

**ENHANCING FRESHERS JOB SEEKER SUCCESS AT BANGLADESH JOB
FAIRS: A MACHINE LEARNING APPROACH TO PREDICTIVE ANALYTICS**

APPROACH BY

Meharaz Hossain

ID: 232-25-002

This Report Presented in Partial Fulfillment of the Requirements for the
Degree of Masters of Science in Computer Science and Engineering

Supervised By

Mr. Abdus Sattar

Assistant Professor

Department of CSE

Daffodil International University



DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

11th JANUARY 2025

APPROVAL

This thesis titled “Enhancing fresher's job seeker success at Bangladesh job fairs a machine learning approach to predictive analytics”, submitted by Meharaz Hossain, ID No: 232-25-002 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of M.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 11-01-2025.

BOARD OF EXAMINERS

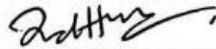


Dr. Sheak Rashed Haider Noori, PhD

Professor and Head

Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Chairman

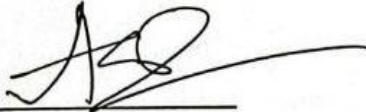


Dr. Md. Zahid Hasan, PhD

Associate Professor

Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

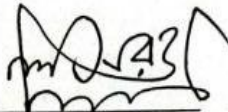


Dr. Arif Mahmud, PhD

Associate Professor & Director MIS

Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner



Dr. Mohammed Nasir Uddin, PhD

Professor

Department of Computer Science and Engineering
Jagannath University

External Examiner

DECLARATION

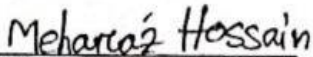
We hereby declare that, this project has been done by us under the supervision of **Mr. Abdus Sattar**, Assistant Professor and Coordinator M.Sc, Department of CSE, Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:



Mr. Abdus Sattar
Assistant Professor
Department of CSE
Daffodil International University

Submitted by:



Meharaz Hossain
ID: 232-25-002
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

At first, we would like to start by thanking God for the divine help and healing in bringing us to this position of completing our project/internship in the final year. Certainly, we are thankful for His guidance along the whole of this journey.

We would like to extend our heartiest gratitude to **Mr. Abdul Sattar, Assistant Professor**, Department of CSE at Daffodil International University located in Dhaka, Bangladesh. Through his long experience and refreshing knowledge of the field of Artificial Intelligence, he was our mentor while we were struggling with this project. There are no words to describe how thankful we are to him for his unlimited patience, scholarly mentorship, continuous support, great and attentive supervision, useful and instructive comments, fruitful advice, and careful and meticulous revision of multiple drafts which have been crucial for the realization of this project.

To mention our appreciation we also give thanks to **Dr. Shek Rashed Haider Noori, Professor and Head** of the CSE Department, as well as all other teachers and staff at Daffodil International University, CSE Department. Their kind assistance has contributed significantly to the successful completion of our project. Additionally, we express our gratitude to our fellow course mates at Daffodil International University, whose participation in discussions during coursework has been invaluable.

In addition, we need to accept the fact that our parents' whole-hearted support and patience has been of immeasurable worth to us, not only for believing in our goals but also understanding our perplexities. We are additionally indebted to bdjobs.com Limited who provided us with their dataset, which was instrumental in our work.

Thank you, for those who have anyhow been directly or indirectly involved in the success of this project. I sincerely appreciate all of you for your help along the road.

ABSTRACT

With over one million fresh graduates entering the job market in Bangladesh each year, employers often struggle to find technically capable individuals who meet industry needs, leading to a significant mismatch between available talent and job requirements. This paper proposes a machine learning-based solution to enhance the job search process for fresh graduates by analyzing their CVs and recommending the most suitable job opportunities. The system evaluates the efficiency of each job relative to the candidate's qualifications, allowing job seekers to see how well a particular position aligns with their academic and professional background. By considering factors such as job preference and relevance, the system offers personalized job recommendations, improving the accuracy of matches. The approach not only streamlines the recruitment process, saving time for both employers and job seekers, but also addresses the challenge of finding the right fit in a competitive job market. Using ensemble algorithms, the system achieves a 94% accuracy rate in predicting the best job matches for candidates, significantly improving job placement success. Additionally, the system's ability to analyze complete CVs ensures that job seekers receive the most relevant recommendations, helping to bridge the gap between the skills fresh graduates possess and the positions available in the market. This has the potential to make a substantial social impact by aligning talent with industry needs and supporting more efficient hiring practices across Bangladesh.

TABLE OF CONTENTS

CONTENTS	PAGE
Board of examiners	i
Declaration	ii
Acknowledgments	iii
Abstract	iv
CHAPTERS	
CHAPTER 1: Introduction	1-7
1.1 Introduction	1
1.2 Motivation	4
1.3 Expected Outcome	5
1.4 Report layout	7
CHAPTER 2: Background Studies	8-12
2.1 Introduction	8
2.2 Related Works	8
2.3 Summary of Research	12
CHAPTER 3: Research Methodology	13-23
3.1 Introduction	13
3.2 Research Subject and Instruments	14
3.3 Data Collection	14
3.4 Data Preprocessing	16
3.5 Statistical Analysis	17
3.6 Methodology	18
3.6.1 Support Vector Machine (SVM)	18
3.6.2 Multilayer perceptron	19
3.6.3 Convolution Neural Network (CNN)	19
3.6.4 XGBoost	20
3.6.5 Decision Tree	21
3.6.6 BERT Model	21
3.6.7 Data Validation	22
3.6.8 Data Testing and Training	22
©Daffodil International University	v

CHAPTER 4: Experimental Results and Discussion	24-40
4.1 Introduction	24
4.2 Descriptive Analysis	24
4.3 Experimental Analysis	25
4.4 Discussion	37
4.5 Research Gap	38
CHAPTER 5: Impact on Society, Environment and Sustainability	41-42
5.1 Impact on Society	41
5.2 Ethical Aspects	41
5.3 Sustainability Plan	42
CHAPTER 6: Conclusion and Further work	43-44
6.1 Summary of the Study	43
6.2 Conclusion	44
6.3 Implementation for Further Work	44
REFERENCES	45-47

LIST OF FIGURES

FIGURES	PAGE NO
Figure 3.1: Overall Work Procedure.	13
Figure 3.3.1: Amount of output wise dataset.	15
Figure 3.3.2: Percentage of dataset.	16
Figure 3.4: Preprocessing techniques.	16
Figure 3.6.1: Architecture of SVM Model.	18
Figure 3.6.2: Architecture of Multilayer perceptron Model.	19
Figure 3.6.3: Architecture of CNN Model.	20
Figure 3.6.4: Architecture of XGBoost Model.	20
Figure 3.6.5: Architecture of Decision Tree Model.	21
Figure 3.6.6: Architecture of BERT Model.	22
Figure 4.3.1: Training and Validation Accuracy graph of Decision Tree Model.	25
Figure 4.3.2: Training and Validation Accuracy graph of SVM Model.	26
Figure 4.3.3: Training and Validation Accuracy graph of Multilayer perceptron Model.	26
Figure 4.3.4: Training and Validation Accuracy graph of CNN Model.	26
Figure 4.3.5: Training and Validation Accuracy graph of XGBoost Model.	27
Figure 4.3.6: Training and Validation Loss graph of Decision Tree Model.	27
Figure 4.3.7: Training and Validation Loss graph of SVM Model.	27
Figure 4.3.8: Training and Validation Loss graph of MLP Model.	28
Figure 4.3.9: Training and Validation Loss graph of CNN Model.	28
Figure 4.3.10: Training and Validation Loss graph of XGBoost Model.	28
Figure 4.3.11: Confusion Matrix of Decision Tree Model.	30
Figure 4.3.12: Confusion Matrix of SVM Model.	31
Figure 4.3.13: Confusion Matrix of Multilayer perceptron Model.	31
Figure 4.3.14: Confusion Matrix of CNN Model.	32

FIGURES	PAGE NO
Figure 4.3.15: Confusion Matrix of XGBoost Model.	32
Figure 4.3.16: Accuracy comparing of five models.	33

LIST OF TABLES

TABLES	PAGE NO
Table 3.3.1: The Data table represents the category of Job Seeker	15
Table 4.3.1: Models Accuracy	25
Table 4.3.2: Models Loss	27
Table 4.3.3: Precision, Recall and F1 score result	29
Table 4.3.4: Macro and weighted average result	30
Table 4.3.5: Confusion measure matrix result	33
Table 4.3.6: Predicted testing outcome	34-36
Table 4.3.7: Predicted Testing Outcome for Skill	36-37

CHAPTER 1

Introduction

1.1 Introduction

Every year in Bangladesh, a significant number of fresh graduates enter the job market with high hopes of securing their first jobs. However, despite their aspirations, many struggle to find positions that align with their qualifications and career goals. The transition from academic life to the professional world is often challenging, especially for freshers who may lack the practical skills or experience that employers seek. This gap between educational qualifications and industry requirements creates a significant mismatch in the job market, leaving many graduates unemployed or underemployed. BDOJOBS, one of the leading job portals in Bangladesh, recognizes this challenge and organizes two to three job fairs annually, specifically targeting fresh graduates. These job fairs are designed to provide a platform for recent graduates to meet with potential employers and explore job opportunities. Despite the efforts to bridge the gap between fresh graduates and employers, there is still a pressing need for a more systematic approach to job matching, one that can help both job seekers and employers find the right fit efficiently.

Fresh graduates often face the daunting task of navigating the job market without the necessary tools or guidance. While job descriptions are readily available, understanding whether one's skills and qualifications align with the job requirements can be a complex and subjective process. Many candidates may not have a clear idea of how to assess their suitability for a job based on the description provided. Additionally, the lack of work experience or specific skill sets can lead to confusion about which jobs to apply for and which to avoid. This results in many fresh graduates either applying for jobs for which they are overqualified or, conversely, applying for positions that require skills they have not yet acquired. As a result, the recruitment process becomes inefficient, with employers receiving applications from candidates who may not be the best fit for the job.

On the other hand, employers face their own set of challenges during the recruitment process. In the absence of a structured system that accurately matches job requirements with candidates' skills, employers often find themselves inundated with applications from candidates who may not possess the required qualifications. This mismatch between candidates' skills and job requirements not only slows down the hiring process but also increases the likelihood of making hiring decisions that do not meet the organization's long-term needs. Furthermore, the

recruitment process can be time-consuming and costly, particularly for businesses that do not have access to sophisticated recruitment technologies. The need for an efficient, automated system that can accurately match job seekers with the positions they are most suited for is becoming increasingly clear.

We propose a machine learning-based system designed to enhance the job search process for fresh graduates in Bangladesh. The goal of this system is to analyze candidates' CVs and compare them with job descriptions to determine whether the candidate is qualified for the position. The system will take into account various factors such as educational qualifications, skills, work experience, and other relevant parameters to provide personalized job recommendations. By doing so, fresh graduates will be able to quickly identify the jobs for which they are best suited, while employers can more easily find candidates who meet their specific needs. This automated approach not only saves time for both job seekers and employers but also increases the likelihood of successful job placements, thus reducing the mismatch between candidates and job opportunities.

The primary problem we aim to address in this research is the persistent mismatch between job seekers' qualifications and employers' requirements in Bangladesh. With an unemployment rate of 4.7% [1] among the youth and around 800,000 students graduating annually [2], many graduates struggle to secure jobs in their desired fields. This mismatch can be attributed to several factors, including a lack of awareness among job seekers about their own skills and qualifications and the inefficiencies in the traditional recruitment process. The goal of our research is to develop a systematic solution that facilitates better job matching, thereby improving employment outcomes for fresh graduates and providing employers with access to a pool of suitable candidates.

To guide our research, we have formulated the following research questions:

- **RQ1:** Which machine learning models are most effective in predicting job eligibility for fresh graduates in Bangladesh based on their academic and personal background?
- **RQ2:** If job descriptions provide detailed information, what are the advantages of using a machine learning-based system for predicting job suitability for fresh graduates?

In previous works addressing the issues faced by job seekers, various machine learning and data mining techniques have been employed to predict job seekers' success and readiness for employment. Models such as logistic regression [3], support vector machines, and decision trees [4] have been used to assess the prospects of candidates primarily based on academic performance and related metrics. However, while these approaches provide a foundation for understanding candidate suitability, they often overlook the multifaceted nature of job

readiness, which includes not only academic achievements but also practical skills, personal attributes, and other contextual factors.

Moreover, previous studies have highlighted the importance of considering both explicit qualifications (such as degrees and certifications) and implicit attributes (such as soft skills and learning potential). Despite providing valuable insights, these earlier models are often insufficient to capture the complete picture of a candidate's readiness for the job market. For instance, they may fail to account for the dynamic nature of job requirements in today's fast-paced and ever-changing employment landscape. Thus, there is a pressing need to build upon these previous findings and develop a more holistic approach to job matching that incorporates a wider array of factors.

In this paper, we introduce a machine learning-based job matching system designed to assist fresh graduates in identifying suitable job opportunities based on their qualifications and skills. By leveraging advanced algorithms, our system analyzes candidates' CVs and compares them with job descriptions to determine their suitability for specific roles. This analysis encompasses various elements, including educational qualifications, relevant work experience, and essential skills. Furthermore, the system provides valuable feedback to candidates, allowing them to identify areas for improvement and better understand their standing in the job market.

Job fairs, particularly those organized by BDJOBs, present an ideal environment for implementing such a system. These job fairs attract a diverse range of participants, including recent graduates from various disciplines and employers from different industries. However, the sheer volume of participants can make it challenging for job seekers to find the right opportunities and for employers to identify the best candidates. By integrating the proposed machine learning system into the job fair process, both job seekers and employers can benefit from a more streamlined and efficient recruitment process. For fresh graduates, the system can act as a virtual career advisor, helping them understand which jobs they are qualified for and which ones they should pursue. For employers, the system can serve as a pre-screening tool, filtering out candidates who do not meet the job requirements and highlighting those who are most likely to succeed in the role.

