**AN INDUSTRY ORIENTED MINI PROJECT**

**ON**

**CAREER GUIDENCE FOR 12th STANDARD STUDENTS**

**USING MACHINE LEARNING**

Submitted in Partial fulfilment of the requirement for the

**BACHELORS OF TECHNOLOGY**

**IN**

**COMPUTER SCIENCE AND DESIGN**

**BY**

1. **AKHILA 22R91A7445**
2. **MAHESHWARI 22R91A7403**
3. **VENU 22R91A7462**

Under the esteemed guidance of

**Mr.T. JAYARAJAN**

**Assistant professor**

**DEPARTMENT OF COMPUTER SCIENCE AND DESIGN**

**TEEGALAKRISHNA REDDY ENGINEERING COLLEGE**

**UGC-AUTONOMOUS**

Approved by AICTE, Affiliated by JNTUH, Accredited by NAAC Grade

Medbowli, Meerpet, Balapur, Hyderabad-500097

(2022-2026)

**DECLARATION BY CANDIDATE**

I am, **P.AKHILA, A.MAHESHWARI,Y.VENU** here by certify that the project entitled “**CAREER GUIDANCE FOR 12th STANDARD STUDENTS USING MACHINE LEARNING**”, under the guidance of **Mr.T.JAYARAJAN**, department of Information Technology, Teegala Krishna Reddy Engineering College for the award of the degree of Bachelor of Information Technology.

This is a record Bonafide work carried out by us and the results embodied in this project have not been reproduced or copied from any source. The results embodied in this project have not been submitted to any other university or institute for the award of any other Degree.

**P.AKHILA - 22R91A7445**

**A.MAHESHWARI -22R91A7403**

**Y.VENU -22R91A7462**

**CERTIFICATE**

This is to certify that the project entitled ‘**CAREER GUIDANCE FOR 12th STANDARD**

**STUDENTS USING MACHINE LEARNING”** is Bonafide work carried out by

**P.AKHILA,A.MAHESHWARI,Y.VENU** in partial fulfillment of the requirement for the award of the degree bachelor of technology in Information Technology. The results presented in this project have been verified and are found to be satisfactory. The results embodied in this project have not been submitted to any other university for the award of any other degree.

**Internal Guide External**

Mr. T. JAYARAJAN

(Assistant Professor)

**Head of the Department Principal**

Dr. J. PREVEEN KUMARDr. K. VENKATA MURALI MOHAN

(Associate Professor)

**ACKNOWLEDGEMENT**

The satisfaction and euphoria that accompanies the successful completion of any task be incomplete without the mention of the mention of the people who made it possible and whose encouragement and guidance have crowed our efforts with success.

I extend our drop sense of gratitude to Principal **Dr. K Venkata Murali Mohan**, Teegala Krishna Reddy Engineering College, Meerpet, for permitting us to undertake this project.

I indicated to **Dr. J. Praveen Kumar** Associate Professor and HOD, Information Technology, Teegala Krishna Reddy Engineering College, Meerpet, for his occident support and guidance throughout our project.

I indicated to our guide **Mr. T. JAYARAJAN**, Assistant Professor, Information Technology, Teegala Krishna Reddy Engineering College, Meerpet, for her occident support and guidance throughout our project. Finally, We Express thanks to one and all who have helped us in successful completing this project. Furthermore, we would like to thank our family and friends for their moral support and encouragement.

