

PERSONIPHY

CSD 334 Mini Project

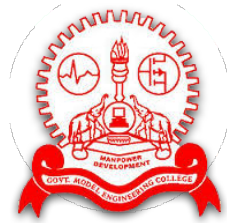
MDL21CS017 Amil P A

MDL21CS042 D Krishnadas

MDL21CS066 Joel G Antony

MDL21CS070 Karthik G Kumar

B. Tech. Computer Science & Engineering



Department of Computer Engineering
Model Engineering College, Thrikkakara
Kochi, 682021

Phone: +91.484.2575370

<http://www.mec.ac.in>

hodcs@mec.ac.in

2024

**Model Engineering College, Thrikkakara,
Dept. of Computer Engineering**



C E R T I F I C A T E

This is to certify that, this report titled ***PERSONIPHY*** is a bonafide record of the **CSD 334** Mini Project presented on May 06, 2024 by

MDL21CS017 Amil P A

MDL21CS042 D Krishnadas

MDL21CS066 Joel G Antony

MDL21CS070 Karthik G Kumar

Sixth Semester B.Tech. Computer Science and Engineering

scholar, under our guidance and supervision, in partial fulfillment of the requirements for the award of the degree, **B. Tech. Computer Science and Engineering** of **APJ Abdul Kalam Technological University**.

Guide

Coordinator

Dr. Sony P
Assistant Professor
Department of Computer Engineering

Mr. Jibi J Puthiyedom
Asst. Professor
Department of Computer Engineering

Head of the Department

Dr. Binu V P
Associate Professor in Computer Engineering

Acknowledgment

This seminar would not have been possible without the kind support and help of many individuals. I would like to extend my sincere thanks to all of them.

First of all, I would like to thank our esteemed Principal, Prof. (Dr.) Mini M G, for her guidance and support in maintaining a calm and refreshing environment to work in, and also, for providing the facilities that this work demanded.

I thank the Head of the Department, Dr. Binu V P, Department of Computer Engineering, Model Engineering College, for all the support and also for providing the facilities the work demanded.

I am highly indebted to our mini project coordinator, Mr. Jibi J Puthiyedom, Asst. Professor, Department of Computer Engineering, Model Engineering College, for his guidance, support and constant supervision throughout the duration of the work as well as for providing all the necessary information regarding the mini project.

I would like to thank my mini project guide Dr. Sony P, Assistant Professor, Department of Computer Engineering, Model Engineering College, for her guidance, support and valuable insights and suggestions, and also for helping me out in correcting any mistakes that were made during the course of the work.

I offer my sincere gratitude to all my friends and peers for their support and encouragement, that helped me get through the tough phases during the course of this work.

Last but not least, I thank the Almighty God for guiding me through, and enabling me to complete the work within the specified time.

Abstract

Personify is a software solution designed to revolutionize human resources by providing data-driven insights into employee personalities. Through a tailored questionnaire approach, Personify offers in-depth analysis of individual traits, including those identified by the Big Five personality framework (OCEAN). This information empowers companies to optimize hiring practices, fostering a culture of synergy and growth. Personify's core functionalities target improved team dynamics through enhanced communication and collaboration based on personality compatibility. Additionally, employers can leverage these insights to customize job roles and responsibilities, maximizing individual potential and employee satisfaction. By deciphering the intricacies of personalities and harnessing this knowledge, Personify paves the way for a more cohesive and productive work environment.

Contents

1	Introduction	iv
2	Report of Preparatory Work	v
2.1	Literature Survey Report	v
2.1.1	Big Five personality traits in the workplace: Investigating personality differences between employees, supervisors, managers, and entrepreneurs [1]	v
2.1.2	Cluster Analysis of Personality Types Using Respondents' Big Five Personality Traits [2]	vi
2.1.3	The Development of Markers for the Big-Five Factor Structure [3]	vi
2.2	System Study Report	vii
3	Proposed Project	viii
3.1	Problem Statement	viii
3.2	Proposed Solution	viii
4	Project Design	ix
4.1	Data Design	x
4.1.1	Level 0 data flow diagram	x
4.1.2	Level 1 data flow diagram	x
4.2	Architecture Diagram	xii
4.3	Component Level Design	xii
4.3.1	User Interface Module	xii
4.3.2	Clustering Module	xiii
4.3.3	Hardware and Software Requirements	xv
5	Experimental Results and Analysis	xvi
5.1	Data Preparation	xvi
5.2	Output	xvii
6	Conclusion	xviii

