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

Project ID: 2023-098

Project Proposal Report

Zoysa E.S. – IT20231200

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 PROCESSIN THE IT INDUSTRY USING MACHINE LEARNING TECHNIQUES

Project ID: 2023-098

Project Proposal Report

Bachelor of Science (Hons) Degree in Information

TechnologySpecializing in Data Science

Department of Information

TechnologyFaculty of Computing

Sri Lanka Institute of Information

TechnologySri Lanka

March 2023

1 DECLARATION

We 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 doesnot contain any material previously published or written by another person except where theacknowledgement is made in the text.

Name	Student ID	Signature
Zoysa E.S.	IT20231200	

The above candidate is carrying out research for the undergraduate Dissertation under mysupervision.

…… 26.03.2023

Signature of the supervisor Date

(Dr. Anuradha Karunasena)

2 ABSTRACT

The primary focus of this research is the information technology sector, with a specific emphasis on the recruitment of IT-related positions. This study aims to streamline the hiring process in organizations through the integration of machine learning techniques, data extraction techniques and natural language processing. This research paper examines the process of using data extracted from GitHub and LinkedIn user profiles to aid in job recruitment. The study analyzes the benefits and limitations of collecting data from these platforms and explores the relationship between the collected data and job performance. The research methodology involves data collection from a sample of job candidates' GitHub and LinkedIn profiles, The objective is to comprehensively evaluate a candidate's technical skills, professional skills and preferences by extracting data from candidate generating a candidate profile. Results show that collecting data from these platforms can provide valuable insights into candidates' skills, experiences, and suitability for the job. However, limitations such as the potential for bias in the data and privacy concerns must also be considered. The findings of this study can assist HR professionals in enhancing their recruitment process by incorporating insights from GitHub and LinkedIn profiles.

Keywords: *GitHub, LinkedIn, Rest API, Machine Learning, IT industry, user profile, professional, data visualization*

3 ACKNOWLEDGEMENT

I am pleased to acknowledge the efforts of everyone who has contributed to the preparation of this proposal report. Firstly, I would like to express my gratitude to my supervisor Dr.Anuradha Karunasena and co-supervisor Dr.Lakmini Abheywardhana for providing me with valuable guidance and support throughout the process. I would also like to extend my appreciation to the research team for their dedication and hard work in conducting the necessary research and analysis. I would like to acknowledge the support and cooperation of the participants who contributed their time and insights to the study. Lastly, I would like to thank all the stakeholders who have shown interest in this proposal and provided feedback to help improve its quality. The success of this report is a result of the collaborative efforts of everyone involved, and I am grateful for their contributions.

4 TABLE OF CONTENT

5 LIST OF TABLES

6 LIST OF FIGURES

7 LIST OF ABBREVIATIONS

DM	Data Mining
API	Application
NLP	Natural Language processing
ML	Machine learning
HR	Human Resource
SDLC	Software Development Life Cycle

8 INTRODUCTION

8.1 CV Analysis and Optimizing the Recruitment Process

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 and abilities. Such a method would not only enhance the overall hiring process, but also help to ensure that the most suitable candidates are selected for the available roles. It is widely recognized that individuals possess a unique combination of personality traits, resulting in diverse personalities. However, there is currently no universally accepted method for assessing these traits during the hiring process, which can present a challenge in accurately evaluating a candidate's potential fit with a company culture or role. This limitation highlights the need for a more comprehensive approach to personality assessment in the hiring process.

8.2 GitHub and LinkedIn

GitHub is a web-based platform that allows developers to store and manage their code repositories. It provides a collaborative environment for developers to work on projects, track changes to the code, and manage versions of their code. GitHub provides developers with several functions, including issue tracking, pull requests, code reviews, and project management tools, that facilitate collaboration on projects.

LinkedIn is the world's preeminent social network for professionals. Members create CVs, list their current and previous job roles, skills and education. The business network is also a recruiting website, with businesses able to create profiles and list current vacancies.
(LinkedIn Usage and Revenue Statistics, n.d.)

8.3 Area of Research

A lot of research has been done to evaluate a candidate's professional skills through different aspects and views and aspects. Further, there has been lots of research done to extract data from LinkedIn and GitHub user profiles separately. This research focuses on analyzing and evaluating a candidate's professional skills. Additionally, this research explores the effectiveness of using data from professional media platforms, such as LinkedIn and GitHub, to assess a candidate's professional skills. So, HR could investigate whether this method provides a more accurate representation of a candidate's skills than traditional methods, such as resumes and interviews, and whether it can help recruiters

identify top talent more efficiently.

8.4 Component Overview

The focus of this research component is to develop a way for identifying candidate professional skills using candidate's LinkedIn and GitHub user profiles. Although these platforms provide basic information along with their UI, they are more focused on representing directly what user has directly input to the platform. This component will provide a solution so that the HR will be able to understand the summary inspection and representation of candidate's professional skills through this embedded tool.

