CV ANALYSIS AND OPTIMIZING THE RECRUITMENT PROCESS IN THE IT INDUSTRY USING MACHINE LEARNING TECHNIQUES

Project ID: 2023-098

Project Proposal Report

De Silva M. – IT20207854

Bachelor of Science (Hons) Degree in Information Technology

Specializing in Data Science

Department of Information Technology

Faculty of Computing

Sri Lanka Institute of Information Technology

Sri Lanka

March 2023

CV ANALYSIS AND OPTIMIZING THE RECRUITMENT PROCESS IN THE IT INDUSTRY USING MACHINE LEARNING TECHNIQUES

Project ID: TMP-23-098

Project Proposal Report

Bachelor of Science (Hons) Degree in Information Technology

Specializing in Data Science

Department of Information Technology

Faculty of Computing

Sri Lanka Institute of Information Technology

Sri Lanka

March 2023

DECLARATION

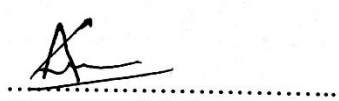
I declare that this is our own work, and this proposal does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to Sri Lanka Institute of Information Technology, the nonexclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Name	Student ID	Signature
De Silva M.	IT20207854	M. De Silva

The supervisor/s should certify the proposal report with the following declaration.

The above candidates are carrying out research for the undergraduate Dissertation under my supervision.



Signature of the Supervisor:
(Dr. Anuradha Karunasena)

26/03/2023
Date:

ABSTRACT

The success and growth of an organization are heavily dependent on the selection of the right candidates. The traditional hiring process, which mainly consists of manual steps such as reviewing resumes and educational records, evaluating technical and professional skills, is both time-consuming and inefficient. In order to effectively address the needs and expectations of employers, it is crucial to implement a more efficient and accurate method for evaluating a candidate's skills, abilities and personality traits. The primary focus of this research is to streamline the hiring process in organizations through the integration of machine learning techniques, data extraction techniques and natural language processing. The objective is to comprehensively evaluate a candidate's technical skills, professional skills, and Big Five Personality traits, thereby providing hiring managers with a thorough understanding of the candidate. This integration of advanced technologies aims to streamline the hiring process and facilitate informed decision-making.

Keywords: candidate, recruitment, big five personality traits, natural language processing, machine learning, skills

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

NLP	Natural Language Processing
FWHR	Facial width-to-height ratio
PPM	Personality Prediction Model
CPP	Candidate Personality Prediction
SVM	Support Vector Machines

LIST OF APPENDICES

1. INTRODUCTION

1.1 Background

Recruitment is a critical process for any organization that aims to hire the right talent to achieve its goals. However, given the current demand for new hires in the IT industry, the traditional recruitment process is often time-consuming, inefficient, and costly. Manual steps such as reviewing resumes and academic transcripts, and conducting interviews to evaluate skills, can be time-consuming, and sometimes lead to hiring the wrong candidate. Moreover, there is a lack of a clearly defined methodology for assessing critical areas such as technical skills and personality traits in candidates and evaluating the extent to which they align with job requirements. Consequently, there is a need for a more streamlined and systematic approach to the recruitment process that accounts for these crucial factors and enables more efficient and effective hiring practices.

To address these challenges, many organizations have started adopting automated recruitment systems that use machine learning, data extraction, and natural language processing techniques to analyze resumes, evaluate skills, and personality traits of the candidates. These automated systems aim to make the recruitment process faster, more accurate, and cost-effective.

As per the Sri Lanka Information and Communication Technology Agency's (ICTA) records [1], the country had set a goal of attaining 200,000 employees in the IT industry alone for the year 2022. This emphasizes the pressing need for a recruitment system that facilitates more efficient and less time-consuming hiring practices for IT-related job positions. The focus of this research is to examine the criteria involved in selecting new candidates and creating an automated system that can effectively evaluate a candidate in all of the above-mentioned criteria to identify and recruit the most qualified candidate for a job position.

1.2 Component Overview

Technical skills, professional skills, academic qualifications are certainly important for job performance, but they are not the only factors that determine whether an individual is a good fit for a particular role or workplace.

Personality traits and characteristics play a crucial role in determining job fit, job satisfaction, and overall success, benefiting both the individual and organization as a whole.

The focus of this research component is to determine the personality traits of a prospective employee with the purpose of evaluating their suitability for a particular job position. This includes determining the existence of the big five traits and aligning the personality traits of the candidate with those anticipated in the job description. Subsequently, a score will be generated to illustrate the degree to which the candidate's personality traits align with the requirements of the job position.