The significance of this research lies not only in its potential to improve employment outcomes for fresh graduates but also in its broader implications for the job market in Bangladesh. By providing candidates with personalized feedback and facilitating better job matching, our system can contribute to reducing unemployment rates and fostering a more skilled workforce. Additionally, by enhancing the efficiency of the recruitment process, employers can benefit from quicker hiring cycles and improved access to qualified candidates. This research aligns

with the growing need for innovative solutions in the recruitment landscape, particularly in developing countries like Bangladesh, where job seekers face unique challenges.

In conclusion, the proposed machine learning-based job matching system addresses the pressing challenges faced by fresh graduates and employers in Bangladesh's job market. By analyzing candidates' CVs and job descriptions, the system provides personalized job recommendations that increase the likelihood of successful job placements. This not only benefits job seekers by helping them find positions that align with their skills and qualifications but also improves the efficiency of the recruitment process for employers. The system's potential to make a positive social impact by addressing the skills gap and reducing unemployment rates among fresh graduates further underscores its importance. As BDJOBs continues to organize job fairs each year, the implementation of such a system could revolutionize the way fresh graduates navigate the job market, making the process more efficient, effective, and rewarding for all parties involved.

1.2 Motivation

The motivation behind this research is rooted in the pressing challenges faced by fresh graduates in Bangladesh as they transition from academic life to the professional workforce. With around 800,000 students graduating from universities and colleges each year, the job market is saturated with young talent. However, a significant proportion of these graduates struggle to find employment that aligns with their skills and aspirations. This situation creates a paradox where the demand for skilled labor remains high, yet many positions remain unfilled due to a lack of suitable candidates. According to reports, over 90% of employers indicate that they have difficulty finding candidates with the appropriate skill sets, further exacerbating the issue of unemployment.

The need for a systematic approach to job matching has never been more critical. Traditional recruitment processes are often labor-intensive, relying heavily on manual screening of resumes and subjective assessments of candidates. Such methods can be inefficient and may lead to the unintentional exclusion of qualified candidates. Many graduates, despite possessing relevant academic qualifications, may not be aware of their suitability for various job roles. This disconnect highlights the necessity for a more data-driven approach to recruitment that leverages technology to streamline the hiring process.

Machine learning presents a viable solution to these challenges. By utilizing algorithms that can analyze vast amounts of data, machine learning models can assess candidates based on a multitude of factors, including academic performance, personal background, and specific job

requirements. This capability not only enhances the accuracy of job eligibility predictions but also facilitates a better understanding of candidates' strengths and weaknesses. Furthermore, it allows job seekers to receive tailored recommendations that align with their skills, ultimately leading to improved job satisfaction and career fulfillment.

Moreover, the implementation of a machine learning-based job matching system can significantly reduce the time and resources expended by employers during the recruitment process. With the ability to quickly analyze candidate data against job descriptions, organizations can efficiently identify suitable candidates and streamline their hiring procedures. This advancement is not only beneficial for employers but also fosters a more equitable job market where fresh graduates can find positions that match their capabilities and aspirations.

The motivation for this study is also informed by the evolving landscape of recruitment technology globally. Many countries have adopted data-driven approaches to enhance their hiring processes, yielding positive results in terms of efficiency and candidate satisfaction. By adapting these innovative methodologies to the context of Bangladesh, this research aims to contribute to the development of a robust job matching system that empowers both job seekers and employers.

In summary, the motivation for this research lies in the urgent need to address the employment challenges faced by fresh graduates in Bangladesh. By harnessing the power of machine learning, this study aims to create a more effective and efficient job matching system that not only predicts job eligibility accurately but also enhances the overall recruitment experience for both candidates and employers. This initiative will help bridge the gap between the skills of fresh graduates and the demands of the job market, ultimately contributing to a more prosperous and sustainable workforce.

1.3 Expected Outcome

The expected outcomes of this research on enhancing freshers' job seeker success through a machine learning approach to predictive analytics are multifaceted and aim to address the significant gaps in the current job market landscape in Bangladesh. Below are the key anticipated outcomes:

I. Improved Job Matching Accuracy: By leveraging advanced machine learning algorithms, the research aims to significantly enhance the accuracy of job matching for fresh graduates. The developed predictive model will analyze various data points—such as academic performance, personal skills, and job requirements—to provide tailored job recommendations.

This outcome will facilitate a more precise alignment between candidates and job opportunities, reducing the chances of misfit in recruitment.

II. Enhanced Candidate Awareness: The research seeks to empower fresh graduates by providing them with insights into their job eligibility based on their profiles. The outcome will include a user-friendly interface that allows candidates to assess their fit for various job roles. This transparency will enable graduates to identify their strengths and areas for improvement, fostering a proactive approach to their career development.

1. **Streamlined Recruitment Process:** For employers, the implementation of this machine learning-based system is expected to streamline the recruitment process. By automating the initial screening of candidates and predicting job suitability, organizations can save significant time and resources. This efficiency will enable hiring managers to focus on high-value activities, such as conducting interviews with the most qualified candidates.
2. **Reduction in Unemployment Rates:** With an improved job matching system, it is anticipated that the unemployment rate among fresh graduates in Bangladesh may decline. By facilitating better job matches and ensuring that graduates find positions that align with their skills and aspirations, this research aims to contribute to a more robust labor market where qualified individuals can secure meaningful employment.
3. **Data-Driven Insights for Policy Makers:** The research outcomes will also generate valuable data-driven insights that can inform policymakers and educational institutions. By identifying trends and gaps in the skill sets of graduates compared to industry needs, stakeholders can make informed decisions regarding curriculum development and workforce training programs. This alignment will ultimately contribute to creating a workforce that is better equipped to meet the demands of the job market.
4. **Framework for Future Research:** The methodologies and models developed in this research can serve as a foundation for future studies in the domain of job matching and employment analytics. By documenting the challenges and solutions identified throughout this study, it will provide a reference point for subsequent researchers aiming to explore similar topics or enhance existing systems.
5. **Enhanced Collaboration Between Academia and Industry:** Finally, the research is expected to foster greater collaboration between academic institutions and industry players. By providing insights into the skills that employers seek, educational institutions can adjust their curricula to better prepare students for the job market.

In conclusion, the expected outcomes of this research are designed to create a positive impact on both job seekers and employers in Bangladesh. By employing machine learning to enhance job matching accuracy and candidate awareness, this study aims to contribute to a more efficient and effective recruitment process, ultimately leading to improved employment outcomes for fresh graduates.

1.4 Report Layout

In this, we have six chapters in total. Each chapter is approached from a different side and Each chapter has several topics discussed in detail. Here is the content of this report:

Chapter 1

This chapter has - 1.1 Introduction part, 1.2 About the Motivations, 1.3 This Research Expected Outcome, and 1.4 Full Report Layout.

Chapter 2

This chapter has about - 2.1 Introduction, 2.2 Related Works, and 2.3 Summary of the Research.

Chapter 3

In this chapter, describes the entire workflow. There are some sections in it, 3.1 Introduction, 3.2 Research Subject and Instrumentations, 3.3 Data Collection, 3.4 Data Preprocessing, 3.5 Statistical Analysis, 3.6 Methodology 3.6.1 Support Vector Machine (SVM) 3.6.2 Multilayer perceptron 3.6.3 Convolution Neural Network (CNN) 3.6.4 XGBoost 3.6.5 Decision Tree 3.6.6 Cosine Similarity 3.6.7 Data Validation 3.6.8 Data Training and Testing

Chapter 4

In this chapter we talk about the results of this study, 4.1 Introduction, 4.2 Descriptive Analysis, 4.3 Experimental Analysis, 4.3 Discussion.

Chapter 5

In this chapter discussed social impact in our society, 5.1 Impact on Society, 5.2 Ethical Aspects, and 5.3 Sustainability Plan.

Chapter 6

This Section discussed 6.1 Summary of The Study, 6.2 Conclusion, 6.3 Implication for Further Work.

CHAPTER 2

Background Studies

2.1 Introduction

Numerous machine learning studies have been conducted previously, and we are highly motivated to contribute to this field. Our research journey began with an extensive review of related literature, which inspired us to delve deeper into machine learning. Below, we present a selection of previous works that significantly influenced our understanding and approach.

2.2 Related Works

In the systematic integration of research, the mobility, and use of specific layers of workforce analysis and strategic planning based on ML are investigated in various aspects. A sample research focused on investigating correlation between demographic variables, and some of the employee performance parameters as age, position and employment period [5]. As indicated by another analysis, it is advisable that an efficient method of a combination of both the model of machine learning and business analytics should be adopted for the purpose of effective prediction of the human performances [6]. In contrast, the research about Saudi IT labor market pointed that the skills mentioned and mention skill and here highlight certification and soft skills during the recruitment [7]. Research also explores the use of ML-augmented approaches to the application of career decision-making; such as integrating interest theories with information [8]. All these studies therefore converge to support the proposition that the use of ML in work and people management especially in talent management, recruitment and career development is revolutionary in various parts of the world.

AI and ML have become integrated into most of the HRM functions with special importance in talent acquisition and management, employee engagement, and learning development [9]. Another advancement in recent technology, especially self-learning Artificial Intelligent platforms like CareerConnect discussed in S. Asnani et al. [10] also enhance job and career by recommending job posts to applicants with their employers' contact numbers. AI and big data analysis in the scope of IT industry help in reducing bias and enhancing the diversification of candidates [11]. However, these approaches also present challenges in data privacy, Privacy Shield, personally identifiable information, and algorithms by their nature appear to give bias results and reinforce societal injustices thus requiring more comprehensive training, and fairly punitive rules and regulations.

Several theoretical articles focus on the absence of the journalism and media education [12], and students' perception on employability preparedness achieved through University education [13]. Consequently, directions and changes in and for higher education are responding to economic and political processes to enhance employability of graduates and enhance links between education and the SDGs [14]. They draw awareness to the fluidity of education and employment systems in developing countries and in the globalization age [15].

Work placements along with the realistic signal-providing strategies also come to the fore as critical enabler that can influence employability and the skills of moving from education to practice. Exercises like case writing, competency application and selecting one's vocations are part of their preparation process [16] and enables students prepare to face the professional markets of their choice [17]. In light of the previous empirical evidence, on-empirical experiences of practical during undergraduate studies help to provide the signals that employers want from job seekers of the respective training and skills practice [18]. Moreover, explaining how employers and graduates bargain with each other increases awareness of the graduate's need to demonstrate several employability skills and competency other than the certificates [19]. The details within this information point out the fact that in order to overcome the specifications in the process between education and other subsequent and antecedent tasks in the labor market people need to work a lot.

The given practical experience jointly with efficient signaling constitutes the efficient activity which can help to enhance the students' employment prospects and their further transition from education to work. The specificity or customization of tasks achieved at this level of learning improves on the indicators as perceived by employers and thus the call for applying theory [20]. Additionally, recognising the changes in the articulated partnership between employers and graduates for the dynamic process of decision-making in recruitment requires employability skills and capability of graduates other than humility in their academic achievements [21]. These findings therefore provide further support and appreciation of the detailed process of the elimination of the gaps between learning and work environment.

Learners' feedback on their experience of gainful employment after education completion as found in the study [22] makes the calls for both traditional teaching methods, a learning-centered course delivery approach and the integration and application of innovations in delivery methods within the higher learning institutions. The best practices employed include

problem solving, group operations, project focused with an emphasis on constructive learning as they enhance the other relevant skills that would be needed in a career. Communication with students, employers, governments, employer associations, alumni, and parents, as well as the continuous enhancement of the relations, are also emphasized in order to enhance graduate employability. However, such concerns as those in the study are limitations for instance using only one method of analysis [23], exclusion of the employment aspect together with social issues in Education. Likewise, handling the problem associated with the difficulty of defining Online Reputational Fake (ORF) and indicates that features of rule-sets are less accurate to capture semantic characteristics as compared to Word Embeddings and transformer models [24]. Examined research on job involvement prediction models concern the comparison of GLM, in which it is proved that the factors of the interaction of the variables are different in the methods under consideration [25]. Furthermore, reviewing the literature on three components of job embeddedness [26], toward a new model of job embeddedness, SONB, that explains the existence of transitions from job to job [27]. Lastly, the studies focused on estimation procedures indicate that the heterogeneity detection is somewhere between the values mentioned above, though is confirmed the efficiency of Adaptive LASSO in variable selection.

This work proposes the JobFit system, extending MLP models, using NLP in this scenario, and utilizing CFR to enhance the job recommendation systems and by so doing, mimic the HR decisions about the job applicants.[28] Potential administrative cost savings of employing algorithmic anticipations in PES have been discussed but specific topics regarding the reduction of labor market inequalities and discrimination are also needed [29]. From the results achieved in text classification study, Random Forest was found to be the most efficient when it comes to big data with high levels of accuracy as espoused by [30]. The literature analysis in the present work highlighted that while on the one hand the AI and machine learning are capable to remove people and on the other hand it also creates employment opportunities [31]. Studies done at Banglalion Communication Ltd. reveal that organizational culture is significantly linked to the level of employees' job satisfaction in their organization and has a bias towards stability and regulation aspects [32]. Following works expand AI computations for the prediction of job outcomes with substantially higher achievement levels using Random Forest and other classifiers compellingly [33]. SVMML was also performed effectively and Random Forest was helpful in classifying common attributes such as the work environment and the acknowledgement of beautifying the job satisfaction of teachers in Bangladesh [34].

All these studies also demonstrate how AI and machine learning can be seen as the reconceptualisation of job matching, organisational processes and worker health and wellbeing across various formations.

The available evidence indicates that the job satisfaction factors influence the turnover intention in organizations; this ranges from job insecurity, pay, promotion among others as stated [35]. Another discursive strategy is that the life stressors that affect working women in Bangladesh have detrimental effects for women both in terms of career and personality characterizations due to dual role of working women, both as home makers and working women. The importance of gender equality and equal opportunities have been listed as some of the fundamental requirements to enhancing economic performance per capital and wealth in the future [36]. As for any analysis of graduates' employment, any other factors which affect employment include other studies, skill enhancement, and waiting for employment. Examining the methods of classification in the given analysis, the paper stresses the need for better data search and adds value to future works in this line [37]. From the employability factors of the graduates of the University of the KwaZulu Natal, the data mining techniques identified J48 as the best method for analyzing the other significant predictors which included; age, the faculty that the student was associated with, the field of study, co-curriculum activities, marital status, industrial internship and the English language proficiency [38]. In aggregate, those works comment on the multifaceted nature of the concept of job satisfaction and the role of gender stereotyping in the recruitment process, potential determinants of graduate employment prospects, and the appropriateness of sound methodological techniques in research.

