STUDENT ASSESSMENT & CAREER PREDICTION SYSTEM

BY

Sakir Hossain Faruque
ID: 203-15-3862

FINAL YEAR DESIGN PROJECT REPORT

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

Supervised By

Mr. Amit Chakraborty
Assistant Professor
Department of CSE
Daffodil International University

Co-Supervised By

Mr. Golam Rabbany
Lecturer
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

July 2024

APPROVAL

This Project titled "**Student Assessment & Career Prediction System**", submitted by **Sakir Hossain Faruque** 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 B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on **13-07-2024**.

BOARD OF EXAMINERS

Dr. Md. Ismail Jabiullah — **Board Chairman**
Professor
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Mr. Abdus Sattar — **Internal Examiner 1**
Assistant Professor
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Ms. Zahura Zaman — **Internal Examiner 2**
Lecturer
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

Dr. Ahmed Wasif Reza — **External Examiner**
Professor
Department of Computer Science and Engineering
East West University

DECLARATION

I hereby declare that this project has been done by me under the supervision of **Mr. Amit Chakraborty, Assistant Professor, Department of Computer Science and Engineering, Daffodil International University**. I also declare that neither this project nor any part of this project has been submitted elsewhere for the award of any degree or diploma.

Supervised by:

Mr. Amit Chakraborty
Assistant Professor
Department of CSE
Daffodil International University

Co-Supervised by:

Mr. Golam Rabbany
Lecturer
Department of CSE
Daffodil International University

Submitted by:

Sakir

Sakir Hossain Faruque
ID: -203-15-3862
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, I express my heartiest thanks and gratefulness to almighty for His divine blessing making it possible for me to complete the final year project/internship successfully.

I am grateful and wish our profound indebtedness to **Mr. Amit Chakraborty**, **Assistant Professor**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "*Machine Learning and Web Integrated System*" to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts, and correcting them at all stages have made it possible to complete this project.

I would like to express our heartiest gratitude to the **Head of the Department of CSE**, for his kind help in finishing my project and also to other faculty members and the staff of the Department of CSE, Daffodil International University.

I would like to thank my entire course mate in Daffodil International University, who took part in this discussion while completing the course work.

Finally, I must acknowledge with due respect the constant support and patience of my parents.

ABSTRACT

In today's world, choosing a suitable career is difficult for undergraduate students. Most of the time, students cannot choose a suitable career despite their skills and abilities. One of the main reasons is proper guidance and enough facility. The student assessment system is one of the artificial intelligence-based systems that helps students choose the right career path with proper guidance and all kinds of facilities. The student assessment system focuses on the difficulties that create barriers to lead a student to choose a suitable field for their life. So that the system introduces a collection of features to solve the difficulties of choosing the right career path. Initially, the system is developed for Computer Science and Engineering students' career selection. By collecting comprehensive student data, the system employs advanced data analysis techniques to predict and recommend personalized career paths aligned with individual skills, interests, and past experiences. The core functionality of the system lies in its ability to assist students in identifying faculty matches based on their unique skill sets and interests. This feature will lead a student to proper career guidance and help them to face real-life challenges. This will create an interactive bonding between students and faculties. Faculty can track student progress and engage them in various competitive tasks which will build their confidence and make their life more enjoyable. There is also a chance to build community between similar skills and interest students. So that they can share knowledge and ideas and invent new things.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

1.1 Overview

Student assessment system where a student gets all kinds of support for their professional and educational life to train a student to make his or her future best. It is an artificial intelligence-based data analysis system where student data will be analyzed for selecting their career. The system is a comprehensive platform designed to assist students in making informed decisions about their academic and career journeys.

At the initial stage of the system, several features make it more efficient for the assessment of the student. Data-driven career prediction is one of the features where student data will be analyzed based on the pre-trained model and then it will predict the best career path for their future dream job. Students will get personalized faculty-matching to find out the suitable faculty based on their skills, interests, and related works. It is a skill-based community-building platform system where a student can share knowledge in a collaborative environment for learning and growth. It is a powerful system in which students can showcase their skill and their education activity to the mentor to make their foundation strong for their professional life. The system's user-friendly interface will boost every system-related activity for every role-based user.

Comprehensive student data collection, predicting career paths, a personalized recommendation for faculty matches, and fostering a guided approach to academic and professional pursuits make the system an exceptional one. The project's overall goal is to use creative technology integration to provide students the power to make informed decisions about their educational and professional career.

1.2 Background and Present State

Background

The system allows students to select their skills and interests, set related work, and view suitable career paths. Beyond their studies, students can enroll under a faculty member who

can assign and create new tasks. Upon task completion, students can submit their work and receive feedback, enhancing their learning experience. The admin has the capability to create master fields, subfields, jobs, and courses, assigning subfields to master fields and associating jobs and courses with the relevant subfields. This comprehensive setup facilitates personalized career guidance, effective task management, and streamlined administrative oversight, fostering a collaborative and structured educational environment.

Present State

In the current system, students have the capability to choose skills and interests aligned with their career aspirations, allowing them to set relevant career goals. They can also enroll under specific faculty members who can assign tasks and create new ones tailored to student needs. Upon task completion, students submit their work for feedback. Meanwhile, administrators oversee the system's functionality, managing master fields, sub-fields, jobs, and courses, ensuring alignment and relevance across educational and career pathways. This integrated approach aims to facilitate personalized learning and career development through structured engagement and feedback mechanisms.

1.3 Problem Statement

In the rapidly evolving landscape of education, students face challenges in identifying their career paths based on their skills and interests. Especially it is a difficult task for Computer Science students to find out suitable career that satisfies their goals and interests. Computer science is a vast field with numerous professional sectors such as artificial intelligence, machine learning, cybersecurity, software development, data science, and more. These numerous options make students confused about selecting the right career path. It seems that the academic courses that are allocated for Computer Science (CS) students may not fulfill the demand of industrial sectors. So, after graduation, it is a difficult task for a student to prepare themselves for an interview for a job sectors. Students are not fully aware of their skill and interests besides lack of opportunities they can not practice their skill-based project. So, after graduation, they can not show their project which will satisfy the modern technology.

Now I will thoroughly discuss the challenges that student face in their academic career.

1. **Unaware Of CS Profession:** It is seen most students do not which profession has the most demand or which profession satisfies their skill and interest. Lack of knowledge about modern industry they do get the opportunity to get employed in a good company.
2. **Limited Faculty Matching Resources:** Students often struggle to find faculties that align with their skills and interests. When a new student admits to a university, he or she does not know about all the faculties of his or her department. Though he or she has enough skills he or she can not utilize them for proper guidance.
3. **Limited Resource:** It is seen that only a few fields of computer science have enough resources on the internet but most of the field has no resources or guidelines. So that students are compelled to choose those resources as a career. Sometimes they feel interested in these professions or sometimes they do not. As a result, many good and high-demandable fields are left behind. Only a few students choose those fields.
4. **Career Guide Model Implementation:** In various journals and conferences people research student career prediction systems. But there is no existing system anywhere. There is no implemented model for career guidance.
5. **Machine Learning Model Accuracy:** The system's performance is strongly dependent on the machine learning model's ability to forecast career pathways. The objective is to create a model that continually learns and adapts to the changing dynamics of the job market and individual students' choices.
6. **Collaborative Relationships:** There is no other existing system that makes collaborative relationships between students and faculty based on skill and interest matching.
7. **Scalability and Accessibility:** As the system grows in popularity, it must accommodate a rising user base. To accommodate a varied spectrum of students and faculty, scalability and accessibility across numerous devices and platforms are critical.
8. **User Interface and Experience Design:** Designing an intuitive and user-friendly interface is vital for the success of the system. The aim is to provide an accessible, user-friendly, and interesting platform for both students and faculty.

1.4 Objectives

1. **Personalized Career Path Prediction:** Develop a system that will predict students' career paths by analyzing student skills, interests, and related works. So that it will help students to make the right decision from the beginning of their graduation life and help them to develop their skills.
2. **Faculty Matching Based on Skills and Interests:** The first students set up their skills, interests, and related works in the system then the system automatically matches their information with faculty experience based on certain criteria. The system prefers students to those faculties that match their given information.
3. **Data Security and Privacy Measures:** User security is ensured in the system. When a student creates an account, he or she must sign up with their university email. If a student signs up without a university email, he or she cannot log in. Besides email verification, JSON web token verification is used for authentication purposes.
4. **Continuous Learning and Adaptation of Machine Learning Model:** In the system, there is a feedback system where students give their feedback about the system and their career preferences. Based on the feedback the model will be trained more accurately.
5. **User-Friendly Interface and Experience Design:** Create an intuitive and user-friendly online interface for both students and faculty.
6. **Scalability and Accessibility:** Ensure that the system is scalable and accessible to a growing user base across several devices and platforms.
7. **Integration with Educational Institutions:** Collaborate with educational institutions to incorporate the technology into existing frameworks, resulting in a smooth adoption process.

1.5 Scope and Limitations

Scope

The project scope includes creating and implementing a complete web and machine learning integrated system that will provide tailored career guidance and promote a collaborative educational community. The essential components and capabilities within the project scope are described below:

1. **Student Data Collection and Analysis:** Gather student-relevant information such as skill, interest, and skill-related works. Based on this information, a student career prediction model was built. To analyze data natural language processing technology is used. After pre-processing of the data model accuracy is checked then the highest accuracy model is integrated with the student assessment system.
2. **Implement Machine Learning Model for Career Path Prediction:** Use machine learning algorithms to develop a prediction model that considers academic skill, interest, and skill-related works. Continuously improve the model to reflect shifting industry trends and student preferences.
3. **Faculty Matching System:** Design and implement a faculty matching system that considers both student and faculty profiles. The system evaluates student's activities and faculty expertise to show students about their faculty suggestions.
4. **User-Friendly Interface:** Create an interface that simplifies data entry for students, offers clear career suggestions, and is easy to navigate. Prioritize accessibility and responsiveness across several devices to improve the user experience.
5. **Data Security and Privacy:** To safeguard the security and integrity of student data, use encryption, secure data storage, and access control techniques. To develop user trust, comply with appropriate data protection rules.
6. **Continuous System Improvement:** At the initial stage, the system will be launched with only a few features. But it will take feedback from users and continuously improve with new features.
7. **Integration with Educational Institutions:** Collaborate with educational institutions to implement the system into their platforms while guaranteeing compatibility with current student databases and academic systems. Provide the required support to ensure a seamless adoption process.

Limitations

- The accuracy of career predictions is dependent on the quality and comprehensiveness of the input data.
- Initial deployment is limited to specific departments and may not reflect the full

diversity of career paths across all fields.

- -Integration with existing educational institution databases and systems can be complex and require significant customization.
- Continuous updates and maintenance are required to keep up with evolving industry trends and student preferences.
- Data privacy concerns need to be meticulously managed to maintain user trust and comply with legal standards.
- The effectiveness of the recommendation engine may vary based on the specific algorithms and models used.
- Limited scalability in the initial phases may restrict the system's ability to handle large volumes of data and users efficiently.

1.6 Report Organization

Chapter 1: It introduces the Student Career Prediction and Assessment System

Chapter 2: Discuss on literature review of related papers and works, comparative analysis with other models and systems, and gap analysis of others models and systems.

Chapter 3: Discussion on model training and system development methodology, data pre-processing, model training, system design, project maintenance, and financial analysis.

Chapter 4: Discussion on implementation overview, train model/ prototype design, system testing and model evaluation.

Chapter 5: Discussion on model performance result and system output

Chapter 6: Discussion on system impact on life, society and environment, ethical aspects and sustainability plan

Chapter 7: Discussion on further suggested work, limitation and conflict of interests and discussion

1.7 Summary

Student Assessment is an artificial intelligence-based web-integrated system where students' data are analyzed based on their skills, interests, and related works. So that system automatically guides them about the most suitable career for them. Besides the system

suggests student about the faculty based on their activity and faculty expertise. So that a student can enroll under a faculty get proper guidelines and make them prepare for the industrial sector. The project aims to address challenges like limited resources, inaccurate career predictions, and lack of collaborative relationships between students and faculty. The project includes comprehensive student data collection, machine learning model implementation, user-friendly interface design, and robust data security measures

CHAPTER 2

Background

2.1 Overview

Student Assessment System is not only a career guidance system but also a skill development roadmap that leads a student to his or her dream job. Other existing systems allow finding out his or her interested technology from a large amount of data. However, the difference between the Student Assessment System and other existing systems is it helps students to find out automatically the best opportunity for them from large amounts of data by using machine learning technology. Besides it helps a new student to find out his or her best suitable faculty based on his or her matching ability and personal interest.

The model used for the system has also good accuracy. Here are two models that are used. One model is used for data processing and the other one is used for predicting output. To pre-process data natural language processing is used. It is seen that most of the paper has only generated a model for career prediction but they do not implement the model anywhere or do not interact with the model with the user. However, the Student Assessment System model is implemented in the system and this model is updating day by day by adopting more data from user feedback.

In this literature review, I will discuss two parts of the project. The first one is about systems similar to the Student Assessment System and the other one is about computer science career prediction and its related works.

2.2 Related Work

Similar Technologies to the Student Assessment System of Computer Science and Engineering Students:

Wang et al. [1] dramatically developed a remarkable web-based teacher evaluation education system. The system incredibly improves administration and test management support. The system employed the Tripple-A model for test analysis. Overall, the use of

technology significantly improves teachers' capacity to administer examinations and other activities too.

Dhalan et al. [2] strongly suggested an incredibly web-based school children's assessment system that allows an intelligent teacher to take students' performance more effectively. The system will autonomously and amazingly generate students’ results and make a noteworthy performance grading sheet. Finally, the writer emphasizes the tremendous potential of e-ATLMS as an alternate tool for incredibly implementing school-based assessments and supporting the teaching and learning process online.

Coleman et al. [3] passionately discussed mind-blowing technology-based exams to test students' environmental awareness and data analysis. The suggested system passionately works to identify high and poor implementers in educational programs, enabling targeted interventions and assistance.

Antal et al. [4] focused on cutting-edge computerized test systems to detect knowledge weaknesses in students and help tutors understand educational issues. The authors highlighted a truly web-based self-assessment system that exceptionally provides students with exam preparation services on various courses. The authors also innovatively informed about test taking on various courses. Based on the test result tutors can excellently track students' progress and evaluate the system in real classroom conditions.

Guzmán et al. [5] successfully implemented an out-of-this-world web-based assessment system that combines the Intelligent Tutoring System (ITS) and the System of Intelligent Evaluation Using Tests for Teleeducation (SIETTE) that integrates web technology with CATS. The system tremendously allows students to engage in self-assessment tests and give feedback about the test.

Supriyanto et al. [6] curiously examined the mind-boggling use of expert systems in student career guidance, with an emphasis on educational, evaluative, academic, and employment recommendations. It emphasizes the life-changing advantages of various systems, including learning success, specialization, enhanced performance, and self-evaluation. Finally, the authors discovered that employing expert systems in educational aid greatly helps students achieve their goals and objectives.

Huang et al. [7] proposed a career counseling system that provides offers to university and college students employment career guidance. The paper showed that after integration of artificial intelligence and ideological and political education with career guidance gives better performance by indicating higher AUC value. Besides authors discussed the recommendation algorithm which suggests jobs for the student with a higher employability rate that makes the system effective for career guidance.

