

ML Based Career Suggestive System for Informal Job Sector Considering Cognitive Skills

Submitted by

Ms. Ayesha Siddika Tonny

18301197

Hafsa

18301205

Md. Tousif Hasan Lavlu

18301190

Abhijit Kumar Ghosh

18301191

Sourojit Roy

18301199

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science and Engineering.

Department of Computer Science and Engineering
Brac University
May 2022

© 2022. Brac University
All rights reserved.

Declaration

It is hereby declared that

1. The thesis submitted is our own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

Ayesha Siddika

Ms. Ayesha Siddika Tonny
18301197

Hafsa

Hafsa
18301205

Touf.

Md. Tousif Hasan Lavlu
18301190

Abhijit Kumar Ghosh

Abhijit Kumar Ghosh
18301191

Sourojit Roy

Sourojit Roy
18301199

Approval

The thesis/project titled “ML Based Career Suggestive System for Informal Job Sector Considering Cognitive Skills.” submitted by

1. Ms. Ayesha Siddika Tonny (18301197)
2. Hafsa (18301205)
3. Md. Tousif Hasan Lavlu (18301190)
4. Abhijit Kumar Ghosh (18301191)
5. Sourojit Roy (18301199)

Of Spring, 2022 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on May 29, 2022.

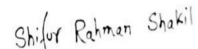
Examining Committee:

Supervisor:
(Member)



Md. Khalilur Rahman, PhD
Associate Professor
Department of Computer Science and Engineering
BRAC University

Cosupervisor:
(Member)



Shifur Rahman Shakil
Manager, Skills Development Programme, BRAC

Program Coordinator:
(Member)

Md. Golam Rabiul Alam, PhD
Associate Professor
Department of Computer Science and Engineering
BRAC University

Head of Department:
(Chair)

Sadia Hamid Kazi
Chairperson and Associate Professor
Department of Computer Science and Engineering
Brac University

Ethics Statement

Our goal is to choose the right career for unprivileged people using an AI model based on the assessment. We are the ones who are representing it. Because the majority of people in Bangladesh live in poverty, it is clear that most individuals are unable to work well as a result of poor decisions. For which financial problems arise. But if they use our system, they can choose the right path and overcome their financial difficulties. There are other NGOs, including BRAC, which are also conducting various training programs without resorting to any structural approach. As a result, many are losing interest in working in their trained trades. In our system, Q&A will select a qualified trade through an AI model with answers and assessments so that there is no problem in going to that trade. This will increase the efficiency of the employee and reduce the loss of the system of change in the workplace. We have ensured total transparency in our evaluation process and have provided a visual interpretation of the model's results.

Abstract

We are striving to build a realistic procedure by which we, particularly the future generation, will be able to choose the right career based on their capacity and interests. A few well-known international firms, including IBM, Unilever, LinkedIn, Accenture, and others, utilize Pymetrics to hire their staff, which is based on cognitive skills in the formal sector. Our work, however, is the first in the informal sector. On our primary collected dataset, we used six distinct algorithms, including Logistic Regression, Decision Tree, Random Forest Classifier, Support Vector Classification, Multilayer Perceptron Classifier, and Extreme Gradient Boosting (XGB), and discovered that Random Forest Classifier and Extreme Gradient Boosting (XGB) are the best for this system, with the accuracy of 57% and 60%, respectively. We've also used MinMaxScaler to enhance our output. After that, we observed that the Random Forest Classifier approach had a nearly 62% higher accuracy. The Extreme Gradient Boosting (XGB) approach, on the other hand, has a precision of 58.6%. After completing our evaluation, we opted to use the Random Forest Classifier for our system instead of MinMaxScaler. Based on these insights, we'll match individuals with employment, smoothing out labor market inefficiencies and leading to considerable boosts in productivity, income, and well-being.

Keywords: Career, Capacity, Interests, cognitive skills, Informal sector, Extreme Gradient Boosting (XGB), Random Forest Classifier, MinMaxScaler, Pymetrics.

Dedication

We want to dedicate this thesis to our wonderful parents. Along with all of the incredible faculties we met and learned from while pursuing our Bachelor's degree. It's been a rewarding experience.

Acknowledgement

First and foremost, we give thanks to the Great Almighty for finishing our thesis.

We would like to give credit to Dr. Md. Khalilur Rhaman, an Associate Professor at BRAC University, for supervising the work that was done on this study. We would like to extend our deepest thanks to our Md. Khalilur Rhaman, Ph.D., Associate Professor, as well as our co-supervisor, Shifur Rahman Shakil Manager, Skills Development Programme, BRAC, for their assistance and patient guidance during the course of the year. Their insightful recommendations, thoughts, support, and efforts have been of tremendous use to us in the conduct of our research.

Finally, without our parents' continuous support, it may not be achievable. We are currently on the verge of graduating thanks to their generous assistance and prayers.

Table of Contents

Declaration	i
Approval	ii
Ethics Statement	iv
Abstract	v
Dedication	vi
Acknowledgment	vii
Table of Contents	viii
List of Figures	x
List of Tables	xi
Nomenclature	xi
1 Introduction	1
1.1 Introduction	1
1.1.1 Motivation	1
1.1.2 Major Contribution	2
1.2 Problem Statement	3
1.2.1 Correlation of theoretical and practical knowledge	3
1.2.2 High illiteracy rate	3
1.2.3 Unconsciousness about training	3
1.2.4 Contributions of BRAC NGO	4
1.2.5 System for employee hiring in the formal sector	4
2 Literature Review	6
2.1 Literature Review	6
3 Working Plan	11
3.1 Research Objectives	11
3.2 Questionnaire	12
3.3 Model Analysis	18
3.3.1 Logistic Regression	18
3.3.2 Decision Tree	19

3.3.3	Random Forest Classifier	19
3.3.4	Support Vector Classifier (SVC)	21
3.3.5	Multilayer Perceptron Classification (MLP)	21
3.3.6	Extreme Gradient Boosting (XGB)	22
3.4	Workflow Diagram	23
4	Data Pre-Processing	24
4.1	Data Collection	24
4.2	Data Pre-Processing	27
4.3	Cleaning	28
4.4	Mapping	28
4.5	Feature Engineering	29
4.6	Train-Test split	29
4.7	Data Analysis	30
5	Model Implementation and Optimization	31
5.1	Evaluating Machine Learning Models	31
5.2	Before Optimization	31
5.3	After Optimization	33
6	Result Analysis	35
7	Application	36
7.1	Model Selection and Application Design	36
8	Conclusion	39
8.1	Challages	39
8.2	Future Work	39
8.3	Conclusion	40
	Bibliography	40

List of Figures

3.2.1 Questionnaires classification and leveling	12
3.3.1 Logistic Regression	18
3.3.2 Decision Tree	19
3.3.3 Random Forest Classifier	20
3.3.4 Multilayer Perceptron Classification	21
3.4.1 Workflow diagram	23
4.1.1 Adaptability responses table part one	24
4.1.2 Adaptability responses table part two	25
4.1.3 Logical ability table.	26
4.2.1 Adaptability Responses dataset's columns with number of null values.	27
4.2.2 Logical Ability dataset's columns with number of null values.	27
4.4.1 Adaptability Responses dataset's unique values.	28
4.4.2 Logical Ability dataset's unique values.	28
4.5.1 Correlation Heatmap for the combined dataset	29
4.7.1 Data Analysis scatter diagram	30
5.2.1 Comparison of algorithm's Log-loss before using MinMaxScaler	32
5.3.1 Comparison between the pre-accuracy and post-accuracy of algorithms	34
6.0.1 Final Result Analysis	35
7.1.1 Web application home page	37
7.1.2 Web application question categories	37
7.1.3 Web application trade page	37
7.1.4 Web application dataflow diagram	38

List of Tables

5.1	Pre-accuracy	32
5.2	Post-accuracy.	33

Chapter 1

Introduction

1.1 Introduction

1.1.1 Motivation

Bangladesh's unemployment rate is anticipated to reach 6% by the end of 2021, according to global macro simulations and Trading Economics analysts. As a result, a lot of young people, many of whom are members of minority and marginalized groups, will be unemployed or poor. The predicted rise in economic disparity and lack of job opportunities could have a detrimental affect on a generation of young people around the world. According to a poll conducted by the Bangladesh Bureau of Statistics in 2019, the Academic Information and Statistics (BANBEIS) has 10.34 million secondary school students, with 53.83 percent of them being female. At the secondary level, the current attrition rate is 36%.

In this digital era, everyone is dependent on the internet for every little thing in their day-to-day lives. From the weather update to our career decision for each information we rely on the internet. During the pandemic, during lockdown time, this dependency on the internet and electric devices is increasing as well as the effects. According to Worldometer elaboration of the latest United Nations data, Bangladesh is the 8th largest populated country in the world and had 167.1 million people by January 2022, which is 2.11% of the world population. Being a developing country, we are trying to be a developed country, so it is very important to make our population educated, self-reliant, and career-oriented. By keeping in mind all these views, our government, along with the telecommunication companies, is trying to make internet facilities available to every single person for a long time. As a consequence, there are now 52.58 million internet users in Bangladesh in January 2022, which is 31.5% of our total population. In the very beginning, most people normally used the internet and websites for their entertainment. But, as proactive citizens, individuals are increasingly putting these resources to good use.

Cognitive skills or abilities are the means by which our brain memories, reasons, maintains attention, resolves issues, thinks, reads, and learns. By assigning accessible material to the relevant regions, our cognitive talents aid us in processing scientific understanding. Apart from being entertaining, the website and internet data sharing processes have now been demonstrated to have the ability to affect a

wider range of cognitive capacities. Currently, we have various types of websites that have dynamic activities that help to be updated, enrich knowledge and also test skills through assessments. These are helping us to consume more and more information in this era. Also, these have repeatedly been demonstrated to improve various elements of cognition, including visual short-memory, multitasking, and spatial abilities.

We live in a time where anybody wants to study or improve their abilities in a certain field since every piece of data is available for free. Any person can choose what vocation to pursue, and cognitive skills operate as a subconscious mind in this scenario. In our day-to-day life, almost 95% of our behavior and activities are controlled by our subconscious mind. Likewise, when it comes to deciding on a suitable career, research shows that around 88% of people choose a career based on their cognitive skills.

1.1.2 Major Contribution

From the beginning of using the website, it was only available for national security, research, and e-commerce company-related work only. But now every small organization, company, startup, educational institution, or even kindergarten is maintaining a proper website. By this, it becomes easier for them to reach large numbers of people in the shortest time as people are more used to this kind of medium. From our context, it is really a good opportunity to communicate with the trainees and give them a clear visualization of the whole process.

In our general education system, we don't have enough direction for the formal and informal sectors' practical life knowledge. A large percentage of this percentage of people is the younger generation who will enter the job sector and maintain our economy in the future. These people stay in dilemmas about their career choice and switch from one job to another frequently, which causes unemployment and also an imbalance in the system. So, if someone is going to choose his or her career before taking training in a specific field, this website-based assessment can help to suggest a suitable career path for him or her.

1.2 Problem Statement

1.2.1 Correlation of theoretical and practical knowledge

It is critical to have both a theoretical and a hands-on education system in order to produce excellent graduates. Tragically, but unfortunately true, the vast majority of Bangladeshi youth have no or just little idea about job sectors. On the other hand, 36% of youth left their education at the secondary level, where they do not have either educational qualifications or skills. They used to change professions frequently since they lacked certain talents. According to the most current publicly-available survey figures, the typical individual changes jobs 12 times over their lifetime (2019) because they can't find the right job for them or lack certain skills [1].

1.2.2 High illiteracy rate

Furthermore, the bulk of our students come from low-income families. The majority of them are in a state of flux as they witness the birth of the day. As a result, individuals are unable to acquire proper education and training, though if they wish to. According to studies, dropout rates begin to increase slowly in Class VI, overuse injuries until Class IX, and then peak at around 60 to 70% in Class X. The students' eventual realization that they are unable to cope with the rigors of completing the test exam in Class X or the nationwide SSC exam, and also their inability to pay the exam costs in many cases, may explain the causes for this quick increase in exam costs. They are mainly between the ages of 12 and 16. Because of not having proper training, they fell behind in the employment market and spent days looking for work.

1.2.3 Unconsciousness about training

If someone is offered the opportunity to be trained, they are confused about which training to choose because their cognitive, behavioral, and personality qualities are unknown. As a result, people are unable to continue working and transferring jobs.

In the twenty-first century, the age of digital globalisation, technical skills are the most crucial way of finding work. Almost every aspect of our lives has been affected by developments. The country's total growth will be expedited only if these typical youth are converted into human capital.

1.2.4 Contributions of BRAC NGO

From the very beginning, BRAC NGO has been providing technical and vocational training with the traditional education stream and is now also offering 28 trades of training for disadvantaged children to make them economically stable in the future [2]. Even if the students fail in general education, they can be hired for the right employment by gaining expertise in a certain craft, which is a great initiative towards development [2]. In the training centers, BRAC provides 28 types of trade training among the participants, and these are fully free at a cost [2]. Unfortunately, in some cases, it fails to provide suitable training to them as the young people are not aware of their cognitive, behavioral, and personality qualities, so they have dilemmas about choosing the training sector. As a consequence, they suffer in training time and, in the end, the training is in vain because they cannot apply that knowledge in the future and have to change their job sector frequently. As a result, all those failures turn them into depression, and for the BRAC, this causes a huge system loss issue.