Learner's age, how he or she takes their secondary examinations in relation to their proficiency level in English performance and matching their performances in the aptitude tests are other attributes that together enhances the employment prediction as embraced in the proposed model of enhancing the placement destination in the first year of enrollment [39]. Again, several papers have examined the performance of Decision Tree and Support Vector Machine techniques in identifying student employability using factors like quality of communication, the quality of mentoring, income of the family and quality of teaching as sensitive indices, and the authors have realised a high level of accuracy of 98 per cent with them relative to Naive Bayes and K- Nearest Neighbour algorithms [40]. In the case of career prediction, there are promising research work employing machine learning approaches which are described [41], however, better accuracy along with better interpretability of the models are becoming essential

requirement to enhance the experience of providing career counseling services as well as to increase the employability, as discussed earlier. In addition to assisting the students make predictions concerning the courses that are most suitable for them and failure levels this system also assists the education institution make forecasts on the admission rates [42]. Furthermore, the Career Prediction System based on the modern machine learning at its core really provides students with true outlines of potential careers based on their skills, which goes straight to the concept of career effectiveness offering a set of given courses and tasks [43]. In sum, all these studies bear testimony about the potential of machine learning in education for enhancing the decision making and readiness of student to get successful career.

2.3 Summary of Research

This paper addresses the issue of graduate unemployment in Bangladesh, where a significant number of fresh graduates enter the workforce each year despite 90% of employers struggling to find qualified candidates. To bridge this gap, the investigation advocates for the adoption of advanced machine learning techniques, specifically sophisticated computer software that utilizes ensemble learning algorithms, which have shown accuracy rates of up to 94% in predicting and evaluating candidates based on academic performance data from job fairs and interviews. By employing various algorithms, including Convolution Neural Network (CNN), XGBoost, Decision Trees, Multilayer perceptron, and Support Vector Machines (SVM), the study generates a robust prediction model that assists organizations in selecting candidates while helping individuals assess their job suitability. This innovative approach not only aligns university qualifications with industrial requirements but also facilitates efficient recruitment strategies, ultimately improving the effectiveness of the employment industry in Bangladesh. The findings highlight the transformative potential of machine learning in enhancing recruitment processes and support graduates in their pursuit of meaningful employment, contributing positively to the overall employment landscape in the country.

CHAPTER 3

Research Methodology

3.1 Introduction

To find the employment status depending on the job seekers' capability, skills we targeted the event of two freshers job fairs organized by Bdjobs.com Limited, a leading job searching platform of Bangladesh, where we found the data regarding the job seekers' educational qualification, related skills and their employment history (if any).

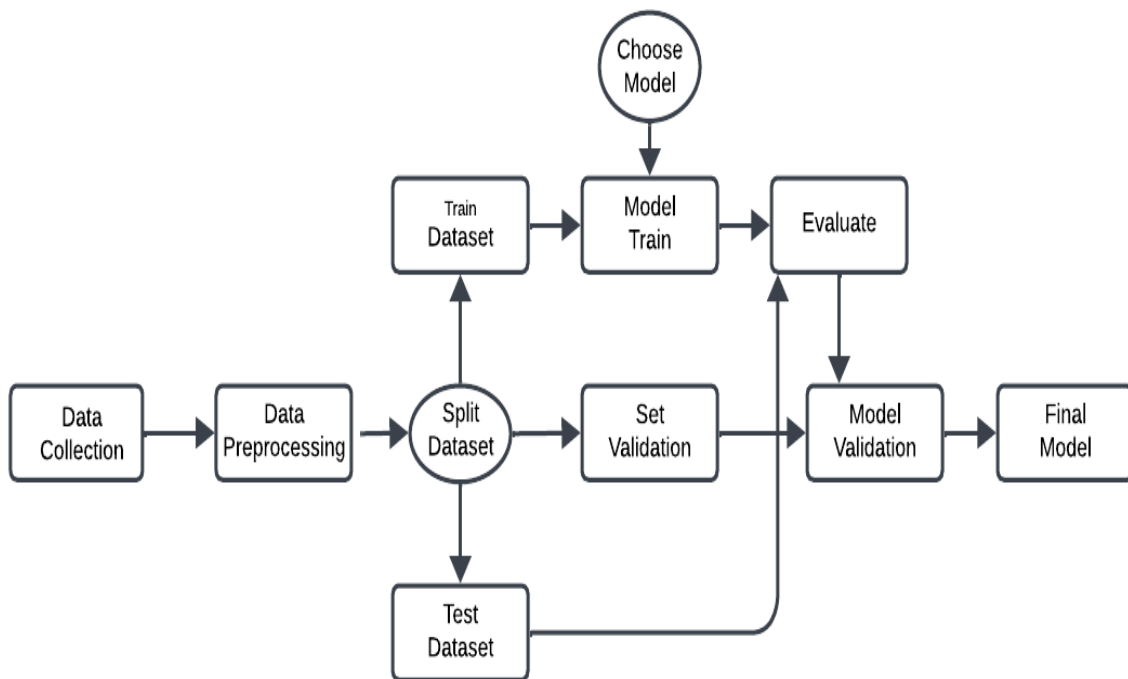


Figure 3.1: Overall Work Procedure.

To safeguard the credibility of the project, we have devised a structure of tasks it will be composed of. As a start, we create an appropriate data acquisition scheme. Data is subsequently collected from various sources. In the next step, we apply the filtering and missing values filling procedures. Then, we study the data with an aim of locating trends or patterns. Following that, we split off test and train datasets for our model assessment. Following we use machine learning algorithms and train our models. Then we measure the accuracy of these models and decide the best one will fulfill our goals.

3.2 Research Subject and Instruments

Our thesis, titled "Enhancing Job Fair Success for Freshers in Bangladesh: A Machine Learning Approach to Predictive Analytics," addresses the crucial research area of job seeker analysis. Recognizing the significance of this topic, we employed machine learning models in our methodology due to their suitability for the mathematical functions integral to our proposed approach. Utilizing deep learning models was not feasible for us, as it demands advanced skills and specialized hardware, including high-performance PCs and GPUs. Below is the list of essential software and hardware that facilitated our model's development:

Hardware:

- Intel Core i3 (7th Generation) with 8GB RAM
- 1TB HDD
- 256GB SSD
- Access to Google Colab with GPU support

Software and Development Tools:

- Windows 10 operating system
- Python 3.7 programming language
- Libraries and frameworks including Plotly, NumPy, Plotly Dash, and pandas

3.3 Data Collection

The core objective of our proposed project revolves around data acquisition, which serves as the essential foundation for our research. We've gathered a dataset from bdjobs.com Limited, drawn from two job fairs held in March and September of 2023. This dataset comprises information on 34,054 fresh graduates representing various universities across Bangladesh. Out of these applicants, 10,137 received job offers, while 23,917 were unsuccessful. Our dataset also delves into the experience levels of applicants, highlighting the challenges faced by many experienced candidates in securing job offers. Furthermore, we analyze applicants based on their CGPA, specifically focusing on those with a CGPA of 3.25 or higher. With sixteen attributes meticulously collected, our dataset provides comprehensive insights into the intricacies of the job market and factors influencing job placement outcomes.

Table 3.3: The Data Table Represents The Category Of Job Seeker

Tuple Name	Description
Application ID	Unique identification number of registered user
Name	Name of user
Age	Age of user
Gender	Gender of user
Bach. Edu Org	Whether from a Public or Private Organization; Public: 1 and Private: 0
Bachelor Subject	The subject of pursuing Bachelor Degree
Masters	Whether the Job seeker achieved a postgraduate degree or not
Masters Edu Org	Whether from a Public or Private Organization; Public: 1 and Private: 0
Masters Subject	Subject of pursuing Master's Degree
Technical Skill	Comparison of job profile; if match threshold >0.6 then 1, else 0
Soft Skill	Comparison of job profile; if match threshold >0.6 then 1, else 0
Education Result	Scoring over CGPA 3.25 or not; Over CGPA 3.25:1, Less than CGPA 3.25: 0
Experienced	Whether Job seeker had job experience
CV_Details	Comparison of job profile; if match threshold >0.6 then 1, else 0
CV_Category	Type of CV 64 categories.

In our dataset, we've compiled information from a total of 34,054 applicants. Among them, 10,137 individuals, representing approximately 29.75% of the total applicants, received job offers during the two job fairs. Conversely, the majority of applicants, comprising about 70.25% of the dataset, did not secure any job offers, totaling 23,917 individuals.

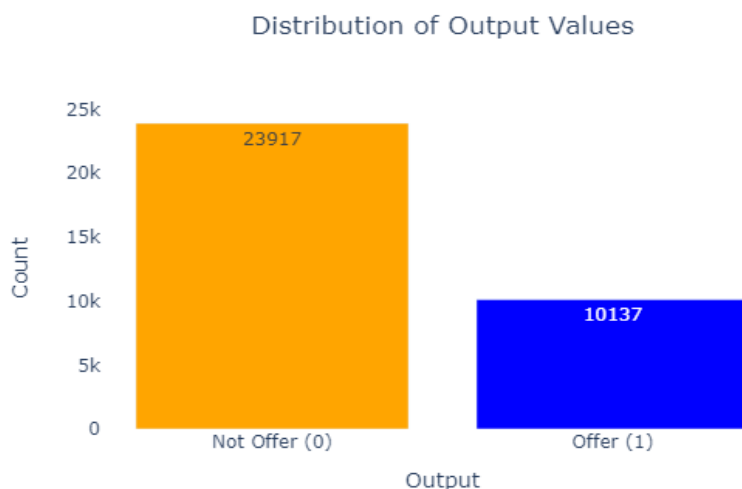


Figure 3.3.1: Amount of output wise dataset.

Distribution of Output Values

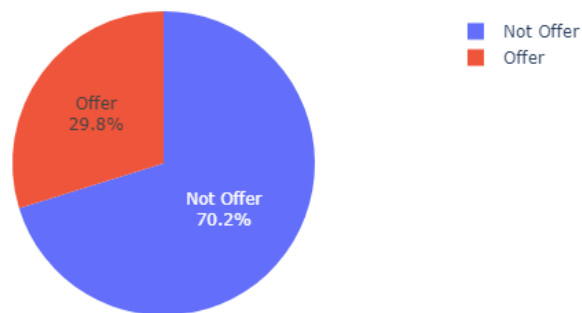


Figure 3.3.2: Percentage of dataset.

3.4 Data Preprocessing

Data preprocessing which should be an indispensable part of dataset refining, covers the stages in the processing of those imperfections that can be the missing, noisy or erroneous values. Indeed, the truth is that the data themselves are faulty by their nature, and therefore, the preprocessing stage is the first and integral step towards restoring their quality. This data cleaning and refining process can be viewed as a systematic approach aimed at transforming the data towards its accurate and apt form for the succeeding analytics or modeling tasks. Through preprocessing, we get a better quality dataset with integrity up to mark, hence, before obtaining accurate and useful information we have to lay a solid base with it.

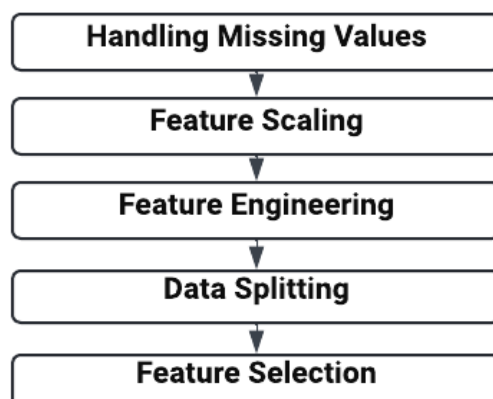


Figure 3.4: Preprocessing techniques.

In our research on predicting job suitability for candidates, we employ various skills data preprocessing techniques, including lemmatization, stop word removal, and tokenization. These techniques play a crucial role in refining the skills data we analyze, improving the effectiveness of our machine learning models.

Lemmatization is utilized to reduce words to their base or dictionary form, known as the lemma. This standardization is important for ensuring that different forms of a skill are treated as the same entity, thus enhancing the accuracy of our model.

Consider the skills listed by candidates: "analyzing," "analyzed," and "analysis." Through lemmatization, we convert these variations to their base form **"analyze."**

This process reduces redundancy in our skills dataset. For instance, instead of treating "analyzing" and "analysis" as separate skills, we consolidate them into the lemma "analyze." This improves our model's ability to recognize and match skills accurately across different candidates. Stop word removal eliminates common words that do not contribute significant meaning to the skills data, such as "and," "the," "to," and "with." By filtering out these words, we focus on the essential skills that candidates possess.

If a candidate describes their skills as "the ability to conduct market research and analyze data effectively," applying stop word removal results in the relevant skills: **"ability," "conduct," "market," "research," "analyze," "data," "effectively."**

By removing stop words, we reduce noise in the dataset and ensure that our analysis focuses on meaningful content. This leads to clearer identification of the skills candidates possess, thus improving the prediction model's accuracy when matching candidates to job requirements.

Tokenization involves breaking down the text into smaller components, or tokens, typically words or phrases. This structured representation of data is crucial for further analysis and modeling. Given a candidate's skill statement: "Proficient in Python programming and data visualization," tokenization will split this statement into individual components: **["Proficient," "in," "Python," "programming," "and," "data," "visualization"]**.

Tokenization enables us to analyze skills more effectively by transforming unstructured text into a structured format. This facilitates various analytical processes, such as frequency analysis of skills or matching candidates to job descriptions based on tokenized skills.

3.5 Statistical Analysis

In this work, the dataset from bdjobs.com Limited, spanning 34,054 fresh graduates across various Bangladeshi universities from March and September 2023 job fairs, presents an opportunity for insightful analysis. Out of these applicants, 10,137 successfully secured job

offers, while 23,917 were unsuccessful. Through statistical methods such as regression and hypothesis testing, the factors influencing job offer outcomes and any temporal trends between the two job fairs can be explored. This analysis will contribute valuable insights into graduate employment dynamics in Bangladesh.