9 LITERATURE REVIEW

9.1 Background Study

In today's digital age, companies are increasingly relying on automated online job recruitment processes to screen potential candidates quickly and efficiently. As part of this process, candidate profiles on social media and professional networking sites such as GitHub and LinkedIn are being analyzed to extract valuable information about their professional qualifications.

The purpose of this research paper is to explore the effectiveness of extracting metadata from candidate profiles on GitHub and LinkedIn and using this information to represent their professional qualifications. The study will focus on the use of automated tools to collect and analyze data from candidate profiles and the potential benefits of using this approach in the recruitment process.

Background:

Traditionally, job recruitment processes involved manually reviewing resumes and conducting interviews to determine a candidate's qualifications. However, with the advent of social media and professional networking sites, recruiters now have access to a wealth of information about candidates that can be used to supplement traditional recruitment methods. Which is called social media recruitment.

A survey done by "ISmartRecruiters" in 2021 https://www.ismartrecruit.com/blog-social-media-recruiting-practices

Figure 1 Social media recruitment in 2021

GitHub and LinkedIn are two popular platforms used by professionals to showcase their work experience, skills, and qualifications. GitHub is a code hosting platform that allows IT developers to collaborate on projects and share code. LinkedIn is a professional

networking site that allows users to showcase their work experience and connect with other professionals. Automated tools can be used to extract metadata from candidate profiles on these platforms, such as programming languages used, projects completed, and endorsements from colleagues. This information can be used to represent a candidate's professional qualifications and can provide valuable insights into their skills and experience.

The utilizing of social media platforms, such as LinkedIn and GitHub, has become increasingly prevalent in the hiring process. Employers are now leveraging these platforms to assess a candidate's professional skills and evaluate their potential fit for a role. In today's highly competitive job market, being able to identify top talent efficiently and accurately has become critical for organizations to maintain a competitive edge.

Extracting data from LinkedIn and GitHub profiles provides a wealth of information about a candidate's skills and experience. LinkedIn profiles, for example, provide details about a candidate's education, work history, endorsements, and recommendations. Meanwhile, GitHub provides insight into a candidate's coding abilities, including the programming languages they know, the projects they have worked on, and the contributions they have made to open-source projects.

While the use of social media platforms in the hiring process has its advantages, it also presents some challenges. For example, recruiters may struggle to analyze large amounts of data and extract meaningful insights. Additionally, there may be ethical considerations related to the use of personal data and privacy concerns that need to be addressed.

This research paper aims to explore the effectiveness of extracting data from LinkedIn and GitHub user profiles to evaluate a candidate's professional skills. Specifically, the paper will investigate whether this approach provides a more accurate representation of a candidate's skills than traditional methods, such as resumes and interviews. Additionally, the paper will discuss the ethical considerations of using personal data for hiring decisions and explore ways to ensure that the data collected is used fairly and transparently. By providing insights into this topic, this paper aims to help organizations improve their recruitment processes and identify top talent more efficiently.

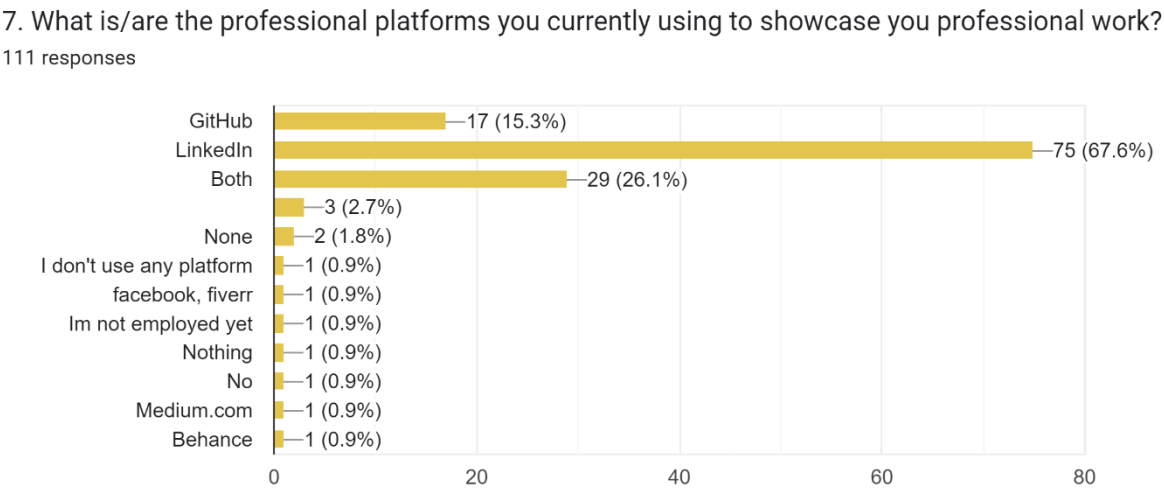

Figure 2 Current Day Professional Media Platforms

Figure 2 shows that most people are using professional social platforms to showcase the professional work they have done. Further this figure shows people are more likely to in GitHub (15.3%) and LinkedIn (67.6%) compared to other platforms. Some of them are using both platforms which represent as 29.1%