1.2.5 System for employee hiring in the formal sector

In this modern world, we can see a number of well-known multinational companies like IBM, Unilever, LinkedIn, Accenture, and some others using Pymetrics to hire their employees to make this process digitized. In 2003, Frida Polli and Julie Yoo founded Pymetrics with the goal of eliminating the age-old employee hiring system [3]. We know that in a normal employee hiring system, applicants have to face various levels in the job recruitment process. Research shows that there are manual application submissions, subjective exams, first-level vivas of the subjective exam passed candidates, and second-level vivas of the shortlisted candidates in an employee hiring system of the maximum companies [3]. Even though some of the companies also follow some other levels of verification, all of these are time-consuming as well as expensive. The co-founder of Pymetrics, Polli, highlighted that some firms invite more than quadruple the number of candidates for in-person interviews than the number of vacancies in their firms. On the other hand, Pymetrics uses machine learning and decades of neuroscience research and cognitive science to connect qualified job applicants with the right positions in businesses and jobs [3]. Here, a gamified procedure is being followed to collect employees' information on their behavioral features along with cognitive skills. The algorithms of Pymetrics continually test for and remove ethnic and gender prejudices, which results in more women and minorities being hired by multinational companies. It also permits companies to reach out to people who cannot afford to pay for expensive higher education. As a result of implementing the advanced technologies, the insurance and asset management firms showed massive recruiting gains, and according to their statistics, 77% of candidates enjoyed the process, 54% were more likely to stay with the firm, and 33% increased sales per year [4]. So we can say, all these processes are necessary for maintaining a healthy balance between market requirements and employee satisfaction, which is a must to use human resources properly for the continued economic growth of any country.

The report of the International Labor Organization (ILO) shows that the informal economy of South Asia is estimated at almost 80% to 90% of the total economy [5] and in our country, it's around 75%. To meet the Sustainable Development Goals (SDGs) by 2030, Bangladesh must emphasize the formal sector, and the most genuine method to do so is through digital technology [6]. For so long, BRAC NGO has been working with this informal sector, and to reduce their system loss and improve the trainee success level, we are proposing to use the cognitive science method, which combines human cognitive skills, behavior, and personal qualities analyzed through ML and AI. We will follow a gamified analysis process so that a participant can face every step through a scenario to relate to reality like Pymetrics technology, however focusing on the informal job sector in Bangladesh, where the pre-historical dataset is absent for the participants to predict the career where s/he would be successful in the future.

Chapter 2

Literature Review

2.1 Literature Review

Career choice is the most important decision for a human being ever. In our subconscious mind, cognitive skills help to decide our career pathway around 88% of the time. In the economy, the most productive age of a human being is 15 to 64 years old means these age-ranged people have a direct impact on the economy of a county which is called the Demographic Dividend of a country, as defined by UNFPA [7]. Our Bangladesh is being criticized worldwide many times because we have a huge population. According to statistics of 2019, 67.61% of the population is 15 to 64 years old but not able to utilize these human resources properly [7]. One of the biggest reasons behind that is that most of the people's career concern is in the formal sector and the opportunities are not enough. Instead of that people should concentrate on the informal sectors, and digital innovations as 75% of the economy depends on that. For that, we have to recognize human behavior, and cognitive skills through cognitive science accurately and sharpen up their soft skills as well as industrial skills.

In the developed countries they emphasized more on a human's cognitive skills to hiring for a job position. They are utilizing human resources with the help of modern technologies and one of these is Pymetrics. Now Pymetrics becomes an AI and ML community for research and industry to unlock human potential in a scientific way. In this system, AI and ML technologies are used to extract data patterns, categories, and useful insights from large raw or trained data set to make suitable decisions or predictions. To collect various data from the candidates, it uses in a gamified process. They are variously gamified with different tasks. Any company or organization can design suitably gamified for their own purpose. According to the [8] source, Pymetrics mainly focuses on two traits in the skills testing process: personality and technical traits. In personal traits, it focuses on nine personality attributes through almost 12 to 16 games and those attributes are human's attention, effort, focus, decision-making ability, emotion, learning, fairness, generosity, and risk tolerance. All these attributes help to create a candidate's personal behaviors and capability image through analysis. In the technical traits, it offers four games to test two technical attributes: quantitative reasoning and numerical agility. These kinds of attributes map a candidate's logical capabilities and problem-solving ability. All these traits and attributes have their own importance from different points

of view for any organization. Pymetrics used Neuroscience research to develop the games, which has a set of 12 to 16 neuroscience mini-games that takes less than one hour to 90 cognitive, social, and emotional data point sets of the participants and maps those in categories at a time [3]. Paper [9] highlights that Pymetrics uses the ethical AI de-biased algorithm to maintain fairness in mapping and categorizing the trained data and trying to build race-blind and gender-blind systems.

Paper [9] introduced cognitive science as the new people science for the future working field. Cognitive science is the scientific study of how the human mind works, functions, and behaves vary from situation to situation. This paper's system is used to evaluate the human potential that has massive effects on society in a painful and progressive way. From history, it is visible that the consequence of ineffective employment is painful, and on the other hand, the use of effective employment helps to grow society's economy in a progressive way. That's why it tries to highlight the need for fundamental changes in employment science, the power of cognitive science, and related insights in this sector. They emphasize the importance of soft skills along with hard or industrial skills to maintain gradual growth in the economic sector. To combine soft and hard skills cognitive science is the main key point for today. Besides this cognitive science can also serve as a strategy to address complex social problems. Right now in society, the most crucial progress barrier is the inequality of races and gender. For that, countries cannot optimize human resources properly because they cannot verify one's ability from a neutral point of view. The prominent technical way of justifying one's ability and cognitive skills is to use Pymetrics' gamified process. According to this paper, in this time of the global pandemic to maintain the workforce and accelerate it in a positive way, everyone must follow cognitive science and use it in the industrial area for a better future.

Paper [10] shows the recruitment case study of Unilever hiring people around the world. Being a multinational company Unilever with over 400 brands had to hire 30,000 people in a year and deal with 1.8 million job applications all over the world for that. To filter this huge number of applicants in a short time and make recruitment easy they use AI automation technology. For that, they partnered with two AI tech-leading firms - Pymetrics and HireVue. Pymetrics provides them with the modern game-based assessments platform during the application process and helps to justify candidates' cognitive abilities and behavioral traits through playing a series of game-based assessments. The interesting fact is that the game-based assessment process does not test for specific answers rather it emphasizes how candidates approach problems in specific situations. On the other hand, through HireVue uses video intelligence in the online interview sessions. Through AI they evaluate verbal and non-verbal hints such as facial expressions, eye movements, body gesture-posture, voice tone details, cloth details, etc. It uses proprietary ML algorithms and Natural language processing to analyze each data point and make individual candidates' profiles which help to find the best-fit candidates for the job positions. Now by using the Pymetrics and HireVue technique Unilever successfully narrowed its applicants from 250,000 to 350 who met the real-life recruitment to make the final choice of 800 candidates.

International Business Machines Corporation (IBM) is one of the most prominent companies all over the world that focus more on the cognitive skills in hiring an employee and applicants have to face Pymetrics cognitive game tests in their job assessment test. Through these game tests, IBM actually collects the data on an applicant's behavior, problem-solving skills, focus, the responsibility of a situation, and adaptability power simultaneously [11]. Besides that, it is the first company to hire disabled people and train them so that they could deal with their disabilities and give their best output to the company. The article [12] shows that after collecting that data, IBM applies Watson Natural Language Understanding using deep learning in their system to extract the real meaning and metadata from raw text data. Through this, they can easily extract the emotion, the sentiment of the writer, and keywords of his/her mindset, and classify and categorize those raw data to analyze them in a better way. Because all the small entities of the raw data are the reflection of a person's psychological state. After hiring it also continuously observe the employee's behavior, and cognitive skills through conducting regular bases workshops and training for their employees which develop the employee's skills too. As a consequence of that, now research shows that IBM's annual revenue is increasing at a 5% rate, benefit of USD 6.13 million over three years [12].

Using Pymetrics many more companies, and organizations are now analyzing their previous data to make new business models to enhance their steps and revenue. So it is surviving as a consultant for companies and organizations to solve their issue in a practical way. Bank now considers it as an HR consultant as it is easy to research and map manpower, excellence, cognitive skills, business solutions, and compare industrial issues. That's why in 2017 in India Axis bank launched its first accelerator program with the help of Pymetrics to incorporate solutions through startup planning efficiency, loan, and tax planning [13]. This will help the lender to prove the conceptual solutions that are now intended to be put into practice in other banks too.

From [10] paper's discussion, it is visible that by implementing Pymetrics both job candidates and the organizations are having benefits simultaneously. It helps to increase the engagement of the candidates by taking exams in gamified form rather than the old age job assessment process. Research shows that some organizations experience over 50% of candidate dropout rates in the old-age employee hiring process. In multinational companies, they hire employees all over the world so for candidates to have a suitable time of giving an assessment or interview becomes harder but by Pymetrics game-based test anyone can give their test according to his/ her preferable time. Besides that, the candidates can have instant results which helps them to understand their qualifications. It is also considered a fair system for the candidates because it acts as an unbiased system by using the ethical AI-enabled de-biased algorithm built-in. As a result, in the job market the percentage of disadvantaged people, and minorities is increasing gradually. From the very beginning, its aim is to build a race-blind and gender-blind system which helps to hold people's faith in the organization effectively [9]. In the organizational sectors, Pymetrics helps to have quick evaluations of the huge number of applicants and find the best-fit candidates for the particular job place. Game-based assessments help to predict their employees' job performance so that they can have a clear view of which employee is better suited for a particular project assignment in the future. It also helps to

map future projects as a consultant which saves time and capital and also suggests taking risks based on scientific research data facts.

Like other systems, this digital skill analysis process also has disadvantages too. As it is totally a technology-based process it is hard to maintain and not everyone can execute this as they are not familiar with it. Also, maintenance cost is higher than the age-old traditional way. Additionally, till now all these Pymetrics technology implementations are done to the high profile job and formal economy areas to achieve higher accuracy but they did pay any attention to the huge informal economy at all, where more than 60% of the world's employed population are in the informal economy sector according to ILO [14].

The global COVID-19 pandemic has triggered a significant re-examination of how human psychology research may be conducted safely and successfully in a new era of digital working and physical separation. This reexamination is being done as a direct result of the global COVID-19 epidemic. Recently, the usage of online, web-based testing has come to the forefront as a potentially beneficial tool for the speedy mass gathering of cognitive data that does not involve a direct connection with a human person. This method may be used to collect cognitive data on a large scale. However, there has been an ongoing debate over the dependability of the data that was collected from research conducted via the internet. This study addresses the benefits and problems presented by the move toward web-based testing in society, and it demonstrates an urgent need to build a uniform data quality assurance system for online studies. Specifically, this study focuses on the benefits and problems presented by the move toward web-based testing in society.

This approach is extensively utilized despite being based on a procedure that is dependent on individual programmers' talents rather than software engineering principles. This approach supports web app development today. Next-generation web apps are becoming more commercially significant, which is increasing their complexity. As a result, maintaining high-quality standards while keeping the development process under control is critical. The project will likely focus on formalizing the approach utilized to create these apps. Because formalizing the technique is likely to be a major emphasis of the study. In this article, we'll focus on testing Web apps, a key step for achieving quality and dependability. Testing is part of the overall development process, which is progressive and iterative. Having a reference model is essential for testing Web applications. This is shown by examining testing processes. Aspects of Web applications (such as the HTTP protocol) are employed in this work, resulting in a higher level of automation than conventional software. The approach presented in this study includes TestWeb, a prototype online research tool, that is used to help with testing in this study. This was UK research. This tool reverse-engineers the Web application's UML model to develop and run test cases that meet the user's criteria. The UML model describes the application's structure. This technology's practical use will be shown through a real-world case study.

All over the world, several national and international NGOs like UNICEF, World Vision, BRAC, a2i, JAAGO Foundation, and many more are working to ensure food, health, education, technical and vocational training, and career guidance programs

to ensure a secure future for the next generations. Through these skills development programs, BRAC noticed that in the training time candidates cannot cope with the elements, and even after completing all the necessary training people face difficulties to implement those in real life. Research paper [9]F shows that the main reason for this kind of lacking is that a trainee cannot recognize his/her talent, capabilities, and cognitive abilities properly. In the informal sector, BRAC Skill Development Program (SDP) is working for most of the children who are school dropouts and do not have minimum educational qualifications. Cognitive skills-based games for choosing a career will be the research for those disadvantaged minor peoples' betterment. This will be a bridge between connecting peoples' cognitive skills with the opportunities for the future generation. Through this gamified process, new trainees will be able to know their strengths and weaknesses which will help them to choose the training sector accordingly. Besides that, BRAC SDP has a pool of successful candidates in several sectors who are in the leading positions in the community. By participating in this gamified process these leaders will be more clear about the important traits to be successful. Then the system will compare the leader's data set with the new trainee's data set to produce approximate accurate suggestions about the career. This, like the formal sector, this informal sector will be beneficial gradually. There will be a continuous feedback loop to monetize the whole process and improve the system algorithm.

Chapter 3

Working Plan

3.1 Research Objectives

The first success indicator will be the job placement rate for those undergoing the AI-based assessment. If this rate approaches the 81% job placement rate of BRAC's existing skills development program a program that involves resource-intensive training and job matching BRAC will deem it a success. BRAC's monitoring and evaluation system will also track the following key impact indicators, measured versus a control group of non-users: self-reported job satisfaction, users' average salary, percentage of high performers in the pool of candidates using the AI model (based on surveys of peers, mentors, and supervisors), advancement and promotions over time, and other self-expressed indicators of empowerment and wellbeing. BRAC will also seek to measure firms' increased productivity and profitability attributable to hiring people who underwent the AI-based assessment. An increase in any of these metrics will indicate a degree of success; an increase in all would be a clear win.