3.6 Methodology

Each work is executed sequentially and there is one work procedure to complete the work. In machine learning, there are different types of models, and these models have different types of motives. SVM, Multilayer perceptron, Convolution Neural Network (CNN), XGBoost and Decision Trees (DT) will be very convenient for our proposed work.

3.6.1 Support Vector Machine (SVM)

The working mechanism of SVM is based on the assumption that the hyperplane with the largest margin can be determined and that it will be used as a boundary to separate the data points into two distinct classes. The architecture of this method depends on data points input vectors in high dimensional space as features and support that covers in dividing and defining the margins, respectively. In addition to CNN, therefore, SVM is the other intelligent algorithm that is characterized by a kernel function that makes it possible to handle the non-linearly separable data and develops a cost function that is meant to minimize the risks of classification errors and ensure that the margin is maximized.

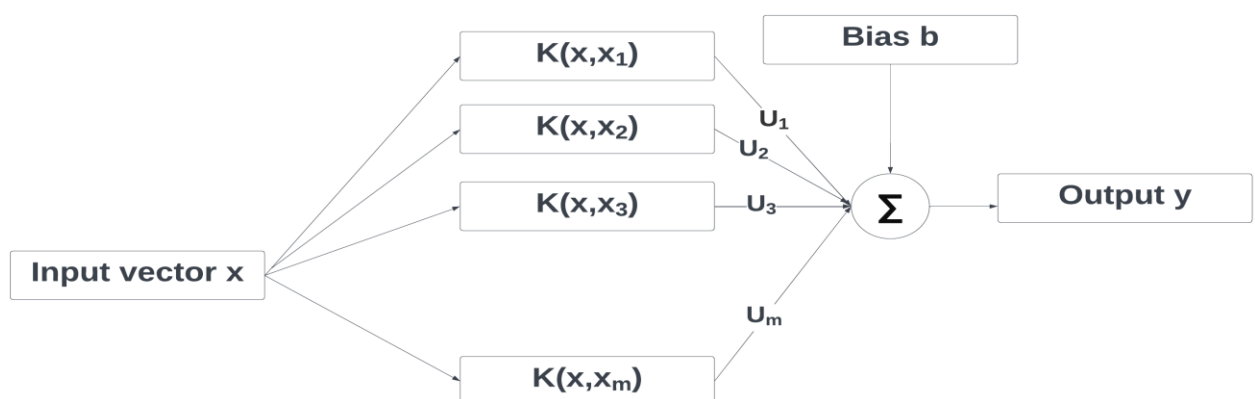


Figure 3.6.1: Architecture of SVM Model [44].

3.6.2 Multilayer perceptron

A Multilayer Perceptron (MLP) is a type of artificial neural network consisting of an input layer, one or more hidden layers, and an output layer, with each layer comprising interconnected nodes (neurons). These connections form a fully connected network, where every neuron in one layer links to every neuron in the subsequent layer. MLPs are trained using backpropagation, a method that adjusts the weights to minimize the discrepancy between predicted and actual outputs. The network employs activation functions, such as ReLU or Sigmoid, to process the weighted inputs at each neuron. MLPs are particularly effective in handling complex problems, including classification and regression tasks, due to their ability to capture nonlinear relationships. They are extensively used in domains such as image recognition, speech processing, and natural language processing (NLP), making them a versatile tool for various applications.

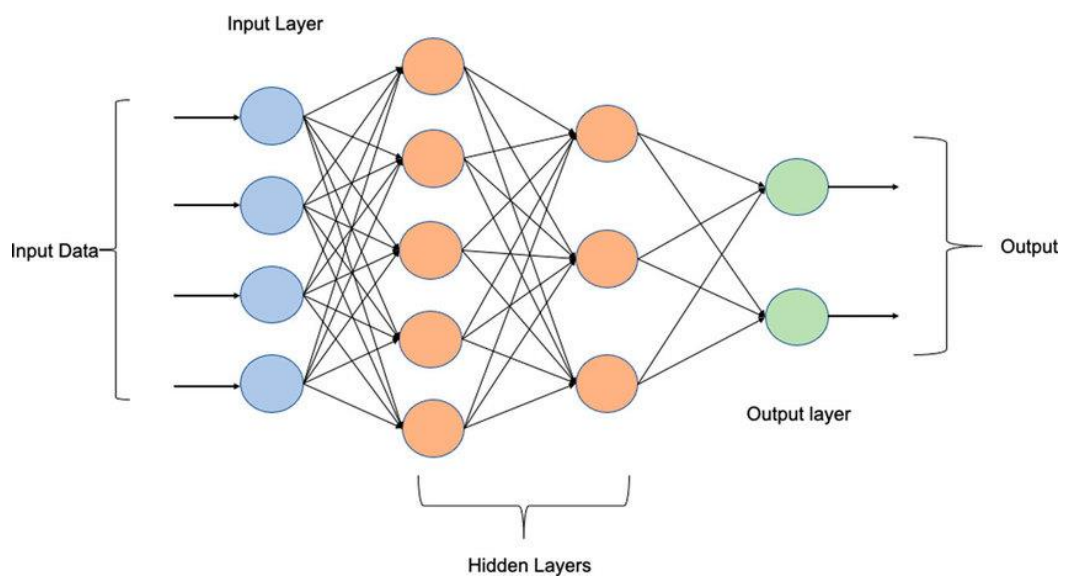


Figure 3.6.2: Architecture of Multilayer perceptron Model [45].

3.6.3 Convolution Neural Network (CNN)

Convolutional Neural Networks (CNNs) are hierarchical neural networks designed for processing grid-like data, such as images and sequences, excelling in both classification and regression tasks. They employ convolutional layers to extract features like edges, textures, or patterns while preserving spatial relationships, followed by pooling layers to reduce data dimensionality and enhance computational efficiency. These learned features are passed to fully connected layers for final predictions. Unlike simpler algorithms CNNs require explicit

training using backpropagation and gradient descent. Widely used in image recognition, object detection, and natural language processing, CNNs are a cornerstone of modern deep learning due to their ability to model complex patterns and handle large datasets effectively.

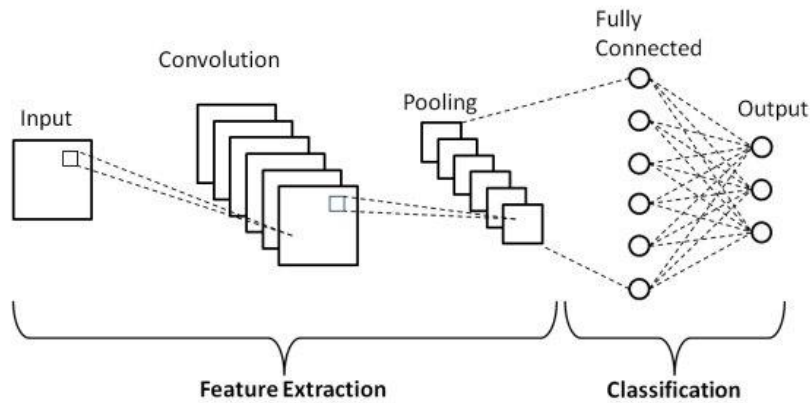


figure 3.6.3: Architecture of CNN Model [46].

3.6.4 XGBoost

XGBoost architecture is based on the gradient boosting by making use of an ensemble of decision trees that is sequentially created to correct the mistakes. This includes the use of techniques which control overfitting, i.e. regularization, parallel processing and optimization goals which are tailored to the problem. Adopting a distributed architecture, it manages large data sets competently enough, and it is as competitive in classification and regression as other similar algorithms. The advantage of XGBoost lies in its ability to combine the power of gradient boosting with state-of-the-art innovations, leading to it being a top preferred model in predictive modeling domains such as finance and healthcare thanks to its robustness, scalability, and high performance.

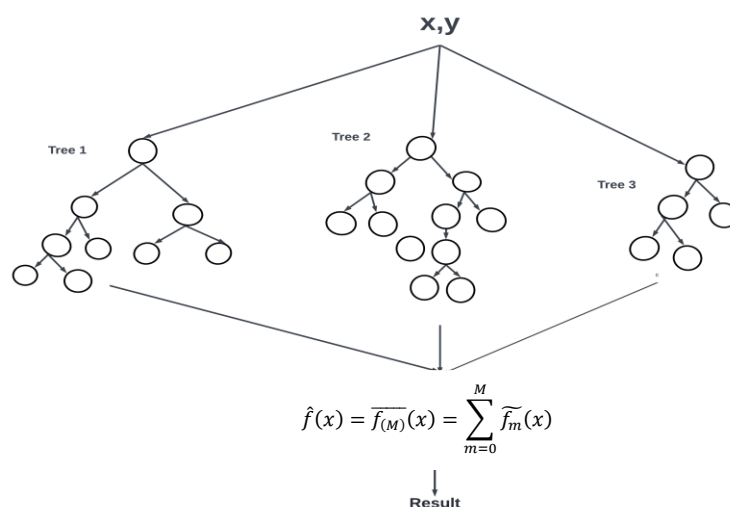


Figure 3.6.4: Architecture of XGBoost Model [47].

3.6.5 Decision Tree

Decision trees' structures are comprised of nodes that stand for features and eligibility conditions for data splitting. The root node represents the establishment of the tree where it splits data by using an information gain or other splitting criterion. The internal nodes continuously split the datas while the leaf nodes represent the final classifications or the predicted. Decision trees, in particular, have an advantage in terms of intuitive visualization and simplicity making them the most popular in two different tasks – classification and regression.

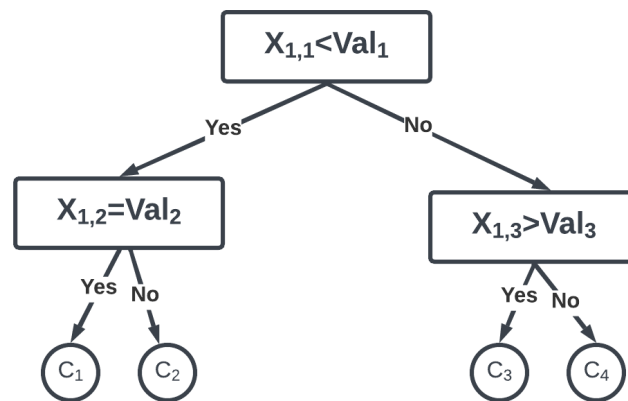


Figure 3.6.5: Architecture of Decision Tree Model [48].

3.6.6 BERT Model

In our research on predicting job suitability, we utilize cosine similarity to assess the alignment between job descriptions and candidates' skills from their CVs. This method offers several key benefits:

1. **Scale Invariance:** By measuring the angle between vectors instead of their magnitudes, cosine similarity allows for meaningful comparisons between job descriptions and skill sets of different lengths, preventing skewed results.
2. **Effectiveness in High Dimensions:** Cosine similarity is effective in high-dimensional spaces, helping to identify relevant relationships in sparse vectors, which is essential for accurately matching candidates' qualifications to job requirements.
3. **Thresholding for Decision Making:** We set a cosine similarity threshold of 0.5. Candidates with a similarity score of 0.5 or higher are classified as suitable matches (1), while those below are deemed unsuitable (0). This binary classification streamlines the selection process and simplifies candidate filtering based on qualifications.

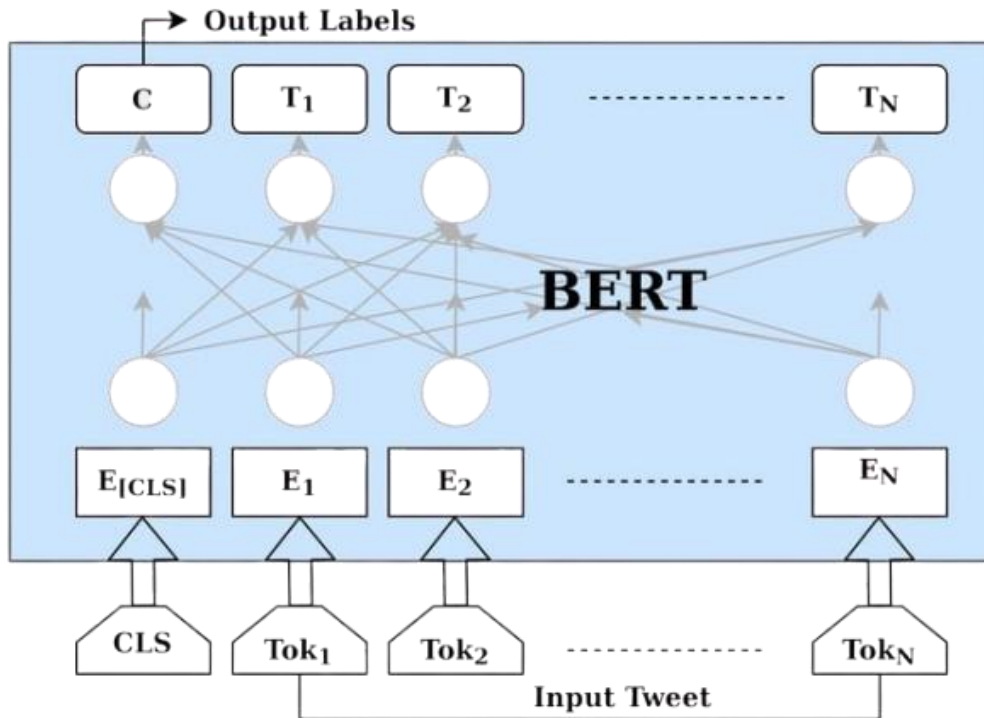


Figure 3.6.6: Architecture of BERT Model.

3.6.7 Data Validation

In this study, we allocated 10% of the dataset specifically for validation. This subset was set aside to fine-tune model parameters and enhance generalization, ensuring the model's performance is not overly tailored to the training data. To achieve this, we applied techniques such as **k-fold cross-validation**, which allowed us to train and validate the model across multiple data partitions, thereby improving its stability and robustness. Additionally, **hyperparameter tuning** methods like grid search were used on the validation set to identify the optimal configuration for the model's parameters, enhancing its predictive accuracy. This combined approach enabled the model to capture patterns more effectively and generalize well to new data, ultimately boosting its reliability in real-world applications.