**Submitted by:**

**P. AKHILA 22R91A7445**

**A. MAHESHWARI 22R91A7403**

**Y. VENU 22R91A7462**

**CONTENTS**

**TOPIC PAGE.NO**

**1.INTRODUCTION 1-4**

1.1. Purpose of the Project

1.2. Problem Statement

1.3. Existing System

1.4. Proposed System

1.5. Scope of the Project

1.6. Design Architecture

1.7. Technologies used

**2. SOFTWARE REQUIREMENTS 5-7**

2.1. Requirements Specification Document

2.2. Functional Requirements

2.3. Non-Functional Requirements

2.4. Feasibility Study

2.5. Software Requirements

2.6. Hardware Requirements

**3. LITERATURE SURVEY 8-9**

3.2. System Design

**4. MODULE DESCRIPTION 10-12**

4.1 Languages used

**5. SYSTEM DESIGNE 13-15**

5.1 Introduction to UML

5.2 UML Diagrams

5.2.1 Use Case Diagram

5.2.2 Sequence Diagram

5.2.3 Activity Diagram

5.2.4 Class Diagram

**6. CODING AND IMPLEMENTATION 16-20**

**7. TESTING 21-23**

**8. OUTPUT SCREENS 24-26**

**9. CONCLUSION 27**

**10. FUTURE ENHANCEMENT 28**

**11. REFERENCES 29-30**

**ABSTRACT**

Most of the students across the world are always in confusion after they complete higher secondary and the stage where they have to choose an appropriate career path. At the age of 18, the students don't have adequate maturity to accurately know about what an individual has to follow in order to choose a congenial career path. As we pass through the stages, we realize that every student undergoes a series of doubts or thought processes on what to pursue after 12th which is the single tallest question. Then comes the next agony whether they have essential skills for the stream they've chosen. Our computerized career counseling system is used to predict the suitable department for an individual based on their skills assessed by an objective test. If one completes their online assessment which we have created in our system, then automatically they will end up in choosing an appropriate course which will also reduce the failure rate by choosing a wrong career path.

1. **INTRODUCTION**
2. **Purpose of the Project**

The primary purpose of career guidance projects for 12th standard students is to help them explore their interests, skills, and values in order to make informed decisions about their future educational and career paths. These projects aim to bridge the gap between academic studies and the world of work, equipping students with the knowledge and tools needed to navigate their post-secondary options and career aspirations.

1. **Problem Statement**

Students often face confusion when selecting a career path after completing higher secondary education. Many students choose courses based on peer influence, parental pressure, or misconceptions, leading to dissatisfaction and career mismatches. There is a need for an effective career counseling system that provides personalized guidance based on students' abilities, interests, and academic performance

1. **Existing system**

Most of the students always in confusion after completed their higher secondary and the stage where they have to choose an appropriate career path. At the age of 18, the students don’t know to choose congenial career path

**Disadvantages:**

* **High Failure Rate:** Many students end up selecting careers they regret later.
* **Lack of Personalized Recommendations:** Existing systems do not consider students' unique skills and abilities.
* **Limited Data Analysis:** No systematic approach to analyzing student performance using data-driven techniques.

1. **Proposed System**

The proposed system enhances the guidance being offered to students by enabling individual counselors to reach at better decisions and accordingly provide assistance to students. Many factors were considered in the system to reach to a decision like ability, skills, interest, Intelligent Quotient, preferences, academic performances in the past, parent occupation and hobbies followed by influence of parents as well as friends.

**Objective and Bias-Free:** Recommendations based on analytical assessment rather than human judgment.

**Efficient and Scalable:** Can serve thousands of students simultaneously.

**Reduces Career Confusion**: Students receive guidance tailored to their strengths and interests.

1. **Scope of the Project**

A career guidance project for 12th standard students aims to help them explore potential career paths, understand their strengths and interests, and make informed decisions about their future education and work. The scope of such a project can be quite broad, encompassing various aspects of career planning, including self-assessment, career exploration, and guidance from experts.

1. **Design Architecture**

**1. Training Data:**

This represents the historical data used to train the ML model. Information about past students (e.g., academic performance in 10th and 12th grades, subjects chosen, extracurricularactivities, interests, aptitude test scores).

**2. Train The ML Algorithm:**

This step involves using the training data to teach an ML algorithm to identify patterns and relationships between student profiles and suitable career paths.

**Classification algorithms:** To categorize students into different career clusters (e.g., engineering, medicine, arts, commerce). Examples include Decision Trees, Random Forests, Naive Bayes, Support Vector Machines.