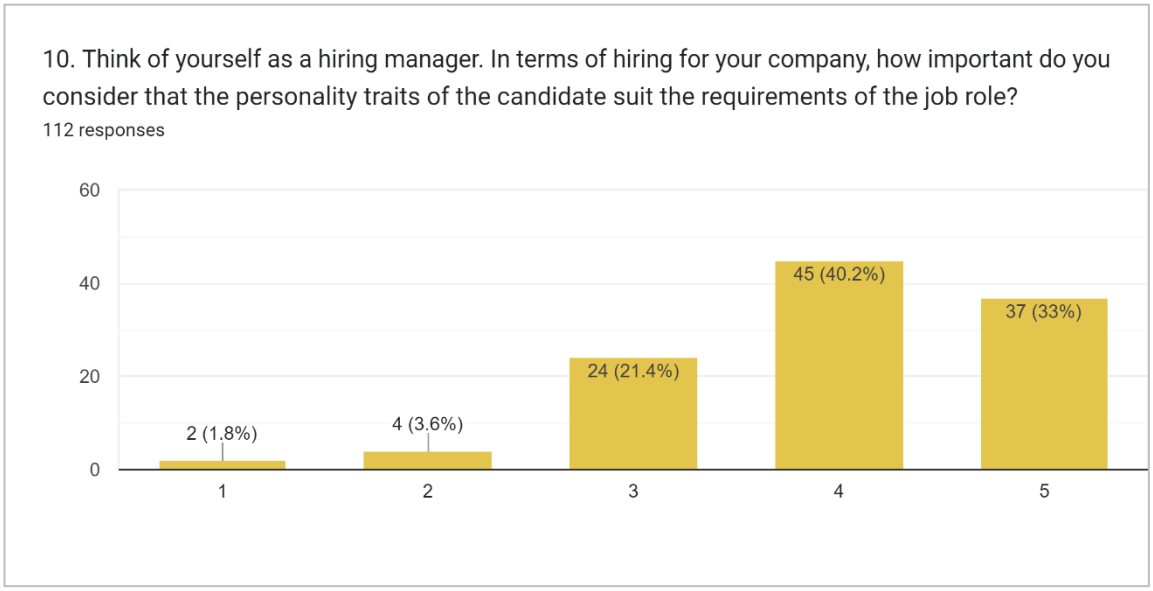


Figure 1.1 : Survey results showing the importance of personality-job fit

Shown above (Figure 1.1) are the results of a survey conducted aimed to gather insights into the level of importance given to personality-job fit in today's society. Based on the obtained results, we can conclude that modern society places a significant emphasis on evaluating the compatibility of an employee's personality with the demands of their job role.

1.2.1 The Big Five Model

The Big Five Model is the preferred personality assessment model for this component, as it has emerged as the dominant model in modern personality research. Over the past five years, the Big Five Model has been featured in 735 scientific publications, surpassing the combined usage of all other assessment models. [2] This statistical evidence suggests that the Big Five Model is currently the most reliable option for predicting personality traits. Hence, it is considered the leading choice for this component.

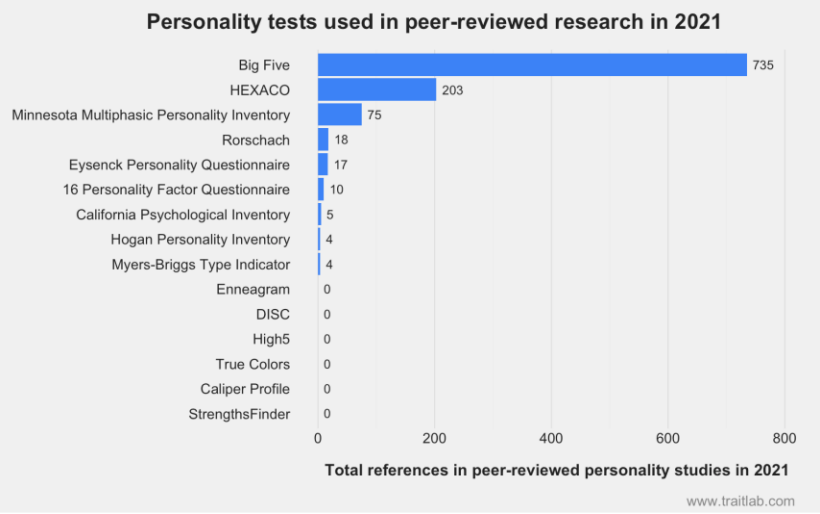


Figure 1.2 : Personality traits used in peer-reviewed research in 2021

Moreover, the Big Five traits are considered to be highly relevant to job performance and organizational success, making it a popular choice for employee selection and development.

Figure 1.3 : The Big Five Personality traits

1.3 Literature Review

In light of the growing interest in the science of psychology and its applications in a variety of fields, researchers have been investigating the use of technology and data analysis to predict an individual's personality traits. Over the past few years, machine learning algorithms and predictive models have been developed to analyze various types of data, such as social media activity, facial expressions, and linguistic patterns, to predict personality traits. This literature review aims to provide an overview of the present state of research on predicting personality traits using various types of data and machine learning algorithms.