According to a study done, based on its global advertising audience reach numbers, LinkedIn has at least **900.2 million** members around the world in January 2023. It suggests that **16.0%** of all people aged 18 and above around the world have an account LinkedIn today. (LinkedIn Statistics and Trends, n.d.) And when it comes to GitHub The total number of developers on GitHub now stands at more than 73 million, with more than 16 million users added in 2021 alone. According to figures released in last year's Octoverse report, this number will increase to 100 million by 2025. The number of first-time contributors increased significantly in the last year to more than 3 million, the highest number of new contributions since 2015. (GitHub Global Growth, n.d.)

.

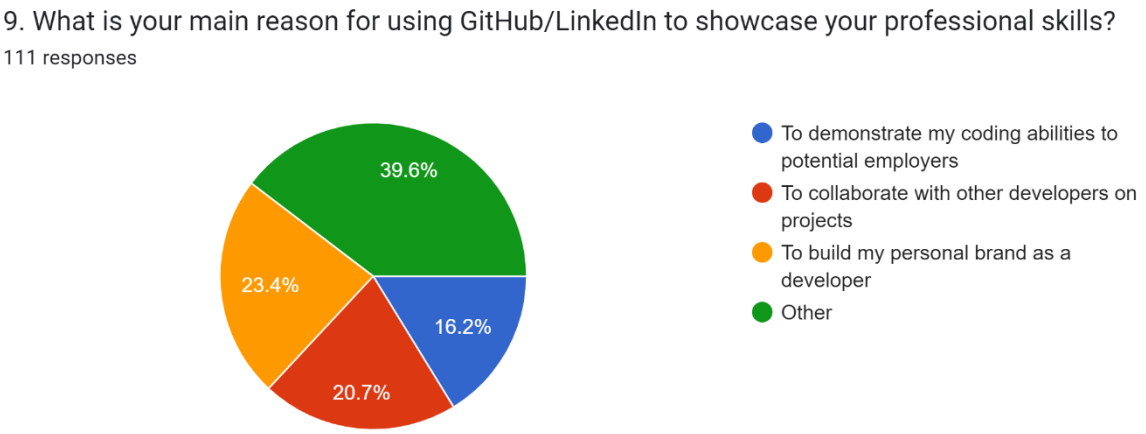

Figure 3 Digital Approach to showcase professional work

Figure 2.2 shows what motivates job seekers to use professional platforms to showcase their work achievements. From the figure it is visible that 23.4% of the people use these platforms to build their personality brand as a developer and 20.7% are using to collaborate with other developers on project.

9.2 Literature Survey

In recent years' competitive job market, identifying the right candidate for a job is crucial for an organization's success. While traditional hiring processes rely heavily on resumes and interviews, these methods may not be enough to fully evaluate a candidate's professional skills. In this literature survey, we aim to critically analyze the existing literature on the use of LinkedIn and GitHub profiles in evaluating professional skills during the candidate hiring process. Specifically, we will review the literature on the types of skills that can be assessed through these platforms, the methodologies used to evaluate skills, and the effectiveness of using these platforms to identify top talent. Through this literature survey, we hope to provide insights into the current state of knowledge on this topic and identify the best practices for using these platforms in the candidate hiring process.

In recent years, the IT industry has experienced a surge in recruitment, resulting in growing need for recruiters and talent scouts to find suitable technical candidates for preliminary assessments. However, they are facing difficulties in identifying potential candidates solely based on resumes.

Research has done in 2022 [1]. This research paper discusses how using novel methods to inspect public GitHub repositories can help manage the influx of candidate profiles during the recruitment process. The manual process of reviewing repositories is both time-consuming and prone to inaccuracies, highlighting the need for a more efficient tool. The paper explores the use of GitHub REST API to gather information on potential candidates from their user profiles.

In [2] it has proposed a pre-screening solution to screen the applicants for the position of Software Engineer where candidates are screen based on a call transcript, GitHub profile, LinkedIn profile , CV, Academic transcript and, Recommendation letters.

Another research has done in 2016 [3] for social coding platforms like GitHub, software developers can display their skills and expertise in specific areas of software development by evaluating their source code contributions. This allows them to showcase their experience and competencies across different projects and programming languages to the community and potential employers. In order to determine the level of contribution, the quantity and continuity of the developer's commits are taken into consideration across various isolated projects over time. By assessing these factors, a clear view of the developer's capabilities can be gained, allowing them to effectively demonstrate their skills and experience to potential employers.