**3. Trained ML Algorithm (Successful Model):**

After training, the ML algorithm becomes a "successful model" capable of making predictions or recommendations on new, unseen data.

**4. New Input Data:**

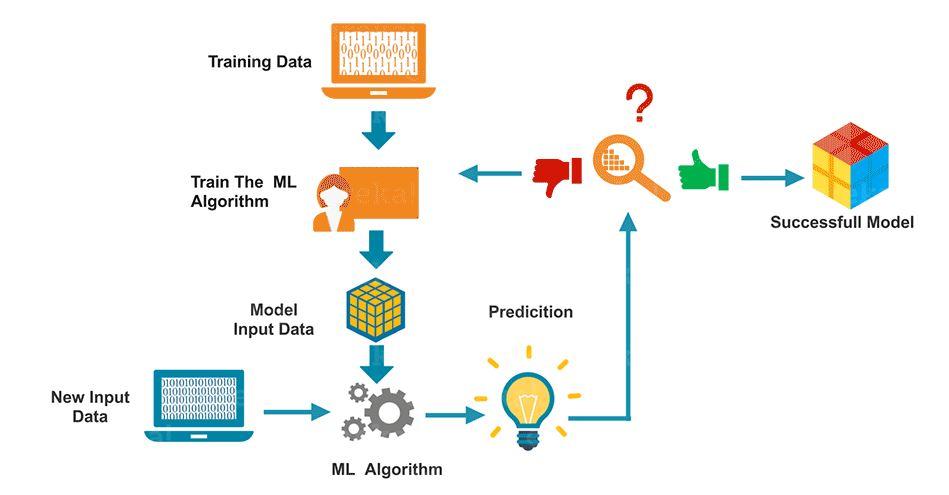
This represents the information provided by a new 12th-standard student who wants career guidance.

**5. ML Algorithm (Prediction):**

The trained ML model takes the "New Input Data" and processes it based on the patterns it learned from the training data. It then generates a "Prediction," which in this context would be a set of recommended career paths or specific career options that align with the student's profile.

**6. Feedback Loop**

If the recommendations are relevant and helpful, it indicates the model is performing well for that student. If the recommendations are not suitable, it provides feedback that can be used to improve the model.



1. **SOFTWARE REQUIREMENTS SPECIFICATIONS**
2. **Requirements Specification Document**

**What is SRS?**

Software Requirements Specification (SRS) is the starting point of the software developing activity. As system grew more complex it became evident that the goal of the entire system cannot be easily comprehended. Hence the need for the requirement phase arose. The software project is initiated by the client needs. The SRS is the means of translating the ideas of the minds of clients (the input) into a formal document (the output of the requirement phase).

**Role of SRS:**

The purpose of the SRS is to reduce the communication gap between the clients and the developers. SRS is the medium though which the client and user needs are accurately specified. It forms the basis of software development. A good SRS should satisfy all the parties involved in the system.

**Scope:**

This document is the only one that describes the requirements of the system. It is meant for the use by the developers, and will also be the basis for validating the final deliver system. Any changes made to the requirements in the future will have to go through a formal change approval process. The developer is responsible for asking for clarifications, where necessary, and will not make any alternations without the permission of the client

1. **Functional Requirements**

Functional requirement should include function performed by a specific screen outline work-flows performed by the system and other business or compliance requirement the system must meet.

* collect and Load Student Dataset
* Data labeling
* Data argumentation
* Label encoding
* Model Training
* Export Student Career models
* Load Student Career models
* Load test image
* Input frame processing

1. **Non-Functional Requirements**

Describe user-visible aspects of the system that are not directly related with the functional behavior of the system.

* Portability
* Reliability
* Usability
* Time Constraints
* Error messages
* Actions which cannot be undone should ask for confirmation
* Responsive design should be implemented
* Space Constraints
* Performance
* Standards
* Ethics
* Interoperability
* Security
* Privacy
* Scalability
  1. **Feasibility Study:**

The feasibility of the project is analyzed in this phase and business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. For feasibility analysis, some understanding of the major requirements for the system is essential.