A series of post-completion surveys will compare the group using the AI model to a similar control group that has used traditional (human) means to identify career tracks. The survey will ask questions like: Is the worker satisfied in their job? Are their earnings higher than the control group? Is the worker a high-performer in the view of their peers, mentors, and supervisors? This data will enable a feedback loop for machine learning, resulting in higher quality recommendations. These results will determine whether the AI model's prediction was "right" according to the following metrics (each compared to a control group):

- Increase in job placement rates
- Self-reported job satisfaction
- Increase in users' average salary
- An increase in the percentage of high performers in the pool of candidates using the AI model
- A progression in the workforce over time, i.e. more frequent advancements and promotions

3.2 Questionnaire

We primarily strive to include cognitive skills in our questionnaire, which comprises Concept Formation, Numeracy, Literacy, Grit Scale, Planning Ability, Focus, Flexibility Scale, and Control Scale skills testing is a method of determining a person's ability in a specific field. The format and focus of the test are defined by the type of skill being evaluated, such as numerical ability vs. teamwork. You can use abstract tasks, surveys, and work samples. Every test should be double-checked to ensure that it accurately assesses the target concept while remaining ethical. Companies gain from skills tests because they help them identify candidates who are likely to succeed in the open position. Having easily available information about applicants' abilities benefits hiring organizations. This could lower the cost of recruiting and increase the desire of employers to hire. Job seekers can demonstrate their abilities and provide feedback through skills testing. Before scaling up a program, implementers should research and test validated nationally recognized skills assessments. The chosen specific trends are the dataset's domain. We created the best-fit assessment list after studying this data and decided to put the trainee's cognitive talents to the test. Concept Formation, Numeracy, Literacy, Grit Scale, Planning Ability, Focus, Flexibility Scale, and Control Scale are the domains here. In a short period, our system will assist a learner to grasp his ability in every assessment element. In this questionnaire we tried to include those types of question where people do not need any prerequisite knowledge but can solve those questions. In our questionnaire we have 40% easy, 30% medium and 30% hard questions for evaluation.

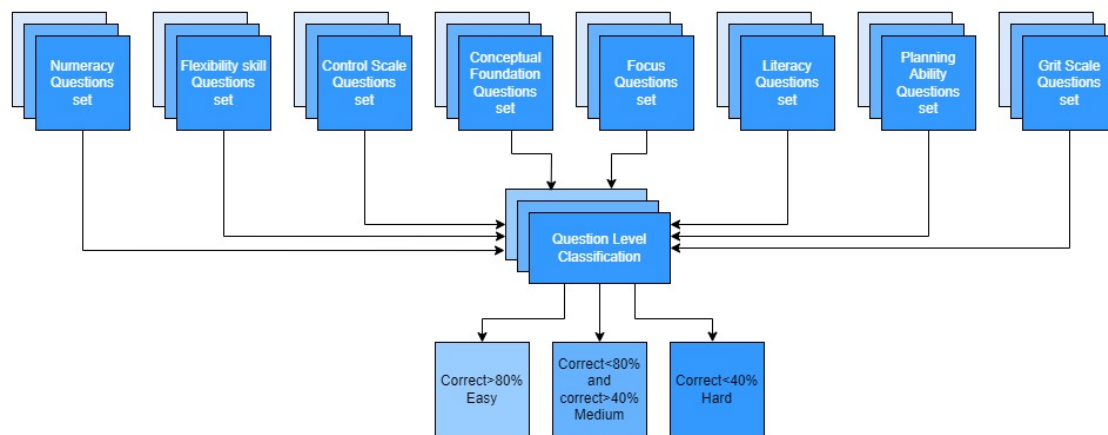


Figure 3.2.1: Questionnaires classification and leveling

Numeracy: This ability helps to recognize and apply math concepts in real-life situations. It involves understanding numbers, algebraic problems, counting, solving number problems, measuring, estimating, sorting, noticing patterns, and so on. Prematurely, the early years of a child's life are a time of rapid learning and development. According to neuroscientists, numbers, patterns, and forms are all recognized by humans from the very beginning of their childhood. They learn by sensing their surroundings and utilizing those concepts in their daily activities. In South Africa, Harambee evaluates candidates' numeracy ability based on their South African Qualifications Authority (SAQA) by comparing different types of math and number problems related to their working trades. This process mostly emphasizes

the ability to manage money management and maintain strategic percentage ratio calculations. Alike this model, we designed a numeracy test to testify to trainees' basic competencies in general math and maintenance issues, and this test follows the concept of Bangladesh National Qualifications Framework (BNQF) levels [15]. In our graphic design trade, assume a candidate is tasked with creating a logo. Now, if he doesn't have any idea about scaling, he won't be able to make the logo appropriately. As a result, his logo will be unappealing. Again, an electrician needs to determine the wire quality and quantity at the same time. Besides that, they also need to calculate the power dissipation of a system. Otherwise, a short circuit may occur, potentially resulting in a catastrophic tragedy. A person conducting tasks in Microsoft Word needs to maintain measurement in various criteria in the documentation so it would be acceptable by the authority and also to make it beautiful in presentation. So because of all these important, a job seeker trainee must have the fundamental numeracy knowledge and also be updated from time to time. For example, in our questionnaire we have Numeracy questions such as "1. একটি ইলেকট্রনিক ট্রেন ঘণ্টায় ১০০ কি.মি. বেগে পূর্ব থেকে পশ্চিমে যাচ্ছে এবং ১০ কি. মি. বেগে বাতাস উত্তর থেকে দক্ষিণে বইছে। ধোঁয়া কোন দিকে যাবে?", " 3. ক * ৫ = ২৫ ক এর মান কত?" ।

Flexibility Skills: Flexibility skill is the thinking ability to focus on the ability to adapt to a new situation, and improvise knowledge in a new sector in general [16]. As we observed, it is a by-born skill of a human and it also measures a person's approach to a goal in the first place without thinking of any options or alternatives. Another similar cognitive skill to this one is adaptability, but both are not the same. For a human, mental flexibility plays an important role when it comes to real-life problem-solving. For that, mental flexibility is considered the pinnacle of a person's excellence. That's why our proposed model emphasizes flexibility over adaptability to testify to the psychological state. When an employee needs to meet several deadlines within the regular work schedule in a short time. At that moment, he should control his anxiety and panic and make a proper road map with his planning ability and maintain it thoroughly to the end. Rahim is a graphic designer. One of the problems with his computer lately is that it works very slowly. So what if his computer works very fast again he can find a potential solution in different ways. This type of problem is related to flexibility skills. Flexibility skills questions, for instance, are included in our questionnaire. "10. Embracing change, maintain composure during adversity. (পরিবর্তনকে মেনে নেয়া, প্রতিকূলতার সময় সুরক্ষা বজায়)", " 17. Arriving on time to an event (নির্ধারিত সময়ে কোন অনুষ্ঠানে পৌঁছানো কে বলে)", " 15. Statement: In a T20 match played between India and Australia, the total runs made by the Indian team were 200. 160 runs out of 200 runs were made by spinners. Conclusion I: 80% of the team consists of spinners Conclusion II: The opening batsmen were spinners" and "14. Each problem consists of three statements. Based on the first two statements, the third statement may be true, false, or uncertain. All the trees in the park are flowering trees. Some of the trees in the park are dogwoods. All the dogwoods in the park are flowering trees. If the first two statements are true, the third statement is _____".

Control Scale: The Control Scale is a self-reported assessment of how candidates respond to novel and perplexing issues. When confronted with new, complex

challenges, it assesses whether candidates behave methodically and thoughtfully or impulsively and haphazardly. This is the method of determining whether or not one's actions have resulted in methodically achieving a goal. To confront a new and difficult challenge, a person must control his impulsivity and illogical thinking and decide on a practical course of action [16]. It also refers to one's own self-control in a reactive situation that does not come naturally or effortlessly. Getting out of the automatic mode of behavior, on the other hand, necessitates work, which is defined as the usage of available energy resources. As a result, practicing self-control might be exhausting, leading to an unintentional reversion to automatic behavior control. The automated control mode is neither repugnant nor damaging in and of itself. Many of our daily actions are delegated to our "autopilot", which relies on previously established behavioral patterns, habits, and reflexes [17]. As a consequence, every employee should have good control of their attitude, according to the article Self-Control Scale AS-36: Construction and Validation Study to maintain a sound environment in the working sector. The Control Scale, for example, can be found in our questionnaire " 4. Multiples on time ", " 7. What's your AQ? ", " 12. Confidence individual person (স্বতন্ত্র ব্যক্তি হিসাবে আপনার নিজস্ব যোগ্যতার উপর আত্মবিশ্বাসকে বলে ___ ", " 13. What is obedience? (আনুগত্য কি?) " .

Focus: Focus is the most important skill and is the driver of excellence in the human brain. We can refer to it as a technique to distinguish relevant from irrelevant information in potentially confusing environments. It relates to whether a person can act deliberately according to the specifications, rules, and regulations that are very important to maintain a particular systematic approach also. It is proven that a focused person tries to maintain a to-do list along with a daily routine and reminders on the calendar so that those help to keep going along a systematic path and do not mismatch any event or task schedule. With this, one can be punctual in real life. In the job market, this adds a plus point for candidates. Focus questions, for example, can be found in our questionnaire " 14. What disruption? (শ্রেণি বিয়ের উদাহরণ", " 15. You should make you. (কোনও ব্যক্তির সাথে কথা বলার সময় তাঁর চোখের দিকে তাকানো উচিত।)", " 19. Respect and cooperation are important in a job and school-related settings. (চাকরী এবং বিদ্যালয় সম্পর্কিত কর্মে সম্মান এবং সহযোগিতা গুরুত্বপূর্ণ।)", "6. নিচের কোন শব্দটি ভিন্ন ? ", "9. Question 3: Which of the shapes given would complete the sequence? " .

In our process, like Harambee's model, a graphics designer candidate needs to answer a bunch of questions related to the color name, uses, and color senses in the scenarios. For example, a question can be asked to choose the color that shows: "Black" and here this problem's correct answer is blue. Candidates are graded based on how long it takes them to answer the question and whether or not they answer it correctly.

Literacy: Literacy is known as the capability to communicate effectively and make sense of the world through reading, writing, speaking, and listening. It enhances child and family health and nutrition, reduces poverty, and broadens life opportunities, all of which contribute to long-term development. A person's ability to read is impaired at every stage of his or her life. It enhances child and family health and nutrition, reduces poverty, and broadens life opportunities, all of which contribute to long-term development. The capacity to remember and interpret literal mate-

rial, as well as recognize non-literal language usage, is assessed in the listening skills section. The World Bank estimates that Bangladesh has a literacy rate of 92.2 percent. However, many young people in our society continue to be denied their basic rights. Even if they obtain an education, they are unable to obtain the employment that they are entitled to. As a result, people continually change jobs in order to make more money. When we examine the causes behind this, we can observe that the youngest people have a strong grasp of scriptural information but a weak grasp of practical knowledge. It's not as if a person understands that he will never be able to adjust to any system because he lacks actual experience. For that, these young people are woefully underprepared. That is why our system will be assessed for literacy so that they can recognize communication gaps. For example, in order to use Microsoft Word, you must be fluent in English; if you are not, you will be unable to use the software effectively. As a result, he will be unable to perform his duties effectively. Again, many times you must study the circuit documentation to understand the circuit's needs. If you don't have the capacity to read all of these materials, you won't know what you'll need for the circuit. Not just scriptural knowledge but also practical understanding is required to comprehend these issues. Literacy questions, for example, are included in our questionnaire. " 1. When your revenue stream _____, you have to look for a new one ", " 18. Their failure was _____ with the way they've been losing for months now ", " 13. Melt: Liquid:: Freeze: ?? " and " 16. Conclusions: [A] Angle is not a pen [B] Angle is a pen ".

Conceptual Foundations: Conceptual foundations test the ability to think abstractly, reason quickly, and solve problems independently of any formerly achieved knowledge. Come across someone who can take newer information and develop ideas, theories, and even problem-solving strategies based on the data built over time with knowledge and experience. The Woodcock-Johnson Test of Cognitive Abilities, Raven's Progressive Matrices, and the Wechsler Intelligence Scale for Children are various ways fluid intelligence may be tested. We can solve problems without prior knowledge and utilize more remarkable conceptual thinking to break through obstacles, riddles, and mental barriers to discover the best answer by using fluid intelligence. One thing is that, suppose thinking about the puzzles or problem-solving strategies, the primary key is fluid intelligence, because it will not be necessary to come up with the optimal solution before, so that level of abstract thinking is mandatory to find the optimal solution. The article, Reducing labor market information frictions with skill certificates: Evidence from South Africa shows that a conceptual foundation plays a great role in the case of introducing new concepts or policies into the system [18]. Where the authority can make a clear concept in the first place to the targeted audiences the acceptability rate becomes higher than others. As our country is a developing digital country, we need to make the fundamental day-to-day technique idea clear in the trainee's mind so that he can adapt to the new system. That's why, nowadays, the authority party emphasizes the importance of fundamental knowledge in their problem-solving. Suppose a person is writing a document in Microsoft Word but only knows how to write, copy, and paste and does not know how to update Microsoft Word in the new version or change the default setting. As an output, if suddenly the system is required to be updated to execute because of having no fundamental knowledge of this software, he or she will not be able to perform any task. The same thing goes for all other systems. For this, our

model will test a trainee’s background concept and help to enrich the knowledge. Conceptual questions, for example, are included in our questionnaire “ 16. Body language is a dead giveaway for how you really feel about someone. (Meaning of Dead giveaway: A clue, detail, action, etc. that obviously reveals something else) (অঙ্গ-ভঙ্গিকে এমন একটি ”Dead Giveaway” হিসাবে গন্য করা হয় যা প্রকাশ করে আপনি অন্যের বিষয়ে কি বোধ করছেন) ”, “ 7. Which figure logically belongs on the spot of the question mark? ”, “ 8. Question 2: Which of the boxes comes next in the sequence? ”, “ 10. Question 4: Which of the boxes comes next in the sequence? ”, “ 11. Question 5: Which of the boxes comes next in the sequence? ”.

