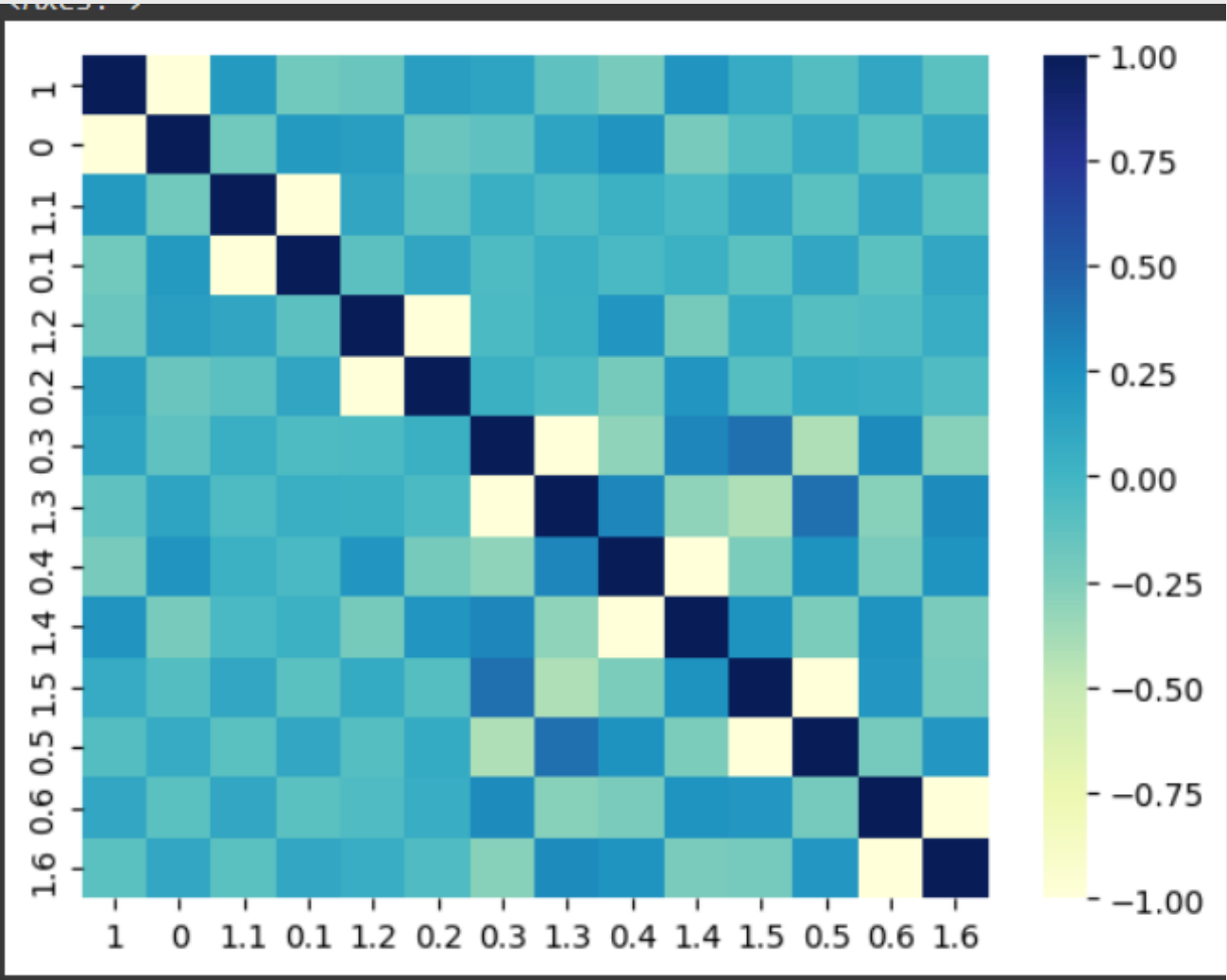


13

Final Outcome



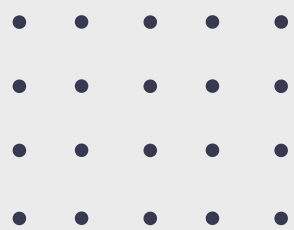
Feature Importance



Heat Map

SVM

KNN



15

Final Outcome

Classification reports

	precision	recall	f1-score	support
Openness	0.50	0.53	0.51	19
Conscientiousness	0.22	0.20	0.21	10
Extraversion	0.47	0.47	0.47	17
Agreeableness	0.33	0.25	0.29	4
Neuroticism	0.25	0.33	0.29	3
accuracy			0.42	53
macro avg	0.36	0.36	0.35	53
weighted avg	0.41	0.42	0.41	53