**Three key considerations involved in the feasibility analysis are**

* Economical Feasibility
* Technical Feasibility
* Social Feasibility

**ECONOMICAL FEASIBILITY**

This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and development of the system is limited. The expenditures must be justified. Thus, the developed system as well within the budget and this was achieved because most of the technologies used are freely available. Only the customized products had to be purchased.

### TECHNICAL FEASIBILITY

### This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands on the available technical resources. This will lead to high demands being placed on the client. The developed system must have a modest requirement, as only minimal or null changes are required for implementing this system.

**SOCIAL FEASIBILITY**

The aspect of study is to check the level of acceptance of the system by the user. This includes the process of training the user to use the system efficiently. The user must not feel threatened by the system, instead must accept it as a necessity. The level of acceptance by the users solely depends on the methods that are employed to educate the user about the system and to make him familiar with it. His level of confidence must be raised so that he is also able to make some constructive criticism, which is welcomed, as he is the final user of the system.

* 1. **Software Requirements**

Programming Language / Platform : Python

IDE : Jupyter

* 1. **Hardware Requirements**

Processor : Intel i3 and above

RAM : 4GB and Higher

Hard Disk : 500GB: Minimum

1. **LITERATURE SURVEY**

**Author:** [Yann Ling Goh](https://www.researchgate.net/profile/Yann-Goh), [Yeh Huann Goh](https://www.researchgate.net/profile/Yeh-Goh-2), [Yip Chun Chieh](https://www.researchgate.net/profile/Yip-Chun-Chieh), [Chen Hunt Ting](https://www.researchgate.net/scientific-contributions/Chen-Hunt-Ting-2174890523)

# Title: Prediction of Students' Academic Performance by K-Means Clustering

**Description:** Schooling system must provide high quality learning opportunities to meet the educational needs and ensuring achievement for every student. All teachers monitor their students’ progress throughout the year, includes formative assessment, questioning, providing feedback, etc. This practice helps teachers continually assess students’ academic performance and evaluate the effectiveness of their teaching. In this paper, k-means clustering method with deterministic model is used to analyze the student's overall performance. The results are important for educators to identify students who are at risk academically and areas where teaching strategies may need adjustment to better meet these students' needs.

**Author:** [Brijesh Kumar Bhardwaj](https://arxiv.org/search/cs?searchtype=author&query=Bhardwaj%2C+B+K), [Saurabh Pal](https://arxiv.org/search/cs?searchtype=author&query=Pal%2C+S)

**Title:** A prediction for performance improvement using classification

**Description:** Now-a-days the amount of data stored in educational database increasing rapidly. These databases contain hidden information for improvement of students' performance. The performance in higher education in India is a turning point in the academics for all students. This academic performance is influenced by many factors; therefore, it is essential to develop predictive data mining model for students' performance so as to identify the difference between high learners and slow learners’ student. In the present investigation, an experimental methodology was adopted to generate a database. The raw data was preprocessed in terms of filling up missing values, transforming values in one form into another and relevant attribute/ variable selection. As a result, we had 300 student records, which were used for by Byes classification prediction model construction. Keywords- Data Mining, Educational Data Mining, Predictive Model, Classification.

**Author: [Gadde Shravya Sree](http://www.internationaljournalcorner.com/index.php/ijird_ojs/search/authors/view?firstName=Gadde%20Shravya&middleName=&lastName=Sree&affiliation=&country=), [Ch. Rupa](http://www.internationaljournalcorner.com/index.php/ijird_ojs/search/authors/view?firstName=Ch.&middleName=&lastName=Rupa&affiliation=&country=)**

### Title: Performance Improvement in Education Sector Using Classification and Clustering Algorithm

**Description:** The ability to predict a student’s performance is very important in educational environments.    Students’    academic performance is based upon diverse factors like personal, social, Psychological and other environmental variables. A very promising tool to attain this objective is the use of Data Mining. Data mining techniques are used to discover hidden information patterns and relationships of large amount of data, which is very much helpful in decision making. A single data contains a lot of information. The type of information is produced by the data and it decides the processing method of data. A lot of data that can produce valuable information, in education sector contains this valuable information. Which helps the education sector to capture and compile low-cost information for this information and communication technology is used. Now-a-days educational database is increased rapidly because of the large amount of data stored in it. The loyal students motivate the higher educationsystems, to know them well; the best way is by using valid management and processing of the students' database.Data mining approach provides valid information from existing student to manage relationships with upcoming students.