The paper titled “DevFlair: A Framework to Automate the Pre-screening Process of Software Engineering Job” [3] *(Research A)* by Jayasekara et al. proposes a system to automate the initial screening of job applicants for software engineering positions.

The writers begin by highlighting the difficulties that recruiters encounter when trying to hire software engineers, especially during the initial screening stage. They suggest that conventional screening methods like reviewing resumes and conducting technical interviews are tedious and can be biased. As a solution, the authors introduce DevFlair, which is a system that streamlines the screening process by examining a candidate's coding contributions online and utilizing artificial intelligence algorithms to forecast their suitability for the job.

The authors describe the design and implementation of DevFlair as a ‘framework designed to automate the initial screening of candidates applying for software engineering jobs’. It works by gathering information from various sources such as social media to determine a candidate's Big Five personality traits, using questionnaires to assess technical skills, and analyzing data from GitHub to assess a candidate's proficiency in different technical areas. The framework then ranks candidates based on these criteria.

In this paper, PPM (Personality Prediction Model) is introduced as a tool for predicting a job applicant's personality traits based on their LinkedIn profile. PPM retrieves the candidate's profile details from LinkedIn using the ProxyCurl API, pre-processes the data, and uses five different models to predict their Big-Five personality traits. The article describes how probabilities are computed for each trait and how they are combined to determine a candidate's overall personality probability.

Overall, the paper presents an interesting approach to addressing the challenges faced by recruiters in the software engineering industry and could potentially save recruiters time and resources while improving the accuracy of their pre-screening process.

Another research “Candidate Selection for the Interview using GitHub Profile and User Analysis for the Position of Software Engineer” [4] *(Research B)*, proposes a pre-screening mechanism to minimize the time taken to conduct interviews for a number of candidates. Similar to the previous research, the authors note that traditional hiring processes can be time-consuming and subjective, and argue that recruiters for software engineering positions expect candidates to have a diverse skill set, including the ability to learn independently and solve problems, which cannot be assessed solely through their resume. Therefore, the proposed solution for pre-screening candidates for the position of software engineer is to analyze the candidates using a range of dimensions, including their GitHub account, academic transcript, letters of recommendation, LinkedIn profile, and personality prediction.

The research mentioned above involves determining a candidate's Big Five personality traits through a phone call interview where the candidate answers open-ended questions. The responses are then transcribed, preprocessed, and analyzed using NLP techniques to predict whether the candidate exhibits the Big Five traits. The algorithms used are Random Forest, SVM, and Logistic Regression, and the study finds that Logistic Regression has the highest accuracy for all five traits. The results are visualized in the form of a radar graph where '1' signifies the presence of a trait, and '0' indicates its absence.

In the research paper titled “Interview Data Analysis using Machine Learning Techniques to Predict Personality Trait” [5] *(Research C),* the author uses prosodic features such as intensity, pitch and frequency to predict five personality traits such as ‘Engaged’, ‘Excited’, ‘Friendly’ ‘Calm’ and ‘Speaking rate’ of interview candidates. The study uses audio-visual recordings from mock interviews conducted with MIT students as the dataset. The results showed that the prosodic features related to intensity played a significant role in predicting traits like "Engaged" and "Excited," while features like pitch and duration of pause were more relevant in predicting "Friendly." Similarly, the study found that prosodic features related to pitch were important in predicting the trait of "Calm." After selecting the top prosodic features, the researchers applied three different regression models to determine the best method for predicting the personality traits. The study concluded that Decision Tree was the best choice for predicting the selected traits using the chosen prosodic features.

The paper titled “Personality Prediction for Online Interview” by S. K. Nivetha, M. Geetha, R. S. Latha, K. Sneha, S. Sobika, and C. Yamuna [6] *(Research D)* explores the application of machine learning algorithms for predicting the personality of job candidates during online interviews. The study focuses on the use of the ‘Big Five Personality Traits’ as the basis for personality prediction.

The authors note that while online interviews have become a popular method for recruitment, they often lack the personal touch of face-to-face interviews, which can make it difficult to accurately assess a candidate's personality.

The paper discusses the use of different methods such as text, image and video analytics to forecast personality traits applying technologies such as support vector machines (SVM) and neural networks.

The study uses two approaches to identify the Big Five personality traits of candidates participating in online interviews. The first approach utilizes the facial height-to-width ratio (FWHR) of the candidates captured in their images, and a CNN-based model is trained to determine the traits. The second approach involves using a questionnaire

consisting of situational questions to assess how a candidate might behave or respond in specific situations, and the responses are analyzed through K-means clustering to predict the presence of Big Five traits.

According to the findings of this research, the K-Means clustering algorithm generated a result with 90% precision. Additionally, the use of FWHR and CNN to analyze a person's personality resulted in an accuracy of approximately 92%.

The paper titled “Predicting Personality Using Answers to Open-Ended Interview Questions” by M. Jayaratne and B. Jayatilleke [7] *(Research E)* aim to investigate the possibility of using open-ended interview questions to predict personality traits of individuals. This research uses NLP and machine learning techniques to predict the personality characteristics of job applicants. The data for the study is gathered from an interview conducted through online chat.