Planning Ability: In every step of life, from waking up in the morning to going to sleep at night or drinking tea, going to work, any kind of formal or informal task means a whole day routine or future decision. Each point needs our planning ability. We can improve our lives by making suitable plans and executing them. In our system, as we are working in the informal economic sector, a trainee will face new challenging situations and give solutions according to his or her planning ability, like Harambee’s system assessment. Harambee gives its candidates a computer game called “Hit15” and the participant has to collect 15 stones into a basket to win this round of the game [19]. On the other hand, in our part of this section of the analysis, a participant will be given a problematic situation, such as an IT technician will be asked to make a budget gaming desktop pc or suggest a gaming laptop having the pre-requisite qualities, or an electrician will be given a building design and have to suggest a proper design of the electric circuit board position, number of electric devices position estimation, the power requirement and distribution of the building, etc. All of the suggestions or answers will be considered as a preferable solution, and we will level these in range to compare them properly with the ML model to find a suitable candidate for this system. We have questions on Planning Ability in our questionnaire, such as “ 2. আপনার সামনে কেউ পানিতে পড়ে গেল, আপনি সাত্তার জানেন না তখন আপনার পরবর্তি পদক্ষেপ কি হবে?”, “ 6. Working together collaboratively with a group to set goals, make decisions, solve problems, and put ideas into action. (লক্ষ্যনির্ধারণ, সিদ্ধান্তগ্রহণ, সমস্যা সমাধান এবং ধারণা গুলিকে কার্যকর করার লক্ষ্যে একটি গোষ্ঠীর সাথে একযোগে কাজ করাকে কি বলে)”, “ 11. Making _____ability”.

Grit Scale: The Grit Scale is a self-reported assessment of applicants’ determination while working on complex challenges. It assesses a person’s willingness to work on strenuous activities until they are completed, as well as their ability to stick to long-term objectives. Even when other qualities like intellect are considered, the grit scale has been found to successfully predict successfully. Angela Duckworth’s Ted-Talk introduces the notion of grit and its predictive potential in the real world, and her book ”GRIT The Power of Passion and Perseverance” delves more into how grit nurtures life success [20]. Candidates were given statements to assess how strongly they agreed or disagreed with them, with options ranging from ”strongly agree” to ”strongly disagree.” Consider the case below: An IT support technician works in a lab. In the lab, the members of a group will help everyone. Give your opinion on any option, either ”strongly agree” or ”strongly disagree.” This type of problem is solved by the grit scale.

Through all these assessment tests, we can help a person overcome the limitation of job search and hiring this two-sided information about work-seekers skills through our system model. As an outcome, the obstacles get reduced for the candidate to get a job, not switch jobs frequently, and also for the hiring authority to hire new candidates in a short time. Grit Scale questions such as these appear in our survey “ 2. An opinion or general feeling about something (কোনও বিষয়ে মতামত বা সাধারণ অনুভূতি) ”, “ 3. When one of my colleagues is in a serious problem and I am putting myself into his/her situation and trying to understand it on a deeper level and trying to help her. What am I showing? ”, “ 5. The drive is accomplished (শুধুমাত্র আত্ম-সম্প্রতিষ্ঠার জন্য নিঃস্বার্থে, কোন কাজ করতে অনুপ্রাণিত করে কোনটি?) ”, “ 8. Discipline is (নিয়মানুবর্তিতা হল)”,”9. Having a positive attitude can improve your chances of getting hired. (ইতিবাচক দৃষ্টি ভঙ্গি থাকলে কাজে নিযুক্ত হওয়ার সম্ভাবনা বেড়ে যায়) ”, “ 20. Responsibility for an employe (কোনও কর্মচারীর দায়বদ্ধতা অর্থ) ”.

3.3 Model Analysis

3.3.1 Logistic Regression

Like other regression analyses, logistic regression is a predictive analysis. Logistic regression is used to explain the data and understand the relationship between one dependent binary variable and one or more nominal, variables, sequence, or ratio-level independent variables. When the variable is categorical, logistic regression is the best regression method to use.

So, the result of linear regression is continuous and can take any value. In the case of logistic regression, however, the predicted result is discrete and constrained to a set of values. Here are the equations:

$$Y = \sigma(z) = \frac{1}{(1 + e^{-z})}$$

$$L(Y, y) = (y \log Y + (1 - y) \log(1 - Y))$$

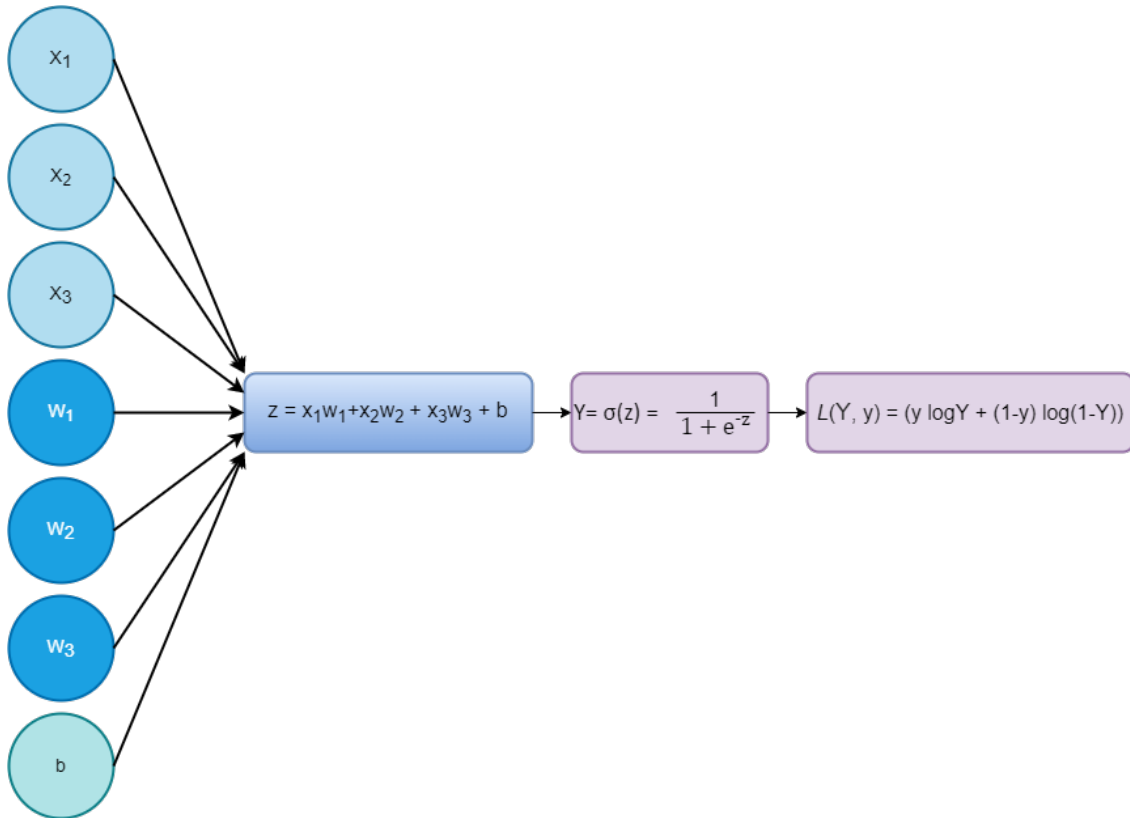


Figure 3.3.1: Logistic Regression

A regression model is a linear regression, while a classification model is a logistic regression. Logistic regression is used in a variety of settings and fields in the real world. Nowadays, for these huge usability scopes in health care, the financial industry, marketing, and other sectors to predict disease, fraud transactions, the targeted audience responds, which is making our day-to-day life easier.

3.3.2 Decision Tree

Creating a succession of splits is the most crucial component of any decision tree. The purpose of the split rule is to divide the data into two purest groups. It indicates that as many elements of a group as feasible should be of the same type.

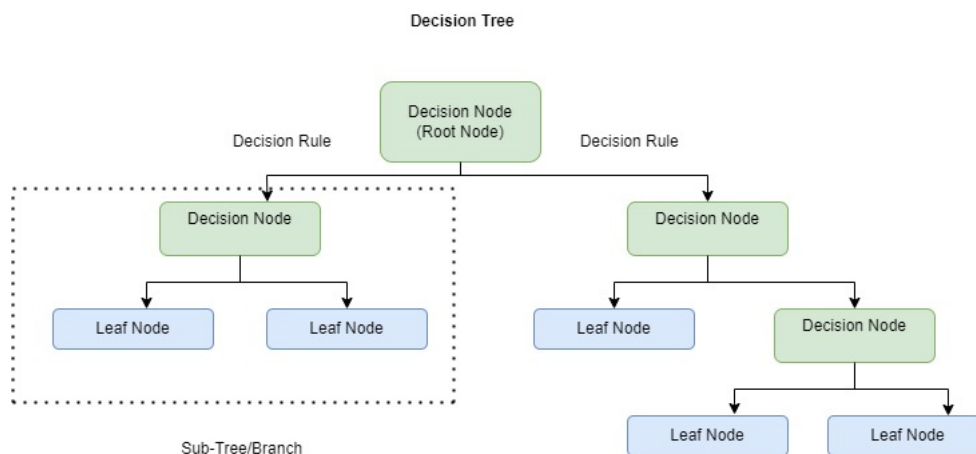


Figure 3.3.2: Decision Tree

For example, if there are 15 green balls in one basket, it is a pure group. In contrast, if a bucket contains three green balls, five blue balls, and two red balls, the group is impure. The entropy value ranges from 0 to 1. The numbers 0 and 1 represent the whole pure group and the maximal impure group, respectively. Purity improves with lower entropy. Calculating the data's entropy is one approach to determine its purity, and the formula is $\text{Entropy} = \sum_i -p_i \log_2 p_i$.

3.3.3 Random Forest Classifier

Random forest is a versatile, convenient machine learning process that utilizes excellent results even when high energy are not adjusted. It can deal with issues like classification and regression, which are prevalent in history's machine learning systems. Because of its simplicity and adaptability, it is among the most widely used techniques.

A supervised learning algorithm known as "bagging" creates a combination of decision trees. The bagging approach's core principle is that combining multiple learning models improves the end result. This method creates a large amount of decision trees, which are then combined to produce a more reliable and accurate results. A random forest with many feature trees is depicted in Figure 3.3.3.

Random Forest adds more unpredictability to the model as the trees grow taller. As an outcome, there is far more significant variation, resulting in a more accurate model. Instead of searching for the most significant characteristic when splitting a node, it searches for the best feature from an arbitrary number of traits.

We can make trees much more random in an arbitrary forest through using random threshold values for each characteristic. As a result, when trying to split a node, the technique only wants to consider a random set of features. It determines this score for each character trait after training and weights the results so that the total relevance equals one.

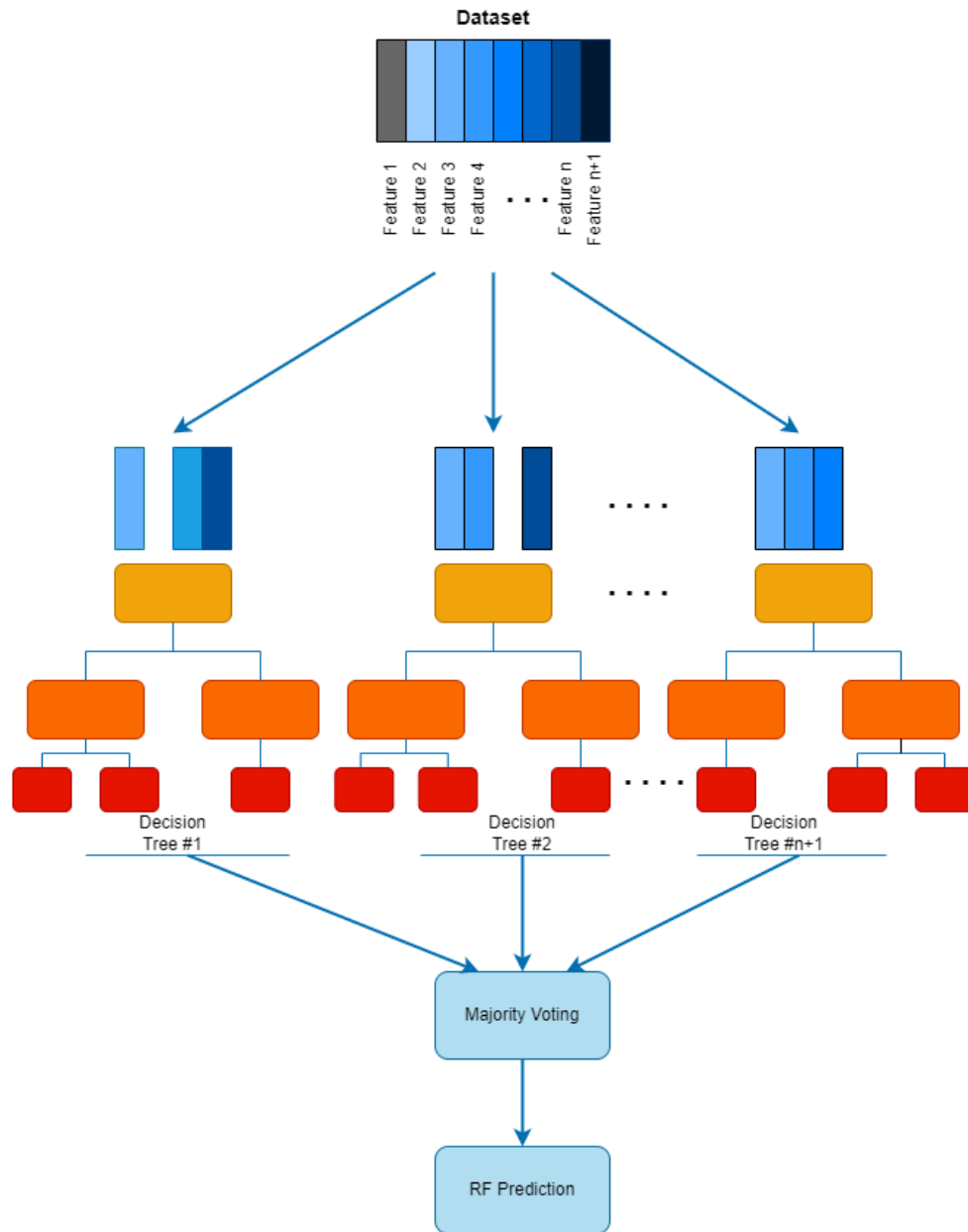


Figure 3.3.3: Random Forest Classifier

3.3.4 Support Vector Classifier (SVC)

A Linear SVC (Support Vector Classifier) is designed to fit the data you supply, providing a "best fit" hyperplane that divides or categorizes your data. SVC approach uses a linear kernel function to classify data and works well with huge datasets. When compared to the SVC model, the Linear SVC contains more parameters, such as penalty normalization ('L1' or 'L2') and loss function. Because linear SVC is based on the kernel linear technique, it cannot be modified. In our work after we've obtained the hyperplane and input some characteristics into our classifier to get the best fit "predicted" class.