In [4] it has study on *"new comer"* candidate to explore new users to the GitHub platform. Research has further track and characterize their initial contributions of 208 newcomer candidates in GitHub using a mixed-methods approach, analyze whether candidates who are new to a particular coding community engage in social coding practices, and examine the types of contributions they make and the projects they focus on.

Another research has done in 2021 [5] as a feasibility study using the use of Facebook data to enhance the recruitment system's performance and determine a candidate's personality through social network analysis. This highlights the significant role of social media platforms like Facebook and LinkedIn in people's lives, where they tend to spend a lot of time posting updates and uploading certifications. The system is designed to verify the identities of individuals by analyzing data from LinkedIn, Facebook, or both. This is achieved by data scraping and a machine learning algorithm that utilizes a previously trained dataset with SVM.

In [6] it analyze the data of LinkedIn in 2021. It used LinkedIn's overly invasive API in order to scrape a massive amount of personal information data. By linking this information with other API sources, the attacker was able to generate a comprehensive list of data, which was then sold illicitly over the internet. This research investigates the impact of a specific event and proposes possible strategies to mitigate its effects. The suggested countermeasures include implementing adequate authorization and authentication procedures, limiting data scraping activities, and utilizing anomaly detection techniques.

Another study has done [7] in 2021 as An examination of different social media platforms reveals that the substantial volume of data generated by users on these platforms is a valuable source of unstructured information that can be used to study public opinion and gauge people's sentiments. This data can help capture views on social events and corporate strategies. The data is collected from different services and programs by utilizing official APIs and scraping tools. To extract valuable insights about individual users' interests, usage patterns, and media reach, machine learning and image processing algorithms are applied to the gathered data. These algorithms help in analyzing the data, thereby providing valuable insights, and enabling businesses and researchers to understand the users' preferences and behaviors on a deeper level.

In [8] it has been resulted of an emergence field of social network mining or sentiment analysis.

The paper outlines a system designed to analyze data from Facebook and Twitter by utilizing scraping tools and applying machine learning techniques to assess user interests, usage patterns, and media reachability. This system offers a solution for researching public opinion and gaining insights into people's sentiments.

10 RESEARCH GAP

According to the literature survey done, the following issues were found as research gaps,

- Limited system tool only to extract raw data.
- Systems use either LinkedIn or GitHub separately.
- Existing systems lack the analyzing both LinkedIn and

So, this research is done to provide a solution to these identified gaps by developing a system with the ability to showcase candidate's professional skills from both GitHub and LinkedIn user profile data.

When we consider existing implementations in this area:

- In Research A [8], the authors have presented user profiling using text and images for topic modelling, sentiment labelling and image labelling from Facebook, Instagram and Twitter.

- In Research B [7], the authors have proposed a personality prediction from scraping data from digital footprints like and classify them using machine learning approach that previously trained dataset using SVM.

- In Research C [2], the authors have proposed a system where data extracted from the CV and the programming Languages obtained from the GitHub will be further analyzed and then displayed as the candidate's insights.

- In Research D [1], the authors have proposed evaluate the students by providing them the assignments on GitHub and asking them to upload their solutions on it so that they can view their code easily.

Featu **reResearch**	Based GitHub	Based on LinkedIn	Extract raw/trivial data	Based on both GitHub and LinkedIn	Evaluate Professional skills
Research A [7]	×	×	✓	×	×
Research B [5]	×	✓	✓	×	×
Research C [2]	✓	×	✓	×	✓
Research D [1]	×	×	×	×	✓
Proposed system	✓	✓	✓	✓	✓

Table 1 Research Gap

11 RESEARCH PROBLEM

The research question for this study is as follows:

- Are CVs reliable? Can we go beyond the cv and assess a candidate?
    - CVs are typically self- reported documents that are created by candidates. It should not be considered as the sole indicator of a candidate's qualifications or suitability for a job.
    - This research will focus on identifying a candidate's true identity analyzing their professional media accounts.
- Can we proofread a candidate's professional preferences before interviewing?
    - This research will deliver candidate's professional preferences depends on posts and repositories shared on GitHub and LinkedIn by doing the topic modeling. It will further help in the interview process.
- Can we assess a candidate's professional skills from the point of view of collaborators?
    - This research will focus on doing a sentiment analysis on given comments and recommendations in LinkedIn. From this it will deliver an attitude towards the candidate from his network

Research problem can be conveying as when there is a incoming flood of possible candidates for position or vacant, finding what is the effectiveness of using automated tools to extract metadata from candidate profiles on GitHub and LinkedIn to represent their professional qualifications in an automated online job recruitment process.

Therefore, it is necessary to develop a feature where HR would be able to find the most suitable candidate productively with less time and more matching. This will enable a HR person to reduce human interference and speed up the manual process by referring to this tool even there are flood of candidates, which is the case in many situations.