Mehraj et al. [8] discussed AI-based schemes in career guidance and counseling, critically assessing their scope and implications. The authors declared that automation and digitalization significantly impact career guidance and career choice. Authors discussed various techniques to solve career guidance and counseling among them Machine Learning, Genetic Algorithms, Fuzzy Logic, and Artificial Neural Networks (ANN) are very popular. Finally, the authors concluded the paper by suggesting that integrating AI and civic education can guide students in their employment journey.

Crişan et al. [9] investigated students' career choices and found that most of the students' knowledge and abilities were not aligned with their expected job which is a significant barrier to taking the right career decision. The paper identified the barriers that students faced regularly to making career decisions and evaluating career advice and counseling services at a university, focusing on the need to distinguish between needs and wants. At last, the authors showed that an increment of career guidance and counseling can help students overcome these situations.

Myla et al. [10] proposed an expert system for students' career guidance which helped a student choose the best career path based on the questionnaire. The paper highlighted the limitations of the traditional career counseling system and showed the effectiveness of the computerized decision support system. The implemented system showed effectiveness in career counseling by assisting students with career selection, considering academic performance, and providing suitable branch recommendations.

There are also similar technologies for student assessment are available in Bangladesh. These technologies increase employability rate in Bangladesh. [38] [39] [40]

Similar Technologies to the Career Prediction Model of Computer Science and Software Engineering Students:

Ade et al. [11] proposed an incremental ensemble learning algorithm for predicting students' career choices. They combine three classifiers (Naïve Bayes, K-Star, SVM) using a voting scheme. From these three classifiers, the accuracy percentage is respectively 89.6%, 89.2%, and 89.2%. They highlight the importance of considering interest, talent, and projected growth in career choice.

Mandalpu et al. [12] discussed STEM and non-STEM career choices which were two broad categories. STEM stands for Science, Technology, Engineering, and Mathematics, while Non-STEM refers to all other career fields outside of these four areas. They organized a survey on US middle school students where they provided career choices of 591 students whereas 466 students belong to NON-STEM and other students belong to STEM fields.

Reddy et al. [13] applied XG Boost and a decision tree to predict the right career based on the student's skill. Here they collected data from students and preprocessing step to gain meaningful information. In the result section they discussed how they predicted careers for undergraduate Engineering students (Computer Science) but they did not show any accuracy percentage of the algorithm which they used. Besides they did not discuss the confusion matrix and other elements which were needed to analyze the performance of the system.

Lent et al. [14] targeted to explore the role of self-efficiency, expectations, interests, social supports, and barriers in every student's academic and career choices. The research was conducted on 1208 students and used some models such as the measurement model and structural model to analyze students' data. Here, the measurement model was used to build the relationship between students' goals, supports, and barriers whereas the structural model predicted educational and career goals based on self-efficiency, expectations, interests, supports, and barriers.

The paper of Rangnekar et al. [15] studied students' aptitude, personality, and background, and the dataset was made to access these three basic principles. The dataset of this paper has 2000 tuples which have included four subjects (Science, Mathematics, English, and Logical Reasoning) and career labels to identify suitable career choices for students. After

finishing all the preprocessing steps authors took 80% data for training and 20% of data for testing from the aptitude and student information section which was a little bit different for the personality section (90% training and 10% testing). There was a total of five machine learning algorithms such as K-nearest neighbors, Stochastic Gradient Descent, Logistic Regression, and Random Forest, Adaptive Boosting whereas K-nearest neighbors and Stochastic Gradient Descent gave high accuracy for career mapping while Logistic Regression and Adaptive Boosting had the high accuracy for student background data.

Computer Science final-year students' project choices and how this impacts their professional careers were explained in the paper of Kabakus et al. [16]. They focused on Github and IEEE to find project topics of the students using Python libraries PyGithub and IEEE Xplore. In this paper, a total of 14 machine learning were used to predict project topics and increase the accuracy level. Among all the algorithms Bayes Net Kleene Star and J48 gave the best accuracy which was 100%.

Casuat et al. [17] investigated the high unemployment rate and showed that there was a mismatch between university-provided curriculum and job market demand. The research predicted employability based on 3000 entries and 12 features using machine learning algorithms such as SVM, Decision Tress, Naïve Bayes, and KNN. Among them, SVM got the highest accuracy which was 92.22%. On the other hand, DT got the lowest accuracy only 56.36%.

Shankdhar et al. [18] showed that choosing a career is difficult for college students who lack self-awareness about their skills and interests. The research data was made based on several attributes such as gender, nationality, age, attendance, participation in extra activities, attentiveness in class, and the number of online courses taken. The authors claimed that the Decision Tree algorithm is the best to predict career but did not mention the accuracy percentage.

Vignesh et al. [19] addressed that after secondary education students faced confusion in choosing the right career path. So the authors suggested a career prediction system that took input from students' skills to help them to choose a preferable career. The proposed system had three modules. In the Prediction Module authors used both supervised and unsupervised learning. Under supervised learning, K-Nearest Neighbor (KNN) and other

algorithms were used for classifying the target values whereas KNN got approximately 94% accuracy. K-Means Clustering an unsupervised learning algorithm was used to cluster n observations and to prefer departments based on students' performance.

How machine learning algorithms can be much more effective in predicting careers was shown by Sinha et al. [20]. In the paper, the authors simply discussed the career prediction factors, data processing system (One hot encoding), and some algorithms (SVM, Decision Tree, XG Boost, etc.) but did not show any accuracy percentage or prediction model which are important steps to analyze the performance of the machine learning algorithms.

Sobnath et al. [21] mainly focused on the challenges faced by disabled students in choosing the right career and using machine learning and big data technologies for career guidance. The research worked on 270,934 students' records and applied missing value handling, label encoding, one-hot encoding, and other related activities to preprocess the dataset. Authors used ML algorithms such as Logistic Regression, Linear Discriminant, Decision Tree Classifer, and Gaussian NB whereas Logistic Regression and Decision Tree Classifer got the highest accuracy 96% to predict the Standard Occupation Classification (SOC).

2.3 Comparative Analysis

Table 2.3.1: Comparative Analysis with Previous Student Assessment and Career Guidance Technologies

SL	Author Name	Proposed Method
01	Wang, Tzu-Hua, Kuo-Hua Wang, Wei-Lung Wang, Shih-Chieh Huang, and Sherry Y. Chen. et al. [1]	• Web-Based Test Analysis (WATA) system for teacher education • The WATA system was designed and programmed based on test theory. • Three steps are followed: Design, Programming of WATA system, and feedback from the user

		• USSW survey was used to collect quantitative data.
02	Dahalan, Hamsiah Mohd, and Raja Maznah Raja Hussain et al. [2]	• Develop e-ATLMS system and its functionality • Taking users' feedback about the system's overall activities • Highlighting the strengths and weaknesses of the e-ATLMS system
03	Coleman, Elaine B., and William R. Penuel. Etal. [3]	• Develop web-based assessment to determine students' environmental awareness and data analysis skills • Data Collection from middle and high school students • Based on students' responses environmental data is interpreted
04	Antal, Margit, and Szabolcs Koncz. et al. [4]	• CBT system generates questions based on students' knowledge • CAT system to analyze student behavior • Computer-based test system • Graphical representation of student knowledge

05	Guzmán, Eduardo, and Ricardo Conejo et al. [5]	• SIETTE system that is used for administrative self-assessment CATS • Item Response Theory (IRT) is a model that helps to select the next item and determine the test finish time
06	Supriyanto, G., Widiaty, I., Abdullah, A.G. and Yustiana, Y.R et al. [6]	• Utilizes a systematic approach to analyze expert systems in educational guidance, evaluation, academic career guidance, and work guidance. • Uses a spreadsheet program for analysis and coding of articles. • Identifies relevant articles using search keywords like "expert systems, education, student guidance career, artificial intelligence." • Extracts key findings and insights from selected articles. • Uses a literature review spreadsheet model for structuring and summarizing information.
07	Huang, Li. et al. [7]	• The system combines algorithmic recommendation with ideological and political education

		• Aggregate the preferences of student and employer • Integrates student career competence and guidance system • Design a similarity-based random negative sampling module to address negative sample disbelief • Includes two steps of similarity calculation between students and units • Presents calculation method for inter-student and interunit similarity
08	Mehraj, Tehseen, and A. Mehraj Baba et al. [8]	• Examine AI schemes in career guidance and counseling • Evaluates the scope of various technologies • Discusses J48 decision tree algorithm and Case-Based Reasoning (CBR) techniques • Acknowledge a system for career selection in education • The system considers gender, IQ, grades, hobbies, and skills • CBR algorithm has higher accuracy (80%) than J48 (50-60%).

		• The system is limited to a specific geographical area and lacks a generic solution
09	Crişan, Claudia, Anişoara Pavelea, and Oana Ghimbuluţ. et al. [9]	• Collect students' career information • Surveying students' professional paths involves gathering insights into their career decisions, values, abilities, competencies, and future intentions. • Participants were given a group briefing before completing the anonymous 30-minute paper-pencil questionnaire. • Mixed methodology for a comprehensive understanding of students' career counseling needs
10	Myla, Saiteja, Surya Teja Marella, A. Swarnendra Goud, S. Hasane Ahammad, G. N. S. Kumar, and Syed Inthiyaz. et al. [10]	• Developed rule-based and artificial intelligence system for career prediction • Utilizes questionnaires to analyze students' mental abilities • Conducts sample tests with completed higher secondary students • Uses expert system's output to allocate branches to students • Represents data through graphs and tables

		• Considers academic performance and provides branch recommendations

Table 2.3.2: Comparative Analysis with Previous Students' Career Prediction System

SL No	Author Name	Used Algorithm	Best Accuracy with Algorithm
11	Ade, Roshani, and P. R. Deshmukh. et al. [11]	Naïve Bayes, K-Star, SVM	Naïve Bayes = 89.6%
12	Mandalapu, Varun, and Jiaqi Gong et al. [12]	Gradient Boosted Tree, Random Forest, Logistic Regression, AutoMLP, Deep Neural Network	Gradient Boosted Tree = 98.83%
13	Reddy, M. M. et al. [13]	XG Boost, Decision Tree	Decision Tree has better performance
14	Rangnekar, Rucha Hemant, Khyati Pradeep Suratwala, Sanjana Krishna, and Sudhir Dhage et al. [15]	K-nearest neighbors, Stochastic Gradient Descent, Logistic Regression, and Random Forest, Adaptive Boosting	**Aptitude:** Stochastic Gradient Descent = 81.035% **Personality (Myers Briggs classification):** K-Nearest Neighbors = 81% **Student Background Data:**

			Adaptive Boosting = 60.09%
15	Kabakus, Abdullah Talha, and Arafat Senturk Et al. [16]	Bayes Net, Naïve Bayes, Multinomial Naïve Bayes, Logistic Regression, Multilayer Perceptron, SVM, kNN, Kleene Star, J48, Random Tree, Random Forest, AdaBoost, Stacking, and Bagging	Bayes Net = 100%, Kleene Star = 100%, J48 = 100%
16	Casuat, Cherry D., Enrique D. Festijo, and Alvin Sarraga Alon et al. [17]	SVM, Decision Tree, Random Forest, Logistic Regression, Naïve Bayes, KNN	SVM = 92.22%
17	Shankhdhar, Ashutosh, Akash Agrawal, Deepak Sharma, Suryansh Chaturvedi, and Mukesh Pushkarna et al. [18]	Decision Tree	Decision Tree has the best accuracy
18	Vignesh S, Shivani Priyanka C, Shree Manju H, Mythili K et al. [19]	KNN, SVM, Naive Bayes	KNN = 94.10%
19	Sinha, Ankita, and Ashish Singh et al. [20]	SVM, Decision Tree, XG Boost	NA

20	Sobnath, Drishty, Tobiasz Kaduk, Ikram Ur Rehman, and Olufemi Isiaq et al. [21]	Logistic Regression, Linear Discriminant, Decision Tree Classifier, Gaussian NB	Logistic Regression = 96% Decision Tree Classifier = 96%

2.4 Open Issue

Several open issues need to be addressed to ensure the effectiveness and reliability of the Student Assessment and Career Prediction System:

- Model Accuracy and Improvement: Continuous evaluation and enhancement of the machine learning models are necessary to ensure the highest accuracy in career predictions. This includes addressing potential biases and incorporating new data to reflect current industry trends and student preferences.
- Data Collection and Privacy: Ensuring the security and privacy of student data is paramount. Implementing robust encryption methods, secure data storage, and strict access controls are crucial. Additionally, compliance with relevant data protection regulations must be regularly reviewed and updated.
- User Interface Optimization: The user interface needs ongoing refinement to improve usability and accessibility. This includes ensuring compatibility with various devices and simplifying the data entry process for students.
- System Scalability: As the user base grows, the system must be capable of handling increased data loads and user interactions without compromising performance. This requires scalable architecture and regular performance testing.
- Integration Challenges: Seamless integration with existing educational institutions' platforms and databases can be complex. Ensuring compatibility and providing adequate support for a smooth adoption process is critical.
- Feedback Mechanism: Establishing an effective feedback mechanism to gather user input and make continuous improvements is essential. This involves addressing user concerns and incorporating new features based on feedback.
- Faculty Matching Accuracy: The accuracy of the faculty matching system needs to be

constantly evaluated and improved to ensure it effectively matches students with appropriate faculty based on their profiles and expertise.

- Technological Advancements: Keeping up with technological advancements and incorporating new tools and methods into the system is necessary for maintaining its relevance and efficiency.
- User Training and Support: Providing adequate training and support for users, both students and faculty, to ensure they can effectively utilize the system is important for maximizing its benefits.

2.5 Summary

In this literature review section, I discussed related works in two parts. In the first part, I discussed Student Assessment and Career Prediction and their similar technologies whereas I have shown other research-based work on student assessment systems and career prediction. I have also shown their proposed methods to develop the technology and their limitations. In the second part, I discussed the Career Prediction System Based on Computer Science Students' Skills, Interests, and Related Works and I reviewed other research-based works. I showed their used algorithms, model accuracy, and their limitations.

CHAPTER 3
METHODOLOGY/ REQUIREMENT ANALYSIS & DESIGN SPECIFICATION

3.1 Overview

The method used for this combined web and machine learning-based technology is diverse. Students' data are collected to train the career prediction model at the initial stage. There are, like, so many attributes that are considered to predict students' careers. This includes students' skills, interests, and their skill-based related works. Then, integrate the generated model with a web-based system. There are also functionalities in the web system such as student skill-based assessment, faculty matching, faculty enrollment, faculty task-assign, etc. To support faculty matching, the system will have an easy-to-use interface that enables students to explore and enroll in faculties that fit their abilities and interests. After enrolling under a faculty, the faculty will be able to view students' profiles and assign tasks to students. The design, goal is to create an efficient and intuitive platform that enables students to make educated careers and academics.

3.2 Proposed Methodology/System Design Proposed Methodology:

Proposed Methodology

The proposed methodology is about the overall research progress to develop a career prediction system for Computer Science students' based on their field of skills, interests, and skill-based related works.