Gradient Boosting

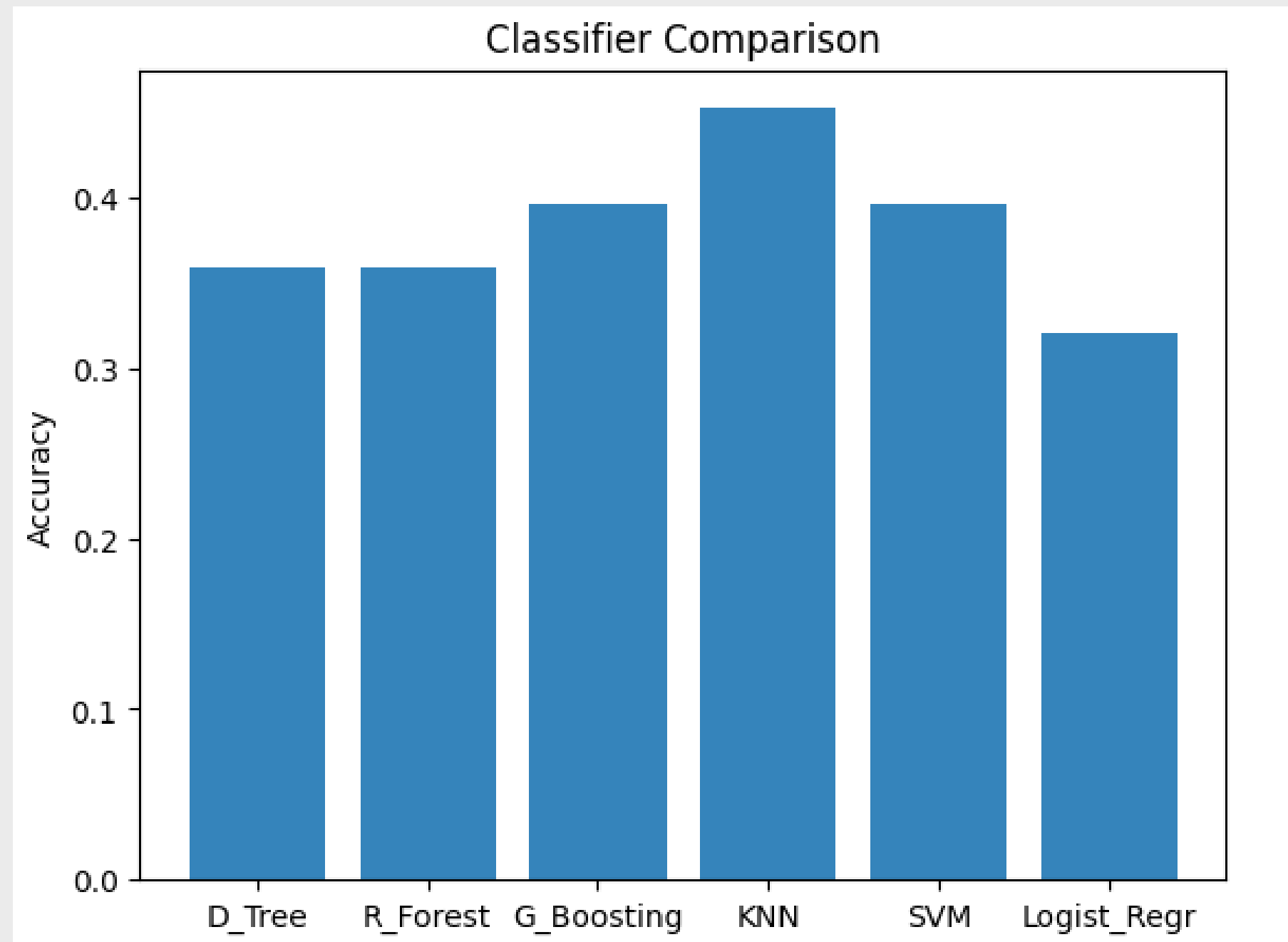
	precision	recall	f1-score	support
Openness	0.46	0.58	0.51	19
Conscientiousness	0.29	0.20	0.24	10
Extraversion	0.53	0.53	0.53	17
Agreeableness	0.00	0.00	0.00	4
Neuroticism	0.00	0.00	0.00	3
accuracy			0.42	53
macro avg	0.25	0.26	0.26	53
weighted avg	0.39	0.42	0.40	53

Decision Tree

Random Forest

	precision	recall	f1-score	support
Openness	0.55	0.63	0.59	19
Conscientiousness	0.42	0.50	0.45	10
Extraversion	0.47	0.47	0.47	17
Agreeableness	0.00	0.00	0.00	4
Neuroticism	0.00	0.00	0.00	3
accuracy			0.47	53
macro avg	0.29	0.32	0.30	53
weighted avg	0.43	0.47	0.45	53

Classifier Comparison



18

Challenges Faced

- Low accuracy rate of 43 percent due to limited dataset size.
- Lack of expertise in facial feature extraction affecting the precision of feature extraction.
- Absence of a personality expert impacting the accuracy of the questionnaire design.
- Questionnaire not filled in the presence of an invigilator, potentially leading to errors in responses.
- Missing facial pictures from some volunteers resulting in reduced sample size and potential bias.



Thank You

Team:
Jayant Kumawat
Jay Prakash
Neha Porwal
Aryan Patel