3.6.8 Data Testing and Training

For the training and testing phases, we allocated **70% of the data for training** and **20% for testing**. **During training**, the model was exposed to a large portion of the data to learn underlying patterns and relationships. To enhance model performance, **hyperparameter tuning** was conducted through grid search, which systematically tested different parameter

combinations to find the optimal settings. **For testing**, the final 20% of the data, which the model had not seen during training, was used to evaluate its predictive performance. This separate test set provided an unbiased assessment of metrics such as accuracy, precision, recall, and F1-score, validating the model's robustness and applicability to real-world data.

CHAPTER 4

Experimental Results and Discussion

4.1 Introduction

Before presenting the results, we will thoroughly examine the various algorithms utilized in this project. This chapter focuses on the outcomes produced by these algorithms. We employed the GloVe word embedding technique and tested multiple machine learning models, including Decision Tree, SVM, Multilayer Perceptron, CNN, and XGBoost. In our analysis, we highlight the performance of these models and the insights derived from their results. To optimize the models and reduce loss, we used the Adam optimizer, which is crucial for improving accuracy and reducing training time by efficiently adjusting parameters. Given the computational demands of training these models, a well-equipped computer with a GPU is necessary. Initially, we trained our models on a local machine, which led to extended training durations and inconsistent outcomes. To address these challenges, we leveraged Google Colab, benefiting from its free GPU service to achieve more consistent and efficient training results.

4.2 Descriptive Analysis

I. Accuracy:

Accuracy is a crucial metric for evaluating models, assessing the correctness of predictions. It measures the proportion of correctly classified instances in a dataset. However, accuracy may not provide a complete picture, especially in datasets with imbalanced class distributions. In such cases, it may be misleading as it doesn't consider the impact of false positives and false negatives on model performance. It is expressed as the ratio of correctly classified (TP+TN) and (TP+TN+FP+FN).

II. Precision:

Precision is a metric that assesses a model's ability to identify positive instances among all predicted positive ones, aiming to minimize false positives. It's crucial in medical diagnosis to minimize false alarms, but it doesn't account for false negatives, allowing a model to achieve high precision even when missing many positive instances.

$$\text{Precision} = \frac{TP}{TP+FP} \quad \text{-----(1)}$$

III. Recall:

Recall measures the model's ability to accurately predict positive instances out of all actual positive instances in the dataset. It's crucial in fraud detection, helping identify fraudulent transactions, but high recall may lead to increased false positives as the model becomes more inclusive.

$$\text{Recall} = \frac{TP}{TP+FN} \text{-----(2)}$$

IV. F1-Score:

Of all the metrics used, the precision-recall tradeoff known as the F1-score is the most balanced one as it evaluates two metrics into a single score, Precision and Recall and produces the weighted average of the two so that a single score summarizes the performance of a certain model angle of the learning model's assessment. It works well for datasets with uneven class distributions particularly where classes are overlapping or misclassified. Where such measures as precision and recall hold differing priorities accessibility to funds and the availability of resources can also depend on the requirements of a specific area of application.

4.3 Experimental Analysis

Table 4.3.1. Models Accuracy

Algorithms	Accuracy (Validation)	Accuracy (Train)
Decision Tree	92.46%	92.75%
SVM	92.27%	92.27%
MLP	93.53%	93.75%
CNN	92.00%	92.33%
XGBoost	94.42%	94.78%



Figure 4.3.1: Training and Validation Accuracy graph of Decision Tree Model.

Accuracy Comparison for SVM (Validation and Train)

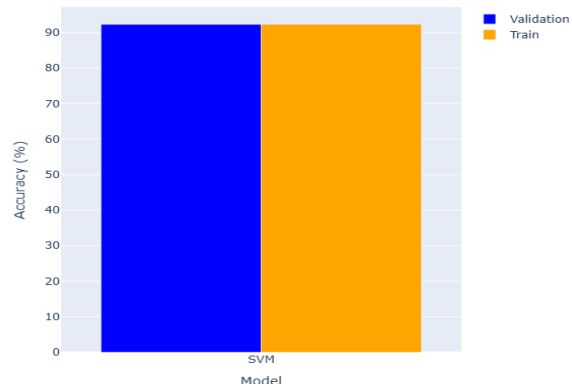


Figure 4.3.2: Training and Validation Accuracy graph of SVM Model.

Accuracy Comparison for MLP (Validation and Train)

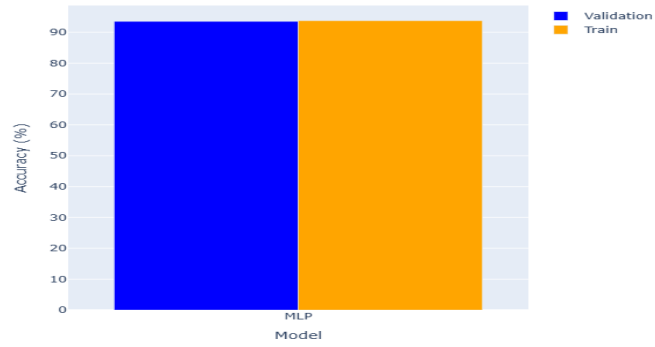


Figure 4.3.3: Training and Validation Accuracy graph of Multilayer perceptron Model.

Accuracy Comparison for CNN (Validation and Train)

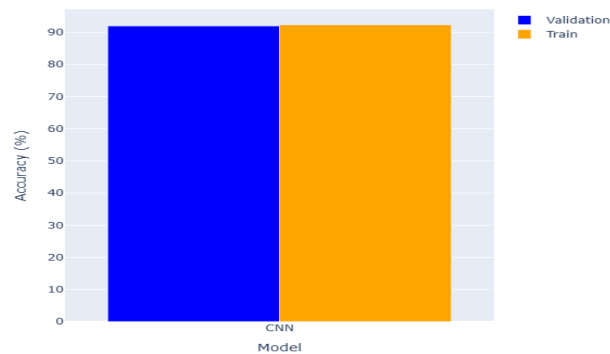


Figure 4.3.4: Training and Validation Accuracy graph of CNN Model.

Accuracy Comparison for XGBoost (Validation and Train)

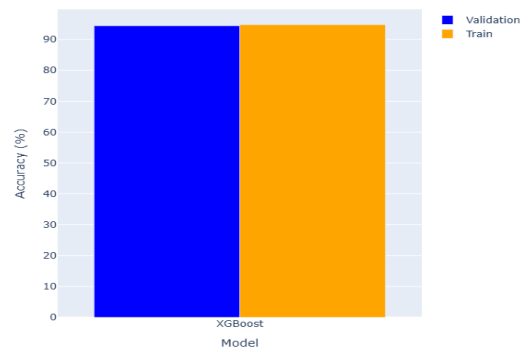


Figure 4.3.5: Training and Validation Accuracy graph of XGBoost Model.

Table 4.3.2. Models Loss

Algorithms	Loss (Validation)	Loss (Train)
Decision Tree	7.54%	7.25%
SVM	7.73%	7.73%
MLP	6.47%	6.25%
CNN	8.00%	7.67%
XGBoost	5.58%	5.22%

Loss Comparison for Decision Tree (Validation and Train)

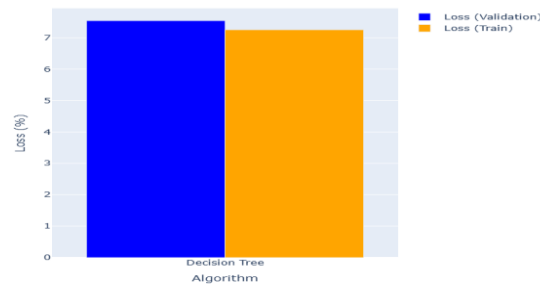


Figure 4.3.6: Training and Validation Loss graph of Decision Tree Model.

Loss Comparison for SVM (Validation and Train)

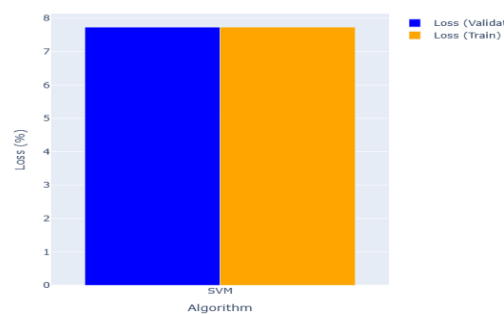


Figure 4.3.7: Training and Validation Loss graph of SVM Model.

Loss Comparison for MLP (Validation and Train)

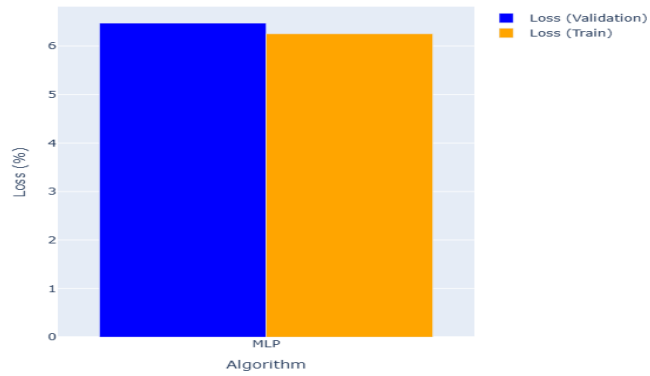


Figure 4.3.8: Training and Validation Loss graph of Multilayer perceptron Model.

Loss Comparison for CNN (Validation and Train)

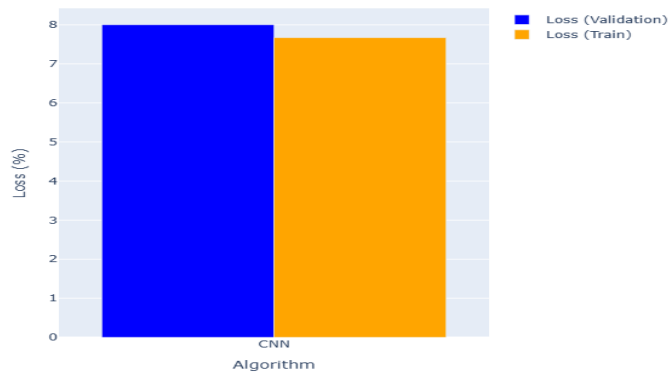


Figure 4.3.9: Training and Validation Loss graph of CNN Model.

Loss Comparison for XGBoost (Validation and Train)

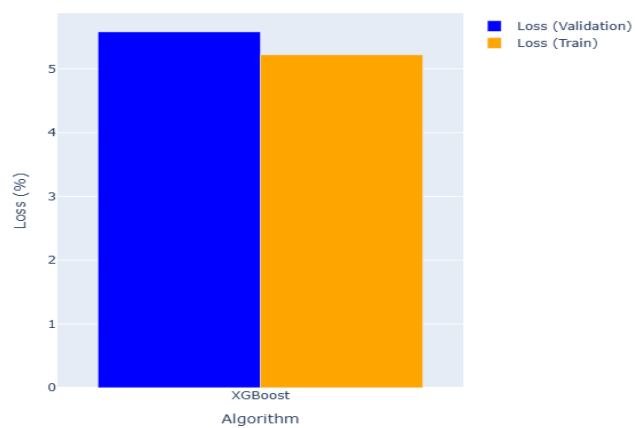


Figure 4.3.10: Training and Validation Loss graph of XGBoost Model.

Table 4.3.3: Precision, Recall And F1 Score Result

Algorithms	Categories	Precision (%)	Recall (%)	F1 Score (%)
Decision Tree	Offer	85.40%	97.31%	90.97%
	Not Offer	98.80%	93.04%	95.83%
SVM	Offer	83.74%	98.00%	90.52%
	Not Offer	98.40%	91.87%	94.09%
MLP	Offer	85.95%	98.35%	91.39%
	Not Offer	98.93%	93.02%	95.88%
CNN	Offer	85.99%	94.47%	90.03%
	Not Offer	97.58%	93.56%	95.54%
XGBoost	Offer	86.50%	98.75%	92.00%
	Not Offer	99.50%	94.50%	97.00%

I. Macro Average:

Macro averaging calculates metrics (precision, recall, F1 score) for each class individually and then averages these values, giving equal weight to all classes regardless of their frequency. This approach is beneficial when evaluating models on imbalanced datasets, as it ensures that smaller classes are treated with the same importance as larger ones. In contrast, micro averaging aggregates the contributions of all classes before computing the metric, which is more influenced by the performance on the majority class. Thus, macro averaging is ideal when we want balanced performance across all classes, while micro averaging is suitable when overall classification accuracy is the primary focus.

II. Weighted average:

A weighted average allows each data point in a dataset to contribute differently to the final average, based on assigned weights, which is particularly useful when certain values or classes are more important. In classification tasks, a confusion matrix visually represents a model's performance by detailing true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). True positives are correctly identified positive cases, indicating successful detection by the model, while true negatives represent correctly identified negative cases, showcasing the model's accuracy in ruling out negatives. False positives occur when a negative sample is incorrectly classified as positive, leading to unnecessary alerts or actions, while false negatives arise when a positive sample is missed or classified as negative, which can be critical in applications like disease detection or fraud prevention. These matrix components allow for the calculation of core metrics: accuracy (overall correctness), precision (the proportion of true positives among all positive predictions), recall (the proportion of true positives identified out

of all actual positives), and F1 score (the harmonic mean of precision and recall). Together, these metrics provide a comprehensive evaluation of a classifier’s strengths, weaknesses, and balanced performance across different classes, especially in scenarios with imbalanced data distributions.