List of Figures

4.1	Overall framework	ix
4.2	Level 0 DFD	x
4.3	Level 1 DFD	xi
4.4	Architecture Diagram	xii
4.5	User-interface Module	xiii
4.6	Clustering Module	xiv
5.1	Dataset Example	xvi
5.2	User Side Result	xvii
5.3	HR side result	xvii

List of Tables

4.1	Hardware and Software Requirements	xv
-----	--	----

Chapter 1

Introduction

The ever-changing landscape of the workplace demands innovative solutions for building high-performing teams and fostering a thriving company culture. Traditional hiring methods heavily rely on resumes and interviews, offering a limited view of a candidate's personality. This can lead to mismatched placements where personalities clash, hindering team dynamics and overall productivity. HR departments lack a systematic approach to understanding employee personalities, making it difficult to build well-rounded teams and create a work environment that caters to diverse personalities. This is where enter Personify, a visionary software platform designed to redefine how companies understand, engage, and empower their employees. With Personify, organizations embark on a journey of discovery, utilizing a tailored questionnaire-based approach to comprehensively understand the personalities of their workforce. From bustling startups to multinational enterprises, companies harness the power of Personify to cultivate harmonious team dynamics and foster a culture of collaboration and innovation. By providing in-depth insights into individual personality traits, the platform becomes a trusted ally for HR departments seeking to refine their hiring processes and optimize workplace efficiency. Drawing on renowned personality assessment models such as the Big Five (OCEAN), Personify empowers companies to make data-driven decisions throughout the employee lifecycle. From recruitment and job customization to performance evaluation and career development, the platform serves as a beacon of insight and guidance in the fast-paced world of talent management.

Chapter 2

Report of Preparatory Work

2.1 Literature Survey Report

2.1.1 Big Five personality traits in the workplace: Investigating personality differences between employees, supervisors, managers, and entrepreneurs [1]

This study investigates the relationship between personality traits and employment status, focusing on entrepreneurs, managers, supervisors, and employees. Using data from the UK Household Longitudinal Study, the researchers employed multivariate and univariate ANOVA analyses to explore differences in the Big Five personality traits across these groups. The results reveal significant differences in Neuroticism, Openness, Agreeableness, Conscientiousness, and Extraversion.

Advantages:

A comprehensive analysis explored how personality traits vary across different employment statuses, ensuring findings' robustness with a large dataset. The study revealed significant differences in traits among those in full-time, part-time, and unemployed roles, highlighting the impact of employment on personality characteristics. By leveraging extensive data, the research provided valuable insights for employers, policymakers, and individuals navigating work-personality dynamics.

Disadvantages:

The study's cross-sectional design hinders causal relationship establishment, while the broad measurement of employment status overlooks intra-occupational differences. Additionally, the exclusive focus on the Big Five traits neglects other pertinent characteristics like emotional intelligence or temperament.

2.1.2 Cluster Analysis of Personality Types Using Respondents' Big Five Personality Traits [2]

This study employed a mixed model approach, integrating k-means clustering analysis for data examination, discriminant analysis for classification, and a multilayer perceptron (MLP) neural network for prediction. The dataset consisted of 19,692 observations collected through an interactive online personality test focusing on the Big Five Personality Traits in 2012. The main finding of the study was that the four identified personality types differed in terms of age, gender, race, English-speaking versus non-English-speaking, right-hand-writing versus left-hand-writing, information source, and the Big Five Personality Traits. Given the nature of the dataset, the study's results can serve as a reference for human personality classification

Advantages: Utilizing a comprehensive mixed model approach, the study combines k-means clustering analysis, discriminant analysis, and multilayer perceptron neural networks. This methodology achieves high accuracy in personality classification and prediction, affirming its effectiveness in the study.