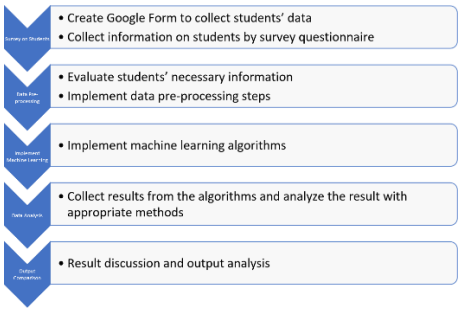

Figure 3.2.1: Methodology of extracting knowledge step by step

Survey on Students: A Google Form is created to collect students' career information. Students are asked several questions such as

1. Their semester
2. The most interesting field they want to pursue for their career
3. Their research field (If they select a career as a researcher)
4. Their higher study field (If they select a career for higher study)
5. The most contributed Computer Science (CS) courses for their career field
6. They are also asked about their skill outside of academic courses which are needed for their chosen career field
7. They engaged in any field to build their career (If they choose yes)

Students' data are collected from various top universities in Bangladesh. I have maintained a snowball process to collect data. The record of the students' information is stored in the Google Excel sheet.

Data Pre-processing: As the data is raw data it is necessary to pre-process the data for getting better results. I followed various steps to pre-process data.

1. **Manual Preprocessing:** After collecting raw data, I found that students selected various they wanted to pursue for their career. It was found that some fields were rarely chosen. So, I categorized those fields under some master fields. To categorize the fields under master field I matched core skills and pieces of knowledge that were needed for the recommended fields. Similar skill and knowledge-based fields were kept under one master field. In Table: 5 I showed the master field and the field under those master fields.

Table 3.2.1: Master Fields, Students' Chosen Fields, and Required Skills for the Chosen Fields

Master Field	Students Chosen Field	Required Skill
Development	• Mobile App Development	• Web Development (WD) • Mobile App Development (MAD)

	• Video Game Development Web Development	• Game Development (GD) • Database Development (DD) • Problem-Solving and Analysis (PSA) • Critical Thinking (CT) • Adaptability (ADP) • Cloud Computing (CC) • Project Management (PM) • Communication • Interpersonal Skills • Network Security (NS) • System Design (SD) • API Knowledge (AK)
Security	• CyberSecurity • Information Security Analysts • Cloud computing and DevOps Tech Policy and Advocacy	• Cybersecurity • Data Analysis and Visualization (DAV) • Network Security • Critical Thinking • Artificial Intelligence (AI) and Machine Learning • Cloud Computing • Project Management • Cloud Computing • Network Topologies • Adaptability • Problem Solving and Analysis

		• Operating Systems and Networking • Cryptograph • Data Engineering (DE) • System Design (SD) • Communication • Interpersonal Skills
Artificial Intelligence	• Artificial Intelligence • Robotics and Automation • Aerospace Industry • Internet of Things Machine Learning	• Problem Solving and Analysis • Data Engineering • Artificial Intelligence (AI) and Machine Learning • Data Analysis and Visualization • Critical Thinking • Database Development • Big Data Technologies (BDT) • Feature Engineering (FE) • Statistical Analysis (SA) • System Design • Cloud Computing • Robotics Knowledge (RK) • Project Management • Internet of Things (IoT) • Communication • Interpersonal Skills • Adaptability

		• Critical Thinking
Software Development and Engineering	• Software Development and Engineering • Project Manager • Consulting and IT Management • System Analyst	• Web Development • Mobile App Development • Artificial Intelligence (AI) and Machine Learning • Problem Solving and Analysis, • Cybersecurity, • Operating Systems and Networking • Big Data Technologies, • Project Management • Database Development, • Data Analysis and Visualization • Data Engineering • System Design • Cloud Computing • Communication • Interpersonal Skills • Adaptability • Critical Thinking
User Experience (UX) and User Interface (UI) Design	• User Experience (UX) and User Interface (UI) Design • Graphic Design Video Editing	• UI/UX Knowledge • Graphic Design • Working with Technology and New Idea Generate

		• Video Editing and Animation
Data Aanalysis	Data Analysis	• Problem Solving and Analysis • Data Engineering • Artificial Intelligence (AI) and Machine Learning • Data Analysis and Visualization • Critical Thinking • Database Development • Communication • Big Data Technologies • Feature Engineering • Statistical Analysis • System Design • Cloud Computing • Critical Thinking • Interpersonal Skill • Adaptability

2. **Handling Garbage Value:** The database or data warehouse stores raw data that may be inconsistent and require cleaning and integration. [22] It was noticed that many students did not read the question on the form properly and provided some skills and interests that were not matched with their chosen career. If there was extensive mismatch between chosen and other related then I removed those data. Sometimes bought a little bit of change to match their career with other fields. Though it was a really hard task I ran this activity every row.
3. **Remove Unnecessary Columns:** Some columns have no impact on students' career choices. So, I removed those columns (Computer Science Courses, Semester).
4. **Marge the Columns:** As the data is text data it is needed to apply Natural Language Processing for other pre-processing steps and get the career prediction output. To apply natural language's library I need to combine the input field in one field and keep the master fields as output. Here input field x = Skill and output field y = Career.

5. **Stop words Remove and Make the Words Lower:** When preprocessing text data, certain words lack significant value and instead contribute redundancies that can decrease model performance. Here I removed stop words and made the words lowercase using NLTK library.
6. **Label Encoding:** Label Encoding is used to assign each category a unique integer label. [31] To label the master fields Label Encoder technique is used.

Table 3.2.2: Master Field Liberalization

Label	Master Field
0	Artificial Intelligence (AI)
1	Data Science (DS)
2	Development (DEV)
3	Security (SEC)
4	Software Development and Engineering (SDE)
5	User Experience (UX) and User Interface (UI) Design (UI/UX)

7. **Vectorization:** Vectorization is a method in computer science and mathematics in which operations are done to complete arrays (vectors) of data rather than individual pieces. It is a key idea in numerical computing that contributes notably to the efficiency of algorithms and calculations, particularly in data science, machine learning, and scientific computing. Text vectorization approaches can significantly improve classification accuracy. [25] This method can enhance classification accuracy. [25] Here vectorization works to calculate how many times a word comes for a particular field.
8. **Data Train Test Split:** Train test split is an important part of data science. The data was split in such a way that 80% of each dataset was allocated for training, while the remaining 20% was reserved for testing.

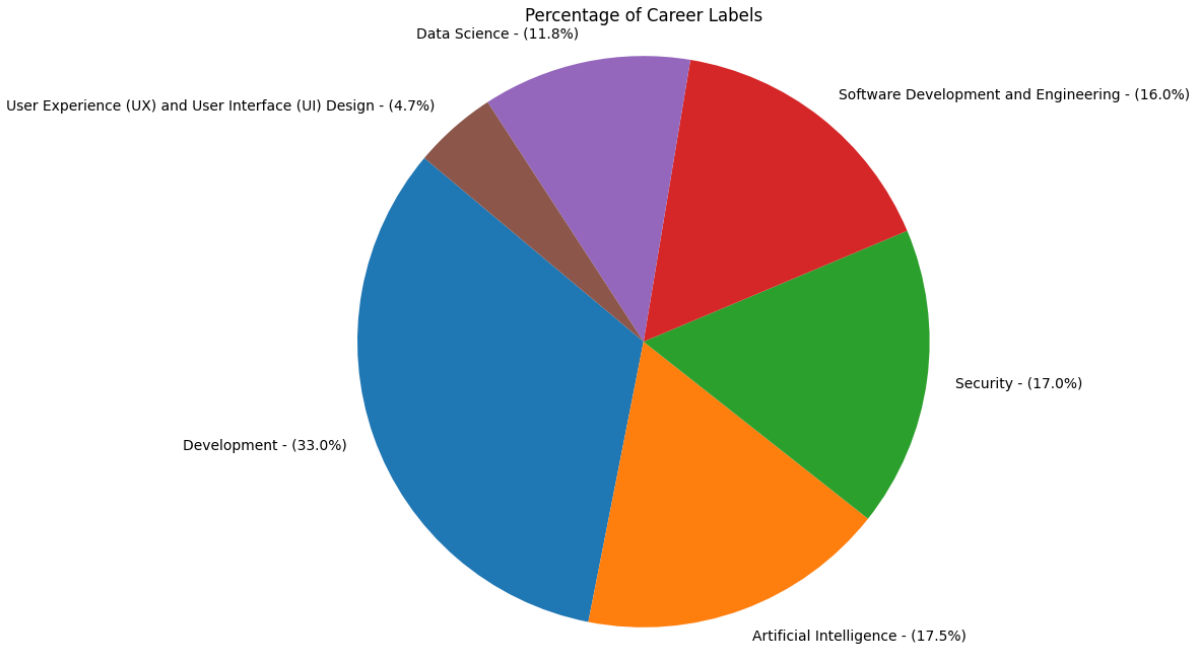

Figure 3.2.2: Number of the percentage of each master field

9. **Data Shuffleing:** In machine learning, shuffling data is important because it provides randomness and removes any inherent order or patterns in the dataset. [26] When data is given in a specified order and stuff, such as all examples from one class followed by another it might result in biased learning during the training phase, which is a big issue. Shuffling the data guarantees that the model sees a broad and random collection of instances throughout each iteration of training, which stops it from learning patterns based on the sequence of the data. So, this thing helps to build a more robust and generalizable model.

Apply Machine Learning Algorithms: ML is the study of utilizing computers to replicate how individuals learn and develop self-improvement techniques for gaining new knowledge and skills. It involves building models and algorithms that allow systems to discover current facts and constantly enhance performance. [27] In this phase, I will discuss implemented machine learning algorithms. Before going to discuss applying machine learning I will give a short overview of True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN). A false positive (FP) is when a result is wrongly

forecasted as positive when it is really negative. [22] A false negative (FN) is when a result is wrongly forecasted as negative when it fact is positive while the TP and TN are correct predictions of the model. [22]

So, the question can arise as to why we need to apply machine learning algorithms in our research as I have been researching CS students' career prediction so it is necessary to learn machines to predict the career of any student based on his or her skill, interest, and skill-based work. For this purpose, I used several machine learning algorithms to analyze the comparison of prediction.

Decision Tree (DT): A supervised ML algorithm that is for both classification and regression tasks. It is similar to a flowchart where core nodes are represented by rectangles and leaf nodes by ovals. [28] Decision Tree learning is a highly effective technique due to its ease of use, comprehensibility, lack of factors, and ability to deal with mixed-type data. [29] As our dataset is already clean and pre-processed DT gives an average performance.

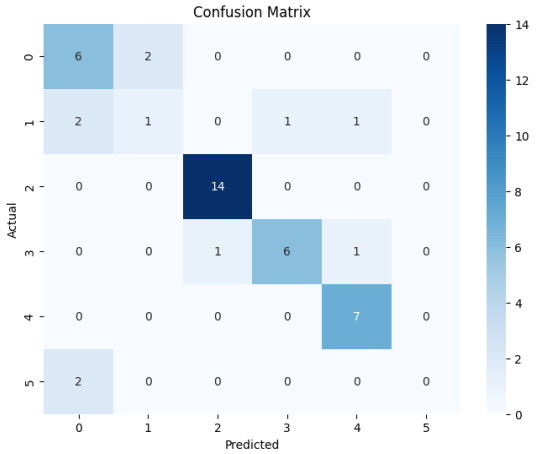

Figure 3.2.3: Confusion Matrix of Decision Tree

The model did not work well for classes 0, 1, and 5 (Figure 3). It is noticed that for class 0 FP is 4 and FP is 2 for class 2 FP is 2 and FN is 4 and lastly, for class 5 TP is 0 and FN is 2. That means the model did not identify properly class 5.

Support Vector Machine (SVM): It is a supervised ML algorithm that uses a separating hyperplane to classify data points into different categories, maximizing the margin between classes for effective linear binary classification. [23] SVM is effective for Natural Language Processing (NLP). In the pre-processed dataset it gives better performance.

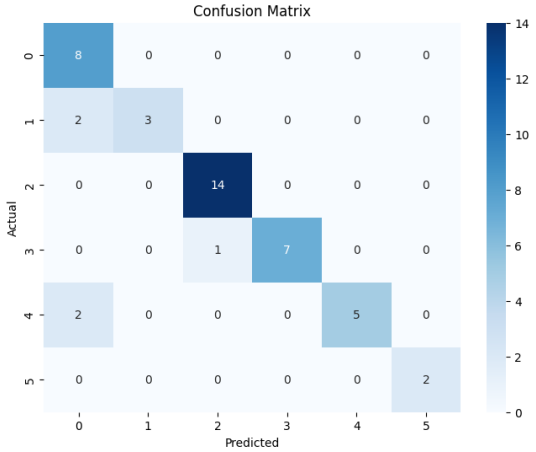

Figure 3.2.4: Confusion Matrix of Support Vector Machine

It is noticeable that the model struggles hard for classes 0, 1, and 4 (Figure 4). For class 0 FP is 4, for class 1 FN is 2 (predicted as class 0) and for class 4 FN is 2 (predicted as class 0).

Logistic Regression (LR): Logistic Regression is a statistical technique used in the field of machine learning for binary and multiclass classification tasks. It gives relatively good performance on the overall dataset.

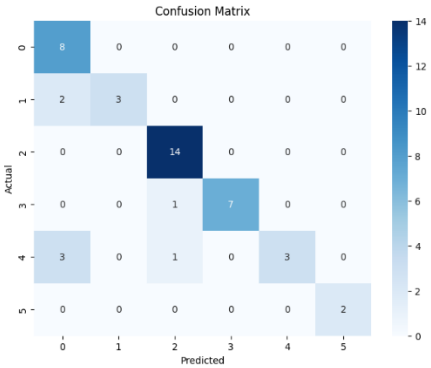

Figure 3.2.5: Confusion Matrix of Logistic Regression

The model appears to struggle with identifying between classes 1 and 0, as seen by the false negatives for class 1 and false positives for class 0 (Figure 5). For class 1 FN is 2 and for class 0 FP is 5. Class 4 also confuses with classes 0 and 2, with occurrences misclassified.

K-Nearest Neighbors (KNN): It is a supervised machine learning algorithm that also supports both classification and regression tasks. It gives an average performance.

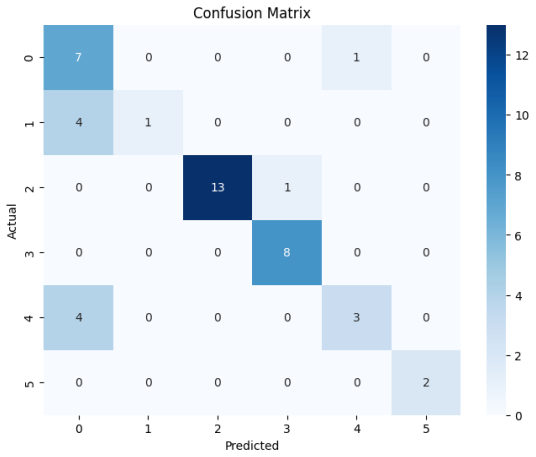

Figure 3.2.6: Confusion Matrix of K-Nearest Neighbors

Class 0 and Class 4 have poor precision and recall, indicating that the model struggles with these two classes (Figure 6). Class 4 has a particularly high rate of false negatives. The model also gives low performance to identify class 1.

Naïve Bayes (NB): NB is a probabilistic ML method that relies on Bayes' theorem. [34] The word "naïve" signifies the idea that the existence or missing of one aspect does not affect the presence or lack of other traits. [30] It gives an average performance on the dataset.