3.3.5 Multilayer Perceptron Classification (MLP)

MLP is an ANN that is completely connected and feedforward (ANN). Any feedforward artificial neural network (ANN) or network with multiple layers of perceptrons can be referred to as MLP (with threshold activation). Multilayer perceptrons with a single hidden layer are the building blocks of "vanilla" neural networks. Input, hidden, and output nodes are all components of an MLP. Every node in the network is a nonlinear neuron, with the exception of the input nodes. Backpropagation is the method of learning utilized by MLPs. In contrast to a linear perceptron, a multilayer perceptron (MLP) contains several layers and a non-linear structure. It divides the data that is nonlinear.

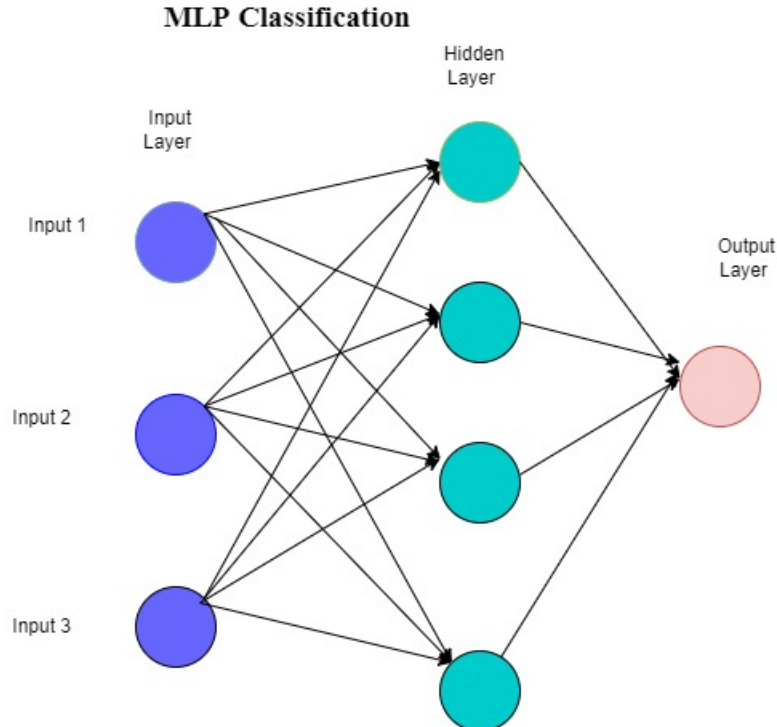


Figure 3.3.4: Multilayer Perceptron Classification

Perceptron Weights Adjustment $\Delta w = \eta \times d \times x$
 $d = \text{Predicted output} - \text{Desired output}$

η = Learning rate, usually less than 1

x = Input database

3.3.6 Extreme Gradient Boosting (XGB)

Extreme gradient boosting, also known as XGB, is a method that performs very similarly to gradient boosting (GB), however it has been demonstrated to be a more effective and accurate algorithm than gradient boosting as a result of a few modifications made to the gradient boosting algorithm.

The goal function for XGB that needs to be reduced with regard to each tree function is not merely the loss function. This is one of the changes that have been made. The regularization function is incorporated into the objective function as well. Each tree function receives a penalty from the regularization function that is calculated based on the number of leaves and the weight that is assigned to each leaf. In addition to that, it possesses a regularization parameter that enables the user to adjust the level of influence that the regularization function ought to have on the overall objective function. Because the gradient boosting approach was used to fit each consecutive tree function to the residuals of the model, it gave the impression that each tree function was learning to fix all of the errors that were being made by the most recent model. Because of this, the likelihood of overfitting was significantly raised. By assigning a penalty to tree functions that are excessively complicated, the regularization function in XGB works to cut down on the likelihood of the model being overfit.

$$L^{(t)} = \sum_{i=1}^n l(y_i, Y^{(t-1)} + f_t(x_i)) + \Omega(f_t)$$

y_i = Real value (level known from the training data set)

Can be seen as $f(x + \Delta x)$ where $x = Y^{(t-1)}$

3.4 Workflow Diagram

We create questionnaires in 8 different categories which are Concept Formation, Numeracy, Literacy, Grit Scale, Planning Ability, Focus, Flexibility Scale, and Control Scale based on psychological study for collecting data. After that we assign specific domains to the questions. For collecting data of our work, we took help from former students who took training from BRAC in Graphic Design, IT Support Technician, MS Office, and Electrical Installation & Maintenance. For the earliest experiment, two different data sets were used for our experiment. In the Adaptability_Responses data set we have 23 features and In the Logical_Ability data set we have 20 features in total. Then we preprocessed our datasets. We do not often come across clean and prepared data. Also, before completing any data-related activity, the data must be cleaned and formatted. As a result, we implemented a data preprocessing procedure.

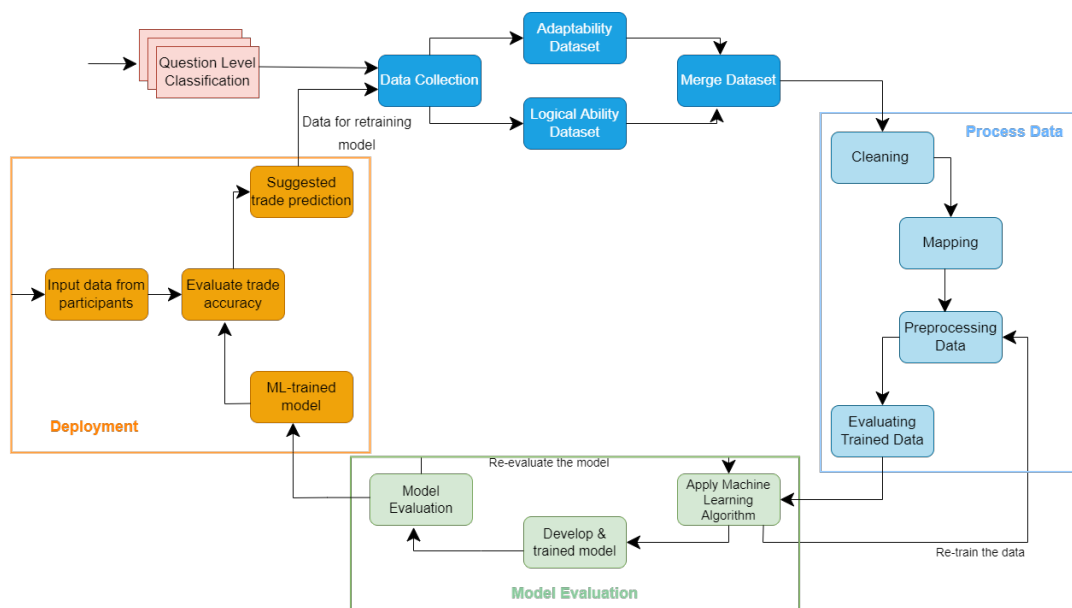


Figure 3.4.1: Workflow diagram

We use SVC, Logistic Regression, Decision Tree, Random Forest Classifier, and MLP Classifier. All of these are used to solve classification problems and predict the trade. After finalizing the model, we will retrain our model and re-evaluate our predictions.

Chapter 4

Data Pre-Processing

4.1 Data Collection

For data collection of our work, we took help from former students who took training from BRAC in Graphic Design, IT Support Technician, MS Office, and Electrical Installation Maintenance. For the earliest experiment, two different data sets were used for our experiment. In the Adaptability_Responses data set we have 23 features in total, namely- ID, Score, An opinion, or general feeling about something. Moreover, the features have categorical values in figure 4.1.1 and 4.1.2.

Serial	Column Name	Count	Data Type
0	ID	573	int64
1	Score	573	object
2	2. An opinion or general feeling about something (কোনও বিষয়ে মতামত বা সাধারণ অনুভূতি)	573	object
3	3. When one of my colleagues is in a serious problem and I am putting myself into his/her situation and trying to understand it on a deeper level and trying to help her. What am I showing?	573	object
4	4. Multiples -----the client introduces prevented us from completing this project on time	573	object
5	5. The drive to do something simply for the reward of feeling good and satisfied once it is accomplished (শুধুমাত্র আত্ম-সন্তুষ্টির জন্য নিঃস্বার্থে, কোন কাজ করতে অনুপ্রাণিত করে কোনটি?)	573	object
6	6. Working together collaboratively with a group to set goals, make decisions, solve problems, and put ideas into action. (লক্ষ্যনির্ধারণ, সিদ্ধান্তগ্রহণ, সমস্যা সমাধান এবং ধারণা গুলিকে কার্যকর করার লক্ষ্যে একটি গোষ্ঠীর সাথে একযোগে কাজ করাকে কি বলে)	573	object
7	7. What's the best way to-----adaptability and improve your AQ?	573	object

Figure 4.1.1: Adaptability responses table part one

8	8. Discipline is....(নিয়মানুবর্তিতা হল _____)	573	object
9	9. Having a positive attitude can improve your chances of getting hired. (ইতিবাচক দৃষ্টি ভঙ্গি থাকলে কাজে নিযুক্ত হওয়ার সম্ভাবনা বেড়ে যায়)	573	object
10	10. Embracing change, maintain composure during adversity.(পরিবর্তনকে মেনে নেয়া, প্রতিকূলতার সময় সুরক্ষা বজায় রাখাকে বলে _____)	573	object
11	11. Making -----a priority is one of the ways to improve adaptability	573	object
12	12. Confidence in your own merit as an individual person (স্বতন্ত্র ব্যক্তি হিসাবে আপনার নিজস্ব যোগ্যতার উপর আত্মবিশ্বাসকে বলে _____)	573	object
13	13. What is obedience? (আনুগত্য কি?)	573	object
14	14. What is an example of a class disruption? (শ্রেণি বিঘ্নের উদাহরণ কী?)	573	object
15	15. You should make eye contact with a person when they are speaking to you. (কোনও ব্যক্তির সাথে কথা বলার সময় তাঁর চোখের দিকে তাকানো উচিত।)	573	object
16	16. Body language is a dead giveaway for how you really feel about someone. (Meaning of Dead giveaway: A clue, detail, action, etc. that obviously reveals something else) (অঙ্গ-ভঙ্গিকে এমন একটি “Dead Giveaway” হিসাবে গন্য করা হয় যা প্রকাশ করে আপনি অন্যের বিষয়ে কি বোধ করছেন) (“Dead Giveaway” এর অর্থঃ একটি ইঙ্গিত, বিশদ, প্রতিক্রিয়া ইত্যাদি যা স্পষ্টতই অন্য কিছু প্রকাশ করে।)	573	object
17	17. Arriving on time to an event (নির্ধারিত সময়ে কোন অনুষ্ঠানে পৌঁছানো কে বলে _____)	573	object
18	18. Their failure was -----with the way they’ve been losing for months now	573	object
19	19. Respect and cooperation are important in job and school-related settings. (চাকরী এবং বিদ্যালয় সম্পর্কিত কর্মে সম্মান এবং সহযোগিতা গুরুত্বপূর্ণ।)	573	object
20	20. Responsibility for an employee means (কোনও কর্মচারীর দায়বদ্ধতা অর্থ)	573	object
21	Unnamed: 21	573	float
22	Trade	573	object

Figure 4.1.2: Adaptability responses table part two

In the Logical_Ability data set we have 20 features in total, namely- ID, Score, Find the missing series 3, 12, 27, 48, 75, 108, ? , Melt : Liquid:: Freeze: ??, Conclusions: [A] Angle is not a pen [B] Angle is a pen, Trade, etc. and among all, the target feature here is ‘ Trade’ which gives the ultimate result of a person going for a specific trade(‘GD’, ‘IST’, ‘MS Office’, ‘EIM’). Moreover, the features have categorical values in figure 4.1.3.

Serial	Column Name	Count	Data Type
0	ID	573	int64
1	Score	573	object
2	একটি ইলেকট্রনিক ট্রেন ঘণ্টায় ১০০ কি.মি. বেগে পূর্ব থেকে পশ্চিমে যাচ্ছে এবং ১০ কি. মি. বেগে বাতাস উত্তর থেকে দক্ষিণে বইছে। ধোঁয়া কোন দিকে যাবে?	573	object
3	আপনার সামনে কেউ পানিতে পড়ে গেল, আপনি সাতার জানেন না তখন আপনার পরবর্তি পদক্ষেপ কি হবে?	573	object
4	ক *৫ = ২৫ ক এর মান কত?, $((২^২)^৩)$ নিচের কোনটি সঠিক?	573	object
5	২০, ২২, ২৫, ২৯, -- পরের সংখ্যাটি কত?	573	object
6	নিচের কোন শব্দটি ভিন্ন?	573	object
7	Which figure logically belongs on the spot of the question mark?	573	object
8	Question 2: Which of the boxes comes next in the sequence?	573	object
9	Question 3: Which of the shapes given would complete the sequence?	573	object
10	Question 4: Which of the boxes comes next in the sequence?	573	object
11	Question 5: Which of the boxes comes next in the sequence?	573	object
12	Find the missing series 3, 12, 27, 48, 75, 108, ?	573	object
13	Melt : Liquid :: Freeze : ??	573	object
14	Each problem consists of three statements. Based on the first two statements, the third statement may be true, false, or uncertain. "All the trees in the park are flowering trees. Some of the trees in the park are dogwoods. All dogwoods in the park are flowering trees. If the first two statements are true, the third statement is-----"	573	object
15	Statement: In a T20 match played between India and Australia, the total runs made by the Indian team were 200. 160 runs out of 200 runs were made by spinners. Conclusion I: 80% of the team consists of spinners Conclusion II: The opening batsmen were spinners.	573	object
16	Statement: In a T20 match played between India and Australia, the total runs made by the Indian team were 200. 160 runs out of 200 runs were made by spinners. Conclusion I: 80% of the team consists of spinners Conclusion II: The opening batsmen were spinners.	573	object
17	Trade	573	object
18	Unnamed: 19	573	float64

Figure 4.1.3: Logical ability table.