Here are some reasons to understand the need to automate your recruitment process.

1. Save time and money.
2. Reach out to a wider pool of candidates.
3. Keep track of candidates
4. Keep the recruitment process organized.

5. Make better hiring decisions.
6. Leverage social site recruiting.
7. Remove bias and build diversity.
8. Build a better workplace culture.

12 OBJECTIVE

12.1 Main Objective

The main objective of this component is to give the opportunity for a HR system to evaluate and confirm candidate's professional skills through professional social media platforms like GitHub and LinkedIn user profiles.

12.2 Sub Objectives

The following sub objectives should be fulfilled to achieve the specified objective.

- Extracting the content of the candidate's LinkedIn user profile.

  This objective aims at analyzing the extracted candidate LinkedIn profile data using the LinkedIn REST API then summarize the name, follower, about, comments and locations etc. This will give the opportunity to understand professionalism through LinkedIn.

- Extracting content of the candidate's GitHub user profile.

  This objective aims at analyzing the extracted candidate GitHub profile data using the GitHub REST API then summarize the repositories. Programming languages, followers, about, comments, collaborators, and code quality etc. This will give the opportunity to understand professionalism through GitHub.

- Apply machine learning algorithm to train extracted data and graphically represent professional skills.

- Extracting the content of the candidate's LinkedIn and GitHub user profile.

- Generate a candidate profile using professional skills.
- Identify candidate's professional preferences.

This objective aims at a Machine learning algorithm to match to evaluate candidate’s professional skills through extracted meta data and visualize them to get a better understanding on candidates’ professionalism.

13 METHODOLOGY

The objective as mentioned in this component is to provide efficient tools to evaluate candidate's professional work as a summary inspection. To achieve this, the first user profile URLs of LinkedIn or GitHub or both should be entered through a chatbot. Then the candidate's profile should be extracted by using REST APIs to metadata, and it will convert to graphical representation of professional skills. Then those visualizations will be used by the HR or employer to take better decision making by analyzing the insights. Finally, these visualizations will be added to the candidate profile as a report.

13.1 System Architecture Diagram

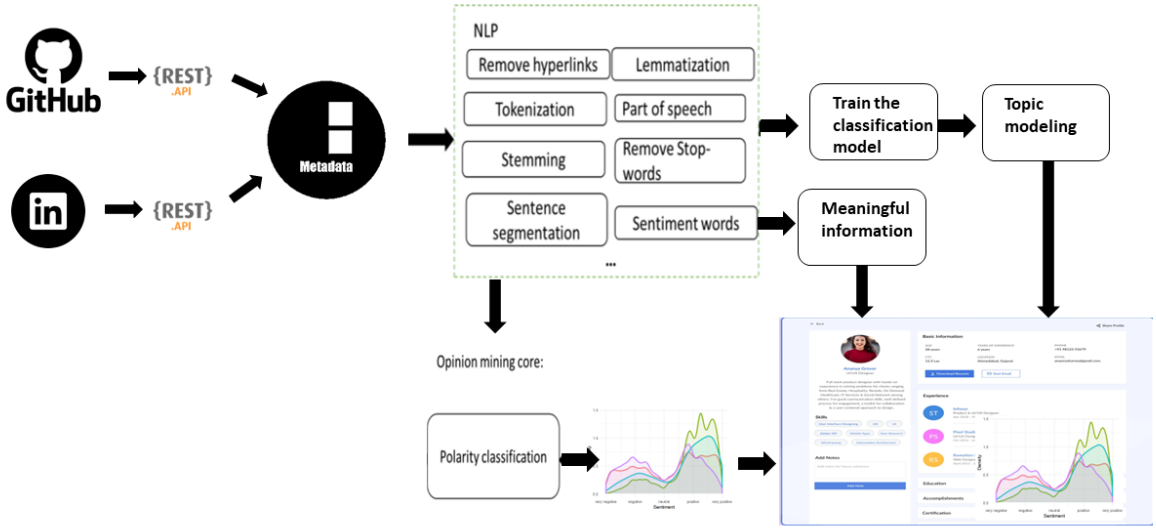

Figure 4 System Architecture Diagram

13.2 Software Solution

SDLC stands for Software Development Life Cycle. It is a framework or methodology that is used in the development of software products, applications, and systems. It is a systematic approach that outlines the various stages or phases of software development from the initial planning and requirements gathering to the final deployment and maintenance of the software. The SDLC process typically involves several stages, including planning, analysis, design, development, testing, deployment, and maintenance. Each stage in the SDLC process has its own objectives, deliverables, and associated tasks, which are typically defined by industry-standard frameworks such as Waterfall, Agile, or DevOps.