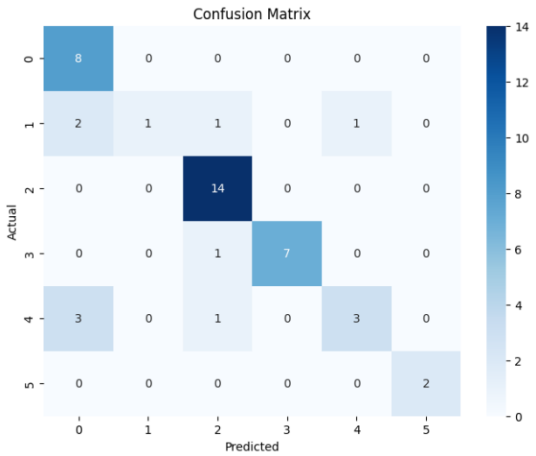

Figure 3.2.7: Confusion Matrix of Naïve Bayes

For class 0 the model had a few false positives (misclassifying as class 1) and false negatives (misclassifying as class 4) (Figure 7). For class 2 the model misclassified some instances as class 1 and class 4. There was one false negative for class 3 (misclassified as class 2). For class 4 the model has false positives (misclassified as class 1) and false negatives (misclassified as class 0 and class 2).

Convolutional Neural Network (CNN): It is a sort of artificial neural network that is designed specifically to process and analyze image input. [32] [33] The input data for both the sets used for training and testing was reshaped by applying the 'reshape ()' function to conform to the CNN model format, assuring appropriate format and compliance with the model shape's input data for effective training and prediction. I split the train and the validation dataset with a proportion of 80:20. We used the ‘stratify’ parameter to balance the distribution of data for each class in the train validation dataset. Then I used a total of six layers (Convolutional Layer, MaxPooling Layer, Dropout Layer, Flatten Layer, Dense Layer (Fully Connected), and Output Layer (Dense with Softmax Activation)). Here I used the ReLU activation function. In the model, there are a total of 50 epochs. That means, that in each epoch, the neural network will process the entire training dataset and update the weights to minimize the discrepancy between what is expected and the actual target values.

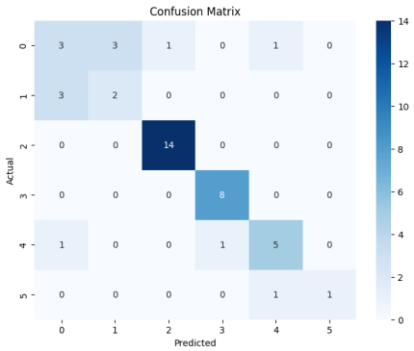

Figure 3.2.8: Confusion Matrix of Convolutional Neural Network

For CNN total TP are 33 and FP are 11.

Multilayer Perceptron (MLP): A multilayer perceptron (MLP) is an artificial neural network made up of many layers of nodes or neurons. [25] Initially I split the dataset into train and validation datasets in the proportion of 80:20 by taking the random state as 42. Similar to CNN here I also used the 'stratify' parameter so that the distribution of data for each class was properly maintained for the training and validation dataset. Then I created TensorFlow datasets for the training, testing, and validation dataset and shuffled each dataset to increase the model performance. In the model, I used four hidden layers, each with a different number of neurons: the first included 128 neurons, the second 64 neurons, the third 32 neurons, and the fourth 16 neurons. Additionally, ReLU activation algorithms were added to each hidden layer. To avoid overfitting, dropout layers were carefully positioned after each hidden layer at rates of 0.5, 0.5, 0.4, and 0.3, respectively.

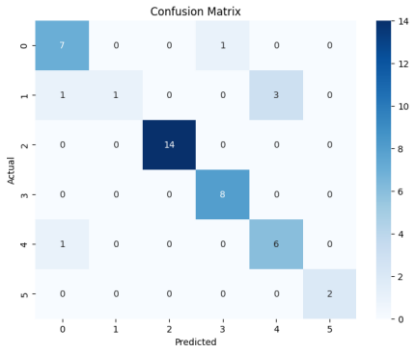

Figure 3.2.9: Confusion Matrix of Multilayer Perceptron

For MLP TP are 38 and FP are 6.

Long Short-Term Memory: It is a recurrent neural network (RNN) architecture that solves the vanishing gradient problem in typical RNNs. [24] [36] The input data for both the set used for training and testing sets was reshaped by applying the 'reshape()' function by the LSTM model formatting. I split the training and the validation dataset with the proportion of 80:20 by using the parameter ‘stratify’ to balance the data distribution for each class. In the model, I used a total of two hidden layers (LSTM layer, and dense layer). I added a dropout layer following the LSTM layer, with a rate of dropout of 0.5, to help regulate the model.

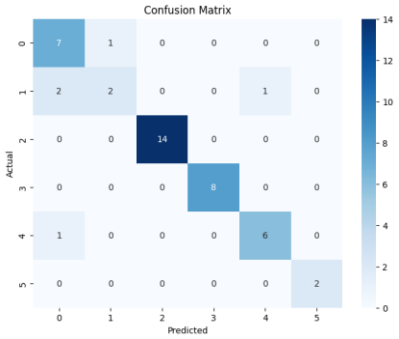

Figure 3.2.10: Confusion Matrix of Long Short-Term Memory

In LSTM total TP are 39 and total FP are 5.

System Design:

In this section, I will highlight the system design and development of the Student Assessment and Career Prediction web and machine learning integrated system.

System design is crucial in the software development process that defines a system's architecture and components. It requires creating a plan of action outlining how the application and data system will satisfy the criteria. To develop the system I followed some steps.

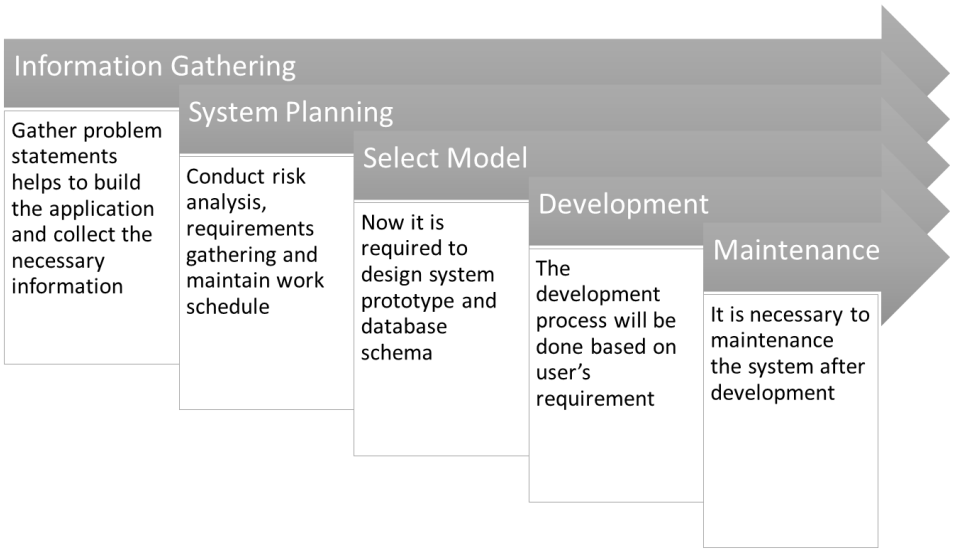

Figure 3.2.11: System Design Steps

Information Gathering: It is a crucial phase to develop a web application that helps to understand the project's requirements, user demands, and corporate objectives. This phase includes collecting necessary information, finding problem statements, choose the platform.

1. **Collect Necessary Information:** As the system is about students' career prediction it is necessary to collect student career-related information such as students' academic performance, skill, interest, and skill-based work. This information is collected through Google Forms.
2. **Problem Statement:** It is recommended to know why the system is developing and its purposes. As the system will work with students and faculty it needs to connect both types of users.
3. **Platform Selection:** It is understood to know which type of user uses the system, and for whom the system is developing. As the Student Assessment and Career Prediction System is developing for students and faculties it is necessary to maintain these users' requirements.

System Planning: It is a crucial phase in developing a multifunctional system. It discussed the project scope, objectives, and strategies of the overall project, risk, and feasibility analysis.

1. **Project Scope:** It defines which feature will fulfill the user's demand. Based on the user's requirements developers must include those types of features and functionalities.
2. **Resource Assessment:** Identify and evaluate necessary items for the project such as budget, storage capacity, timeline, etc.
3. **Risk Assessment:** Analyse potential risks and challenges such as finding similar kinds of systems available in the market, system weaknesses, and user satisfaction. So the first task of risk assessment will be to point out the vulnerable sites of the system and fix those parts.
4. **Feasibility Study:** It addresses the viability of the project. It considers technical, operational, economic, legal, and scheduling factors.
5. **Project Timeline:** As the Student Assessment and Career Prediction is a big project it needs a long time (around three to four months to develop).

Model Selection and System Architecture

Model Selection: Model selection is an important step for software development that ensures the success and efficiency of the development process. As it is a large-scale project that needs frequent changes the agile model will be the best choice for this project.

1. **Iterative Development:** Agile supports iterative development so that the project will be developed incrementally.
2. **Flexibility to Changes:** Agile allows for alterations and modifications to requirements even later in the development process. Since the project is related to students' career prediction and it is also related to data analysis it needs to be updated over time. Agile adaptability will help if there are any frequent changes.

3. **Continuous Feedback:** Agile emphasizes regular feedback from users. This is a beneficial part of the project to collect users' feedback if they require any feature that helps them to use the system more comfortably.

4. **Collaboration and Communication:** The model will build a collaborative relationship between users and developers. Besides it helps to communicate with users to understand their needs and requirements.

5. **Adaptive Planning:** Agile's adaptive planning suits project requirements. The project plan may change over time in this case Agile's adaptive planning will be the best option.

System Architecture: In this phase, I will discuss the system frontend design and workflow.

1. System's Activity Diagram: An activity diagram is a UML (Unified Modeling Language) that gives a clear idea about the workflow of the system. In this section, I will show the system diagram (Figure 11) of the Student Assessment and Career Prediction System.

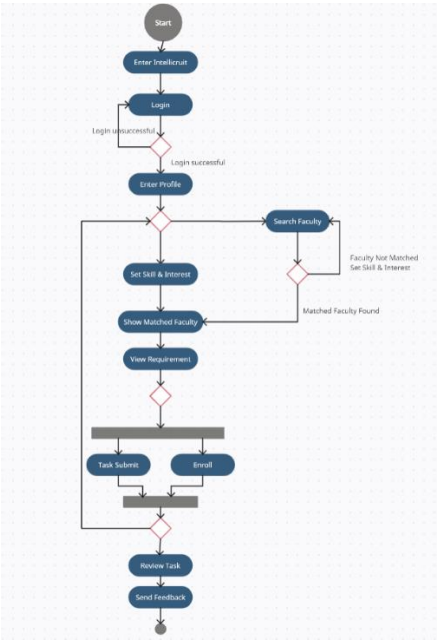

Figure 3.2.12: Activity Diagram

2. System Prototype: It is important to design a system prototype before developing one. The system prototype gives a total overview of the system. I have designed some prototypes of the Student Assessment and Career Prediction System in Figma.

Development: I will discuss coding and used technologies in this section.

Technologies: In this system, several technologies are used both in frontend and backend.

1. **Frontend Technologies:** Next JS (version 14, app router), React, TypeScript, Axios, Tailwind, ANT Design, DaisyUI, Eslint, Redux, React Hook form, JSON Web Token, Axios, Redix Toolkit, JSON Web Token Decode, ReactDom, Yup Validation.
2. **Backend Technologies:** Node JS, TypeScript, Prisma, PostgreSQL, Axios, Eslint, Husky, Prettier, Express, Cors, Nodemailer, Bcrypt, Cookie-Parser, JSON Web Token, Dotenv, Embedded Javascript Template (EJS), HTTP Status, Multer, Winston, Zod Validation.
3. **Python Server:** Flask, Jsonfy, Joblib

Coding: The coding part is related to developing both frontend and backend. Here I have developed both frontend and backend parallely. I have followed the MVC (Model, View, Controller) pattern to develop the backend.

1. **API Integration:** Several API routes are used to maintain data flow. API interconnects both frontend and backend. When a user submits data from the client side it goes through API to the backend (Figure 16). When data is received from the server data will be pre-processed to store in the database. To connect data storage to the server there are also several API routes.

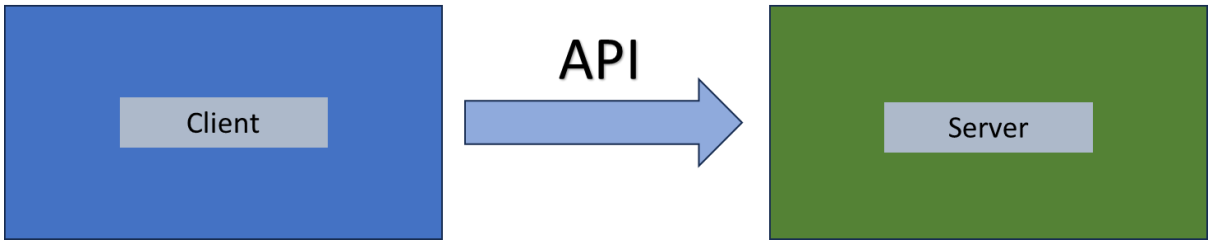

Figure 3.2.13: Connect Client and Server

2. **Maintain Authorization:** When a student creates his or her account, he or she must give his or her email address, and authentic university name. To complete the registration the system will send an email to the specific student's email address. When a student checks the email and goes to the link then his or her email verification will be complete. If the student provides the wrong email address or the outside email address of his or her institute then the system will automatically reject to create an account. This process is created so that outside users can not create student accounts on this website. When logging into the system student must provide his or her email address and password. If the student is not verified then he or she can not log in.

   When a faculty member creates an account, he/she has to wait for some time till the administrator verifies his/her account. After the verification is done then the faculty member will get an email.

3. **Maintain Authentication:** In the system, every role-based user has a different dashboard. So that it adds more efficiency to maintain system performance. To maintain authentication, I used the JSON secret key for every role-based user. So that one user can not enter another user's profile or dashboard.

4. **Integration of Python Server:** When a student inputs his or her skill, interest, and skill-based work the student's data will be sent to the Python server to analyze his or her career. When the career prediction is done the response data will again be sent to the client-side through the Node JS server (Figure 17). It is need to mention I have already implemented a pre-trained model to the Python server.

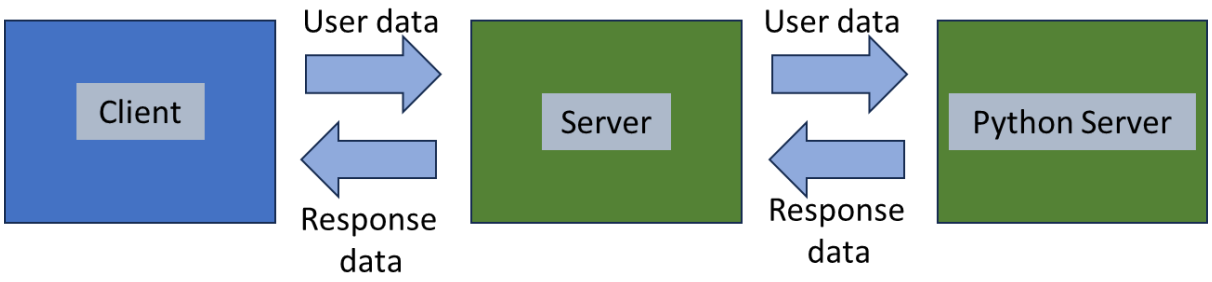

Figure 3.2.14: Communication Between Client and Python Server

Maintainance: It refers to the ongoing process of the system that includes system enhancement, the system continues to meet its requirements, can adapt to changing requirements, and remains efficient. Now I will discuss system maintenance.