4.2 Data Pre-Processing

We added that the SVC, Logistic Regression, Decision Tree, Random Forest Classifier, and MLP Classifier are the supervised machine learning techniques in our system. All are used to tackle categorization issues (sorting data into categories). As a result, firstly, we make a psychological questionnaire and level the questions with specific domains. These questions will be well arranged on a website with levels. In that case, from the user interface, the system takes input after that, transfers the Bangla questions' answers into English numeric, and then preprocesses the dataset and extracts features. These things happen in the preprocessing of the dataset.

The datasets are saved as CSV format files, and the CSV files are read into a pandas data frame for preprocessing. The datasets were not filled with null or zero values initially. So we did not need to replace or remove them.

ID	Score	Null Count
2.	An opinion or general feeling about something (কোনও বিষয়ে মতামত বা সম্মান অনুভূতি)	0
3.	When one of my colleagues is in a serious problem and I am putting myself into his/her situation and trying to understand it on a deeper level and trying to help her. What am I showing?	0
4.	Multiplies -----the client introduces prevented us from completing this project on time	0
5.	The drive to do something simply for the reward of feeling good and satisfied once it is accomplished (শুধুমাত্র আনন্দ-সন্তুষ্টির জন্য নিয়মিত, কোন কাজ করতে অনুপ্রাণিত করে কোনটা)	0
6.	Working together collaboratively with a group to set goals, make decisions, solve problems, and put ideas into action.	0
7.	What's the best way to----adaptability and improve your AQ?	0
8.	Discipline is.....(নিয়মানুবর্তিতা স্বা)	0
9.	Having a positive attitude can improve your chances of getting hired. (ইতিবাচক দৃষ্টি থাকলে কাজে নিযুক্ত হওয়ার সম্ভাবনা বেড়ে যায়)	0
10.	Embracing change, maintain composure during adversity.(পরিবর্তনকে মেনে নেয়া, প্রতিভুলতার সময় সুরক্ষা বজায় রাখতে বলে)	0
11.	Making -----a priority is one of the ways to improve adaptability	0
12.	Confidence in your own merit as an individual person (কতই ব্যক্তি হিসাবে আপনার নিজস্ব যোগ্যতার উপর আত্মবিশ্বাসকে বলে)	0
13.	What is obedience? (অনুগত্য কি?)	0
14.	What is an example of a class disruption? (শ্রেণি বিঘ্নের উদাহরণ কী?)	0
15.	You should make eye contact with a person when they are speaking to you.(কোনও ব্যক্তির সাথে কথা বলার সময় তাঁর চোখের দিকে তাকানো উচিত)	0
16.	Body language is a dead giveaway for how you really feel about someone.	0
17.	Arriving on time to an event (নির্ধারিত সময়ে কোন অনুষ্ঠানে পৌঁছানো কে বলে)	0
18.	Their failure was -----with the way they've been losing for months now	0
19.	Respect and cooperation are important in job and school related settings. (চাকরী এবং বিদ্যালয় সম্পর্কিত কর্মে সম্মান এবং সহযোগিতা গুরুত্বপূর্ণ)	0
20.	Responsibility for an employee means (কোনও কর্মচারীর দায়বদ্ধতা অর্থ)	573
Unnamed: 21	Trade	0
dtype:	int64	

Figure 4.2.1: Adaptability Responses dataset's columns with number of null values.

ID	Score	Null Count
একটি ইলেকট্রনিক ট্রেন ঘণ্টায় ১০০ কি.মি. বেগে পূর্ব থেকে পশ্চিমে যাচ্ছে এবং ১০ কি. মি. বেগে বাতাস উত্তর থেকে দক্ষিণে বইছে। ধোঁয়া কোন দিকে যাবে?	0	0
আপনার সামনে কেউ পানিতে পড়ে গেল, আপনি সাজার জামেন না তখন আপনার পরবর্তী পদক্ষেপ কি হবে?	0	0
ক *৫ = ২৫ ক এর মান কত ?	0	0
((২^২)^৩) নিচের কোনটি সঠিক ?	0	0
২০, ২২,২৫,২৯, -- পরের সংখ্যাটি কত?	0	0
নিচের কোন শব্দটি ভিন্ন ?	0	0
Which figure logically belongs on the spot of the question mark?	0	0
Question 2: Which of the boxes comes next in the sequence?	0	0
Question 3: Which of the shapes given would complete the sequence?	0	0
Question 4: Which of the boxes comes next in the sequence?	0	0
Question 5: Which of the boxes comes next in the sequence?	0	0
Find the missing series 3, 12, 27, 48, 75, 108, ?	0	0
Melt : Liquid :: Freeze : ??	0	0
Each problem consists of three statements. Based on the first two statements, the third statement may be true, false, or uncertain.	0	0
Statement: In a T20 match played between India and Australia,.....The opening batsmen were spinners	0	0
Conclusions: [A] Angle is not a pen [B] Angle is a pen	0	0
Trade	0	0
Unnamed: 19		572
dtype:	int64	

Figure 4.2.2: Logical Ability dataset's columns with number of null values.

4.3 Cleaning

We had to drop some columns from the datasets that contained null information. Moreover, features with too many missing values were dropped. That is why we dropped the ‘Unnamed: 21’ column from the Adaptability_Responses dataset and the ‘Unnamed: 19’ column from the Logical_Ability dataset.

4.4 Mapping

The Python map() method returns a map object after applying a function to all components of an iterable. We can loop through the elements of a Python map object since it is an iterator. We can also use their factory functions to transform map objects into sequence objects like lists and tuples.

As in both datasets, we have values as objects and we need to map them so we can represent the rows in numerical values.

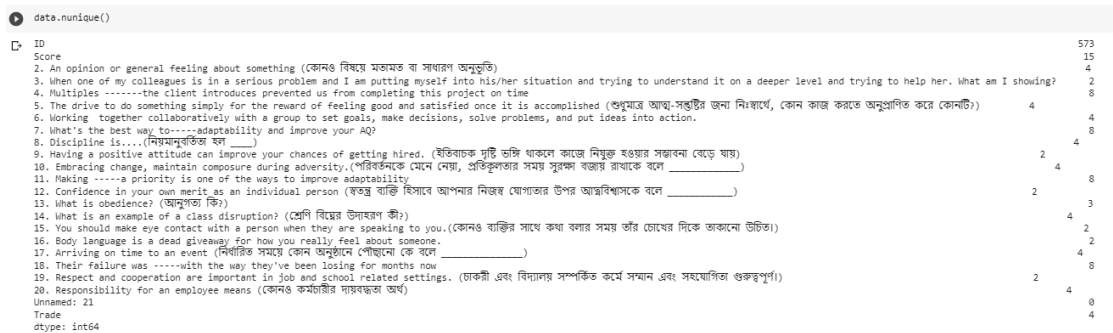


Figure 4.4.1: Adaptability Responses dataset’s unique values.

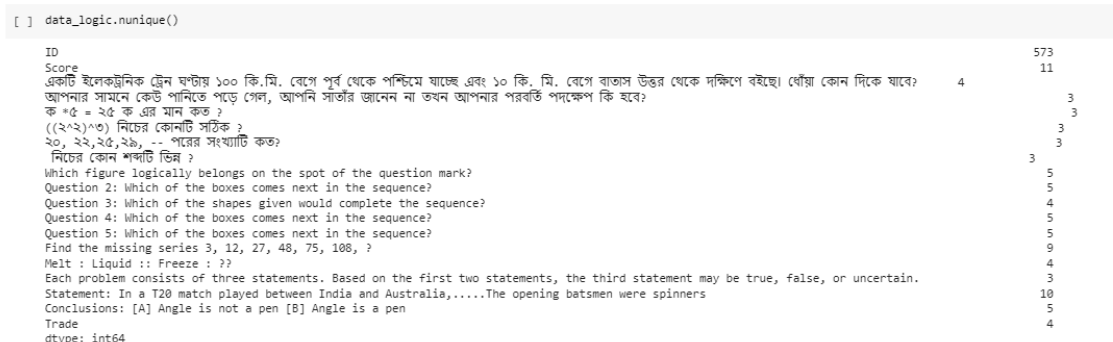


Figure 4.4.2: Logical Ability dataset’s unique values.

We can see how many unique values each column has in the diagram above. We mapped them properly because we want to implement ML algorithms and the machine does not understand anything except numbers.

4.5 Feature Engineering

Before applying algorithms, we merged the two datasets based on “ID”. This column is unique in both of the datasets. After merging the datasets, we had 37 columns. We do not need the ID column, so we dropped it.

To assess the impact of each feature on the target value, we used the correlation discovery technique to determine relationships between each feature in both datasets. Features with little effect on class prediction were observed so that they might be given less weight in the final model. The `corr()` function was used to extract the correlation matrix.

Listing: Correlation Function and Correlation Heatmap
`sns.heatmap(data.corr())`

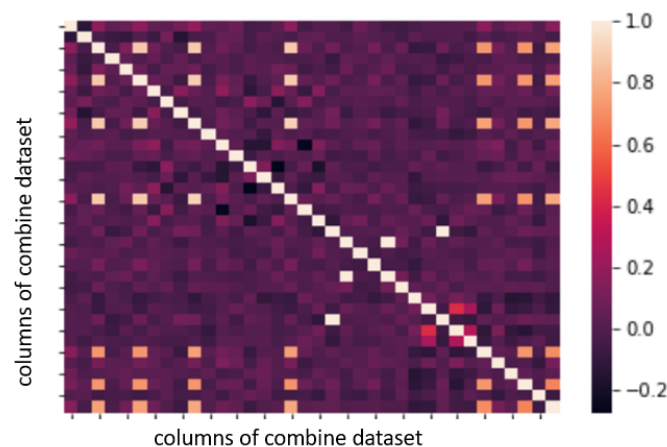


Figure 4.5.1: Correlation Heatmap for the combined dataset .

In the correlation heatmap, we cannot determine which features should be removed, as this is too messy to understand. That is why we will just show the correlation values against each column. Listing : Correlation Function

```
corrmat = data.corr ( )  
corrmat.style.background_gradient(cmap='coolwarm')
```

From the correlation, we can understand that we cannot remove any feature. Moreover, each feature is an important part of cognitive skills.

4.6 Train-Test split

The independent features are labeled "X" while the dependent column is labeled "Trade" as y, and the data is separated for training and testing the models. This was accomplished using a train-test split. The data was split into two groups: 20 for the test set and 80 for the training set.

Our model is trained using the training set. The test set, which represents unseen data for the trained model, was used to assess accuracy.

The data is divided into two categories: training and testing. The training data is used to train the model. The model's predictions are then put to the test using test data. If the model's predictions aren't good enough, the hyperparameters are tweaked to make it operate better on the test set.

4.7 Data Analysis

After integrating the two datasets, we had to employed previously described 6 distinct algorithms to train the machine learning models. Later, we need to utilize MinMaxScaler to improve our outcomes.

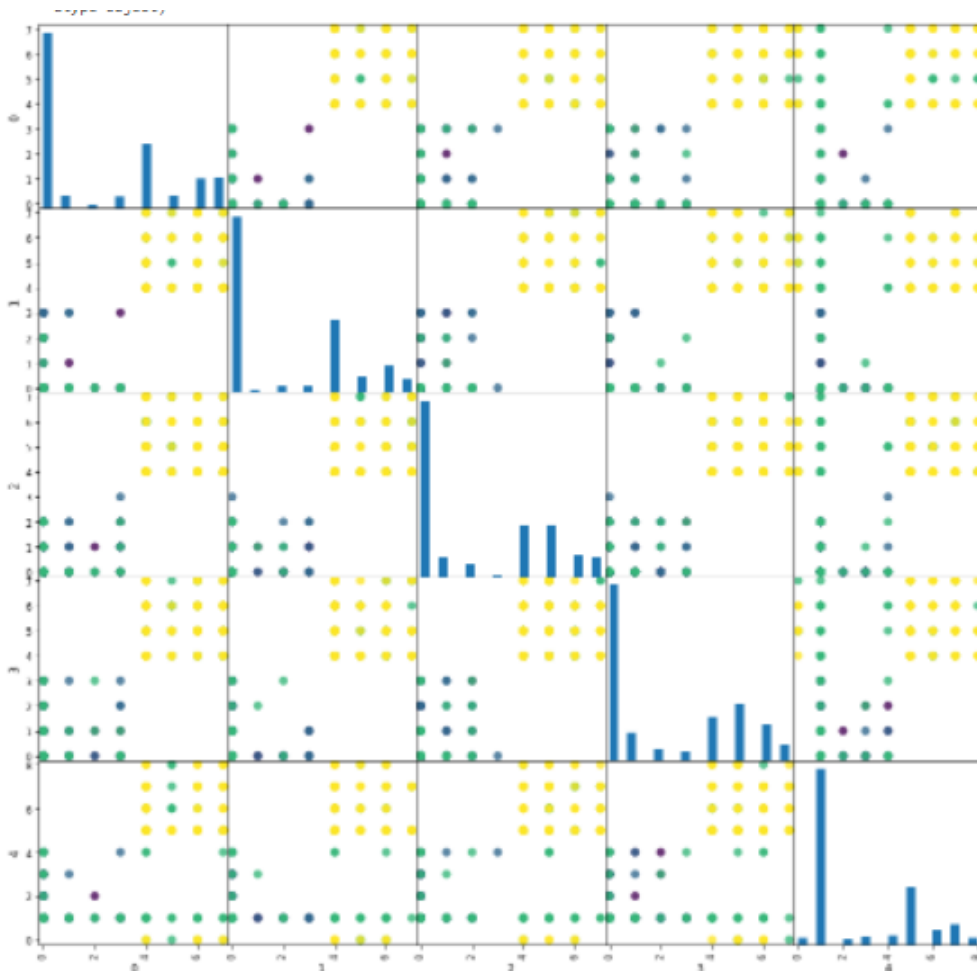


Figure 4.7.1: Data Analysis scatter diagram

From the scatter diagram we can understand that we do not have overlapped data. As a result, we need to compromise all the features.