Disadvantages: The study attains high accuracy in personality classification and prediction, underscoring the efficacy of its methodology. However, reliance on self-report data from a single online personality test may introduce biases and limit generalizability. Additionally, the exclusive focus on the Big Five traits overlooks other pertinent characteristics like emotional intelligence or temperament.

2.1.3 The Development of Markers for the Big-Five Factor Structure [3]

This study introduces a new set of Big-Five factor markers that aim to improve upon previously developed scales over 25 years ago by Norman (1963). The new marker set can be considered an alternative to the scales in the NEO and the Hogan personality inventories. It provides several advantages, such as (a) increasing reliability by using larger marker sets and (b) decreasing subject testing time by using smaller sets. The study also explores various alternative sets of markers and compares bipolar scales, antonym pairs, and unipolar variables. The findings suggest that factor scores based on subject-standardized responses to the new 100 unipolar variables provide quite univocal markers of each of the Big-Five domains. The availability of this easily administered set of factor markers is hoped to encourage investigators of diverse theoretical viewpoints to communicate in a common psychometric tongue.

Advantages:

A comprehensive structural representation of personality traits aims to capture the full spectrum of human personality through a well-defined system. This structure should be robust across various data sets, ensuring its applicability in different contexts. Importantly, it should also be scalable, allowing for the use of either extensive or focused personality marker sets while maintaining reliability and efficiency..

Disadvantages:

The system might not fully capture every nuance within each personality domain, and re-

lying on elements from the original Cattell variables could restrict its effectiveness in some situations.

2.2 System Study Report

The literature survey scrutinizes studies in personality psychology, focusing on the exploration of the Big Five personality traits. Firstly, a study introduces a shift in trait measurement by proposing a set of markers for the Big Five, aiming to refine upon scales established decades ago. This approach promises reliability and streamlines testing procedures, ensuring efficiency in personality assessment. Through exploration of marker sets and comparisons, the study underscores the importance of standardized responses in capturing nuances within the Big Five domains. By providing a framework, this research fosters communication among researchers from diverse orientations, advancing our understanding of human personality dynamics.

Furthermore, another study delves into the interplay between personality traits and employment statuses, ranging from entrepreneurial ventures to managerial roles. Drawing upon a dataset from the UK Household Longitudinal Study, this research illuminates disparities in personality traits across occupational landscapes. These findings shed light on the manifestation of personality traits within job contexts and offer insights for employers, policymakers, and individuals navigating workforce dynamics. Additionally, a study employs a mixed model approach, combining clustering analysis and neural network methodologies to dissect personality types grounded in the Big Five traits. Despite the findings, the studies acknowledge limitations such as biases in self-report data and the focus on the Big Five traits, indicative of complexities in personality research. Nevertheless, the collective findings enrich our understanding of personality psychology, highlighting the evolution of this field.

Chapter 3

Proposed Project

3.1 Problem Statement

Traditional hiring methods heavily rely on resumes and interviews, offering a limited view of a candidate's personality. This can lead to mismatched placements where personalities clash, hindering team dynamics and overall productivity. HR departments lack a systematic approach to understanding employee personalities, making it difficult to build well-rounded teams and create a work environment that caters to diverse personalities.

3.2 Proposed Solution

Personify presents a comprehensive solution to the challenges faced by HR departments in understanding and harnessing the unique personalities of their employees. Through its tailored questionnaire-based approach, the platform offers a streamlined method for gathering insightful data on individual traits and preferences. Leveraging advanced algorithms and renowned personality assessment models such as the Big Five (OCEAN), Personify provides organizations with actionable insights into their workforce, enabling them to make informed decisions across various aspects of talent management.