1. **System Ongoing Testing:** This includes system bug fixing, identifying any issue, regular monitoring, and taking corrective action based on user feedback.
2. **Adaptive Maintenance:** It ensures that the system remains compatible with the evolving environment. It addresses environmental changes, such as upgrades of operating systems, websites, or external APIs.
3. **Data Analysis Model Updates:** Because the machine learning model used to forecast undergraduate career choices depends on data, maintenance entails updating the model with fresh data to increase its accuracy and relevance. The model might require to be retrained regularly.
4. **Security Updates:** It is necessary to check system security regularly to reduce harmful security threats.

Documentation Maintenance: Keeping documentation up-to-date is important for both the developer and the user. If any changes are made to the system, the documentation should be updated immediately.

3.3 Hardware/ Software Requirement

Hardware Requirements:

1. **Server Hardware:**
    - High-performance server with multiple CPU cores (e.g., 8-core processor or higher)
    - At least 32 GB of RAM
    - Solid State Drives (SSD) for fast data access and retrieval, with a minimum of 1 TB storage capacity
    - Reliable backup storage solution
2. **Client-Side Hardware:**
    - Modern laptop computers with at least dual-core processors
    - Minimum 4 GB of RAM
    - Reliable internet connection for web-based interactions
    - Various devices (desktops, laptops, tablets, smartphones) to ensure cross-platform compatibility

Software Requirement:

- Front End Technologies: Next JS (version 14, app router), React, TypeScript, Axios, Tailwind, ANT Design, DaisyUI, Eslint, Redux, React Hook form, JSON Web Token, Axios, Redix Toolkit, JSON Web Token Decode, ReactDom, Yup Validation.

- Back End Technologies: Node JS, TypeScript, Prisma, PostgreSQL, Axios, Eslint, Husky, Prettier, Express, Cors, Nodemailer, Bcrypt, Cookie-Parser, JSON Web Token, Dotenv, Embedded Javascript Template (EJS), HTTP Status, Multer, Winston, Zod Validation.

- Python Server: Flask, Jsonfy, Joblib

3.4 Project Management and Financial Analysis

Project Management:

- Project Planning: Begin by generating a detailed project plan that includes tasks, deadlines, completion dates, and resource allocation. In this step, I selected the most suitable programming language for frontend and backend.

- Gathering Requirements: Conduct extensive research and gather detailed requirements from stakeholders (such as system administrators, organizers, and customers) to ensure that the platform's functionalities fulfill their needs.
- Agile Development Methodology: Use an agile methodology for iterative development, which allows for adjustments and modifications based on feedback collected during the development process.
- Design and Prototyping: To ensure usability and user-friendliness, begin with wireframes and prototypes that depict the platform's functionality and interface.
- Data Collection: After designing I collected students' career information from various universities' students. As it is an intelligent system I trained the machine to predict the user's input.
- Development and Testing: Following the creation and integration of features and functions, there are extensive testing processes (unit, integration, and user acceptance tests) to ensure functionality and quality.
- Deployment and Launch: After a formal launch and marketing plan, implement the platform in stages or iterations to ensure a smooth transition and little disruption.
- Monitoring and Maintenance: Establish a mechanism for continuing monitoring, problem resolution, upgrades, and post-launch user support to ensure peak performance and user satisfaction

Finance:

- Budgeting: Create a comprehensive plan that accounts for infrastructure, marketing, people, software, hardware, and development costs.
- Resource Allocation: Distribute funds efficiently throughout project phases, taking into account the most critical developmental stages and priorities.
- Cost Management: Regularly monitor and regulate project costs to guarantee budget adherence, discovering cost-cutting opportunities without sacrificing quality.
- Risk Management: Identify financial risks and prepare contingency measures to reduce unexpected costs or delays throughout development.

- Return on Investment (ROI): Consider prospective ROI in light of market research, projected user base, revenue streams (transactions, subscriptions, and advertising), and long-term viability.
- Funding Strategies: Consider financial options like as grants, investors, or partnerships to provide the funds needed for development, marketing, and ongoing operations.
- Financial Reporting: To monitor and assess financial performance, maintain open and accessible financial records, publish regular financial reports, and compare actual expenditure to budgeted amounts.

- Within the allotted resources and strategic objectives, the successful development, launch, and sustainability of "STUDENT ASSESSMENT & CAREER PREDICTION SYSTEM" is ensured by striking a balance between solid financial management and effective project management procedures.

3.5 Summary

Methodology or requirement analysis and design specification is mainly discussed in this chapter 3. The overall section is mainly two parts. In one part I have discussed the model and, in another part, I have discussed the system where the model is integrated. In each section, I clarified the details of the model and the system.

Chapter 4

IMPLEMENTATION

4.1 Overview

The implementation of the Student Assessment and Career Prediction System involves developing a comprehensive web-based platform integrated with machine learning algorithms. This system collects and analyzes student data, including academic performance, interests, and skills, to provide personalized career guidance. The backend is built using Node.js and Express.js, with PostgreSQL and Prisma for database management, ensuring robust data handling and scalability. Python, along with libraries such as Scikit-Learn and TensorFlow, is used for developing and integrating the predictive models. The frontend is designed using modern frameworks like React.js for an intuitive user experience. Security measures, including SSL/TLS encryption and secure authentication, are implemented to protect user data. Continuous feedback and improvements ensure the system evolves to meet user needs and industry trends.

4.2 Train Model/ Prototype Design

Train Model

The model was trained using 80% of the available data, which was divided into training and testing datasets. This training data was used to teach the machine learning model to feature extraction and relationships within the data. Various algorithms, including SVM, DT, LR and Neural Networks, were employed to train the model. During this phase, the model learned from the provided data, adjusting its parameters to minimize prediction errors. This process involved multiple iterations and tuning of hyperparameters to optimize the model's performance. After training, the model was validated using the remaining 20% of the data to evaluate its accuracy and effectiveness in predicting student career paths.

Prototype Design:

I have designed system prototype in Figma

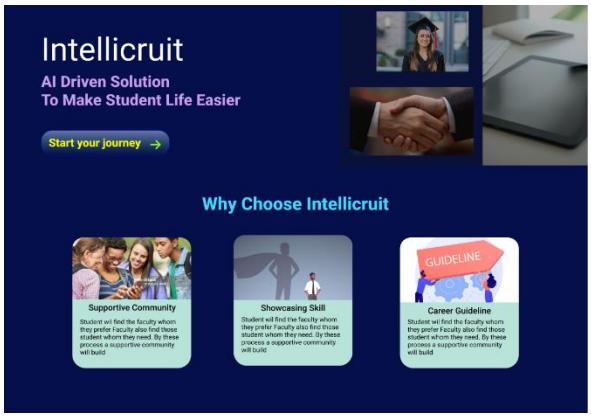

Figure 4.2.1: Home Page

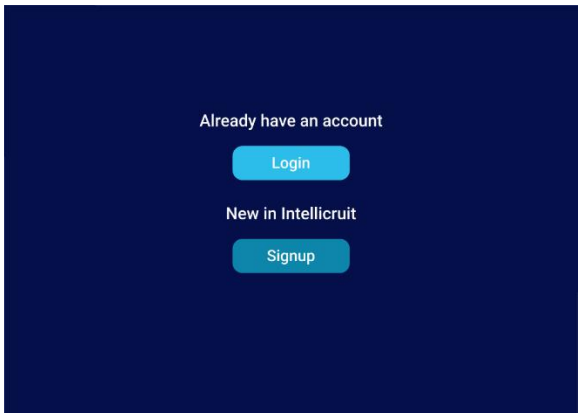

Figure 4.2.2: Signup Login Page

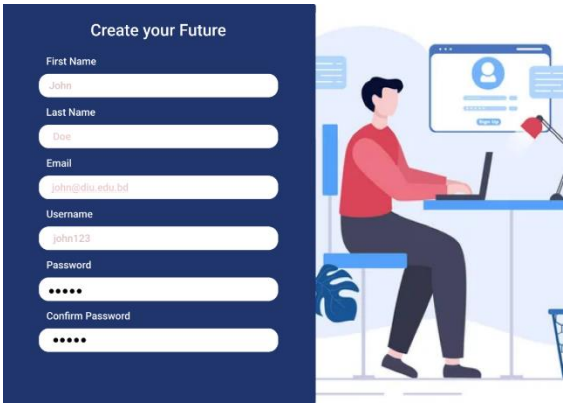

Figure 4.2.3: Signup Page

Figure 4.2.4: Profile Page

4.3 System Testing/ Model Evaluation

System Testing

The testing implementation for the "Student Assessment and Career Prediction System" involves a systematic approach to ensure the platform's functionality, performance, security, and user satisfaction. This is a summary of the testing protocols:

Testing Types:

Unit Testing: Individual components, features, and modules are tested to ensure they function correctly. Each algorithm and model used in career prediction, such as Naïve Bayes, Gradient Boosted Tree, and Deep Neural Networks, undergo unit testing to validate their accuracy and performance on small data sets.

Integration Testing: The interaction between various modules, such as the student assessment module and the career prediction module, is tested to ensure seamless integration. This ensures that the data flow between assessments and predictions is accurate and efficient.

User Acceptance Testing (UAT): Stakeholders, including educators and career counselors, are involved in scenario-based testing to ensure that the system meets real-world needs. They test the system's usability, effectiveness of predictions, and overall user experience to provide feedback for improvements.

Testing Procedures:

Test Case Development: Comprehensive test cases are developed to cover all functionalities, including edge cases and user scenarios. This involves creating test cases for various student responses and career predictions to ensure the system handles diverse inputs accurately.

Execution and Reporting: Test cases are executed systematically, and results are recorded to generate detailed reports on passed and failed scenarios. This helps in identifying any issues in the system's functionality and performance.

Regression Testing: After bug fixes or new feature implementations, regression testing is performed to ensure that existing functionalities remain unaffected. This is crucial to maintaining the stability and reliability of the system over time.

Testing Environments:

Development Environment: Developers test individual components and functionalities in their local development environments. This includes testing each machine learning model and its integration with the assessment system.

Staging Environment: Integration and overall system functionalities are tested in an environment that mirrors the production environment. This ensures that the system performs well under conditions similar to those it will face when deployed.

User Testing Environment: Actual users, such as students and educators, perform UAT in a controlled testing environment. Their feedback is crucial for identifying usability issues and making necessary improvements to enhance the user experience.

By adhering to these testing environments and procedures, the "Student Assessment and Career Prediction System" ensures functionality, security, reliability, and user satisfaction before deployment. This systematic testing approach helps in reducing risks and delivering a high-quality user experience.

Table 4.3.1: System Testing Based on Functionality

SL	**Test Condition**	**Given Input**	**Expected Output**	**Actual Output**	**Result**

01	User Log in	Email: sakir15-3862@diu.edu.bd Password: 123456	Login Successful	Login Successful	Pass
02	User Log in	Email: Sakir15-3892@diu.edu.bd Password: 123456	Login Failed	Login Failed	Pass
03	Student Select Skills	Database, Web Development, Mobile App Development	Successfully Added	Successfully Added	Pass
04	Enroll Faculty	Select faculty and enrolled	Successfully Enrolled	Successfully Enrolled	Pass
05	Faculty Create Task	Add Title, Add Description	Successfully Added	Successfully Added	Pass
06	Assign Task	Select students and assign	Successfully Assigned	Successfully Assigned	Pass
07	Student Complete Task	Add solution	Successfully Added	Successfully Added	Pass
08	Feedback Provide	Add status, Add feedback	Successfully Added	Successfully Added	Pass
09	Create Master Field	Add Title	Successfully Added	Successfully Added	Pass
10	Create Sub Field	Add Title, Add Description	Successfully Added	Successfully Added	Pass

Model Evaluation

The model evaluation involved assessing the trained machine learning model using the remaining 20% of the data, designated as the testing dataset. This phase aimed to measure the model's accuracy, precision, recall, and overall performance in predicting student career paths.

Various metrics were calculated to determine how well the model generalizes to unseen data. Cross-validation techniques were employed to ensure robustness and reliability. The results indicated high accuracy, particularly with algorithms like Gradient Boosted Trees and Decision Trees, confirming the model's effectiveness in making accurate predictions. Any discrepancies or errors were analyzed to identify potential areas for improvement, ensuring the model's continuous refinement and enhancement.

Table 4.3.2: Macro and weighted average of the classifiers

Classifier Name	**Macro Average**				**Weighted Average**			
	Precision	**Recall**	**F1-Score**	**Support**	**Precision**	**Recall**	**F1-Score**	**Support**
DT	0.58	0.62	0.59	44	0.72	0.77	0.74	44
SVM	0.93	0.86	0.88	44	0.92	0.89	0.89	44
LR	0.92	0.82	0.83	44	0.89	0.84	0.83	44
KNN	0.85	0.74	0.73	44	0.84	0.77	0.76	44
NB	0.86	0.75	0.75	44	0.83	0.80	0.77	44
CNN	0.78	0.79	0.78	44	0.80	0.80	0.79	44
MLP	0.94	0.88	0.88	44	0.94	0.91	0.90	44
LSTM	0.79	0.79	0.78	44	0.81	0.82	0.81	44

Summary

The implementation of the "Student Assessment and Career Prediction System" involved a

comprehensive process of designing and developing a web-based platform integrated with machine learning models. Key steps included gathering and analyzing student data, implementing various machine learning algorithms to predict career paths, and creating a user-friendly interface for seamless data entry and career guidance. The backend was built using Node.js and Express, while the frontend focused on responsive design principles for easy navigation. PostgresSQL and Prisma were utilized for database management, ensuring efficient data handling. Rigorous testing, including unit, integration, and system testing, was conducted to ensure robustness, security, and accuracy. Continuous feedback and iterative improvements were incorporated to enhance system functionality and user satisfaction.

CHAPTER 5
RESULT AND ANALYSIS

5.1 Overview

The "Student Assessment and Career Prediction System" demonstrated significant success in accurately predicting student career paths through the application of machine learning algorithms. The analysis of test results indicated that the models achieved high accuracy rates, with some algorithms performing exceptionally well in specific contexts. The implementation of an intuitive, user-friendly interface facilitated seamless user interaction and data entry, enhancing the overall user experience. The system's ability to provide personalized career guidance based on comprehensive data analysis was validated through positive user feedback and rigorous testing, confirming its potential as a valuable tool for students and educational institutions.

5.2 Experimental/ Simulation Result

Experimental Result

Accuracy is the fraction of correct classifications compared to the total number of instances and is determined by the confusion matrix. [23] The formula for accuracy is as follows:

From the above equation (1), we can see that the calculation of accuracy.

$$\text{Accuracy}=\frac{TN+TP}{TN+FP+FN+TP} \dots\dots\dots (1)$$

Table 5.2.1: Model Accuracy Percentage

Name of Algorithm	Accuracy
Decision Tree	77.27%
Support Vector Machine	88.63%
Logistic Regression	84.09%

K-Nearest Neighbors	77.27%
Naïve Bayes	79.55%
Convolutional Neural Network	77.27%
Multilayer Perceptron	84.09%
Long Short-Term Memory	84.09%

From the above table it is noticeable that Support Vector Machine (SVM) has the highest percentage 88.63%. Multilayer Perceptron (MLP), Logistic Regression (LR), Long Short-Term Memory (LSTM) have also a good accuracy compared with SVM of over 80%. Decision Tree (DT), K-Nearest Neighbors (KNN) and Convolutional Neural Network (CNN have the same accuracy at 77.27%. Naïve Bayes (NB) an average accuracy of 79.55%.