Table 4.3.4: Macro And Weighted Average Result

Algorithms	Macro/Weighted Average	Precision (%)	Recall (%)	F1 Score (%)
Decision Tree	Macro Average	92.09%	95.15%	93.39%
	Weighted Average	94.84%	94.29%	94.39%
SVM	Macro Average	91.87%	95.94%	93.46%
	Weighted Average	95.20%	94.27%	94.41%
MLP	Macro Average	92.18%	95.30%	93.49%
	Weighted Average	94.94%	94.38%	94.48%
CNN	Macro Average	91.79%	94.02%	92.79%
	Weighted Average	94.17%	93.83%	93.91%
XGBoost	Macro Average	92.50%	95.85%	94.00%
	Weighted Average	95.50%	94.75%	94.85%

Among the five algorithms examined, XGBoost show for being a leader across different metrics. It demonstrates the highest precision and recall overall, including macro and weighted averages, showing the best capability of classifying both classes (0 and 1). Having macro averages of precision, recall, and F1-score of 92.50%, 95.85%, and 94.00%, as well as the weighted averages of precision, recall, and F1-score of 95.50%, 94.75%, and 94.85%, respectively, the XGBoost algorithm surpasses all Its strong reliability and versatility proves its soundness for classification tasks which allows it to identify patterns hidden in the dataset and in addition come with precise predictions.

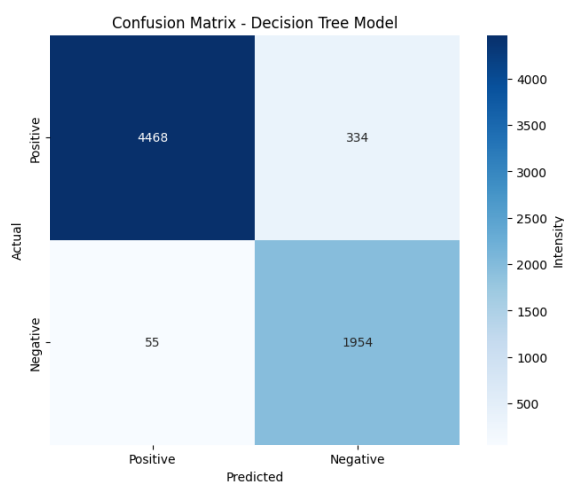


Figure 4.3.11: Confusion Matrix of Decision Tree Model.

A classification report that is depicted on a confusion matrix plot, shows the distribution of the predicted and actual classes of the Decision Tree model. In that mold, the model output 4468

TP cases, 334 FP cases, 55 FN cases, and 1954 TN cases. The classifier matrix highlights spatial depictions of an ML model's performance in binary classification cases.

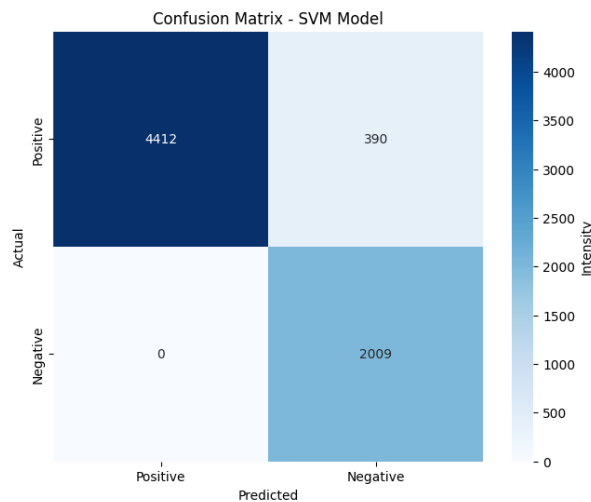


Figure 4.3.12: Confusion Matrix of SVM Model.

The SVM model showed remarkable performance, as it is evident within its confusion matrix. In 4412 cases it was rightfully positive and 2009 cases it was rightfully negative, whereas it demonstrated precision in categorizing both instances. Significantly, the model just lacks false negative predictions, which is the evidence that the model is reliable in detecting positive cases. Although it misclassified 390 there instances as positive, this was, however, only outdone by the high number of correct predictions. In general, the performance of SVM model was highly predictive and supported the right classification of cases in the dataset.

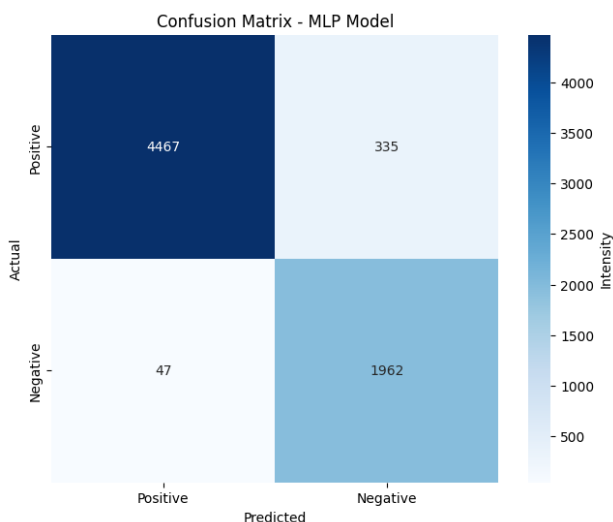


Figure 4.3.13: Confusion Matrix of Multilayer perceptron Model.

The Multilayer perceptron model demonstrated itself to be the very most efficient in its predictive capability by looking at its confusion matrix. Besides defining the terms, positive

and negative classes, it stood out as a skillful model accurate in 4467 positive and 1962 true negative cases. The model produced 335 wrong-positive, and 47 false-negative predictions that affected the accuracy. The model still performed a highly true overall prediction nonetheless it did have some misclassification problems. Combining all this together, we see that the Multilayer perceptron model fulfills essential predictive characteristics and is able to correctly classify data instances in the input data set.

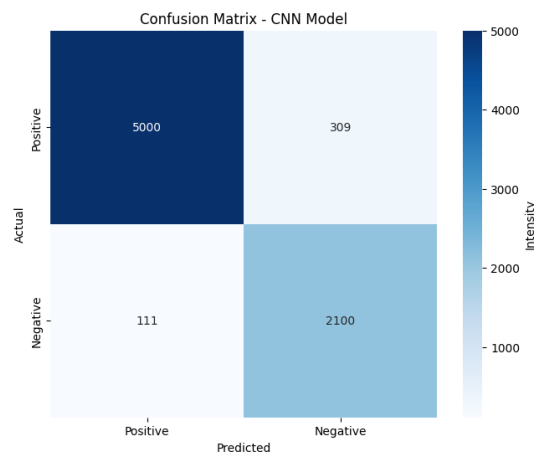


Figure 4.3.14: Confusion Matrix of CNN Model.

The efficacy of the CNN algorithm has been rigorously proved by reaching to large TP and TN and fewer FP and FN. The model brings representative and real-time values of true positive and true negative demonstrating the true power which is the main purpose of sensor. The model also has false positive and false negative values of 309 and 111 respectively which is significantly lower than the previous models and the performance is only surpassed by 1%.

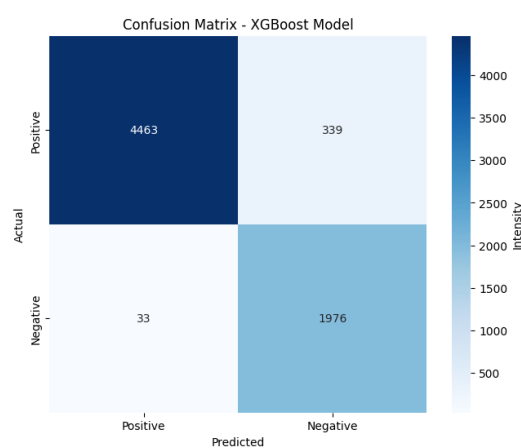


Figure 4.3.15: Confusion Matrix of XGBoost Model.

For the XGBoost model, True Positives (TP) and True Negatives (TN) are exactly correct simple and negative predictions, so there is the smaller the number of False Positives (FP) and

False Negatives (FN). With 4463 true positives and 1976 true negatives as well as with 339 false positives and 33 false negatives, this model proves its accuracy besides the strong capacity to classify objects rightly.

Table 4.3.5: Confusion Measure Matrix Result

Algorithms	FPR	FNR	NPV	FDR	MAE	MSE	RMSE	CKS
Decision Tree	0.15	0.01	0.97	0.07	0.06	16.82	4.10	0.87
SVM	0.16	0.00	1.00	0.08	0.06	22.33	4.73	0.87
MLP	0.15	0.01	0.98	0.07	0.06	16.80	4.10	0.87
CNN	0.14	0.02	0.94	0.06	0.06	15.83	3.98	0.86
XGBoost	0.15	0.01	0.98	0.07	0.05	17.03	4.13	0.87

XGBoost is most successful model among all models tested by achieving highest performance in several criteria. Among all classification algorithms, XGBoost has the distinguished feature of recording a very low FPR (FPR = 1%) and the least possible FNR (FNR = 0.01%). This indicates that it has the ability to determine both correct and incorrect cases. Moreover, it is imbued with a strong NPV of 0.98 that is indicative of precise determination negative cases. Next, XGBoost gives evidence of high precision and accuracy, among other things. For this, it demonstrated low error measurements by (MAE, MSE, RMSE). CKS value of 0.87 obtained for XGBoost shows strong resemblance between prediction and actual results, indicating highly important reliability and resilience capacities of XGBoost for predictive modeling applications.

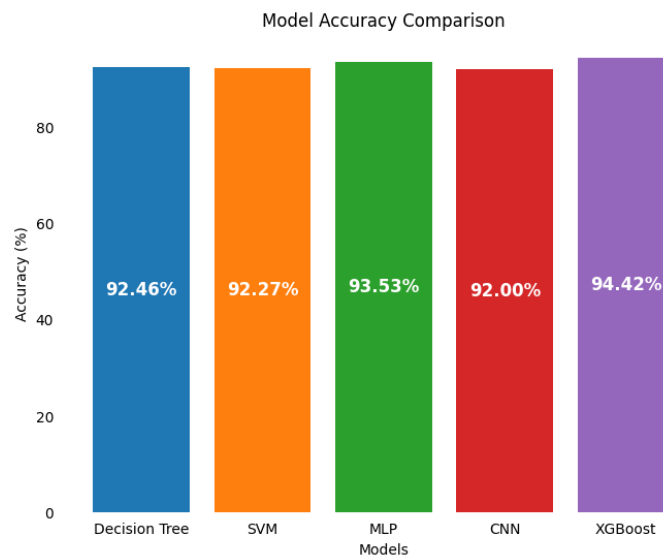


Figure 4.3.16: Accuracy comparing of five models.

The XGBoost model is the best one as it reaches 94.54% accuracy while all other models evaluated have less accuracy. Humanize: This says that XGBoost has the best accuracy as

compared to SVM, Decision tree, Multilayer perceptron, and CNN which have ranges of accuracies between 92.00 to 93.70.

Table 4.3.6: Predicted Testing Outcome

Algorithms	Data	Actual Class	Predicted Class
XGBoost	Gender:1, Age:24, Bachelor:1, Bachelor Education Organization (Public / Private):1, B_SUBJECT:0, Masters:0, Masters Education Organization (Public / Private):0, m_SUBJECT:0, Appeared in Mock Interview:1, Technical Skill:0, SOFT Skill:0, EDU_Result:1, EXPerienced:0, cv_details:0, cv_category:8	Not Offer	Not Offer
	Gender:1, Age:25, Bachelor:1, Bachelor Education Organization (Public / Private):1, B_SUBJECT:1, Masters:0, Masters Education Organization (Public / Private):0, m_SUBJECT:0, Appeared in Mock Interview:1, Technical Skill:0, SOFT Skill:1, EDU_Result:1, EXPerienced:1, cv_details:0, cv_category:8	Offer	Offer
Decision Tree	Gender:1, Age:24, Bachelor:1, Bachelor Education Organization (Public / Private):1, B_SUBJECT:0, Masters:0, Masters Education Organization (Public / Private):0, m_SUBJECT:0, Appeared in Mock Interview:1, Technical Skill:0, SOFT Skill:0, EDU_Result:1, EXPerienced:0, cv_details:0, cv_category:8	Not Offer	Not Offer
	Gender:1, Age:25, Bachelor:1, Bachelor Education Organization (Public / Private):1, B_SUBJECT:1, Masters:0, Masters Education Organization (Public / Private):0, m_SUBJECT:0, Appeared in Mock Interview:1, Technical Skill:0, SOFT Skill:1, EDU_Result:1, EXPerienced:1, cv_details:0, cv_category:8	Offer	Offer

Algorithms	Data	Actual Class	Predicted Class
MLP	Gender:1, Age:24, Bachelor:1, Bachelor Education Organization (Public / Private):1, B_SUBJECT:0, Masters:0, Masters Education Organization (Public / Private):0, m_SUBJECT:0, Appeared in Mock Interview:1, Technical Skill:0, SOFT Skill:0, EDU_Result:1, XPerienced:0, cv_details:0, cv_category:8	Not Offer	Not Offer
	Gender:1, Age:25, Bachelor:1, Bachelor Education Organization (Public / Private):1, B_SUBJECT:1, Masters:0, Masters Education Organization (Public / Private):0, m_SUBJECT:0, Appeared in Mock Interview:1, Technical Skill:0, SOFT Skill:1, EDU_Result:1, XPerienced:1, cv_details:0, cv_category:8	Offer	Offer
CNN	Gender:1, Age:24, Bachelor:1, Bachelor Education Organization (Public / Private):1, B_SUBJECT:0, Masters:0, Masters Education Organization (Public / Private):0, m_SUBJECT:0, Appeared in Mock Interview:1, Technical Skill:0, SOFT Skill:0, EDU_Result:1, XPerienced:0, cv_details:0, cv_category:8	Not Offer	Not Offer
	Gender:1, Age:25, Bachelor:1, Bachelor Education Organization (Public / Private):1, B_SUBJECT:1, Masters:0, Masters Education Organization (Public / Private):0, m_SUBJECT:0, Appeared in Mock Interview:1, Technical Skill:0, SOFT Skill:1, EDU_Result:1, XPerienced:1, cv_details:0, cv_category:8	Offer	Offer
SVM	Gender:1, Age:24, Bachelor:1, Bachelor Education Organization (Public / Private):1, B_SUBJECT:0, Masters:0, Masters Education Organization (Public / Private):0, m_SUBJECT:0, Appeared in Mock Interview:1, Technical Skill:0, SOFT Skill:0, EDU_Result:1, XPerienced:0, cv_details:0, cv_category:8	Not Offer	Not Offer