This research suggests that algorithms can determine a candidate's personality objectively by analyzing their answers, thus eliminating any subjective biases that arise from human interviewer assessments.

The paper emphasizes the importance of job-personality suitability as if greatly affects job performance, job satisfaction and duration of employment. According to the research, individuals tend to experience greater job satisfaction when their personality aligns with their chosen profession. However, the research also highlights that, candidates may find a personality test cumbersome, and it adds an additional expense to the hiring process. Therefore, the authors propose that interview responses are a more appropriate way of predicting a candidate's personality.

The approach utilized in this study involved developing a regression model to estimate a score for each of the six characteristics in the HEXACO model by analyzing the written responses provided by the candidates in the form of open-ended questions. In the HEXACO-based self-rating questions, the candidate rates themselves on a 5-point scale.

The final scores for the individual traits are calculated but taking the average over all the responses for each of the 6 traits.

In conclusion, this literature review suggests that several methods, such as social media platforms, FWHR, prosodic features, and responses to open-ended questions, have been utilized for candidate personality prediction. However, further improvements are needed to make these methods more reliable and ensure more accurate results.

1.4 Research Gap

Based on the literature survey done above, the following were identified as research gaps,

- Although systems have been introduced to assess the big five traits of a candidate, there has been no system to assess how well the candidate's personality traits would align with the personality traits required for the job role.

- Majority of the studies have been done generally and not focusing on a specific industry. This research is done specifically for hiring candidates for jobs in the IT industry focusing specifically on identifying the personality traits that are needed for success in various job roles within the IT sector.

Table 1.1 : Summary of research gap

Research / Feature	A	B	C	D	E	Proposed System
Compares candidate personality traits with expected traits	✗	✗	✗	✗	✗	✓
Uses results of 2 methods to determine personality traits	✗	✗	✗	✓	✓	✓
Focused on the entire IT industry	✗	✗	✗	✗	✗	✓
Uses the Big Five Model	✓	✓	✗	✓	✗	✓
Determines the presence and level of each trait	✓	✗	✓	✓	✓	✓

2. RESEARCH PROBLEM

Organizations are increasingly using personality prediction methods to improve their recruitment and selection processes. These methods can help them gain valuable insights into a candidate's personality traits, which can be useful in finding the right fit for a job. However, there is still a need to investigate how to effectively apply these methods to achieve desired outcomes. Organizations need to find accurate and reliable assessment methods that align with their specific needs and goals.

In addition to identifying accurate assessment methods, it is also crucial for organizations to understand how to assess the fit between a candidate's personality and a job role. A good personality-job fit can have a significant impact on an employee's performance, job satisfaction, and retention rate. Therefore, it is important for organizations to develop strategies for assessing personality-job fit to ensure that new hires can succeed in their roles.

Given the above-mentioned considerations, this research problem aims to study how personality prediction methods can enhance recruitment and selection processes. The study will identify the best assessment methods for predicting candidate personality traits and explore strategies for evaluating personality-job fit. The results will be useful for organizations looking to improve their recruitment and selection processes and increase the success of their workforce.

3. RESEARCH OBJECTIVES

2.1 Main Objective

The primary objective of this research is to streamline the hiring process by using automation to handle the recruitment tasks, with the goal of identifying the ideal candidate for a particular job based on their technical skills, academic qualifications, and necessary personality traits.

The research objective of this specific component is to develop a system to assess the compatibility between a candidate's personality and the personality traits necessary for a particular job role.

2.2 Specific Objectives

- **Develop an algorithm that identifies the presence of Big Five Personality Traits of candidates.**

This objective aims to determine the presence and degree of big five traits in a prospective employee allowing the hiring manager to gain insight into the candidate's personality.

- **Develop an algorithm to assess how well the personality traits of the candidate aligns suit the job role.**

This objective aims at determining the personality-job fit of a prospective employee. The candidate's personality traits will be evaluated against the personality traits deemed necessary by the hiring manager for the particular job position to generate a personality score that shows how compatible their personality traits are with the job requirements.

4. METHODOLOGY

The diagram depicted below illustrates the architecture of the component. As mentioned previously, the primary objective of this component is to evaluate a candidate's Big Five Personality traits and determine how well their traits align with the required personality traits for a particular job position.

The candidate will be given a questionnaire consisting of questions that provide insights into their personality. These questions are of two types; short self-rating questions where the candidate will rate themselves on a scale of ‘strongly agree’ to ‘strongly disagree’, and open-ended questions that encourage the candidate to provide thoughtful answers.

The use of two question types in a personality questionnaire is justified by the complexity and multidimensional nature of personality. Simple self-rating questions may not be adequate in accurately predicting personality traits, and should be complemented by other assessment methods. Conversely, relying exclusively on open-ended questions may disadvantage those who struggle with language fluency and verbal expression, thus, self-rating questions can be employed to address this issue.