Chapter 4

Project Design

The general framework diagram is shown in figure 4.1

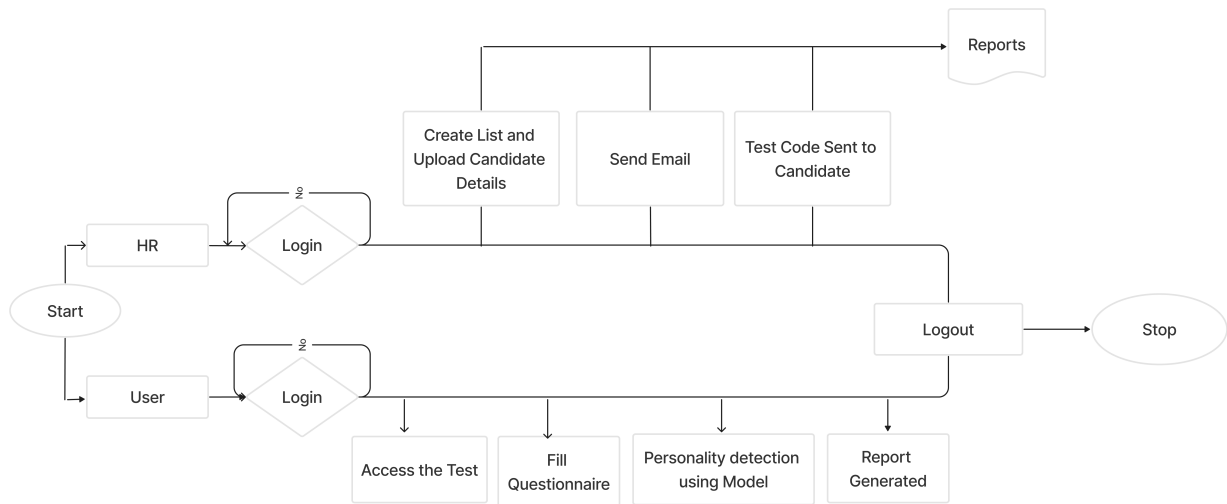


Figure 4.1: Overall framework

The proposed system takes input through webcam or CCTV camera, which were implemented in the public places. After taking the input, it converts into a number of frames. In each frame, it detect persons in that frame and draw a box around them. After drawing the box, the system will calculate the distance between two boxes. Then it will compare the distance with minimum safe distance. If the distance between two boxes are greater than the minimum distance, then it will marked in red color, which indicates that both persons are violating the social distancing. Also it will result total number of persons in that frame,

and the number of persons who are violating social distancing. This process will repeat until the input stop.

4.1 Data Design

The data flow diagram for the proposed system is represented in the following sections.

4.1.1 Level 0 data flow diagram

Level 0 data flow diagram shows the basic overview of the whole system. The system facilitates a testing process where HR uploads tests, candidates take them online, and results are delivered. Figure 4.2 shows the Level 0 data flow diagram.

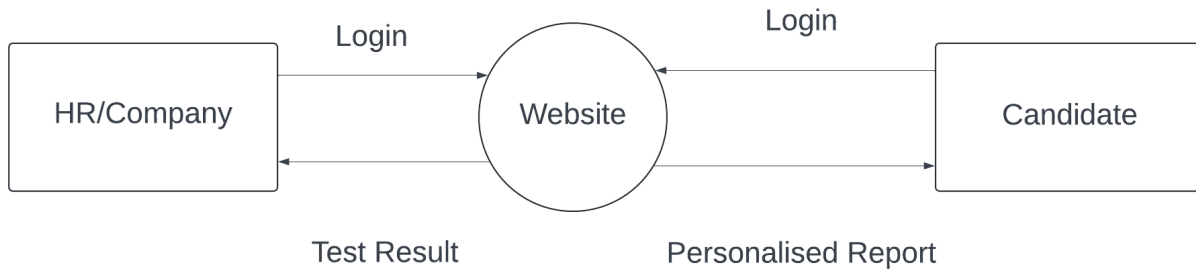


Figure 4.2: Level 0 DFD

4.1.2 Level 1 data flow diagram

In the Level 1 data flow diagram, system is broken down into subprocesses. Here the dataset is divided into training set and testing set. Training set is used for training the pedestrian detection model and testing data is used to test the accuracy of trained model.