1. **System Design**

System design is transition from a user-oriented document to programmers or data base personnel. The design is a solution, how to approach to the creation of a new system. This is composed of several steps. It provides the understanding and procedural details necessary for implementing the system recommended in the feasibility study. Designing goes through logical and physical stages of development, logical design reviews the present physical system, prepare input and output specification, details of implementation plan and prepare a logical design walkthrough.

The database tables are designed by analyzing functions involved in the system and format of the fields is also designed. The fields in the database tables should define their role in the system. The unnecessary fields should be avoided because it affects the storage areas of the system. Then in the input and output screen design, the design should be made user friendly. The menu should be precise and compact.

1. **MODULE DESCRIPTION**

**Numpy:**

Numpy is a general-purpose array-processing package. It provides a high-Performance multidimensional array object and tools for working with these Arrays. It is the fundamental package for scientific computing with Python.

**It contains various features including these important ones:**

* A powerful N-dimensional array object
* Sophisticated (broadcasting) functions
* Tools for integrating C/C++ and FORTRAN code
* Useful linear algebra, Fourier transforms, and random number capabilities

Besides its obvious scientific uses, Numpy can also be used as an efficient multi-dimensional container of generic data. Arbitrary data-types can be defined using Numpy which allows Numpy to seamlessly and speedily integrate with a wide variety of databases.

**Pandas:**

Pandas are an open-source Python Library providing high-performance data manipulation and analysis tool using its powerful data structures. The name Pandas is derived from the word Panel Data – an Econometrics from Multidimensional data. In 2008, developer Wes McKinney started developing pandas when in need of high performance, flexible tool for analysis of data. Prior to Pandas, Python was majorly used for data munging and preparation. It had very little contribution towards data analysis. Pandas solved this problem. Using Pandas, we can accomplish five typical steps in the processing and analysis of data, regardless of the origin of data — load, prepare, manipulate, model, and analyze. Python with Pandas is used in a wide range of fields including academic and commercial domains including finance, economics, Statistics, analytics, etc. Standard Python distribution doesn&#39;t come bundled with Pandas module. A lightweight alternative is to install NumPy using popular Python package installer, pip. pip install pandas

**Matplotlib:**

Matplotlib is a plotting library for the Python programming language and its numerical mathematics extension NumPy. It provides an object- oriented API for embedding plots into applications using general- purpose GUI toolkits like Tkinter, wxPython, Qt, or GTK+. There is also a procedural &quot;pylab&quot; interface based on a state machine (like OpenGL), designed to closely resemble that of MATLAB, though its use is discouraged. [3] SciPy makes use of Matplotlib. Matplotlib was originally written by John D. Hunter, has an active development community, [4] and is distributed under a BSD-style license. Michael Droettboom was nominated as matplotlib&#39;s lead developer shortly before John Hunter&#39;s death in August 2012, [5] and further joined by Thomas Caswell. As of 23 June 2017, matplotlib 2.0.x supports Python versions 2.7 through 3.6. Python3 support started with Matplotlib 1.2. Matplotlib 1.4 is the last version to support Python 2.6. Matplotlib has pledged to not support Python 2 past 2020 by signing the Python 3 Statement.

**Sklern**

Scikit-learn is a library in Python that provides many unsupervised and supervised learning algorithms. It’s built upon some of the technology you might already be familiar with, like NumPy, pandas, and Matplotlib!