Chapter 5

Model Implementation and Optimization

5.1 Evaluating Machine Learning Models

After integrating the Logical and Adaptability datasets, we have implemented Logistic Regression, Decision Tree, Random Forest Classifier, Support Vector Classification, Multilayer Perceptron Classifier, and Extreme Gradient Boosting algorithms to train the machine learning models. Our predicted accuracy varies from one to another model. And our predicted accuracy is not quite good for every algorithm. For that reason we need to optimize and preprocess our data for better prediction. The goal of our model is to create one that works well and produces accurate predictions. But due to low datasets, our predicted accuracy was not good enough.

So, our model requires further refinement in order to provide more accurate predictions by preprocessing our data and Optimized algorithm which works with low datasets. Both the process of preparing data before fitting a model and the process of updating an existing model can be considered optimization tasks.

5.2 Before Optimization

Before using MinMaxScaler, our trained data's values were far scattered. From the matrix given below, we can see that the difference between two values is in the integer unit. As in our datasets, we have multiple answer options like option A, B, C, D for most of the questions. But for a few questions, we have 0 to 9 options after mapping them starting from 0. So, datasets individual features values differ from each other.

$$\begin{bmatrix} 1 & 0 & 1 & \dots & 1 & 1 & 0 \\ 1 & 1 & 0 & \dots & 1 & 1 & 3 \\ 0 & 0 & 0 & \dots & 1 & 1 & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 2 & 0 & 0 & \dots & 1 & 3 & 4 \\ 1 & 1 & 7 & \dots & 1 & 9 & 1 \\ 1 & 1 & 7 & \dots & 1 & 9 & 0 \end{bmatrix}$$

Our accuracy for the Extreme Gradient Boosting approach was higher before we used the MinmaxScaler algorithm. We had 60% accuracy using the Extreme Gradient Boosting technique. On the other hand, the Random Forest Classifier reached nearly 57%. Except for these two machine learning models, others have nearly 51%.

Accuracy before Optimization		
Algorithm Name	Accuracy	Log loss
Logistic Regression	0.495652	0.942036
Decision Tree	0.53913	15.917871
Random Forest Classifier	0.573913	0.83984
Support Vector Classification	0.513043	0.91855
Multilayer Perceptron Classifier	0.513043	1.487001
Extreme Gradient Boosting (XGB)	0.6	0.832764

Table 5.1: Pre-accuracy

The log-loss of the Extreme Gradient Boosting (XGB) approach and the Random Forest Classifier algorithm are smaller in our Log-loss vs Algo graph. This corresponds to 0.832764 and 0.83984, respectively. As a result, in order to acquire the greatest outcome from our model, we sought to increase our accuracy.

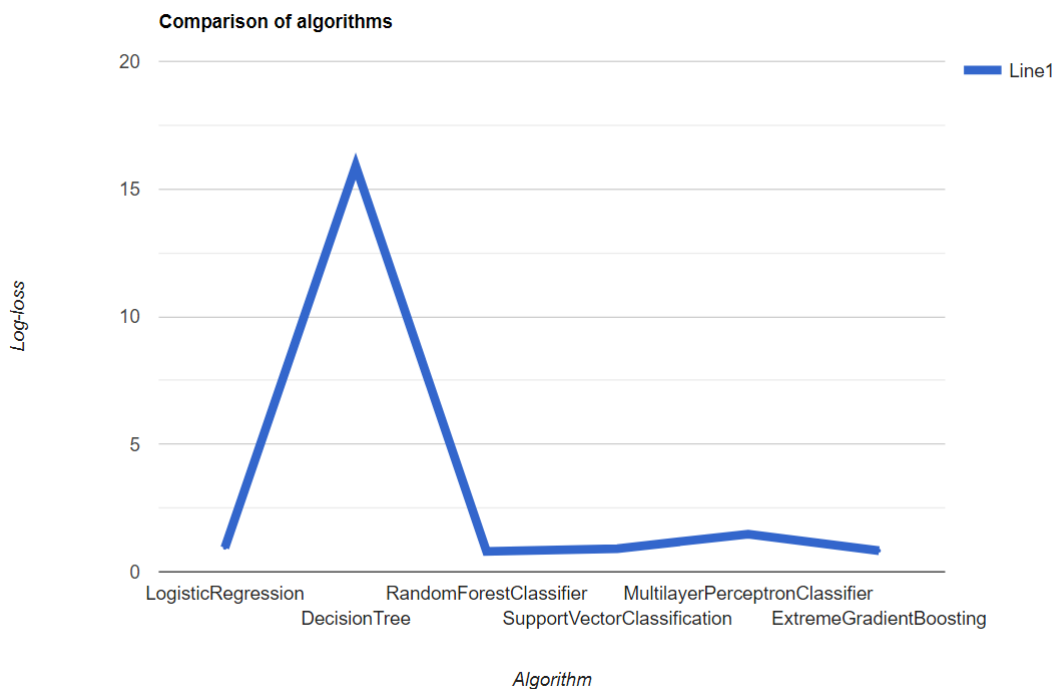


Figure 5.2.1: Comparison of algorithm's Log-loss before using MinMaxScaler

5.3 After Optimization

We can observe that the difference between two numbers is in the integer unit after we've preprocessed our data. Furthermore, after applying MinMaxScaler, this estimator scales and translates each feature independently such that it falls inside the preset range of the training set, which is zero to one. The difference between the data in trained datasets is shown in the matrix below.

$$\begin{bmatrix} 0.33 & 0 & 0.14 & \dots & 0.5 & 0.11 & 0 \\ 0.33 & 1 & 0 & \dots & 0.5 & 0.11 & 0.75 \\ 0 & 0 & 0 & \dots & 0.5 & 0.11 & 0.25 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0.667 & 0 & 0 & \dots & 0.5 & 0.33 & 1 \\ 0.33 & 1 & 1 & \dots & 0.5 & 1 & 0.25 \\ 0.667 & 0 & 0.42 & \dots & 0.42 & 0.5 & 0 \end{bmatrix}$$

MinMaxScaler preserves the shape of the original distribution. It has no influence on the original data's information. MinMaxScaler does not reduce the importance of outliers.

Accuracy after Optimization		
Algorithm	Accuracy	Log loss
Logistic Regression	0.465517	0.946978
Decision Tree	0.500000	17.269388
Random Forest Classifier	0.620690	0.799127
Support Vector Classification	0.517241	0.950979
Multilayer Perceptron Classifier	0.448276	1.820905
Extreme Gradient Boosting (XGB)	0.586207	0.845644

Table 5.2: Post-accuracy.

So After using the MinMaxScaler, Random Forest Classifier's accuracy increases to 62.06%. Where previously it was only 57.39%. So because of using MinMaxScaler we are getting 4.67% more accuracy. On the other hand, the accuracy of the Extreme Gradient Boosting technique is 58.6%. We can see the comparison between pre accuracy and post accuracy of all algorithms. Where only Random Forest Classifier has higher accuracy than pre accuracy which is 62.06%. In figure 5.3.1, we can see the comparison between pre accuracy and post accuracy of all algorithms.

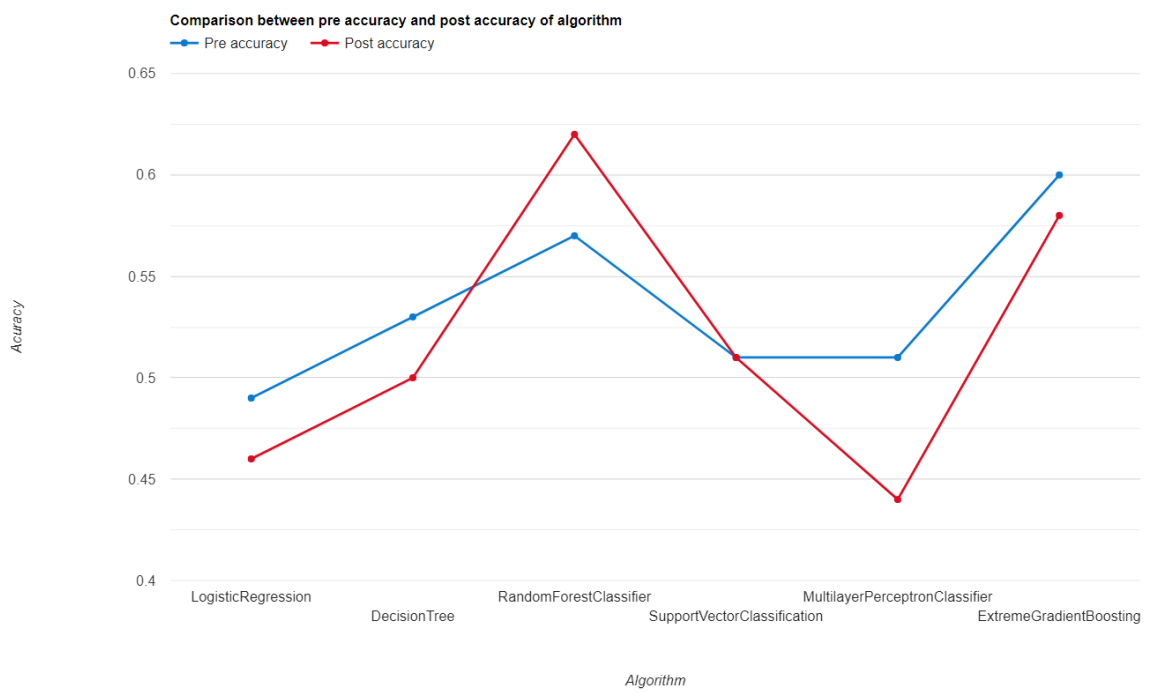


Figure 5.3.1: Comparison between the pre-accuracy and post-accuracy of algorithms

Chapter 6

Result Analysis

After evaluating Logistic Regression, Decision Tree, Random Forest Classifier, Support Vector Classification, Multilayer Perceptron Classifier, and Extreme Gradient Boosting (XGB), finally, we discovered that Random Forest Classifier and XGB provide promising results, with 62.06 percent and 60 percent, respectively. We have two small datasets, logical ability and adaptability, with a total of 573 participants in four different trades. We have four trades, namely IT Support Technician, MS Office, Graphics Design, and Electrical Installation and Maintenance, which are our target variables. Our target variable is definitely not a binary, as we can see. Moreover, if we can create sub categories of those 4 trades in 2 categories, then our accuracy will be boosted because of binomial classification. But our optimal goal is to predict the appropriate trade for the participants when they participate to evaluate themselves. According to our target variable, 62.06 percent accuracy utilizing Random Forest Classification is quite promising. Although the Extreme Gradient Boosting algorithm has provided 60% accuracy, which is close to the Random Forest Classifier's accuracy.

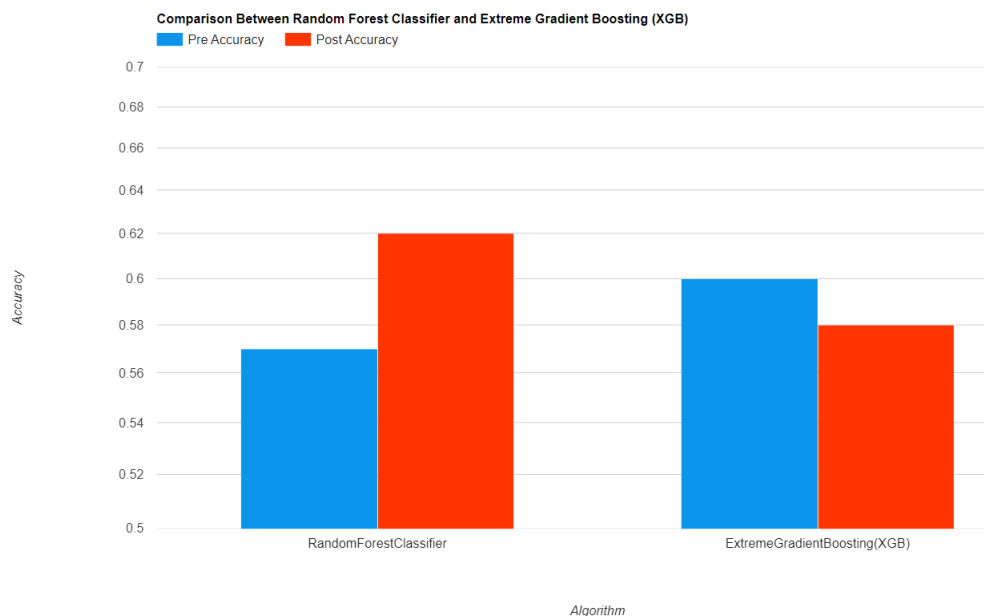


Figure 6.0.1: Final Result Analysis

Chapter 7

Application

7.1 Model Selection and Application Design

All actions involved in the creation and maintenance of a website are included in web development. These responsibilities, which include developing various sorts of websites or online apps, necessitate coding and the use of web programming languages. We will use a website to conduct the assessment for our thesis project. The home page appears when we first visit the website. There will be several buttons and a dashboard on our home page. Home, about, trading, login, and contacts are some of the buttons. The home button will serve as the website's home page. The evaluation will be completed using a play button. The person who presses the play button will be taken to the login page. Then comes the bit about. The description of the idea will be presented here in the about section. After then, trades will take place. There will be four transactions in our trade area with descriptions if we look at the trades section. MS Office, IT Support Technician, Electronic Installation and Materials, and Graphic Design are among the trades available. The user must choose a trade when he/she enters the trade option. When the user selects a trade, he/she will be sent to the login screen. He/she can log in with a username and password if he/she already has one; otherwise, he/she must establish one. It will display the pages of the question section in two parts when you log in. The cognitive section is one, and the trade section is the other. It will take him/her to the question page if he/she clicks the parts. The question will have four alternatives in the question section. After answering all of the questions, the evaluation will be compared to a pre-saved dataset and stored in databases utilizing Machine Learning. The user's trade selection outcome will appear in the profile dashboard area. Finally, the contacts button is used to contact the authorities. The phone number, email address, and location will all be listed on the contact page.