The HR creates and uploads candidate details, potentially sending email notifications. The website manages candidate logins and grants access to the appropriate test based on a unique code (potentially emailed to the candidate). Candidates answer the test questions and submit their responses. The website prepares the data, potentially utilizing a K-means algorithm for analysis or scoring (the details of this processing are not shown here), and delivers the results through an API response.

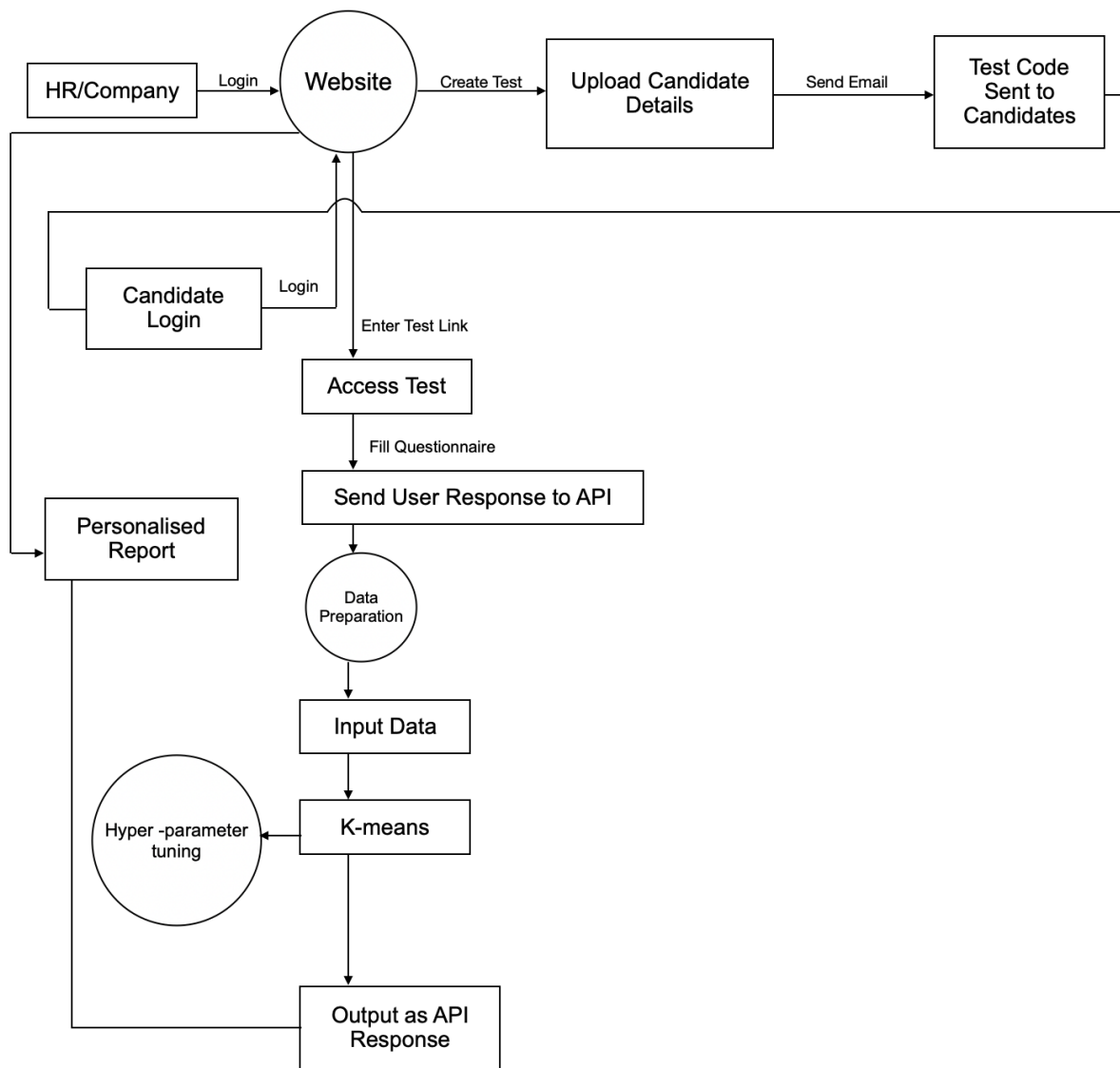


Figure 4.3: Level 1 DFD

4.2 Architecture Diagram

The architecture design diagram is shown in figure 4.4

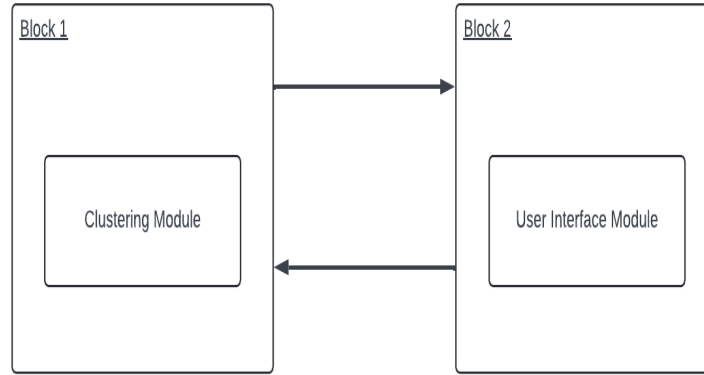


Figure 4.4: Architecture Diagram

4.3 Component Level Design

The system is divided onto three modules.

- User Interface Module
- Clustering Module

The different modules identified and its functionalities in the proposed approach are as follows.

4.3.1 User Interface Module

This module has three subsections. First section is used to collect input from user and second section is used to display the results.

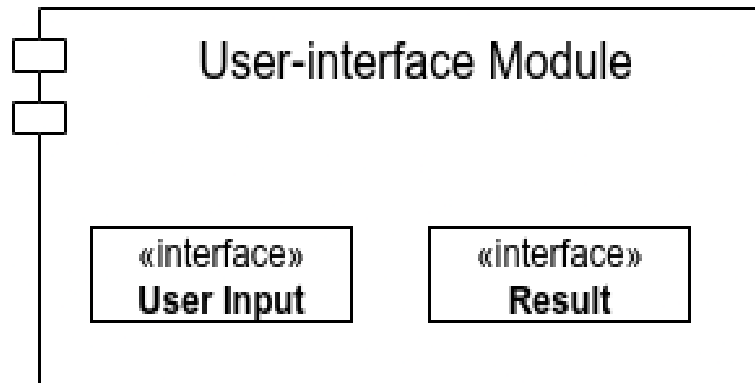


Figure 4.5: User-interface Module

The algorithm is as follows