The candidate's Big Five personality traits will be evaluated and assessed based on their answers to the provided questions. The responses to the self-rating questions will be fed into a supervised learning model to determine the presence and level of each trait. NLP techniques will be utilized to analyze the text from the open-ended questions and extract relevant keywords that indicate the intensity of the Big Five traits. The combination of these two methods will be used to predict the candidate's Big Five traits.

Subsequently, the hiring manager will provide the desired personality traits for the specific job position, which will be used as input. The candidate's Big Five traits will then be compared to the personality requirements for the job role, resulting in a personality score that indicates how closely the candidate's traits align with the

company's desired traits for the role. This score will reflect the degree to which the candidate's personality is a suitable match for the position.

3.1 Architecture Diagram

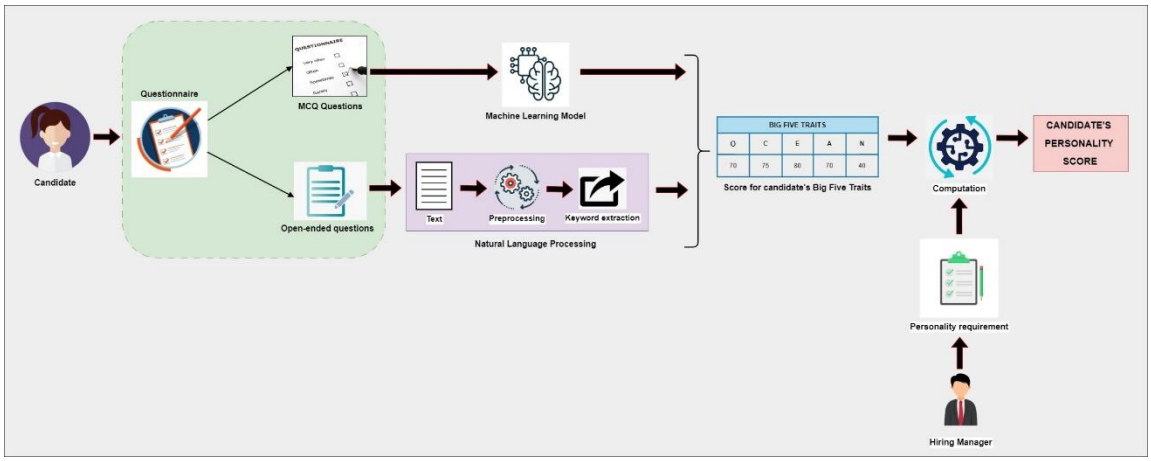


Figure 3.1 : Component Architecture diagram

5. PROJECT REQUIREMENTS

4.1 Functional Requirements

- The system should have the ability to analyze the candidate's responses to the questionnaire and represent the presence and degree of big five traits in a prospective employee.
- The system should represent how well a candidate's personality traits align with the personality traits expected for the job position.

4.2 Non-functional Requirements

- Accuracy – The predicted personality traits should be accurate and match with the candidate's personality.
- Performance – The system should be able to predict the 'Big Five Personality Traits' of the candidate and represent them in a way understandable to the user.
- Usability – The user interface of the system should be user friendly and easy to use.
- Compliance – The system should comply with all applicable laws and regulations. The candidate's confidential information should be kept private and secure.
- Reliability – The system's performance should be consistent and reliable. The system must remain stable and not experience any crashes during usage.

4.3 User Requirements

The intended user of this component is the IT company's hiring manager. The user requirements for the component are as follows,

- The hiring manager should be enter the candidate responses to the personal questionnaire into the system.
- The system should allow the hiring manager to enter the personality traits necessary for the particular job position.

- The system should enable the hiring manager to determine the degree to which a candidate's personality traits align with the required personality traits for the job position.

4.4 Feasibility Study

4.4.1 Technical Feasibility

When creating a system to automate the recruitment process, it is important to consider its technical feasibility. To ensure the successful development of this system, the project team must possess adequate knowledge in areas of artificial intelligence, machine learning, and natural language processing. Additionally, it is crucial to ensure the confidentiality and privacy of candidate data in compliance with data protection laws.

In summary, the proposed automated recruitment system can be technically feasible if the project team has the necessary tools and knowledge.

4.4.2 Economic Feasibility

When creating an automated recruitment system, it is crucial to consider its economic feasibility. The proposed system aims to increase efficiency and reduce costs associated with the recruitment process. By implementing the system, the company can save time and resources that would typically go towards conducting long interviews to evaluate technical skills, professional skills, academic qualifications and personality traits of candidates. Additionally, the system can expedite the recruitment process by quickly identifying qualified candidates, resulting in a shorter time-to-hire. The system can also enable HR departments to focus on crucial tasks, such as employee development and retention, leading to a more productive workforce.