**The functionality that scikit-learn provides include:**

* Regression, including Linear and Logistic Regression
* Classification, including K-Nearest Neighbors
* Clustering, including K-Means and K-Means++
* Model selection
* Preprocessing, including Min-Max Normalization

As you move through Codecademy’s Machine Learning content, you will become familiar with many of these terms. You will also see scikit- learn (in Python, sklearn) modules being used. For example: sklearn.linear\_model.LinearRegression() is a Linear Regression model inside the linear\_model module of sklearn. The power of scikit-learn will greatly aid your creation of robust Machine Learning programs.

* 1. **Languages used**

**python**

The Python programming language is an Open Source, cross-platform, high level, dynamic, interpreted language.

The Python 'philosophy' emphasizes readability, clarity and simplicity, whilst maximizing the power and expressiveness available to the programmer. The ultimate compliment to a Python programmer is not that his code is clever, but that it is elegant. For these reasons Python is an excellent 'first language', while still being a powerful tool in the hands of the seasoned and cynical programmer.

**Python is a very flexible language. It is widely used for many different purposes. Typical uses include:**

* Web application programming with frameworks like Zope, Django and Turbo gears
* System administration tasks via simple scripts
* Desktop applications using GUI toolkits like Tkinter or wxPython (and recently Windows Forms and IronPython)
* Creating windows applications, using the Pywin32 extension for full windows integration and possibly Py2exe to create standalone programs
* Scientific research using packages like Scipy and Matplotlib

1. **SYSTEM DESIGN**
   1. **Introduction to UML**

UML stands for Unified Modeling Language. UML is a standardized general-purpose modeling language in the field of object-oriented software engineering. The standard is managed, and was created by, the Object Management Group.

The goal is for UML to become a common language for creating models of object-oriented computer software. In its current form UML is comprised of two major components: a Meta-model and a notation. In the future, some form of method or process may also be added to; or associated with, UML.

The Unified Modeling Language is a standard language for specifying, Visualization, Constructing and documenting the artifacts of software system, as well as for business modeling and other non-software systems.

The UML represents a collection of best engineering practices that have proven successful in the modeling of large and complex systems.

The UML is a very important part of developing objects-oriented software and the software development process. The UML uses mostly graphical notations to express the design of software projects.

* 1. **UML Diagrams**

**5.2.1 Use Case Diagram**

A use case diagram in the Unified Modeling Language (UML) is a type of behavioral diagram

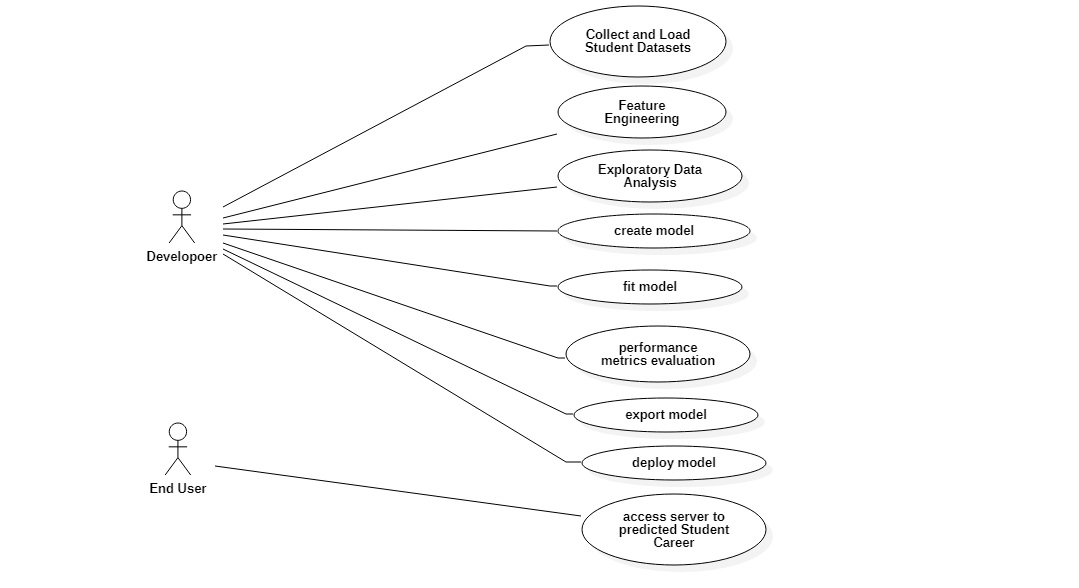
defined by and created from a Use-case analysis. Its purpose is to present a graphical overview