Algorithm 1 User-interface Module Algorithm

- Step 1: Start
 - Step 2: Read user response to all the questions
 - Step 3: Display the results
 - Step 4: Stop
-

The third subsection involves presenting all candidates' results to the hiring manager. The algorithm is as follows

Algorithm 2 Result-Retrieval Algorithm

- Step 1: Start
 - Step 2: Retrieve results of all candidates for a particular test
 - Step 3: Display the results
 - Step 4: Stop
-

4.3.2 Clustering Module

In the clustering model, K means clustering algorithm is used to train the model. The dataset is divided into two set. One is used for training purpose and other is used for testing purpose.

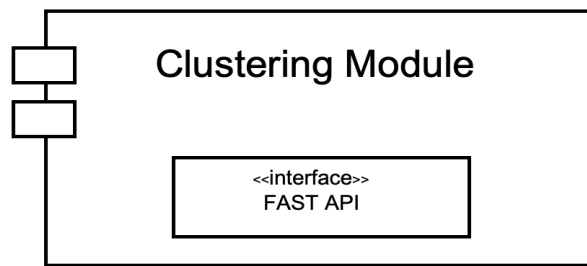


Figure 4.6: Clustering Module

The algorithm is as follows

Algorithm 3 Clustering Module Algorithm

Step 1: Start

Step 2: Import the dataset

Step 3: Split the dataset into training set and testing set. 70% of dataset is training data and remaining 30% is testing data

Step 4: Decide the optimal number of clusters using Elbow method

Step 5: Train the model using K means clustering algorithm

Step 5.1 K-MEANS (D , k , ϵ):

Step 5.1 Initialize the number of clusters (k) and the maximum number of iterations(e).