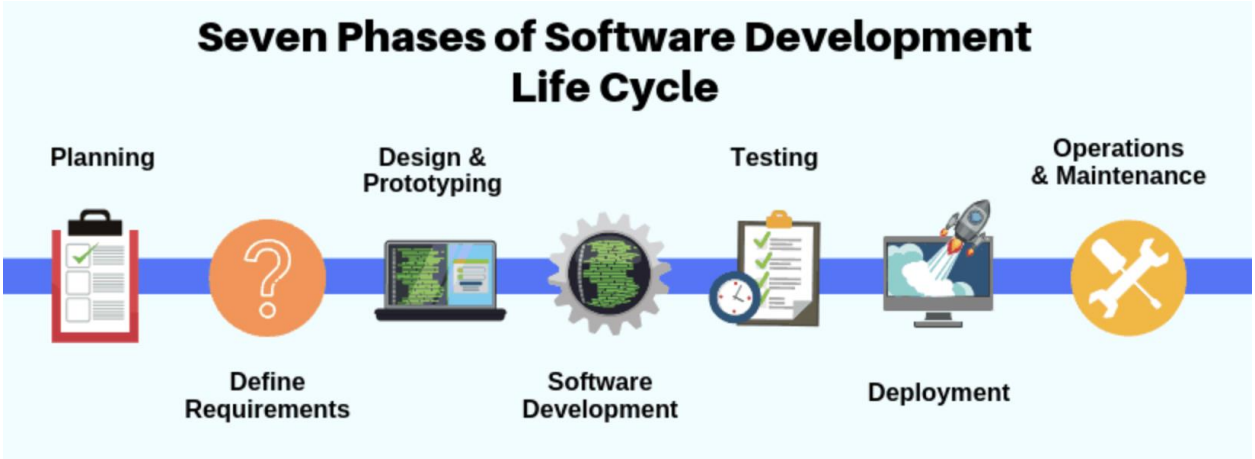

Figure 5 SDLC Illustration

13.2.1 Requirement gathering

This stage includes a gathering of requirements which are to be used to get a clear idea and plan what the system should be and what needs to be implemented. To gather information for this stage a survey was conducted on a set of university students.

13.2.2 Feasibility study

- **Schedule Feasibility**

To deliver the best results a timeline should be created at the beginning and the workshould be done sticking to the deadlines in the timeline. This should include a proper work breakdown structure.

- **Technical Feasibility**

The proposed system should be compatible with the existing technologies and HR should be able to use the system using commonly used computer systems. The developing solution will be built so that it will be technically feasible.

- **Operational Feasibility**

The systems that have been already proposed should be analyzed to identify the gaps that they have not been able to fill, and the proposed system should be built including solutions to these gaps to improve the performance of the system.

13.2.3 Design and implementation

A sketch of the suggested strategy to accomplish the sub objectives should be made during the design process. The project can be moved to the coding step, commonly referred to as the system's implementation once the designing phase is finished.

13.2.4 Deployment and maintenance

Deployment and maintenance processes require careful planning and execution. Initially, the tool needs to be deployed on a reliable server that can handle large volumes of data extraction requests. It is important to ensure that the tool is properly configured and secured to prevent unauthorized access to sensitive information. Regular maintenance and updates are also necessary to ensure that the tool remains functional and up to date with the latest changes in the APIs of both platforms. Additionally, regular testing and monitoring are

critical to identify and resolve any issues that may arise. Overall, proper deployment and maintenance of the tool is essential to ensure that it can effectively support the job recruitment process and deliver the desired results.