Based on the reasons discussed above, it is evident that the economic feasibility of the proposed system is positive because the system's benefits can outweigh the initial costs, resulting in cost savings and increased efficiency for the company.

4.4.3 Schedule Feasibility

When developing a system to automate recruitment, it's crucial to consider schedule feasibility. The team needs to take into account factors such as personnel, technology and resource availability when creating a development schedule for a system. Additionally, the team must consider potential setbacks or delays that may arise during the development process. In summary, creating a schedule that is achievable, has adequate resources, and careful implementation planning can ensure schedule feasibility for the proposed automated recruitment system.

6. TOOLS AND TECHNOLOGIES AND ALGORITHMS

The proposed tools, technologies and algorithms for the implementation of this component are as follows,

- **IDE**:

    - **Visual Studio Code (VS Code)** – VS Code is a free source code editor developed by Microsoft. It is a lightweight and versatile tool that can be used for a variety of programming tasks such as writing code, debugging, and version control. The key benefits of using VS Code include,

        - ✓ Extensibility - VS Code has a vast library of extensions that can be downloaded and installed to add new functionality to the editor.

        - ✓ User interface - VS Code has a clean and intuitive interface that makes it easy to navigate and customize.

        - ✓ Built-in support for Git - VS Code allows developers to easily manage version control directly from the editor.

- **Model development**:

    - **Python using Google Colab** - Python is simple to learn and has the ability to develop models quickly. Python requires less coding and has some great libraries for natural language processing.
      Google Colab offers access to powerful computing resources like GPUs and TPUs that can greatly enhance the speed of training machine learning models. Additionally, Colab comes pre-installed with popular machine learning libraries such as TensorFlow, Keras, and PyTorch. This eliminates the need to manually install these libraries and ensures compatibility with other tools and frameworks.

- **Libraries**:
    - **Scikit-learn (Sklearn) -** Scikit-learn is a freely available machine learning library created using the Python programming language. It offers a diverse range of machine learning models such as clustering, classification, and regression.
    - **Natural Language Toolkit (NLTK) -** NLTK is a commonly used Python library for handling human language data that is accessible for free. It offers a variety of tools and applications for natural language processing (NLP) tasks, including tokenization, stemming, tagging, parsing, and semantic reasoning.
    - **Matplotlib** – Matplotlib is a popular open-source plotting library for Python. It provides a range of tools for creating high-quality plots, charts, and graphs.

- **Frontend:**
    - **Flask** – Flask is a web framework written in Python that can be used to create a user interface for Python models, making it easier to interact with them. With Flask, developers can create a web application that allows users to input data, which is then passed to the Python model for processing. The output from the model can then be displayed back to the user via the Flask interface. This approach provides a user-friendly way to access the functionality of Python models, allowing them to be used in various fields such as data analysis, machine learning, and artificial intelligence.

- **Database:**
    - **Azure Cosmos DB -** Azure Cosmos DB is a fully-managed database that supports both NoSQL and relational data models, designed specifically for modern application development. The decision to select a NoSQL

database as our database solution was based on the requirement to store unstructured data extracted from CVs, academic transcripts, and GitHub and LinkedIn profiles.

- **Deployment:**
    - **Azure App Service -** Azure App Service is a cloud-based platform that can be used to host Python Flask projects. By using Azure App Service, developers can deploy their Python Flask applications to the cloud, making them accessible to users from anywhere with an internet connection. Azure App Service provides a fully managed environment, with features such as automatic scaling, load balancing, and easy integration with other Azure services. Azure App Service is an ideal platform for hosting Python Flask projects. With its powerful features and easy-to-use interface, it provides a hassle-free way to deploy and manage Python Flask applications in the cloud.

Note: The tools, technologies, and algorithms proposed above are subject to change based on any challenges or problems that arise during the development process.