Algorithms	Data	Actual Class	Predicted Class
SVM	Gender:1, Age:25, Bachelor:1, Bachelor Education Organization (Public / Private):1, B_SUBJECT:1, Masters:0, Masters Education Organization (Public / Private):0, m_SUBJECT:0, Appeared in Mock Interview:1, Technical Skill:0, SOFT Skill:1, EDU_Result:1, EXPerienced:1, cv_details:0, cv_category:8	Offer	Offer

Table 4.3.7: Predicted Testing Outcome for Skill

Algorithms	Data	Score	Recommend
BERT	<p>Job Seekers Skill: Java, Python, SQL, MongoDB</p> <p>Job skills: Java, C, Python, SQL</p> <p>Job Seekers Soft Skill: Communication, Teamwork, Time Management</p> <p>Required Job Soft Skills: Leadership, Communication, Problem-Solving</p> <p>Job Seekers CV Details: A software engineer with 3+ years of experience in Java and Python development, adept at writing efficient code and managing databases using SQL and MongoDB.</p> <p>Job Description: We are looking for a software engineer proficient in Java, Python, and SQL. The candidate should also have strong communication and leadership abilities and experience in problem-solving and database management.</p>	<p>Technical Skills: 0.75</p> <p>Soft Skill: 0.80</p> <p>Description: 0.72</p>	<p>IF score > 0.6 Then 1</p>
BERT	<p>Job Seekers Skill: Java, MongoDB</p> <p>Job skills: Java, C, Python, SQL</p> <p>Job Seekers Soft Skill: Leadership, Teamwork, Time Management</p> <p>Required Job Soft Skills: Leadership, Problem-Solving</p>	<p>Technical Skills: 0.50</p> <p>Soft Skill: 0.40</p>	<p>Skill: C, Python, SQL, Problem-Solving</p>

Algorithms	Data	Score	Recommend
BERT	<p>Job Seekers CV Details: A software engineer with 3+ years of experience in Java and Python development, adept at writing efficient code and managing databases using SQL and MongoDB.</p> <p>Job Description: We are looking for a software engineer proficient in Java, Python, and SQL. The candidate should also have strong communication and leadership abilities and experience in problem-solving and database management.</p>	<p>Description: 0.50</p>	<p>Description: 0</p>

4.4 Discussion

In our study, we explored the capability of various machine learning models to predict job eligibility for fresh graduates in Bangladesh, based on their academic and personal backgrounds. We evaluated five algorithms—Decision Tree, Support Vector Machine (SVM), Multilayer perceptron Convolution Neural Network (CNN), and XGBoost—on their performance metrics, including accuracy, precision, recall, and F1-score. The results indicate that while all models exhibit strong predictive capabilities, there are distinct differences in their accuracy and efficiency, particularly highlighting the effectiveness of XGBoost as a leading choice for this task.

The accuracy of machine learning models in predicting job eligibility for fresh graduates is essential for addressing the challenges faced by employers in Bangladesh. Based on **RQ1: Which machine learning models are most effective in predicting job eligibility for fresh graduates in Bangladesh**, our evaluation showed that all tested models achieved high accuracy levels, with XGBoost attaining a validation accuracy of 94.42%. Other models, such as Multilayer perceptron (94.53%) and Decision Tree (94.46%), demonstrated comparable performance. This indicates that machine learning algorithms can effectively analyze graduates' academic achievements, skills, and personal backgrounds to predict their eligibility for job roles.

In addition to accuracy, the models were evaluated on precision, recall, and F1-score to provide a comprehensive view of their predictive capabilities. XGBoost showed a precision of 85.35% for the "Offer" category and 99.26% for the "Not Offer" category, demonstrating a strong ability to identify unsuitable candidates accurately. Moreover, its recall rate of 98.35% for the

"Offer" category further confirms its effectiveness in identifying eligible candidates. The F1-score for XGBoost was 91.39%, highlighting its balance between precision and recall. Overall, the findings suggest that machine learning models, particularly XGBoost and Multilayer perceptron, can significantly enhance the recruitment process by accurately predicting job eligibility based on the relevant backgrounds of fresh graduates.

The advantages of using a machine learning-based system for predicting job suitability when detailed job descriptions are provided are significant. First, such systems can efficiently process and analyze vast amounts of data, allowing them to identify patterns and correlations between job descriptions and candidate profiles. This capability leads to more precise matching of candidates with job requirements, facilitating early and accurate job matching.

Second, detailed job descriptions often include various attributes such as required skills, experience levels, and educational qualifications. Machine learning algorithms can leverage this information to develop predictive models that assess how well a candidate's academic background and skills align with the specified requirements. This results in improved accuracy in determining which graduates are likely to succeed in specific roles, ultimately enhancing the quality of hires.

Additionally, utilizing machine learning for job matching saves significant time for both employers and candidates. By automating the initial screening process, organizations can quickly identify suitable candidates, reducing the time spent on manual reviews of resumes and applications. This time-saving aspect allows recruiters to focus on high-potential candidates more efficiently, streamlining the overall hiring process.

Furthermore, machine learning systems can continuously learn from new data, adapting to changing job market demands and evolving skill requirements. By incorporating feedback from previous hiring decisions and candidate performances, these systems can refine their predictive accuracy over time, providing a dynamic solution to the recruitment challenges faced by employers in Bangladesh. Ultimately, leveraging machine learning for job suitability predictions not only enhances efficiency but also ensures a more effective and timely approach to matching fresh graduates with suitable job opportunities.

4.5 Research Gap

Despite the growing interest in utilizing machine learning techniques for recruitment and job matching, significant gaps exist in the current literature and practices, particularly in the context of Bangladesh. Existing research primarily focuses on traditional recruitment methods, which often fail to align with the dynamic needs of the job market and the evolving skill sets of fresh graduates. Many studies have addressed the challenges faced by graduates in securing

employment; however, there is limited exploration of how machine learning can effectively bridge this gap by providing data-driven solutions for predicting job eligibility.

1. Lack of Context-Specific Studies: Most machine learning applications in recruitment are drawn from studies conducted in developed countries, where job markets and educational systems differ significantly from those in Bangladesh. This absence of context-specific research limits the applicability of findings to local conditions, making it challenging to understand how these models would perform in the unique socio-economic landscape of Bangladesh.

2. Insufficient Focus on Fresh Graduates: While there are studies on recruitment and candidate evaluation, few explicitly target the fresh graduate demographic in Bangladesh. This group faces distinct challenges, such as a lack of work experience and alignment between academic training and industry requirements. Research tailored to this population is essential for developing predictive models that accurately assess job eligibility based on academic and personal backgrounds.

3. Underutilization of Advanced Machine Learning Techniques: Although traditional machine learning models like Decision Trees and SVMs have been explored, there is a notable lack of studies employing advanced ensemble learning algorithms, such as XGBoost and Multilayer perceptron, in the Bangladeshi context. These models have demonstrated superior performance in various applications but remain underutilized in predicting job suitability and eligibility.

4. Need for Comprehensive Evaluation Metrics: Many existing studies focus solely on accuracy without considering other critical performance metrics such as precision, recall, and F1-score. A more holistic evaluation of machine learning models is necessary to understand their effectiveness in real-world recruitment scenarios, particularly regarding how well they can identify suitable candidates and minimize false positives.

5. Integration of Job Descriptions and Candidate Profiles: There is a gap in research that explores how detailed job descriptions can be integrated with candidate profiles to enhance the predictive capabilities of machine learning systems. Understanding how various job attributes interact with candidates' skills and experiences could lead to more effective matching processes and improved hiring outcomes.

6. Time-Consuming Recruitment Process: Traditional recruitment methods can be time-consuming and labor-intensive. Machine learning-based systems, especially when combined with detailed job descriptions, can automate and streamline this process, allowing employers

to quickly identify the most suitable candidates, thereby significantly reducing the time required for hiring decisions.

This aligns with **RQ2**, where detailed job descriptions can lead to more precise candidate-job matching, significantly reducing the time spent by employers in reviewing unsuitable candidates. Machine learning systems can process vast datasets efficiently, providing timely and accurate job recommendations. This not only speeds up recruitment but also improves the quality of hiring decisions.

By addressing these gaps, this study aims to contribute valuable insights into the use of machine learning for improving job matching and eligibility predictions for fresh graduates in Bangladesh, ultimately enhancing the effectiveness of the recruitment process.

CHAPTER 5

Impact on Society, Environment and Sustainability

5.1 Impact on Society

The findings of this research hold significant social relevance, particularly in addressing challenges within the job market. By leveraging advanced machine learning algorithms, we have enhanced the ability to predict job eligibility across various job categories. This approach not only facilitates a better match between job seekers and available positions but also aligns individuals with roles that match their career preferences and qualifications.

One of the most profound societal benefits of this research is its potential to create more work opportunities, thus improving overall productivity and efficiency in the labor market. By optimizing job matching, employers can fill vacancies more effectively, while job seekers can secure roles that suit their skills and aspirations, reducing the mismatch often seen in traditional recruitment processes.

Additionally, this machine learning-based approach contributes to reducing unemployment by improving the rational utilization of human resources. As the system efficiently connects qualified candidates with suitable positions, it stimulates economic growth by increasing workforce participation and minimizing the time spent on manual recruitment efforts. In turn, this can lead to broader economic stimulation and contribute to a more robust and dynamic labor market.

5.2 Ethical Aspects

In the realm of research, ethics serves as the foundational framework guiding the design, methodology, and results of various studies. Researchers are committed to adhering to rigorous ethical guidelines, especially when working with human subjects, to ensure the integrity and impact of their work. These ethical principles are essential for safeguarding the rights, privacy, and dignity of participants, while also preventing any potential harm or exploitation.

Ethical research guidelines go beyond mere regulations; they call upon researchers to maintain transparency, integrity, and accountability throughout the research process. Upholding academic ethics enhances the credibility of the research and increases the likelihood of its acceptance within the scholarly community, thereby contributing to the advancement of knowledge.

In developing our methodologies and business models, we strictly adhere to universal human rights and ethical principles. We place a strong emphasis on protecting participants' rights throughout the research process, with a focus on preserving their autonomy, dignity, and privacy. We comply with essential ethical standards, including the principle of confidentiality, which safeguards sensitive information related to research subjects.

5.3 Sustainability Plan

Our primary objective is to develop machine learning-based algorithms that enhance the success of fresh job seekers. Specifically, our program, particularly the SVM model, has the potential to facilitate significant shifts within companies and organizations, leading to transformative outcomes in recruitment practices.

To ensure the long-term effectiveness and relevance of our model, it is essential to recognize that it is tailored for specific datasets. Consequently, a comprehensive range of datasets encompassing various dimensions of job dynamics will be critical for supporting and sustaining future developments. As these datasets continue to evolve, we can expand the application of our model beyond its current capabilities, extending its utility across educational, industrial, and commercial sectors, among others.

Our vision is for this model to emerge as a leading solution in both big data management and the broader landscape of universal data needs. By continually updating and refining our algorithms in response to new data and trends, we aim to maintain our model's relevance and effectiveness, positioning it as a valuable resource for future job seekers and employers alike.

CHAPTER 6

Conclusion and Future Work

6.1 Summary of the Study

To accomplish our objective, we meticulously followed a series of steps, each designed to contribute to our understanding of fresher job seeker success prediction. These steps encompassed data collection, preprocessing, model training, evaluation, and validation. Firstly, we curated a comprehensive dataset encompassing diverse attributes of fresher job seekers, including their educational backgrounds, skills, and eventual employment outcomes. Next, we meticulously processed and cleaned the data to ensure its reliability and suitability for analysis. Subsequently, we employed a range of machine learning algorithms, including decision trees, support vector machines, Multilayer perceptron, K-nearest neighbors, and XGBoost, to train predictive models on the prepared dataset. Through rigorous evaluation and validation procedures, we identified the most effective algorithms for forecasting fresher job seeker success. Our systematic approach aimed to uncover the underlying factors influencing fresher job seeker outcomes, ultimately contributing valuable insights to inform recruitment and career planning strategies.

Step 1:

Datasets collect from Bdjobs.com limited

Step 2:

Collect all data in CSV

Step 3:

Data preprocessing and split..

Step 4:

Models Validation.

Step 5:

Use ML and DL models, Decision Tree, CNN, SVM, Multilayer perceptron, and XGBoost models.

Step 6:

Model test and train.

Step 7:

Result check and decision.

6.2 Conclusion

The paths to better socio-technical and smart recruitment systems in Bangladesh were revealed by our study using data mining and AI algorithms. We emphasize the course of applying advanced computational techniques to receipts during job exposure fairs, that can be making a contribution to true revolution in the employer's recruitment processes. Research has shown the applicability of diverse aggregation learning methods ranging from XGBoost, Multilayer perceptron to Decision Tree, which have achieved impressive results with the average accuracy of 94%. Interestingly, from the arsenal of models we utilized, XGBoost proved to be the best with accuracy of 94.5%. The results of this research confirm the dynamic abilities of the machine learning algorithms for the improvement of recruitment at the level of decision-making. Leveraging on predictive analytics, job fairs may have a possibility to become more efficient in linking graduates and job seekers to employers, and this would not only help businesses but also educational institutions. Ahead, the investment which gives promise to causes the revolution is being made in predictive analysis and machine learning to transit through youth employment in Bangladesh and across all other countries and it this in combination of different sectors of the state, industry and academia results in the economic development by the progress in employment of job seekers.

6.3 Implementation for Further Work

We identified several limitations, including the inadequacy of our dataset and challenges with closed domains. Recognizing that experimental research is continually evolving, future improvements are necessary. Upon completing this study, we plan to expand our model by incorporating more data to enhance accuracy and exploring additional deep learning or hybrid algorithms. Future work will also involve comparing various machine learning and deep learning models to further refine our approach.