Precision refers to the percentage of correctly classified cases from those presented as positive. [23] The precision equation (2) is given below:

$$Precision = \frac{TP}{TP + FP} \dots\dots\dots (2)$$

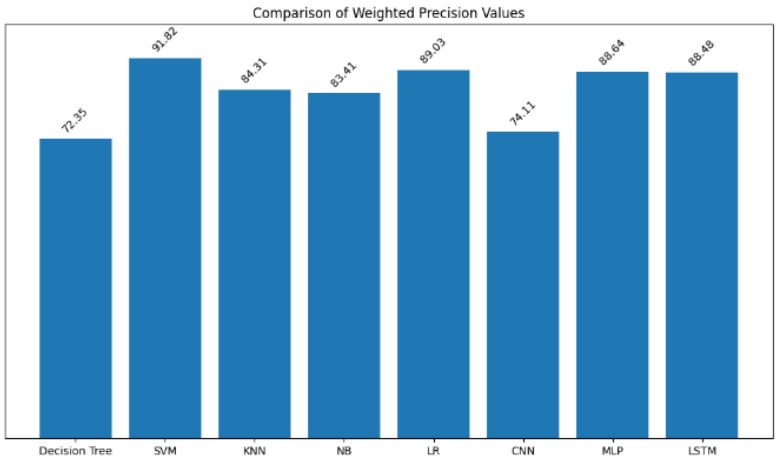

Figure 5.2.1: Model Precision Percentage

SVM has the highest precision rate (91.82%) whereas DT has the lowest precision rate (72.35%). (Figure 18) MLP, LR, LSTM have also good rates of precision compared with SVM respectively 88.64%, 89.03% and 88.48%. (Figure 5.2.1)

The recall is the fraction of correctly classified cases that are positive. The equation (3) of recall is given below:

$$Recall = \frac{TP}{TP + FN} \dots\dots\dots (3)$$

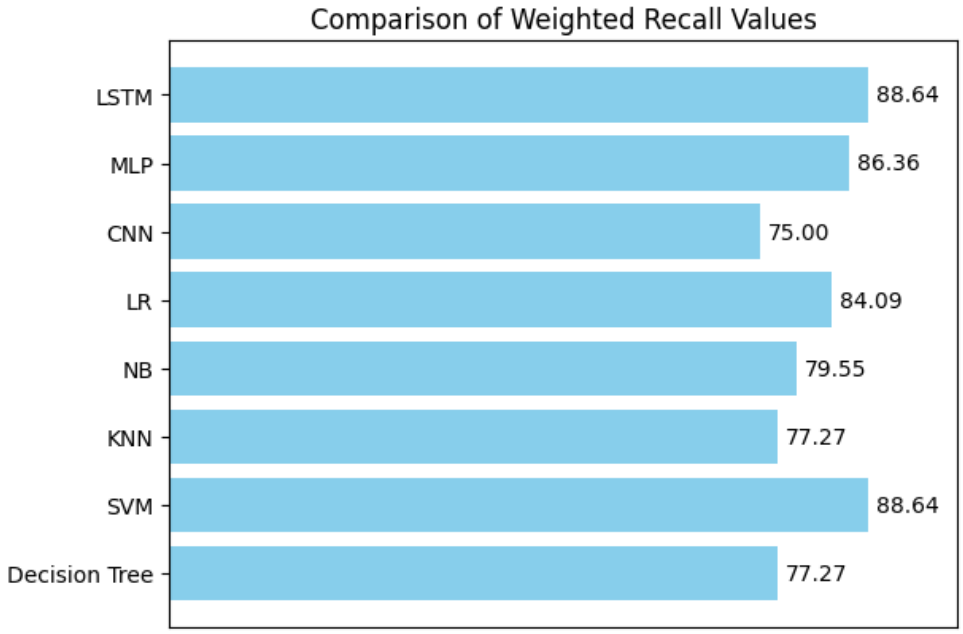

Figure 5.2.2: Model Recall Percentage

As it is seen that SVM and LSTM has the highest recall percentage (88.64%). (Figure 5.2.2). MLP and LR have also good recall percentages compared with SVM respectively 86.36% and 84.09%.

The F-measure is a retrieval measure derived from TP, FP, FN, recall, or accuracy values. I have calculated the F-measure by using the precision and recall values of the models shown in Eq. (4).

$$F - measure = \frac{2*Precision*Recall}{Precision+Recall} \dots \dots \dots (4)$$

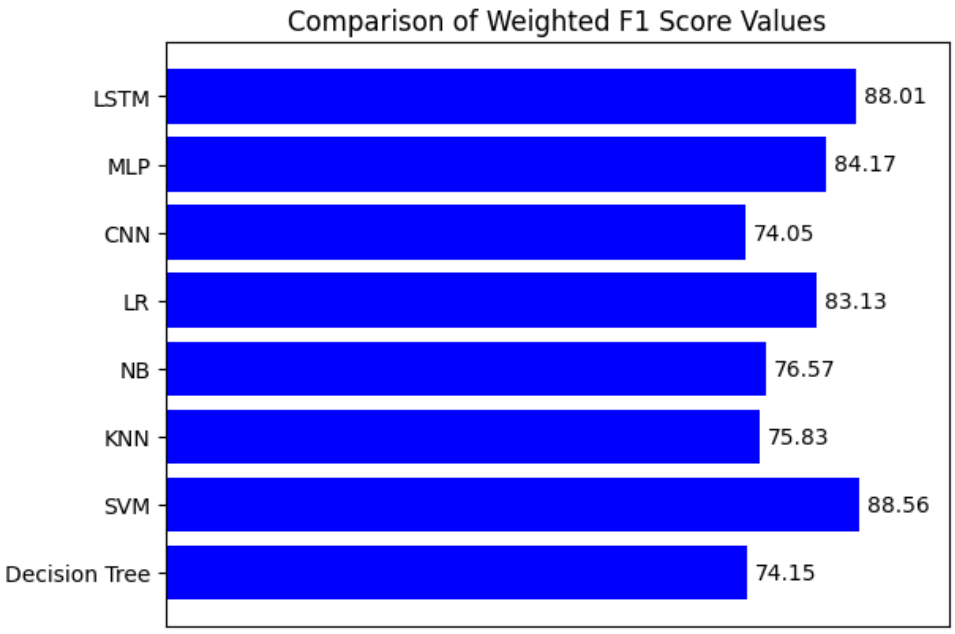

Figure 5.2.3: Model F-measure Percentage

As it can be seen SVM has the highest F1 score which is 88.56%. (Figure 5.2.3)

Simulation Result

Figure 5.2.4: Landing Page Design.

This is the home page of Student Assessment and Career Prediction System. Here show features of the system as well users can navigate to login page by clicking start journey.

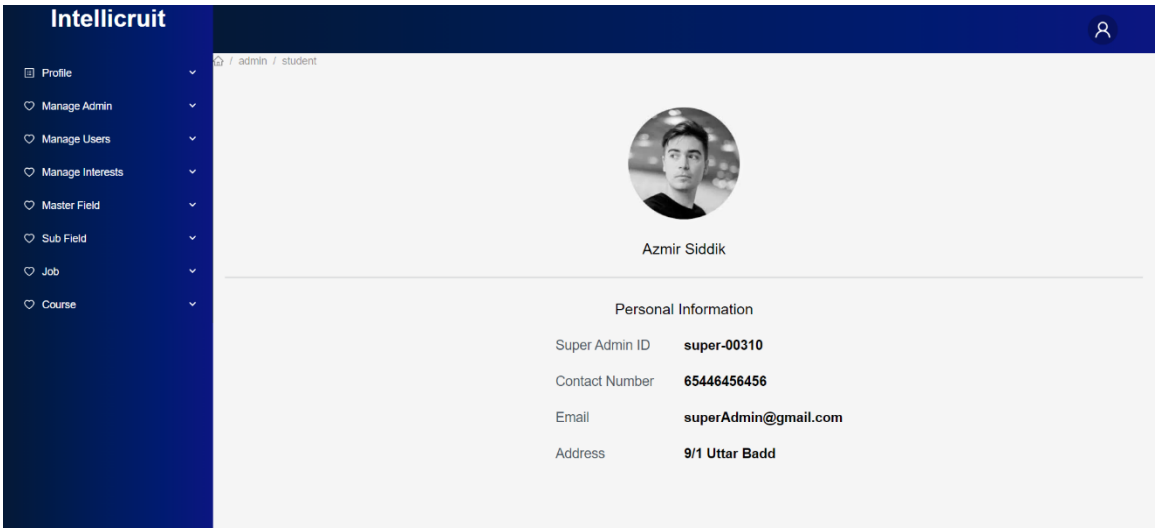

Figure 5.2.5: Super Admin Dashboard Design.

This page is super admin dashboard of Student Assessment and Career Prediction System. Here there are several options are available for the super admins. Super admin can maintain other users and admins. Besides he can maintain the whole system.

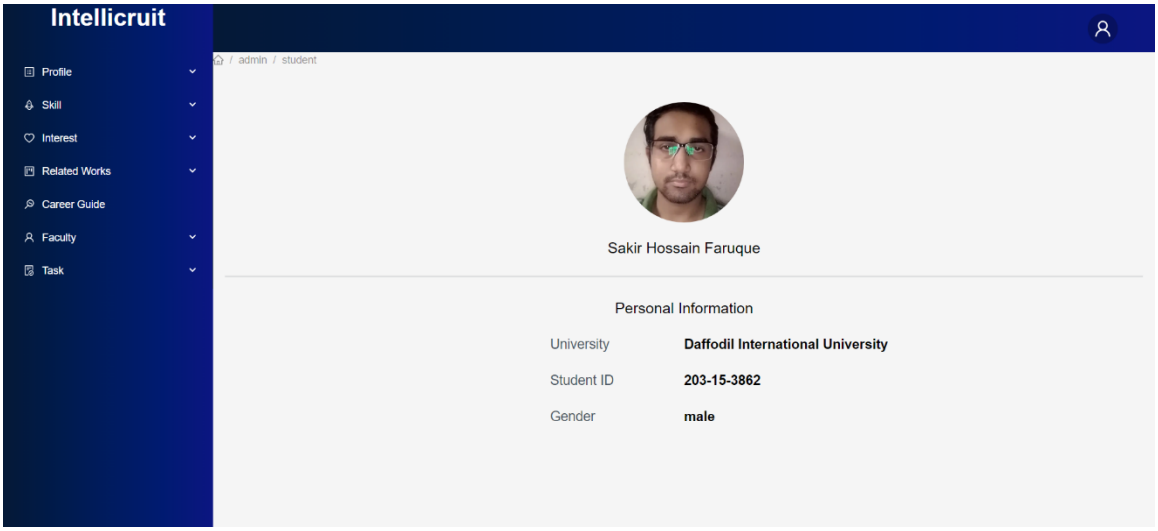

Figure 5.2.6: Student Dashboard Design

This the student dashboard of " Student Assessment and Career Prediction System". By using the dashboard students can set up their skill, interest and related works. After setting

their skills, interests and related works they can view their suitable career and navigate to job and course related to their career.
Besides student can enroll under a faculty and get assigned tasks by faculty.

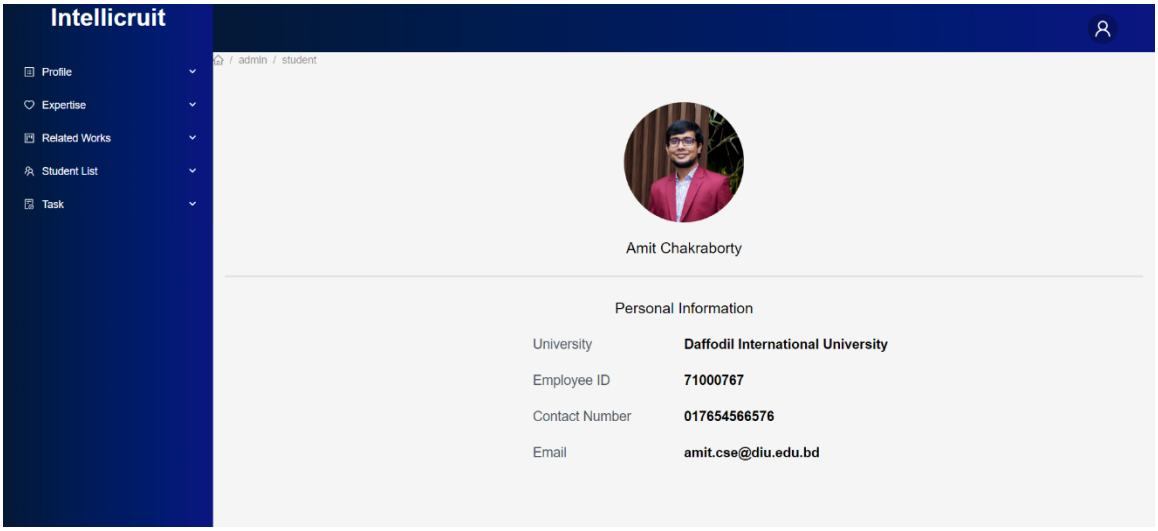

Figure 5.2.7: Faculty Dashboard Design.

This page is faculty dashboard or faculty profile page of Student Assessment and Career Prediction System. Here a faculty can set his or her expertise and activity. Faculty can view enrolled students, create task and assign task to the enrolled students. Besides faculty can provide feedback over the completed task.

5.3 Performance/ Comparative Analysis

Performance

My study consists of testing various neural network structures like CNN, MLP, and LSTM models. Each model presents different positives and negatives, with CNN showing a bit of overfitting as well as lower accuracy levels when compared to MLP and LSTM models. However, it is important to mention that our research does not only look into NN models; we also investigate conventional ML methods like DT, SVM, KNN, LR, and NB. Out of these traditional ML models, SVM is found to have the best performance, and LR is not far behind. KNN also gives an impressive performance, but it scores a little less than SVM and LR. NB demonstrates relatively lower performance when compared with LR and KNN, whereas decision tree models are more or less average in terms of performance. Our CP system utilizes the SVM model to predict

career paths. Three inputs are analyzed, determining machine-predicted outputs based on our research. This method provides users with percentage-based predictions. So that users can identify their most probable career trajectories.

Table 5.3.1: Human giving input and machine-predicted output based on the SVM model

SL	Input	Machine Predicted Output
1	AI, ML, DS, WD, SDE, GD	SDE: 56.94%, AI: 20.05%, DS: 17.00%, UI / UX: 2.68%, DEV: 1.70%, SEC: 1.63%
2	ML, DS, CS, ISA	SEC: 52.13%, SDE: 21.38%, AI: 13.99%, DS: 10.78%, UI / UX: 1.41% DEV: 0.32%
3	AI, ML, DS, DSA, TPA	AI: 49.45%, DS: 40.58%, SDE: 7.98%, SEC: 0.99%, SDE: 0.98%, UI / UX: 0.87%, DEV: 0.15%

Finally, table 5.3.1 outlines the SVM model's forecast on the optimal career for CS and SWE students based on their skills, interests, and relevant work areas. Output as a percentage allows individuals to easily determine the most likely career path. Overall, our study highlights the reasonable selection and evaluation of the model as an essential aspect of research, considering issues such as dataset properties and complexity. Although the NN models MLP and LSTM display excellent results, investigating the use of classical ML algorithms such as SVM or LR will help to better understand when using a small database. In summation, our study has validated that both NN and traditional ML models are feasible in predicting students' career aspiration. In future studies, we can develop more precise answers for students through NLP improved career predication.