7. WORK BREAKDOWN STRUCTURE

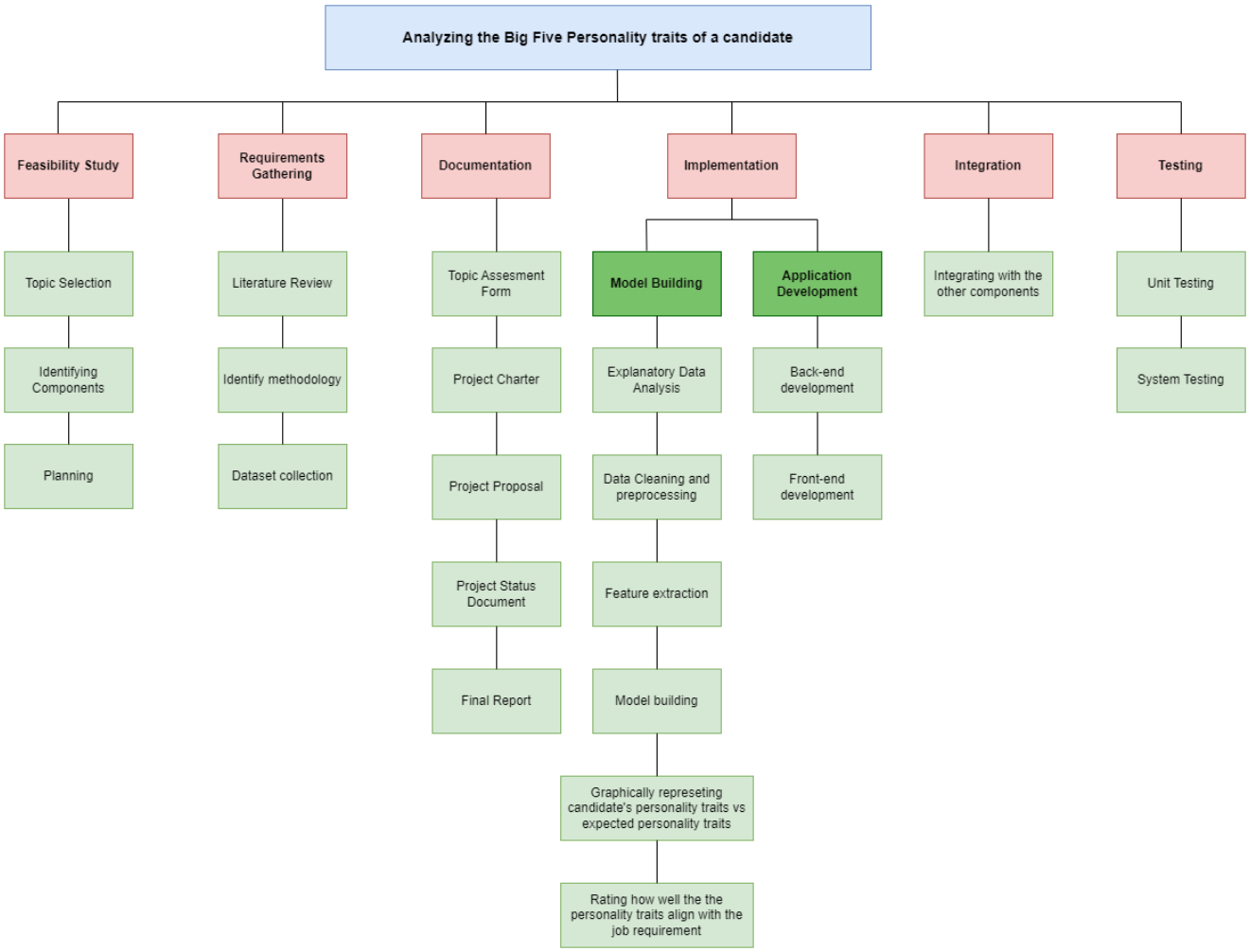


Figure 7.1 : Work breakdown diagram

8. GANTT CHART

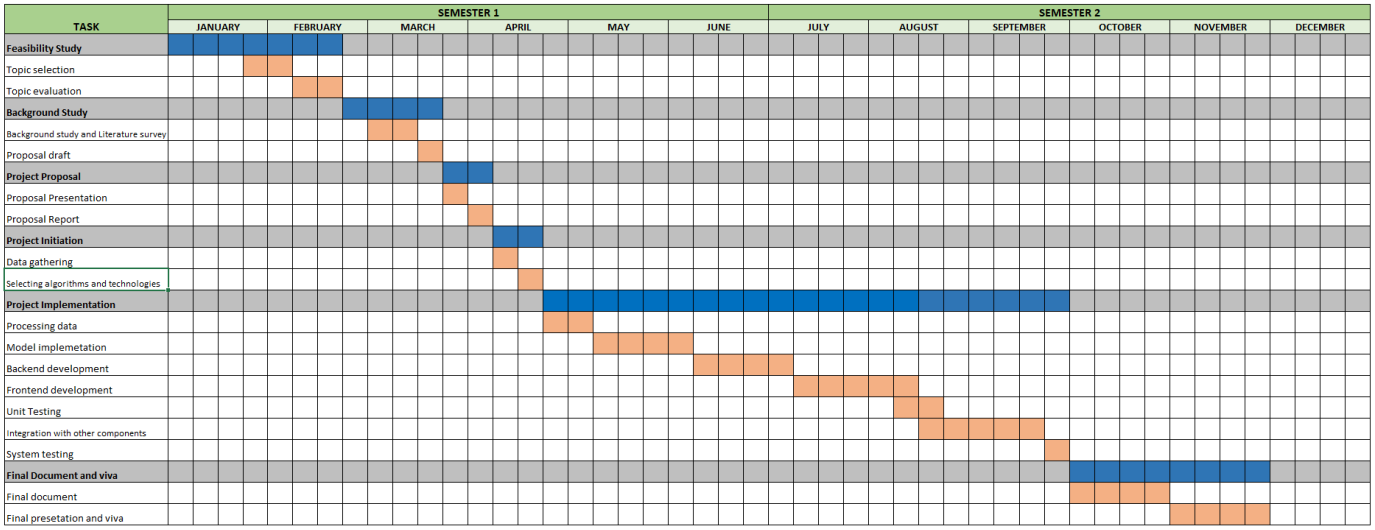


Figure 8.1 : Tentative Gantt Chart

9. COMMERCIALIZATION OF THE PRODUCT

‘Intellihire’ is a recruitment software created by ‘SMMS Software Solutions’ to simplify the hiring process for IT companies. With the increasing demand for skilled IT professionals, the recruitment process has become time-consuming and resource-intensive. Intellihire helps to streamline the process by using advanced machine learning algorithms to analyze candidate profiles, job requirements, and other relevant data to identify the most suitable candidates for a particular job.

Intellihire also features an intuitive user interface that makes it easy for hiring managers to navigate and manage the recruitment process. Additionally, the system can be customized to meet the specific needs of different organizations.

To reach potential customers, Intellihire will be advertised on social media platforms such as LinkedIn, Facebook and Instagram. Interested organizations can request a trial version or purchase the software directly from SMMS software solutions.

The personality prediction feature of Intellihire minimizes the effort required by hiring managers to evaluate a candidate's personality through informal conversations or HR interviews. This can save time in the hiring process by quickly providing insights into a candidate's personality traits and how they may fit into a particular role or team. Additionally, by ensuring that candidates are a good fit for the role and the company culture, a candidate personality prediction system can help reduce turnover rates and increase employee retention improving the hiring process of the organization and ultimately leading to a more productive and satisfied workforce.