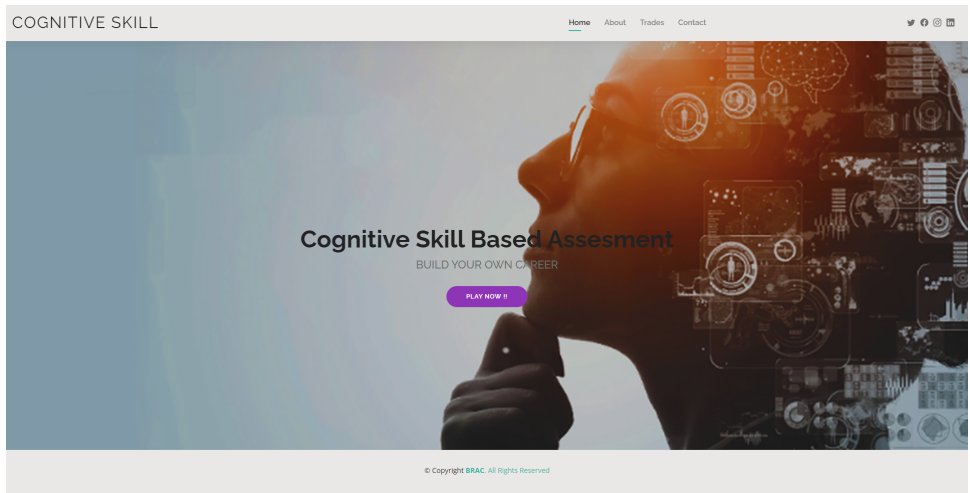


Figure 7.1.1: Web application home page

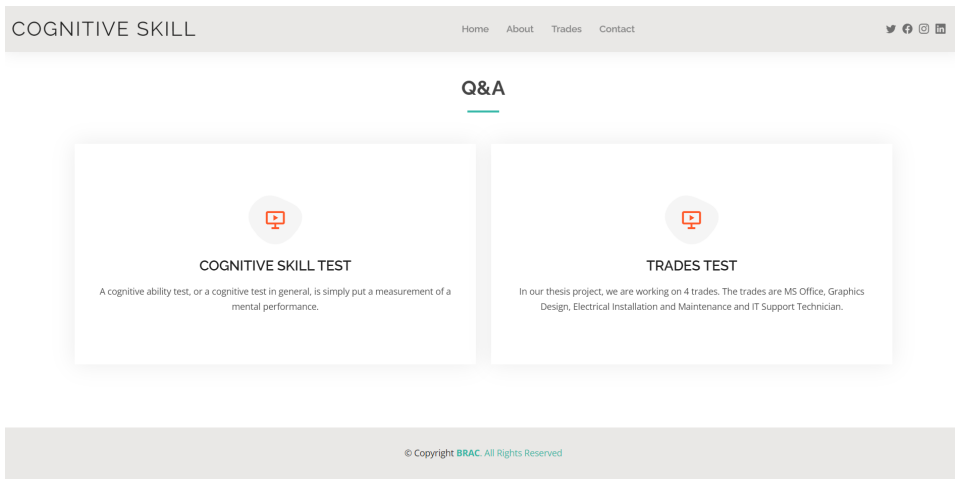


Figure 7.1.2: Web application question categories

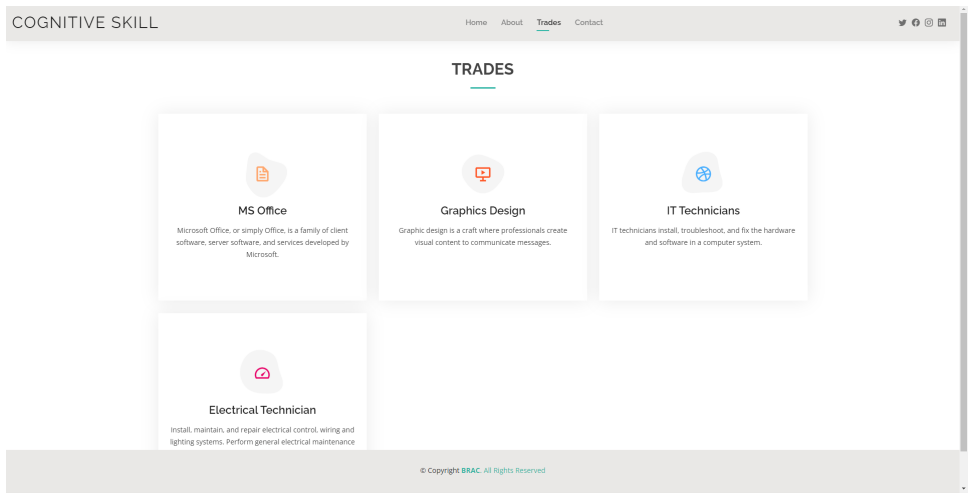


Figure 7.1.3: Web application trade page

We utilized PHP, HTML, CSS, and JavaScript to build our website. We used MySQL to create the database. We basically created the entire website and database on the backend section. All of the user's information will be saved on the server. An admin interface will be available where the administrator may manage the database and add or change questions. If a problem arises, the administrator will investigate and resolve the issue as soon as possible. This is how our web application will work.

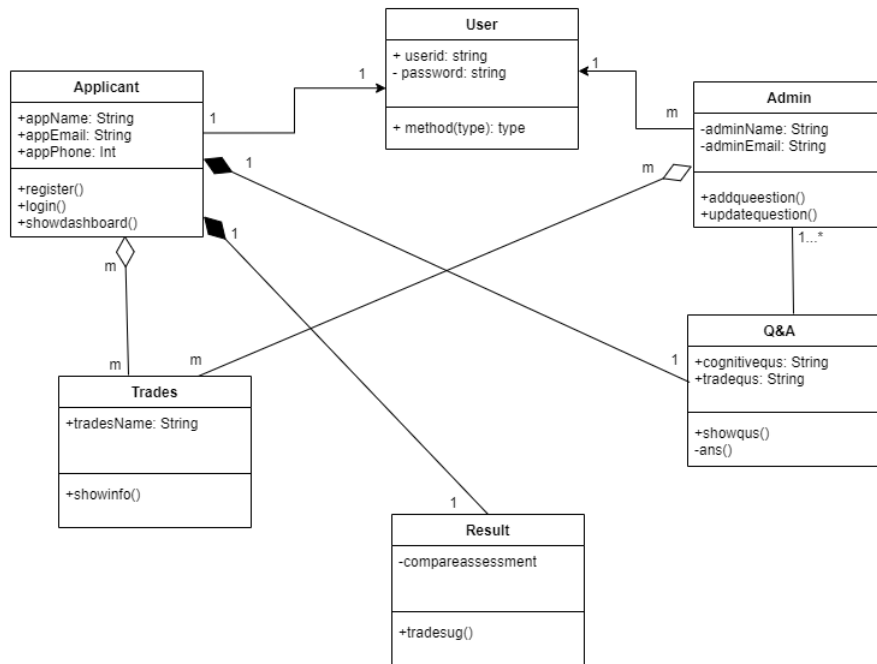


Figure 7.1.4: Web application dataflow diagram

This is our proposed model for this project. It will be developed in the future.

Chapter 8

Conclusion

8.1 Challenges

Our research's objective has been accomplished. However, in order to achieve our goals, we had to overcome some problems or hurdles. First and foremost, there were no available datasets that were comparable to our work, therefore we had to collect primary data. We implemented six algorithms after collecting datasets and obtained an accuracy that was not adequate. As a result, MinMaxScaler was used for further research. Finally, we were able to achieve our desired outcome.

8.2 Future Work

Every study project has space for improvement, and ours is no exception. Our original plan was to collect data from Bangladeshis, graduates, and freshers, but because of the epidemic, we can only collect data from graduates through BRAC NGO. As a result, our primary goal in the future is to compile a database of graduates and freshers. The next plan is to collaborate with the BRAC NGO to obtain a better understanding of data, such as additional data access and a better understanding of all the features that go into the predictions. Following that, we intend to collect a much larger dataset in order to make much more efficient and accurate forecasts. Since the main rationale of machine learning and ensemble models is to predict properly and effectively based on larger datasets, this is a good fit. Finally, we want to test our system in a variety of real-world circumstances. We'll aim to test and validate the prediction software with more trustworthy data, such as career recommendations for people in a given location based on whatever dataset is available. We hope that our research will help us bring about new improvements in the informal job sector and that it will improve the lives of many people in the future.

8.3 Conclusion

The fact that the old system of evaluating human potential has a massive effect on our society is unforgettable, painful, and time-consuming for both employees and organizations. Still, there is discrimination in the system against women and disadvantaged people, which causes a huge lack in our economy. To achieve the 17 Sustainable Development Goals (SDGs) by 2030, a developing country like Bangladesh must begin thinking about the poor's economic problems immediately [6]. BRAC NGO is a pioneer in this field, working with disadvantaged village women and children to provide training and loans so that they can establish their own business or startup in the future and not have to rely on the official economy for employment. As BRAC is providing training on a regular basis and observing that there is still a huge gap between coping with the skill development training staff and career satisfaction in the future, that's why we are trying to introduce a new gamified research model so that this gap can be reduced effectively. We observed that, like our proposed gamified system, JobFlare is one of the platforms helping their over 4,500 clients worldwide by offering leveraged predictive tools to identify quality job candidates through various brain games with the help of modern cognitive science [21]. Besides that, monopoly companies and organizations such as Unilever, LinkedIn, and Accenture are also shifting their employee hiring processes to various Pymetrics games, which are based on neuroscience along with AL and ML [3]. The result of this shift to a technology-based platform is stunning. As their sales rate per year is increasing linearly, consumer attachment is growing step by step, and as the candidates are enjoying the modern process, they are more likely to stay at the same firm [4]. As a result of switching jobs frequently, intentions are removed from their minds, so they are now more confident about their work, which is helping to maintain stability in the whole economy. But all of these are implemented in the formal economy section only. They are not focusing on the informal economy side which is a big reason for lagging behind in the developing countries. Therefore, we are researching and recommending to implement our Cognitive Skill-Based Game for Choosing Career for BRAC NGO, where we are trying to combine cognitive science with human behavior, and personal qualities to rescue all those hassles gradually and create a bridge between new trainees and their career pathways in the informal economy sector. We hope that eventually this modern system will be accepted by other NGOs and informal organizations also.

Bibliography

- [1] S. Mamun, “Covid-19: Over 45% secondary school students may drop out,” *Dhaka Tribune*, vol. 7, 2020.
- [2] J. S. Roy, N. Das, and T.-U. Islam, “Vocational training: Integral policy measure for reducing youth unemployment,” 2022.
- [3] J. B. Diaz and S. F. Young, “Consulting psychology journal: Practice and research,” 2021.
- [4] Y. Kong, C. Xie, J. Wang, H. Jones, and H. Ding, “Ai-assisted recruiting technologies: Tools, challenges, and opportunities,” in *The 39th ACM International Conference on Design of Communication*, 2021, pp. 359–361.
- [5] M. I. Hossain, “Covid-19 impacts on employment and livelihood of marginal people in bangladesh: Lessons learned and way forward,” *South Asian Survey*, vol. 28, no. 1, pp. 57–71, 2021.
- [6] “Undp in bangladesh. 2022. sustainable development goals.,” 2022.
- [7] “En.wikipedia.org. 2022. demographic dividend - wikipedia.,” 2022.
- [8] “Preplounge.com. 2022. how to pass the pymetrics games (2022) | preplounge.com.,” 2022.
- [9] S. Raptis, “Using statistical significance and prediction to test long/short term public services and patients cohorts: A case study in scotland,” *International Journal of Mathematical and Computational Sciences*, vol. 16, no. 4, pp. 18–27, 2022.
- [10] “Teseteo.com. 2022,” 2022.
- [11] “Job description and resume examples | a collection of job descriptions for both hiring managers and job seekers; resume examples, employment, and career resources. 2022. ibm hiring process: Job application, interview, and employment | job description and resume examples.,” 2022.
- [12] N. Amin, T. Yother, M. Johnson, and J. Rayz, “Exploration of natural language processing (nlp) applications in aviation,” *The Collegiate Aviation Review International*, vol. 40, no. 1, 2022.
- [13] “Ask.careers. 2022. axis bank to practice learnings from accelerator project - ask.careers,” 2022.
- [14] “Ilo.org. 2022. more than 60 per cent of the world’s employed population are in the informal economy.,” 2022.

- [15] K.-P. Thai, H. J. Bang, and L. Li, “Accelerating early math learning with research-based personalized learning games: A cluster randomized controlled trial,” *Journal of Research on Educational Effectiveness*, vol. 15, no. 1, pp. 28–51, 2022.
- [16] A. Colgan, S. Rochford, and K. Burke, “Implementing public service reform: Messages from the literature,” 2016.
- [17] E. Necka, “Self-control scale as-36: Construction and validation study,” *Polish Psychological Bulletin*, vol. 46, no. 3, pp. 488–497, 2015.
- [18] JobFlare, “Labour markets migration reducing labour market information frictions with skill certificates: Evidence from south africa,” 2020.
- [19] “Harambee work for work: Planning ability. (n.d.). assessment-reports,” 2020.
- [20] “Harambee work for work: Grit. (n.d.). assessment-reports,” 2020.
- [21] —, “Hiring with game-based assessments,” *Hiring With Game-Based Assessments*, vol. 4, 2021.