Comparative Analysis

A comparative analysis of the "Student Assessment and Career Prediction System" reveals several strengths and opportunities for improvement in relation to similar existing platforms. Unlike traditional systems, our project integrates advanced machine learning algorithms and deep learning models, offering personalized career

guidance based on a comprehensive analysis of student data, including skills, interests, and academic performance. This approach is more tailored and dynamic compared to static career assessment tools. The system's user-friendly interface, designed for ease of navigation and data entry, provides a superior user experience. Additionally, the integration with educational institutions and the continuous improvement based on user feedback set this system apart from competitors. However, there are challenges in balancing robust security measures with usability, maintaining scalability, and ensuring real-time data synchronization. These areas will require ongoing attention and development to maintain a competitive edge in the educational technology market.

Summary

The "Student Assessment and Career Prediction System" effectively utilized machine learning models to predict student career paths with high accuracy, demonstrated through comprehensive testing and analysis. The system's predictive capabilities were validated by dividing the data into an 80:20 split, using 80% for training and achieving notable precision in predictions. User feedback highlighted the platform's ease of use, intuitive interface, and valuable career guidance. Overall, the analysis confirmed the system's robustness, scalability, and potential to significantly benefit students and educational institutions by providing tailored career recommendations based on detailed student data analysis.

CHAPTER 6

Impact on Society, Environment, and Sustainability

6.1 Impact on Life

- Enhanced Career Guidance:
    - Students receive personalized career recommendations, aiding them in making informed decisions about their future career paths.
    - Helps in identifying strengths, interests, and suitable career options early on.
- Improved Educational Outcomes:
    - Facilitates better alignment between student capabilities and academic pursuits, potentially leading to higher student satisfaction and performance.
    - Encourages students to focus on their strengths and areas of interest, leading to more fulfilling educational experiences.
- Support for Educators and Institutions:
    - Provides educators with insights into student capabilities and preferences, allowing for more targeted teaching strategies.
    - Assists institutions in tailoring educational programs to meet the needs of their students more effectively.
- Increased Accessibility:
    - Makes career guidance accessible to a broader range of students, including those in remote or underserved areas.
    - Ensures that all students have access to high-quality career planning resources.
- Data-Driven Decisions:
    - Enables students to make career choices based on data-driven insights rather than guesswork.
    - Provides a reliable framework for students to assess their skills and interests in relation to various career paths.
- Empowerment and Confidence:

- Empowers students by giving them a clear understanding of their potential career trajectories.
- Boosts student confidence by providing concrete evidence of their strengths and capabilities.

➢ Long-Term Benefits:

- Contributes to the creation of a more skilled and appropriately employed workforce.
- Supports long-term career satisfaction and success by aligning educational efforts with career outcomes.

➢ Facilitation of Lifelong Learning:

- Encourages continuous skill development and learning, preparing students for dynamic and evolving job markets.
- Promotes the importance of lifelong learning and adaptability in career planning.

6.2 Impact on Society & Environment

The implementation of the "Student Assessment & Career Prediction System" has the potential to bring about significant positive societal changes in multiple ways:

- Enhanced Educational Efficiency: By simplifying the assessment process and providing timely feedback, the system improves efficiency for educators and students alike.
- Personalized Learning Paths: The system's ability to predict career trajectories based on individual student data helps tailor educational experiences to each student's strengths and interests.
- Informed Career Choices: By providing data-driven career predictions, the system helps students make informed decisions about their future, aligning their academic pursuits with potential career paths.
- Increased Accessibility: The platform's user-friendly design ensures that all students, including those with disabilities, have equal access to its features. This promotes inclusivity and ensures that no student is left behind due to barriers in technology access.

- Support for Educators and Administrators: The system aids educators in identifying students' strengths and areas for improvement, allowing for targeted interventions.
- Economic Impact: By aligning educational outcomes with market needs, the system helps produce a workforce that is better prepared for current and future job markets. This can lead to reduced unemployment rates and a more robust economy.
- Promotion of Lifelong Learning: The system encourages a culture of continuous learning and self-assessment. By providing students with tools to evaluate their progress and potential career paths, it fosters a mindset of lifelong learning and adaptability.
- Technological Advancement: The integration of machine learning and artificial intelligence in educational assessment promotes the adoption of advanced technologies in the educational sector. This not only enhances the educational experience but also prepares students for a technology-driven world.

The "Student Assessment & Career Prediction System" significantly impacts society by enhancing educational efficiency, promoting personalized learning, supporting informed career choices, increasing accessibility, aiding educators and administrators, contributing to economic growth, promoting lifelong learning, advancing technology, fostering social equity, and encouraging sustainable practices.

Impact on the Environment

The "Student Assessment & Career Prediction System" can positively impact the environment in several ways:

- Reduction in Paper Usage: The digitization of student assessments, records, and career predictions eliminates the need for printed exams, report cards, and career counseling materials. This transition to digital documentation conserves paper and reduces waste, contributing to a more sustainable educational environment.
- Energy Efficiency in Learning: By promoting online assessments and career guidance, the system minimizes the need for physical infrastructure such as examination halls and counseling offices. This reduction in physical space usage can lead to decreased energy consumption for heating, cooling, and lighting in educational institutions.

- Optimized Resource Allocation: Through data-driven insights, the system helps educators and administrators optimize resource allocation, such as targeting specific areas for improvement in student performance. Efficient resource management can reduce unnecessary expenditures and the environmental impact associated with overuse of educational materials and facilities.
- Supporting Remote Learning: The system's ability to provide assessments and career guidance online supports remote learning, reducing the need for commuting to educational institutions. This reduction in travel decreases carbon emissions associated with daily commutes by students and staff.
- Sustainable Career Guidance: By incorporating career prediction models that highlight sustainable career paths and industries, the system can guide students towards environmentally friendly career choices. This promotes a culture of sustainability and environmental responsibility among the future workforce.
- Minimized Use of Physical Resources: The implementation of a cloud-based infrastructure for the system reduces the need for on-site servers and data centers, which can be resource-intensive. Cloud services often utilize energy-efficient practices and renewable energy sources, further reducing the environmental footprint.
- Data-Driven Environmental Awareness: The system can include modules that educate students about environmental issues and the importance of sustainability. By integrating environmental education into the career prediction process, students are more likely to make informed and eco-friendly career choices.
- Encouraging E-Learning Platforms: The system encourages the use of e-learning platforms for continuous education and career development, reducing the need for physical materials and resources. This shift supports a more sustainable approach to lifelong.

6.3 Ethical Aspects

The development and implementation of the "Student Assessment & Career Prediction System" necessitate careful consideration of various ethical aspects to ensure fairness, privacy, and integrity throughout its use:

- Ensuring strong security and privacy protocols is paramount to safeguard user data, including student personal information, assessment results, and career predictions. Compliance with data privacy laws and regulations (such as GDPR) is essential, as is obtaining explicit consent from users before collecting and utilizing their data.
- Transparency and user consent are critical components. Providing users with clear information about data usage, privacy policies, and terms of service is essential. Obtaining explicit consent before collecting and using user data ensures that users are fully aware of how their information is being handled.
- Fair and unbiased recommendations are crucial to the system's integrity. Ensuring that career suggestions and assessments are impartial, equitable, and not influenced by any hidden or commercial agendas is necessary to maintain trust and fairness.
- Accessibility for all users must be ensured by designing the platform to accommodate users with disabilities and adhering to accessibility guidelines. This ensures inclusivity and equal access to the system's features and benefits for all users.
- Responsible AI and personalization should be a guiding principle. Using machine learning and AI algorithms to provide personalized career recommendations must be done ethically, ensuring that user privacy is respected and that the algorithms do not unduly influence user choices.
- Accountability and transparency in system practices are essential. Avoiding dishonest practices and maintaining integrity and transparency in how assessments and predictions are made, as well as in customer support and services, helps build trust with users.
- Consideration of bias in algorithmic decision-making is vital. Ensuring that the recommendation system is free from biases related to gender, ethnicity, or other discriminatory factors is critical to providing fair and equitable career advice.
- Ethical advertising and marketing practices must be followed. Ensuring that any promotional activities or advertisements related to the platform are honest, not misleading, and adhere to ethical standards is necessary to maintain the platform's credibility.

- Respect for cultural sensitivities is important in providing career advice. Recognizing and respecting diverse cultural backgrounds and norms ensures that the system is inclusive and culturally sensitive in its recommendations and content.

By integrating these ethical principles into the design and operation of the "Student Assessment & Career Prediction System," the platform can prioritize user privacy, fairness, transparency, accessibility, and ethical practices, thereby fostering trust and integrity among its stakeholders and users.

6.4 Sustainability Plan

Developing a sustainability strategy for the "Student Assessment & Career Prediction System" involves implementing practices to ensure long-term sustainability, reduce environmental impact, and promote ethical and responsible behavior. Here's an outline:

Environmental Sustainability:

- Efficient Resource Usage: Optimize server performance to minimize energy consumption by utilizing efficient algorithms and scalable cloud services.
- Digital Transformation: Encourage the shift from paper-based assessments to digital formats to reduce paper waste and the overall carbon footprint associated with traditional educational methods.

Responsible Data Handling:

- Privacy Compliance: Regularly update privacy policies to comply with international data protection laws, ensuring user privacy and data protection.
- Data Encryption and Security Measures: Continuously enhance data encryption and security protocols to protect user data from breaches and unauthorized access.
- Data Minimalism: Adhere to data minimization principles, collecting and retaining only the necessary amount of user data to ensure the platform's functionality and performance.

Social Responsibility:

- Enhance Accessibility: Regularly assess and improve the platform's accessibility features to accommodate users with disabilities.

- Community Engagement: Support educational initiatives and programs that benefit local communities, fostering cross-cultural understanding and promoting equitable access to education.

Continuous Improvement:

- Integration of User Feedback: Actively collect and incorporate user feedback to refine features and ensure they meet user needs and preferences.
- Innovation and Adaptation: Stay abreast of technological advancements and integrate new features to enhance user experience and overall system efficiency.
- Monitoring and Reporting: Continuously monitor and report on compliance with privacy regulations, user satisfaction, and the system's environmental impact to stakeholders.

Ethical Business Practices:

- Transparent Practices: Maintain transparency regarding pricing, services, and business practices to build user trust and loyalty.
- Fair Partnerships: Establish and uphold ethical business standards by forming fair and responsible partnerships with educational institutions, service providers, and other stakeholders.

Employee Engagement:

- Training and Awareness: Foster a culture of responsibility by providing employees with training on ethical practices and sustainability.
- Well-being Programs: Offer programs that support employee well-being and work-life balance, promoting a healthy and productive work environment.

The purpose of this sustainability plan is to guide the "Student Assessment & Career Prediction System" in maintaining ethical standards, environmental awareness, and continuous improvement. This plan ensures that the platform's growth is aligned with conscientious and sustainable values while meeting the needs of students, educators, and other stakeholders.

Summary

The implementation of a Student Assessment and Career Prediction System has a profound impact on society, environment, and sustainability. By guiding students towards career paths that align with their skills and interests, the system fosters a more skilled and satisfied workforce, which in turn boosts economic productivity and social well-being. Environmentally, it can promote careers in sustainability and green technologies, encouraging eco-friendly practices and innovations. From a sustainability perspective, the system supports lifelong learning and continuous skill development, ensuring that individuals can adapt to evolving job markets and contribute to sustainable economic growth. This holistic approach not only enhances individual lives but also drives societal progress and environmental stewardship.

CHAPTER 7
CONCLUSION AND FUTURE WORK

7.1 Conclusion

The implementation of the Student Assessment and Career Prediction System has demonstrated its potential to significantly enhance the educational landscape by providing tailored career guidance based on individual student data. Utilizing advanced machine learning models, the system accurately predicts career paths, helping students make informed decisions about their futures. This innovative approach not only improves student satisfaction and success rates but also aligns with the dynamic needs of the job market, promoting a more skilled and adaptable workforce. Furthermore, the integration of data security measures and user-friendly interfaces ensures a trustworthy and accessible platform for all users, fostering a collaborative and forward-thinking educational community.

7.2 Further Suggested Works

In the future, this system can be expanded to include students from all departments within a university, providing comprehensive career guidance across various fields of study. Additionally, the system can be enhanced to support job search functionalities, allowing students and other users to find suitable job opportunities based on their skills, interests, and academic performance. Incorporating features like real-time job market analysis, employer matching, and continuous model improvement based on user feedback will further increase the system's effectiveness and relevance, ensuring it remains a valuable tool for career development and job placement.

7.3 Limitations/ Conflict of Interests

Despite the promising outcomes, the Student Assessment and Career Prediction System has certain limitations and potential conflicts of interest. Currently, the system is primarily tailored for specific departments, and expanding it to encompass all university departments requires additional customization and validation to ensure accuracy and relevance across diverse fields. There may also be concerns regarding data privacy and the ethical use of

student information, necessitating stringent measures to protect user data and maintain trust. Furthermore, while the system aims to provide unbiased career guidance, there is a potential risk of algorithmic bias influencing predictions, which needs ongoing monitoring and adjustment. Lastly, collaboration with external job markets and institutions introduces potential conflicts of interest that must be managed transparently to uphold the integrity and impartiality of the career guidance provided.

Reference

Conference/Journal Papers:

[1] Wang, T.H., Wang, K.H., Wang, W.L., Huang, S.C. and Chen, S.Y., 2004. Web-based Assessment and Test Analyses (WATA) system: development and evaluation. *Journal of Computer Assisted Learning*, *20*(1), pp.59-71.

[2] Dahalan, H.M. and Hussain, R.M.R., 2010. Development of web-based assessment in teaching and learning management system (e-ATLMS). *Procedia-Social and Behavioral Sciences*, *9*, pp.244-248.

[3] Coleman, E.B. and Penuel, W.R., 2000. Web-based student assessment for program evaluation. *Journal of Science Education and Technology*, *9*, pp.327-342.

[4] Antal, M. and Koncz, S., 2011. Student modeling for a web-based self-assessment system. *Expert Systems with Applications*, *38*(6), pp.6492-6497.

[5] Guzmán, E. and Conejo, R., 2005. Self-assessment in a feasible, adaptive web-based testing system. *IEEE transactions on Education*, *48*(4), pp.688-695.

[6] Supriyanto, G., Widiaty, I., Abdullah, A.G. and Yustiana, Y.R., 2019, December. Application expert system career guidance for students. In *Journal of Physics: Conference Series* (Vol. 1402, No. 6, p. 066031). IOP Publishing.

[7] Huang, L., 2022. The Establishment of College Student Employment Guidance System Integrating Artificial Intelligence and Civic Education. *Mathematical Problems in Engineering*, *2022*.

[8] Mehraj, T. and Baba, A.M., 2019. Scrutinizing artificial intelligence based career guidance and counselling systems: an appraisal. *International journal of interdisciplinary research and innovations*, *7*(1), pp.402-411.