of the functionality provided by a system in terms of actors, their goals (represented as use

cases), and any dependencies between those use cases. The main purpose of a use case diagram

is to show what system functions are performed for which actor. Roles of the actors in the

system can be depicted.

****

1. **Sequence Diagram**

Sequence Diagrams Represent the objects participating the interaction horizontally and time

vertically. A Use Case is a kind of behavioral classifier that represents a declaration of an

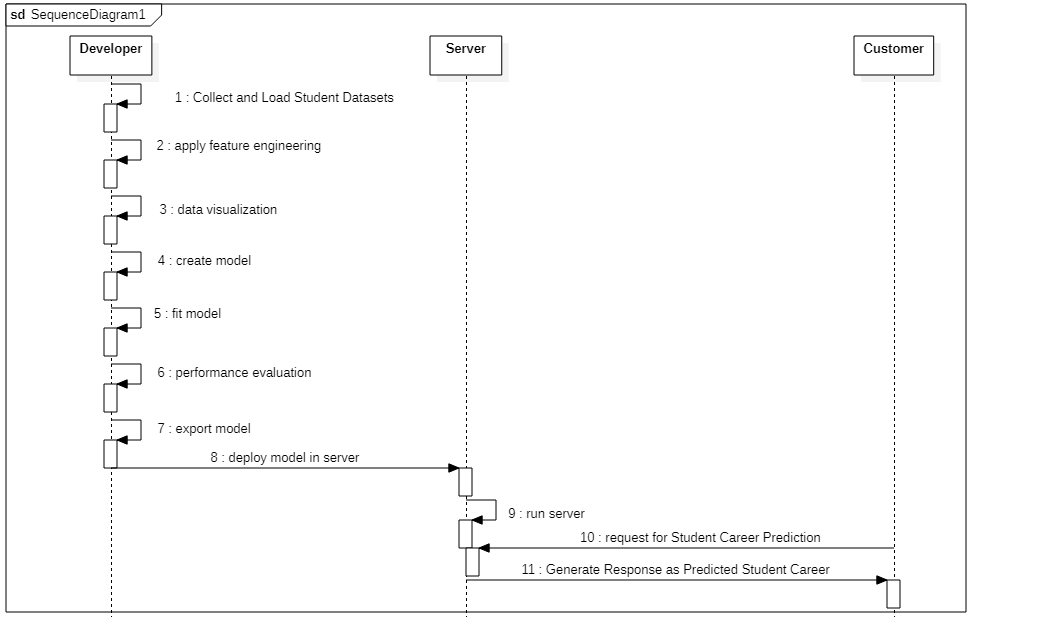
offered behavior. Each use case specifies some behavior, possibly including variants that the

subject can perform in collaboration with one or more actors. Use cases define the offered

behavior of the subject without reference to its internal structure. These behaviors, involving

interactions between the actor and the subject, may result in changes to the state of the subject

and communications with its environment.