Figure 9.1 : Logo - Intellihire Software

Figure 9.2 : Logo - SMMS Software Solutions

REFERENCES

[1] "ICTA - News," [Online]. Available: https://www.icta.lk/news/sri-lanka-aiming-200000-ict-workforce-by-2022/.

[2] "TraitLab - Big Five," [Online]. Available: https://www.traitlab.com/blog/personality-test-comparison#:~:text=Big%20Five,-Research%20results%20in&text=In%20the%20last%20five%20years,Extraversion%2C%20Agreeableness%2C%20and%20Neuroticism..

[3] "R. T. R. Jayasekara, K. A. N. D. Kudarachchi, K. G. S. S. K. Kariyawasam, D. Rajapaksha, S. L. Jayasinghe and S. Thelijjagoda, "DevFlair: A Framework to Automate the Pre-screening Process of Software Engineering Job Candidates,"".

[4] "R. G. U. S. Gajanayake, M. H. M. Hiras, P. I. N. Gunathunga, E. G. Janith Supun, A. Karunasenna and P. Bandara, "Candidate Selection for the Interview using GitHub Profile and User Analysis for the Position of Software Engineer,"".

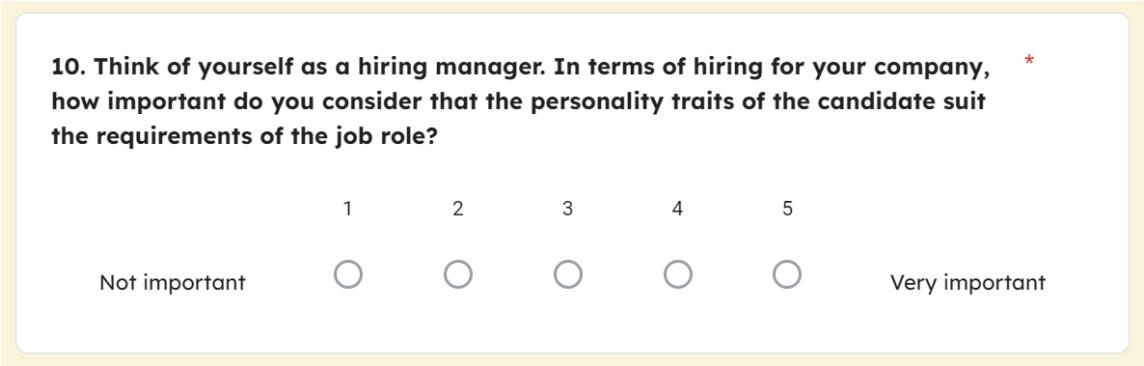
[5] "S. Chopra and S. Urolagin, "Interview Data Analysis using Machine Learning Techniques to Predict Personality Traits," 2020 Seventh International Conference on Information Technology Trends (ITT), Abu Dhabi, United Arab Emirates, 2020".

[6] "S. K. Nivetha, M. Geetha, R. S. Latha, K. Sneha, S. Sobika and C. Yamuna, "Personality Prediction for Online Interview," 2022 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2022".

[7] "M. Jayaratne and B. Jayatilleke, "Predicting Personality Using Answers to Open-Ended Interview Questions,"".

APPENDICES

Appendix A: Survey questionnaire related to personality prediction.



Appendix B: Turnitin similarity score