Step 5.2 Randomly initialize k centroids

Step 5.3 Repeat the following steps until the centroids no longer change or a maximum number of iterations is reached:

Step 5.3.1 Assign each data point to the closest centroid.

Step 5.3.1 Recompute the centroid of each cluster as the mean of the data points assigned to that cluster.

Step 6: Stop

4.3.3 Hardware and Software Requirements

The recommended hardware and software requirements are specified in table 4.1. All devices used are owned and softwares used are open source.

Table 4.1: Hardware and Software Requirements

OS	Windows 8,10,11 Linux
Software	Anaconda Python, Sciikit, Numpy,OS
CPU	Intel Core i5 6th Generation processor or higher or an AMD equivalent processor
RAM	8GB
Disk Storage	40 GB of free disk space

Chapter 5

Experimental Results and Analysis

5.1 Data Preparation

The dataset used is The Big Five personality traits, also known as the five-factor model (FFM) and the OCEAN model, is a taxonomy, or grouping, for personality traits. When factor analysis (a statistical technique) is applied to personality survey data, some words used to describe aspects of personality are often applied to the same person. For example, someone described as conscientious is more likely to be described as "always prepared" rather than "messy". This theory is based therefore on the association between words but not on neuropsychological experiments. This theory uses descriptors of common language and therefore suggests five broad dimensions commonly used to describe the human personality and psyche.

I :	EXT1	EXT2	EXT3	EXT4	EXT5	EXT6	EXT7	EXT8	EXT9	EXT10	..
0	4.0	1.0	5.0	2.0	5.0	1.0	5.0	2.0	4.0	1.0	..
1	3.0	5.0	3.0	4.0	3.0	3.0	2.0	5.0	1.0	5.0	..
2	2.0	3.0	4.0	4.0	3.0	2.0	1.0	3.0	2.0	5.0	..
3	2.0	2.0	2.0	3.0	4.0	2.0	2.0	4.0	1.0	4.0	..
4	3.0	3.0	3.0	3.0	5.0	3.0	3.0	5.0	3.0	4.0	..

Figure 5.1: Dataset Example

An example of a dataset used for this project is shown in figure 5.1. This dataset contains personality trait scores for various candidates. Each data point represents an individual's scores across different personality dimensions (e.g., Openness, Conscientiousness).

5.2 Output

The following figures showcase the user experience on the Personiphy website. Figure 5.2 specifically focuses on the user’s end result, displaying personality trait scores and a corresponding pie chart for visual representation. Additionally, the website offers a 'Know Your Traits' section, providing users with insights into their own personalities.