13.3 Work Breakdown Structure

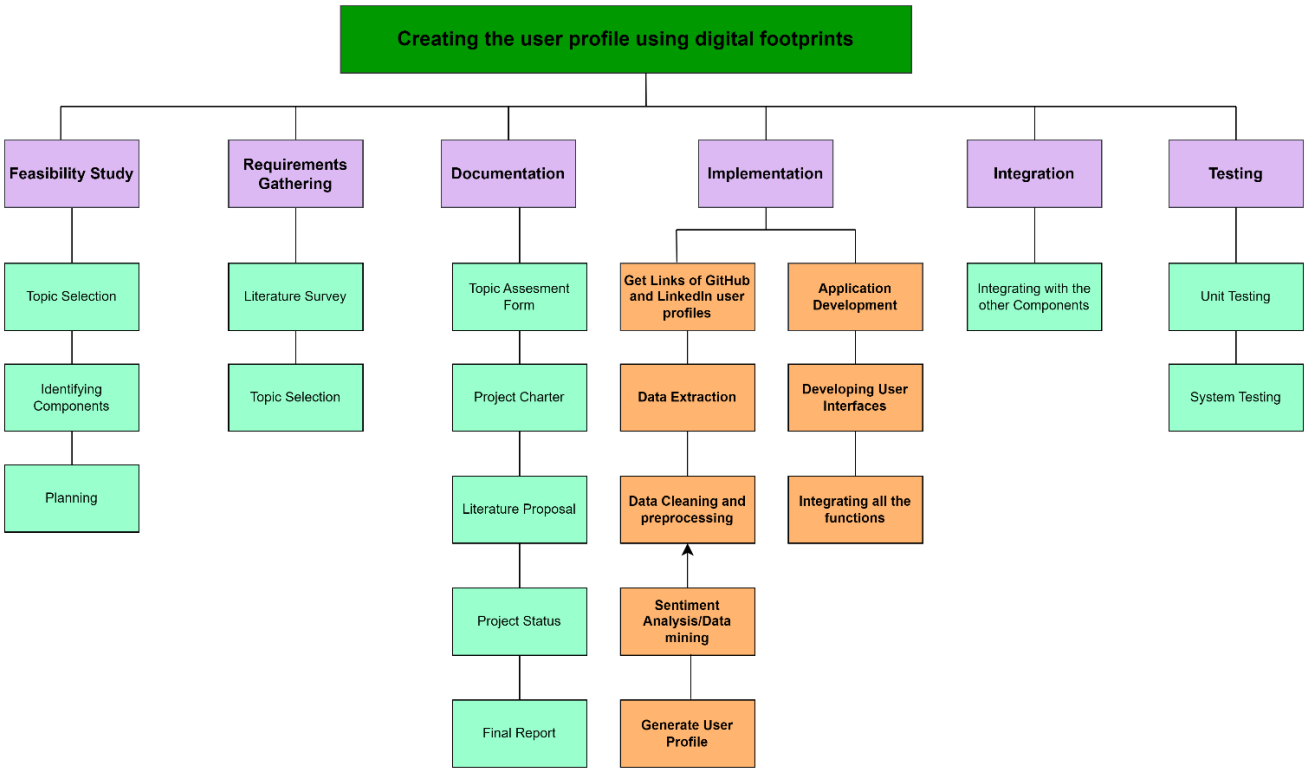

Figure 6 Work Breakdown Structure

14 PROJECT REQUIREMENTS

14.1 Functional Requirements

- The tool should be able to extract data from LinkedIn and GitHub user profiles.
- Ability to showcase professional skills and preferences through extracted data.

14.2 Non-Functional Requirements

- Availability: The system must be available when needed.
- Reliability: The system must be reliable and meet the requirements of the user.
- Usability – the system must be easy to use and understand.
- Scalability – the system should be able to scale up or down as needed.
- Compatibility – The system should be compatible with web browsers and OS.

14.3 User Requirements

- Assess the professional skills of the candidate in an efficient manner.
- Find a candidate professional identity.
- This tool should be secure and follow best practices for data privacy and security.
- Representation of candidate user profile in understandable way.