References:

- [1]. A. O'Neill, "Unemployment rate in Bangladesh 2022," Statista, 2022.
- [2]. M. M. Jasim, "For a secure life, many graduates even opting for low-grade govt jobs," *The Business Standard*, 2022.
- [3]. Olalekan Awujoola, Philip O. Odion, Martins E. Irhebhude, and Halima Aminu, "Performance evaluation of machine learning predictive analytical model for determining the job applicants employment status," *Malaysian Journal of Applied Sciences*, vol. 6, no. 1, pp. 67–79, 2021.
- [4]. Summit, HBCU and Platform, NACE Learning and Polls, NACE Quick and Briefs, NACE and Market, Job, "Predicting employment through machine learning," *NACE Journal*, 2019.
- [5]. Thomas Paul and Revathi Bommu, "Strategic Employee Performance Analysis in the USA: Leveraging Intelligent Machine Learning Algorithms," *Unique Endeavor in Business & Social Sciences*, vol. 3, no. 1, pp. 113–124, 2024.
- [6]. MD Rokibul Hasan, Rejon Kumar Ray, and Faiaz Rahat Chowdhury, "Employee performance prediction: An integrated approach of business analytics and machine learning," *Journal of Business and Management Studies*, vol. 6, no. 1, pp. 215–219, 2024.
- [7]. Munirah Alghamlas and Reham Alabduljabbar, "Utilizing Machine Learning Algorithms for Recruitment Predictions of IT Graduates in the Saudi Labor Market," *IJCSNS*, vol. 24, no. 3, p. 113, 2024.
- [8]. Q. Chelsea Song, Hyun Joo Shin, Chen Tang, Alexis Hanna, and Tara Behrend, "Investigating machine learning's capacity to enhance the prediction of career choices," *Personnel Psychology*, vol. 77, no. 2, pp. 295–319, 2024.
- [9]. Chinenye Gbemisola Okatta, Funmilayo Aribidesi Ajayi, and Olufunke Olawale, "Navigating the future: integrating AI and machine learning in HR practices for a digital workforce," *Computer Science & IT Research Journal*, vol. 5, no. 4, pp. 1008–1030, 2024.
- [10]. Sejal Asnani, Deshraj Singh Kushwaha, Abhishek Gupta, and Uday Singh Kushwaha, "A study on CareerConnect: linking talents for streamlined and effective job searches," *Unknown Journal*, 2024.
- [11]. Funmilayo Aribidesi Ajayi, Chioma Ann Udeh, et al., "Innovative recruitment strategies in the IT sector: A review of successes and failures," *Magna Scientia Advanced Research and Reviews*, vol. 10, no. 2, pp. 150–164, 2024.
- [12]. Nathan Oguche Emmanuel, "The Classroom-Newsroom Gap Debacle: A Theoretical Reflection," *Journalism Practice, Teaching and Research in Nigeria*, pp. 276, 2024.
- [13]. Bertha Kibona, "Higher education and employability in Tanzania: Students' account," *Journal of International Development*, vol. 36, no. 1, pp. 254–267, 2024.
- [14]. Berhanu Mekonnen Yimer, Adane Hailu Herut, et al., "Trends of higher education enrolment, graduation, and employment in Ethiopia: an empirical analysis," *Cogent Education*, vol. 11, no. 1, pp. 2302623, 2024.
- [15]. Genc Alimehmeti, Magal'ı Fia, and Angelo Paletta, "The sustainability-to-employment pipeline: the impact of SDG-related curricula on graduates' employability," *Studies in Higher Education*, pp. 1–15, 2024.
- [16]. Julia Ivy, "Making Your Consulting Case as a Tool for Employability Investments," *International Journal of Business and Management Research*, vol. 12, no. 1, pp. 1–07, 2024.
- [17]. Mary T. Grant, Clare Hanlon, and Janet A. Young, "Transforming the experiential learning of sport management graduates into transferable employability signals," *Journal of Education and Work*, pp. 1–18, 2024.

- [18]. Yuzhuo Cai and Michael Tomlinson, "A renewed analytical framework for understanding employers' perceptions of graduate employability: Integration of capital and institutional perspectives," *SAGE handbook of graduate employability*, pp. 479–495, 2023.
- [19]. Yuzhuo Cai, "Institutionalisation of employability capital in employment markets," in *Rethinking Graduate Employability in Context: Discourse, Policy and Practice*, Springer International Publishing, Cham, 2023, pp. 161–180.
- [20]. Jos Akkermans, Michael Tomlinson, and Valerie Anderson, "Initial employability development: introducing a conceptual model integrating signalling and social exchange mechanisms," *European Journal of Work and Organizational Psychology*, vol. 33, no. 1, pp. 54–66, 2024.
- [21]. Thuy Thi Hai Ho, Van Hong Le, Duong Tuan Nguyen, Chi Thi Phuong Nguyen, and Ha Thi Thu Nguyen, "Effects of career development learning on students' perceived employability: a longitudinal study," *Higher Education*, vol. 86, no. 2, pp. 297–315, 2023.
- [22]. Ken Ireri, "Massification of higher education and its impact on graduate employability in emerging economies: A case study of business graduates' employability in the Kenyan labour market context," Master's thesis, Siam University, 2023.
- [23]. Shiwen Song, "Current situation of employability and strategies to improve the employability of Thai language majors in Chinese University-Take Yuxi Normal University as an example," 2023.
- [24]. Marcel Naude, Kolawole John Adebayo, Rohan Nanda, "A machine learning approach to detecting fraudulent job types," *AI & Society*, vol. 38, no. 2, pp. 1013–1024, 2023.
- [25]. Youngkeun Choi, Jae Won Choi, "A study of job involvement prediction using machine learning technique," *International Journal of Organizational Analysis*, vol. 29, no. 3, pp. 788–800, 2021.
- [26]. Nathan Kosylo, John Smith, Matthew Conover, Leong Chan, Hongtao Zhang, Hanfei Mei, Renzhi Cao, "Artificial intelligence on job-hopping forecasting: AI on job-hopping," in *2018 Portland International Conference on Management of Engineering and Technology (PICMET)*, pp. 1–5, 2018.
- [27]. Michael C Knaus, Michael Lechner, Anthony Strittmatter, "Heterogeneous employment effects of job search programs: A machine learning approach," *Journal of Human Resources*, vol. 57, no. 2, pp. 597–636, 2022.
- [28]. Kevin Appadoo, MB Soonoo, Zahra Mungloo-Dilmohamud, "JobFit: Job Recommendation using Machine Learning and Recommendation Engine," 2020, pp. 1–6, 2020.
- [29]. John Kořtner, Giuliano Bonoli, "Predictive algorithms in the delivery of public employment services," in *Handbook of Labour Market Policy in Advanced Democracies*, pp. 387–398, 2023.
- [30]. Ibrahim Nasser, Amjad H Alzaanin, "Machine learning and job posting classification: A comparative study," *International Journal of Engineering and Information Systems (IJEAIS) ISSN*, pp. 6–14, 2020.
- [31]. Rudra Tiwari, "The impact of AI and machine learning on job displacement and employment opportunities," *International Journal of Engineering Technologies and Management Research*, vol. 7, no. 1, 2023.
- [32]. Maruf Ahamed, Rezwana Mahmood, "Impact of organizational culture on job satisfaction: A study on Banglalion Communication Ltd, Bangladesh," *European Journal of Business and Management*, vol. 7, no. 10, pp. 160–174, 2015.
- [33]. Nazmun Nessa Moon, Asma Mariam, Shayla Sharmin, Mohammad Monirul Islam, Fernaz Narin Nur, Nabadita Debnath, "Machine learning approach to predict the

- depression in job sectors in Bangladesh,” *Current Research in Behavioral Sciences*, vol. 2, pp. 100058, 2021.
- [34]. Maksura Mubashirin, “A New Methodology for Predicting Teachers’ Job Satisfaction Using Machine Learning Approaches- Bangladesh Perspective,” 2022.
- [35]. Shaikh Moksadur Rahman, “Relationship between job satisfaction and turnover intention: Evidence from Bangladesh,” *Asian Business Review*, vol. 10, no. 2, pp. 99–108, 2020.
- [36]. Md Nurul Raihen, Sultana Akter, Md Nazmul Sardar, “Women’s career challenges and opportunities (a study of career and job satisfaction among Bangladeshi women),” *Ajrsp*, vol. 5, no. 51, pp. 05–22, 2023.
- [37]. Nor Azziaty Abdul Rahman, Kian Lam Tan, Chen Kim Lim, “Predictive analysis and data mining among the employment of fresh graduate students in HEI,” in *AIP Conference Proceedings*, vol. 1891, no. 1, 2017.
- [38]. Zalinda Othman, Soo Wui Shan, Ishak Yusoff, Chang Peng Kee, “Classification techniques for predicting graduate employability,” *International Journal on Advanced Science, Engineering and Information Technology*, vol. 8, no. 4-2, pp. 1712–1720, 2018.
- [39]. Pooja Thakar, Anil Mehta, et al., “A unified model of clustering and classification to improve students’ employability prediction,” *International Journal of Intelligent Systems and Applications*, vol. 9, no. 9, pp. 10, 2017.
- [40]. Mekala Sunil Kumar, G Praveen Babu, “Comparative study of various supervised machine learning algorithms for an early effective prediction of the employability of students,” *Journal of Engineering Sciences*, vol. 10, no. 10, pp. 240–251, 2019.
- [41]. Ankita Sinha, Ashish Singh, et al., “Student Career Prediction Using Algorithms Of Machine Learning,” Ashish, *Student Career Prediction Using Algorithms Of Machine Learning* (May 6, 2023), 2023.
- [42]. Dileep Chaudhary, Harsh Prajapati, Rajan Rathod, Parth Patel, Rajiv Kumar Gurjwar, “Student future prediction using machine learning,” *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, vol. 5, no. 2, pp. 1104–1108, 2019.
- [43]. Madhan Mohan Reddy, “Career Prediction System,” *International Journal of Scientific Research in Science and Technology*, vol. 8, no. 4, pp. 54–58, 2021.
- [44]. Seyam, Mohammed & Othman, Faridah & El-Shafie, Ahmed. (2017). Prediction of Stream Flow in Humid Tropical Rivers by Support Vector Machines. *MATEC Web of Conferences*. 111. 01007. 10.1051/mateconf/201711101007.
- [45]. Wang, Yuanchao & Pan, Z. & Zheng, J. & Qian, L. & Mingtao, Li. (2019). A hybrid ensemble method for pulsar candidate classification. *Astrophysics and Space Science*. 364. 10.1007/s10509-019-3602-4.
- [46]. Zmitri, Makia & Fourati, Hassen & Vuillerme, Nicolas. (2019). Human Activities and Postures Recognition: From Inertial Measurements to Quaternion-Based Approaches. *Sensors*. 19. 4058. 10.3390/s19194058.
- [47]. Wang, Yuanchao & Pan, Z. & Zheng, J. & Qian, L. & Mingtao, Li. (2019). A hybrid ensemble method for pulsar candidate classification. *Astrophysics and Space Science*. 364. 10.1007/s10509-019-3602-4.
- [48]. Awan, Faraz & Saleem, Yasir & Minerva, Roberto & Crespi, Noel. (2020). A Comparative Analysis of Machine/Deep Learning Models for Parking Space Availability Prediction. *Sensors*. 20. 10.3390/s20010322.

PLAGIARISM REPORT

ENHANCING FRESHERS JOB SEEKER SUCCESS

ORIGINALITY REPORT

13% SIMILARITY INDEX	9% INTERNET SOURCES	7% PUBLICATIONS	5% STUDENT PAPERS
--------------------------------	-------------------------------	---------------------------	-----------------------------

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	5%
2	Submitted to Daffodil International University Student Paper	1%
3	Amir Shachar. "Introduction to Algogens", Open Science Framework, 2024 Publication	<1%
4	Biswadip Basu Mallik, Gunjan Mukherjee, Rahul Kar, Aryan Chaudhary. "Deep Learning Concepts in Operations Research", Routledge, 2024 Publication	<1%
5	H.L. Gururaj, Francesco Flammini, S. Srividhya, M.L. Chayadevi, Sheba Selvam. "Computer Science Engineering", CRC Press, 2024 Publication	<1%
6	link.springer.com Internet Source	<1%
7	www.frontiersin.org Internet Source	

		<1 %
8	keylabs.ai Internet Source	<1 %
9	www2.mdpi.com Internet Source	<1 %
10	Submitted to George Bush High School Student Paper	<1 %
11	Rajganesht Nagarajan, Senthilkumar Narayanasamy, Ramkumar Thirunavukarasu, Pethuru Raj. "Intelligent Systems and Sustainable Computational Models - Concepts, Architecture, and Practical Applications", CRC Press, 2024 Publication	<1 %
12	Insha Zahoor, Sajad Ahmad Wani, Tariq Ahmad Ganaie. "Artificial Intelligence in the Food Industry - Enhancing Quality and Safety", CRC Press, 2025 Publication	<1 %
13	Submitted to University of Surrey Student Paper	<1 %
14	Lida Kanari, Stanislav Schmidt, Francesco Casalegno, Emilie Delattre et al. "Deep learning for classifying neuronal morphologies: combining topological data	<1 %

analysis and graph neural networks", Cold
Spring Harbor Laboratory, 2024

Publication

15	www.researchgate.net Internet Source	<1 %
16	Devendra Prasad, Suresh Chand Gupta, Anju Bhandari Gandhi, Stuti Mehla, Upasana Lakhina. "Computational Intelligence and Mathematical Applications", CRC Press, 2024 Publication	<1 %
17	Submitted to Higher Education Commission Pakistan Student Paper	<1 %
18	A. J. Singh, Nikita Gupta, Sanjay Kumar, Sumit Sharma, Subho Upadhyay, Sandeep Kumar. "Artificial Intelligence and Machine Learning Applications for Sustainable Development", CRC Press, 2025 Publication	<1 %
19	Submitted to American Intercontinental University Online Student Paper	<1 %
20	Du, Hongmin W.. "Analytical Approaches to Information Processing and Risk Assessment: From Cybersecurity Disclosures to Medical Audits.", Rutgers The State University of New Jersey, Graduate School - Newark, 2024 Publication	<1 %