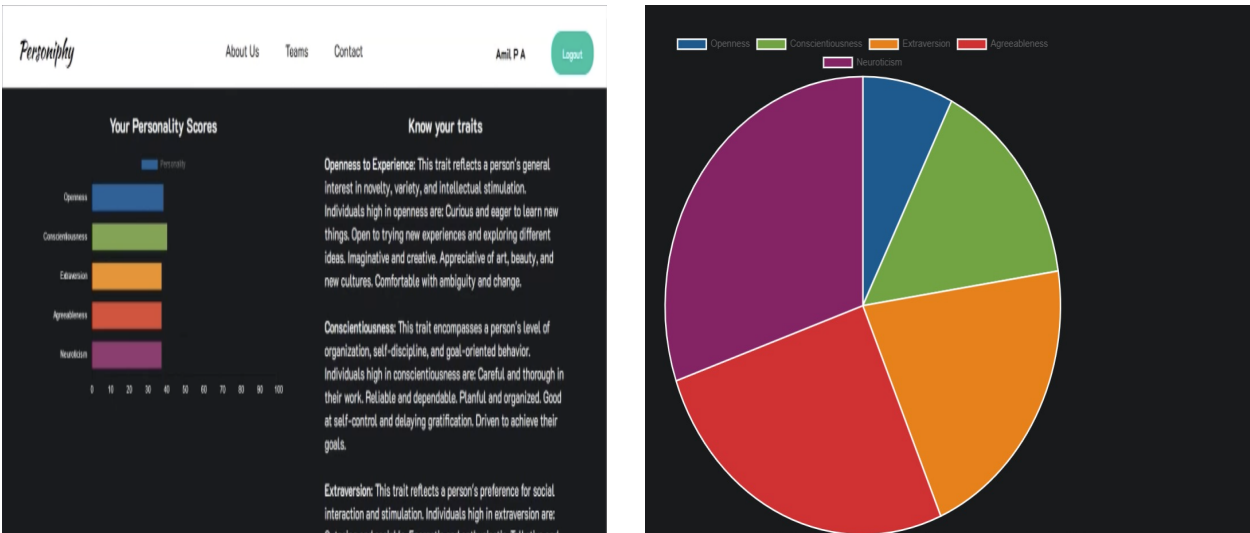


Figure 5.2: User Side Result



Figure 5.3: HR side result

Figure 5.3 presents the HR department’s output, which includes individual candidate test results. This data allows for cross-referencing and validation of the testing outcomes.

Chapter 6

Conclusion

In conclusion, the project "Personiphy" stands as a comprehensive solution for personality assessment and workforce management. Beyond merely facilitating the hiring process, Personiphy offers a range of features for ongoing employee development and team building activities, all rooted in deep insights derived from personality assessments.

By incorporating feedback from users and stakeholders, Personiphy demonstrates a commitment to continuous improvement, both in terms of effectiveness and usability. This dedication ensures that Personiphy remains a valuable tool for optimizing workforce management processes in organizations of all sizes and industries. As organizations navigate the complexities of talent management, Personiphy serves as a trusted partner, enabling informed decision-making and fostering a culture of growth and collaboration.

References

- [1] Weixi Kang, Kreisha Lou Guzman, and Antonio Malvaso. Big five personality traits in the workplace: Investigating personality differences between employees, supervisors, managers, and entrepreneurs. *Frontiers in Psychology*, 14:976022, 2023.
- [2] Jennifer Chi and Yeong Nain Chi. Cluster analysis of personality types using respondents’ big five personality traits. *International Journal of Data Science*, 4(2):116–135, 2023.
- [3] Lewis Goldberg. The development of markers for the big five factor structure. *Psychological Assessment*, 4:26–42, 03 1992.
- [4] David Holman and David Hughes. Transactions between big-5 personality traits and job characteristics across 20 years. *Journal of Occupational and Organizational Psychology*, 94, 02 2021.
- [5] Assem Talasbek, Azamat Serek, Meirambek Zhaparov, Seong-Moo Yoo, Young-Kab Kim, and Geun-Ho Jeong. Personality classification experiment by applying k-means clustering. *International Journal of Emerging Technologies in Learning (iJET)*, 15(16):pp. 162–177, Aug. 2020.
- [6] Dess Mardan Basnet and Murari Prasad Regmi. The correlational study of the personality traits, organizational commitment and self-efficacy of saving and credit co-operative societies (saccos) employees of kathmandu. *Open Journal of Business and Management*, 7(2):311–324, March 7 2019.
- [7] Hussain Ahmad, Mohd Zubair, Alam Khan, and Anam Habib. A systematic literature review of personality trait classification from textual content. *Open Computer Science*, 10, 07 2020.
- [8] Christopher Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. 10 2007.
- [9] Mohammad Awwad and Rana Al-Aseer. Big five personality traits impact on entrepreneurial intention: the mediating role of entrepreneurial alertness. *Asia Pacific Journal of Innovation and Entrepreneurship*, ahead-of-print, 05 2021.

-
- [10] Sebastiaan Rothmann and Elize Coetzer. The big five personality dimensions and job performance. *South African Journal of Industrial Psychology*, 29:68–74, 10 2003.