15 GANNT CHART

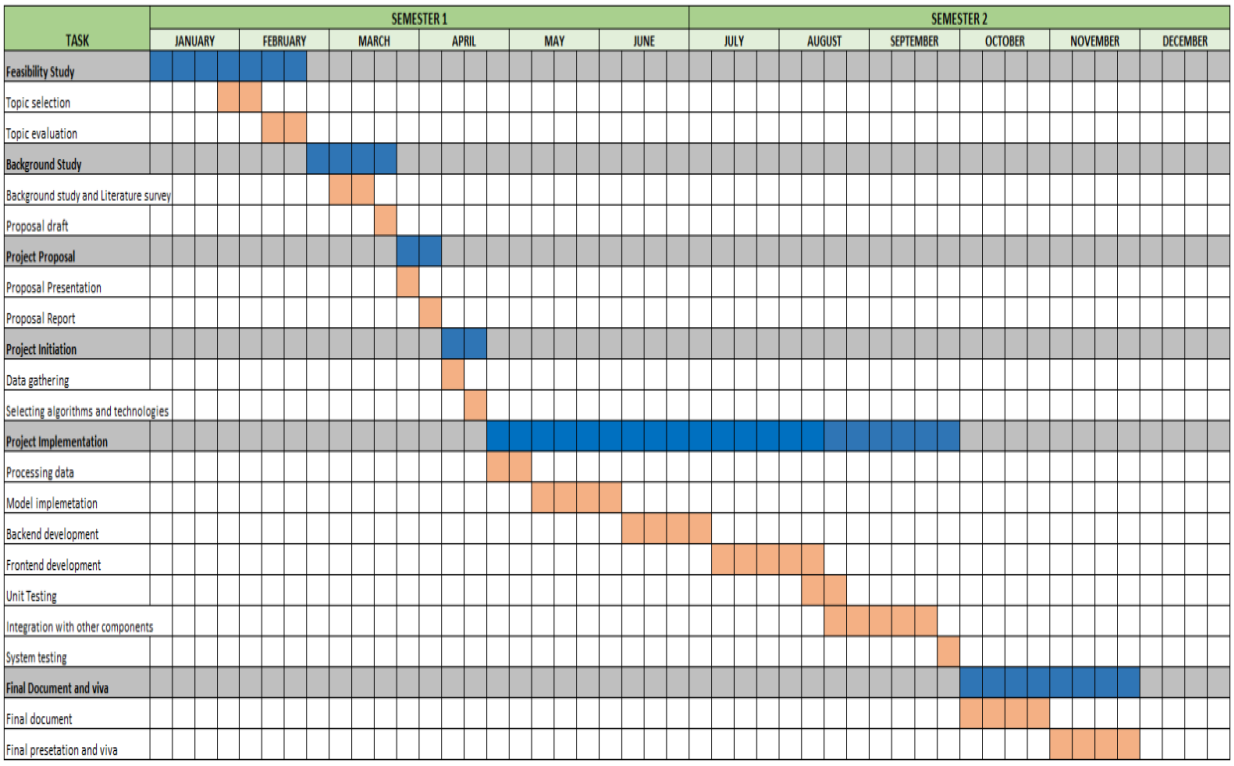

Figure 7 Gannt Chart

16 COMMERCIALIZATION

Commercialization idea for a software tool that extracts data from GitHub and LinkedIn profiles in the job recruitment process is to offer it as a subscription-based service to recruitment agencies and HR departments of organizations. The software tool can be marketed as a time-saving solution that enables recruiters to gather valuable candidate information easily and efficiently from multiple platforms in one centralized location. The service can be offered at various tiers, with different pricing and features based on the volume of data extraction and level of customization required by the client. The software tool can also be marketed as a value-added service to complement existing recruitment software solutions.

This software tool will,

- Reach a wider audience with various backgrounds.
- It is modern, easy, inexpensive, and targeted.
- Gain more information about candidates.
- This component will have two versions.
- Free version - basic features
- Premium version - advanced features
- Target audience would be HR in the IT industry.

17 REFERENCES

GitHub Global Growth. (n.d.). Retrieved from Techmonitor: https://techmonitor.ai/technology/software/github-users-microsoft-thomas-dohmke

LinkedIn Statistics and Trends. (n.d.). Retrieved from DATAREPORTAL: https://datareportal.com/essential-linkedin-stats

LinkedIn Usage and Revenue Statistics. (n.d.). Retrieved from Business of Apps: https://www.businessofapps.com/data/linkedin-statistics/

[1] S. Gupta, B. Gupta, and S. Gupta, “A Novel Method for Technical Candidate Assessment using Github Repository Inspection Automation,” in *2022 International Conference on Computer Communication and Informatics (ICCCI)*, Coimbatore, India: IEEE, Jan. 2022, pp.

1–5. doi: 10.1109/ICCCI54379.2022.9740986.
[2] 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,” in *2020 2nd International Conference on Advancements in Computing (ICAC)*, Malabe, Sri Lanka: IEEE, Dec. 2020, pp. 168–173. doi: 10.1109/ICAC51239.2020.9357279.
[3] F. Chatziasimidis and I. Stamelos, “Data collection and analysis of GitHub repositories and users,” in *2015 6th International Conference on Information, Intelligence, Systems and Applications (IISA)*, Corfu, Greece: IEEE, Jul. 2015, pp. 1–6. doi: 10.1109/IISA.2015.7388026.
[4] 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,” in *2022 4th International Conference on Advancements in Computing (ICAC)*, Colombo, Sri Lanka: IEEE, Dec. 2022, pp. 288–293. doi: 10.1109/ICAC57685.2022.10025337.
[5] E. Constantinou and G. M. Kapitsaki, “Identifying Developers’ Expertise in Social Coding Platforms,” in *2016 42th Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*, Limassol, Cyprus: IEEE, Aug. 2016, pp. 63–67. doi: 10.1109/SEAA.2016.18.
[6] I. Rehman, D. Wang, R. G. Kula, T. Ishio, and K. Matsumoto, “Newcomer Candidate: Characterizing Contributions of a Novice Developer to GitHub,” in *2020 IEEE International Conference on Software Maintenance and Evolution (ICSME)*, Adelaide, Australia: IEEE, Sep. 2020, pp. 855–855. doi: 10.1109/ICSME46990.2020.00110.
[7] S. M. Patil, R. Singh, P. Patil, and N. Pathare, “Personality prediction using Digital footprints,” in *2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS)*, Madurai, India: IEEE, May 2021, pp. 1736–1742. doi: 10.1109/ICICCS51141.2021.9432380.
[8] H. K. S M, S. Hegde, S. G, S. M, S. R, and S. L. N, “User Interest Prediction based on Social Network Profile with Machine Learning,” in *2021 6th International Conference for Convergence in Technology (I2CT)*, Maharashtra, India: IEEE, Apr. 2021, pp. 1–6. doi: 10.1109/I2CT51068.2021.9418126.

18 APPENDICES