[9] Crişan, C., Pavelea, A. and Ghimbuluţ, O., 2015. A need assessment on students' career guidance. *Procedia-Social and Behavioral Sciences*, *180*, pp.1022-1029.

[10] Myla, S., Marella, S.T., Goud, A.S., Ahammad, S.H., Kumar, G.N.S. and Inthiyaz, S., 2019. Design decision taking system for student career selection for accurate academic system. *International Journal of Scientific and Technology Research*, *8*(9), pp.2199-2206.

[11] Ade, R. and Deshmukh, P.R., 2014, August. An incremental ensemble of classifiers as a technique for prediction of student's career choice. In *2014 First International Conference on Networks & Soft Computing (ICNSC2014)* (pp. 384-387). IEEE.

[12] Mandalapu, V. and Gong, J., 2019. Studying Factors Influencing the Prediction of Student STEM and Non-STEM Career Choice. In EDM.

[13] Reddy, M.M., 2021. Career Prediction System. International Journal of Scientific Research in Science and Technology, 8(4), pp.54-58.

[14] Lent, R.W., Lopez Jr, A.M., Lopez, F.G. and Sheu, H.B., 2008. Social cognitive career theory and the prediction of interests and choice goals in the computing disciplines. Journal of Vocational Behavior, 73(1), pp.52-62.

[15] Rangnekar, R.H., Suratwala, K.P., Krishna, S. and Dhage, S., 2018, August. Career prediction model using data mining and linear classification. In 2018 fourth international conference on computing communication control and automation (ICCUBEA) (pp. 1-6). IEEE.

[16] Kabakus, A.T. and Senturk, A., 2020. An analysis of the professional preferences and choices of computer engineering students. Computer Applications in Engineering Education, 28(4), pp.994-1006.

[17] Casuat, C.D., Festijo, E.D. and Alon, A.S., 2020. Predicting students' employability using support vector machine: a smote-optimized machine learning system. International Journal, 8(5).

[18] Shankhdhar, A., Agrawal, A., Sharma, D., Chaturvedi, S. and Pushkarna, M., 2020, February. Intelligent decision support system using decision tree method for student career. In 2020 International Conference on Power Electronics & IoT Applications in Renewable Energy and its Control (PARC) (pp. 140-142). IEEE.

[19] Vignesh, S., Priyanka, C.S., Manju, H.S. and Mythili, K., 2021, March. An intelligent career guidance system using machine learning. In 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS) (Vol. 1, pp. 987-990). IEEE.

[20] Sinha, A. and Singh, A., 2023. Student Career Prediction Using Algorithms Of Machine Learning. Ashish, Student Career Prediction Using Algorithms Of Machine Learning (May 6, 2023).

[21] Sobnath, D., Kaduk, T., Rehman, I.U. and Isiaq, O., 2020. Feature selection for UK disabled students' engagement post higher education: a machine learning approach for a predictive employment model. IEEE Access, 8, pp.159530-159541.

[22] Agarwal, S., 2013, December. Data mining: Data mining concepts and techniques. In 2013 international conference on machine intelligence and research advancement (pp. 203-207). IEEE.

[23] Kumar, D. and Arora, S., 2011. Performance analysis of various data mining algorithms: A review. International Journal of Computer Applications, 32(6), pp.9-16.

[24] Almarabeh, H., 2017. Analysis of students' performance by using different data mining classifiers. International Journal of Modern Education and Computer Science, 9(8), p.9.

[25] Krzeszewska, U., Poniszewska-Marańda, A. and Ochelska-Mierzejewska, J., 2022. Systematic comparison of vectorization methods in classification context. Applied Sciences, 12(10), p.5119.

[26] Muralidhar, K. and Sarathy, R., 2006. Data shuffling—A new masking approach for numerical data. Management Science, 52(5), pp.658-670.

[27] Wang, H., Ma, C. and Zhou, L., 2009, December. A brief review of machine learning and its application. In 2009 international conference on information engineering and computer science (pp. 1-4). IEEE.

[28] Priyam, A., Abhijeeta, G.R., Rathee, A. and Srivastava, S., 2013. Comparative analysis of decision tree classification algorithms. International Journal of current engineering and technology, 3(2), pp.334-337.

[29] Su, J. and Zhang, H., 2006, July. A fast decision tree learning algorithm. In Aaai (Vol. 6, pp. 500-505).

[30] Sueno, H.T., Gerardo, B.D. and Medina, R.P., 2020. Multi-class document classification using support vector machine (SVM) based on improved Naïve bayes vectorization technique. International Journal of Advanced Trends in Computer Science and Engineering, 9(3).

[31] Krstinić, D., Braović, M., Šerić, L. and Božić-Štulić, D., 2020. Multi-label classifier performance evaluation with confusion matrix. Computer Science & Information Technology, 1.

[32] Cohen, O., Malka, O. and Ringel, Z., 2021. Learning curves for overparametrized deep neural networks: A field theory perspective. Physical Review Research, 3(2), p.023034.

[33] Li, L., Xiao, L., Wang, N., Yang, G. and Zhang, J., 2017, December. Text classification method based on convolution neural network. In 2017 3rd IEEE International Conference on Computer and Communications (ICCC) (pp. 1985-1989). IEEE.

[34] Lakhotia, S. and Bresson, X., 2018, October. An experimental comparison of text classification techniques. In 2018 International Conference on Cyberworlds (CW) (pp. 58-65). IEEE.

[35] Kamath, C.N., Bukhari, S.S. and Dengel, A., 2018, August. Comparative study between traditional machine learning and deep learning approaches for text classification. In Proceedings of the ACM Symposium on Document Engineering 2018 (pp. 1-11).

[36] Bai, X., 2018, September. Text classification based on LSTM and attention. In 2018 Thirteenth International Conference on Digital Information Management (ICDIM) (pp. 29-32). IEEE.

Dataset

[37] [dataset] Hossain Faruque, Sakir; khushbu, sharun akter (2024), "Student Career Dataset", Mendeley Data, V1, doi: https://data.mendeley.com/datasets/4spj4mbpjr

Website

[38] Career Prediction System available at https://github.com/hrugved06/Career-Prediction-System last accessed: 30/6/2024 at 3:38pm

[39] Blended Learning Center available at https://elearn.daffodilvarsity.edu.bd/?redirect=0 last accessed: 29/6/2024 at 6 pm

[40] Skill Job available at https://skill.jobs/ last accessed: 25/6/2024 at 10 am

Appendix A

Course Outcomes, Complex Engineering Problems (EP) and Complex Engineering Activities (EA) Addressing

Title: STUDENT ASSESSMENT & CAREER PREDICTION SYSTEM

Student ID: 203-15-3862

CO Description for FYDP

CO	CO Descriptions	PO
Phase -I		
CO1	Integrate recently gained and previously acquired knowledge to identify a **web based travel management** problem for the Final Year Design Project (FYDP)	**PO1**
CO2	Analyze different aspects of the goals in designing a solution for this FYDP	**PO2**
CO3	Explore diverse problem domains through a literature review, delineate the issues, and establish this goals for the FYDP	**PO4**
CO4	Perform economic evaluation and cost estimation and employ suitable project management procedures throughout the development life cycle of the FYDP	**PO11**
Phase -II		
CO5	Design and develop technical solutions and system components or processes that meet specified requirements, ensuring compliance with public health and safety standards, as well as considering cultural, socioeconomic, and environmental factors in this FYDP	**PO3**
CO6	Choose and apply appropriate methodologies, resources, and contemporary engineering and IT technologies to address complex engineering processes, encompassing prediction and modeling, while adhering to relevant constraints in this FYDP	**PO5**
CO7	Analyze societal, health, safety, legal, and cultural considerations, along with associated responsibilities, in the context of professional engineering practice and the resolution of this problem, employing logical reasoning guided by contextual understanding.	**PO6**
CO8	Comprehend and evaluate the enduring sustainability and impact of professional engineering endeavors in addressing intricate engineering challenges within social and environmental frameworks.	**PO7**
CO9	Implement ethical principles and adhere to professional standards and norms in this FYDP	**PO8**
CO10	Capable of operating proficiently both individually and as a team member or leader across diverse teams and interdisciplinary settings in this FYDP.	**PO9**
CO11	Proficiently communicate with the engineering community and broader society regarding complex engineering endeavors, including the ability to comprehend and generate comprehensive reports and design documentation, as well as provide and receive clear instructions throughout this FYDP.	**PO10**

CO12	Acknowledge the importance of self-directed and life-long learning within the evolving landscape of technology, and possess the readiness and capability to engage in lifelong learning endeavors.	**PO12**

Addressing CO (1 to 8), Knowledge Profile (K), Attainment of Complex Engineering Problems (EP), and Attainment of Complex Engineering Activities (EA)

Addressing CO (1 to 8), Knowledge Profile (K), Attainment of Complex Engineering Problems (EP):

SN	EP Definition	Attainment	CO	Justification (with Knowledge Profile)	References
1.	**EP1: Depth of Knowledge required**	Yes	CO1, CO2, CO3, CO5, CO6, CO7 and CO8	**K1 (Theory-based Understanding)** In the proposed methodology section, I have discussed how can I perform students' career predictions so that I need to go to students to know their careers and then collect information. Based on their information I have to pre-process data and generate a model. **K2 (Conceptually-based mathematics, numerical analysis, statistics, and formal aspects of computer and information science)** In the proposed methodology I have discussed on confusion matrix of different models and showed the variation of the models. In the model input-output analysis, I have also shown the accuracy, precision, recall, and f-measure of the proposed models.	Page no: 21-24 Section: 3.2 Page: 29-35 Section: 3.2

				K3 (Engineering Fundamentals) In the proposed methodology in the data pre-processing section, I discussed vectorization, applying feature engineering and data train-test split In the system I have discussed student data storage on a local server using PostgresSQL, authorization, and authentication. **K4 (Specialist Knowledge)** In the FYDP I collected data from students and then pre-processed data for the machine I trained the classifiers and select the best classifier to predict students' career. To develop the system, I used SQL database and other necessary technologies.	**Page: 21-24. 41-43** **Section: 3.2**
				K5 (Design) In chapter 3 proposed methodology system design section I showed the system prototype and system working activity. **K6 (Technology)** In chapter 3 proposed methodology I showed technology used for frontend, backend, and python server	**Page: 46-48** **Section: 4.2** **Page: 42-43** **Section:**

				In the machine learning implementation section, I used several supervised learning algorithms and neural networks **K8 (Research)** In chapter 2 Literature Review section I showed related works based on the career prediction model. Chapter 2 I showed similar technologies as Student Assessment System	**3.3** **Page: 8-13** **Section: 2.2**
2.	**EP2: Range of Conflicting Requirements**	Yes	CO2, and CO7	In the project, I aim to gather career-related information from students. However, I often face challenges in data collection as many students seem uninterested in participating. This lack of participation not only slows down my data collection process but also extends the time required for data preprocessing. To address this, I am exploring new strategies to engage students and streamline my data collection and preprocessing methods.	**Page: 22** **Section: 3.2**
3.	**EP3: Depth of analysis required**	Yes	CO2, and CO6	For predicting students' careers, I selected the most accurate model through accuracy analysis. I also incorporated features from similar online systems to enhance my system's effectiveness.	Page: 43-47 Section: 5.2
4.	**EP4: Familiarity of Issues**	Yes	CO8	To develop students' career and assessment system I needed to study similar journal and tried to find out their gap in their research works.	**Page: 11-13** **Section: 2.3, 2.4**

5.	**EP5: Extends of application codes**	Yes	CO5	In chapter 3 system design I showed step step-wise process of software development coding strategy	Page: 30 Section: 3.2
6.	**EP6: Extends of stakeholders involved and conflicting requirements**	Yes	CO8	In Chapter 3 project management, and Chapter 1 project scope I showed how I will communicate with users and connect users with system	Page: 4-5, 33 Section: 1.5, 3.4
7.	**EP7: Interdependence**	Yes	CO5	In my project, I initially trained various classifiers, then selected the optimal model for integration into our web application. Subsequently, I developed additional features to enhance the application's functionality.	**Page: 19-32** **Section: 3.2**

Addressing CO11 with Complex Engineering Activities (EA) [Some or all of the following]:

SN	EA Definition	Attainment	CO	Justification	References
1.	**EA1: Range of resources**	Yes	CO11	This system's development relies on rigorous financial resource allocation for infrastructure, marketing, and development, and successful execution using tools such as PG Admin, Visual Studio Code, and numerous programming languages, frameworks, and databases.	**Page no:** 29-30, **Section:** 3.2

2.	**EA2: Level of interaction**	Yes		The system interconnected both students and faculty to maintain one to one communication and share knowledge	Page: 47-49 Section: 5.2
3.	**EA3: Innovation**	Yes		The Student Assessment and Career Prediction System is a complete system which interconnected both students and faculty. By this system student will get career guideline.	Page: 53-57 Section: 3.2
4.	**EA4: Consequences for society and the environment**	Yes		This project prioritises social impact and environmental sustainability by improving trip efficiency, promoting sustainable travel practices, and ensuring ethical considerations such as strong security, privacy, and transparent business practices, thereby fostering societal well-being and environmental stewardship.	**Page no:** 62-66, **Section:** 6.1, 6.2, 6.4
5.	**EA-5: Familiarity**	Yes		To develop students' career and assessment system I needed to study similar journal and tried to find out their gap in their research works.	**Page: 13-19** **Section: 2.3, 2.4**

Addressing CO (4, 9, 10, and 12):

SN	COs	Attainment	Justification	References
1	CO4	Yes	The project collects CO4 by combining financial management strategies such as budgeting, resource allocation, cost management, and risk mitigation, assuring efficient fund distribution, monitoring expenditures, and finding	**Page no:** 1-6, **Section:** 1.2, 1.3, 1.4, 1.5

			areas for improvement while retaining quality.	
2	CO9	Yes	The project adheres to ethical norms by prioritising user privacy, ensuring openness in data usage, making fair recommendations, respecting cultural sensitivities, and encouraging accessibility, trust, transparency, and integrity among stakeholders.	**Page no:** 65-67, **Section:** 6.3
3	CO10	No	N/A	N/A
4	CO12	Yes	The project's implementation phase emphasizes lifelong learning through diverse testing protocols, automated frameworks, and CI/CD pipelines, ensuring ongoing improvement and adaptation to emerging tech trends, with results analysis refining the platform for enhanced effectiveness.	**Page no:** 46-48, **Section:** 4.2

Appendix B

Faruque, S.H., Khushbu, S.A. and Akter, S., 2024. Unlocking Futures: A Natural Language Driven Career Prediction System for Computer Science and Software Engineering Students. *arXiv preprint arXiv:2405.18139*.

Student Assessment and Career Prediction System

ORIGINALITY REPORT

18%	9%	7%	9%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Daffodil International University Student Paper	7%
2	dspace.daffodilvarsity.edu.bd:8080 Internet Source	5%
3	"Proceedings of International Conference on Trends in Computational and Cognitive Engineering", Springer Science and Business Media LLC, 2021 Publication	1%
4	"Advances in Visual Informatics", Springer Science and Business Media LLC, 2021 Publication	<1%
5	Savira Dwia Nadela, Kaiqian Yin. "Application of data science in career path and employability analysis from secondary and higher education students: a systematic review", Research Square Platform LLC, 2023 Publication	<1%
6	Iweriebor, Lawrence Ejime. "Approach to Medicare Provider Fraud Detection and Prevention", Capitol Technology University,	<1%