****

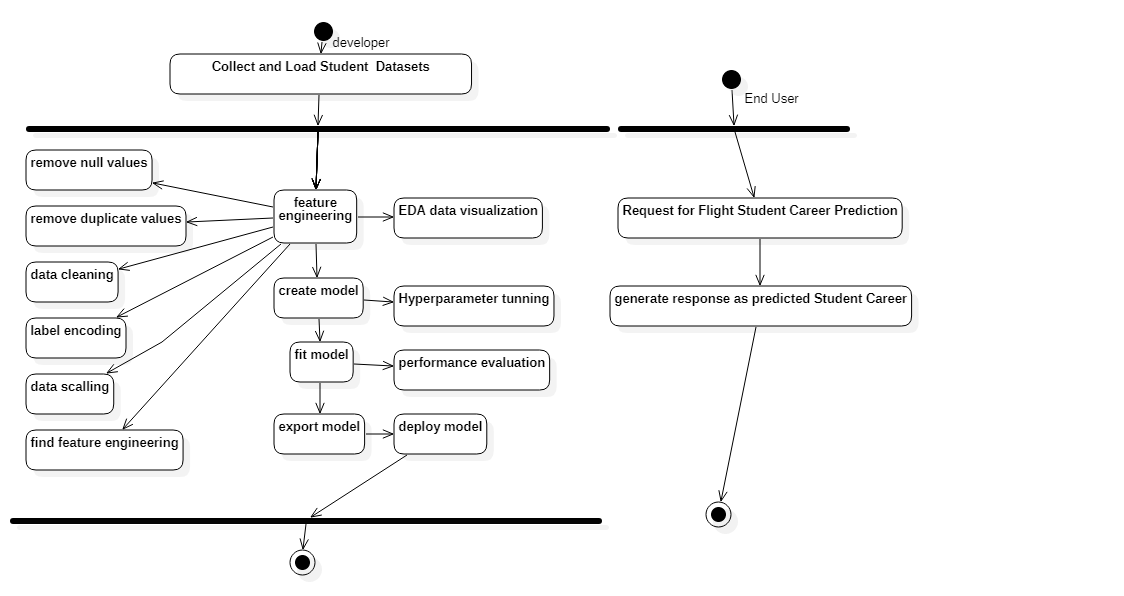
* + 1. **Activity Diagram**

Activity diagrams are graphical representations of Workflows of stepwise activities and actions

with support for choice, iteration and concurrency.In the Unified Modeling Language, activity

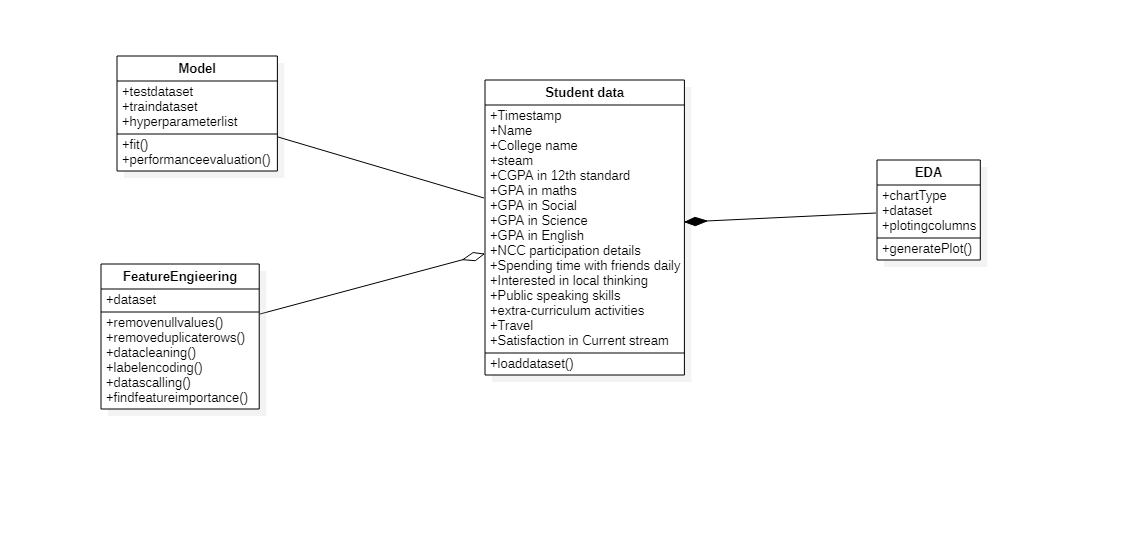
diagrams can be used to describe the business and operational step-by-step workflows of

components in a system. An activity diagram shows the overall flow of control

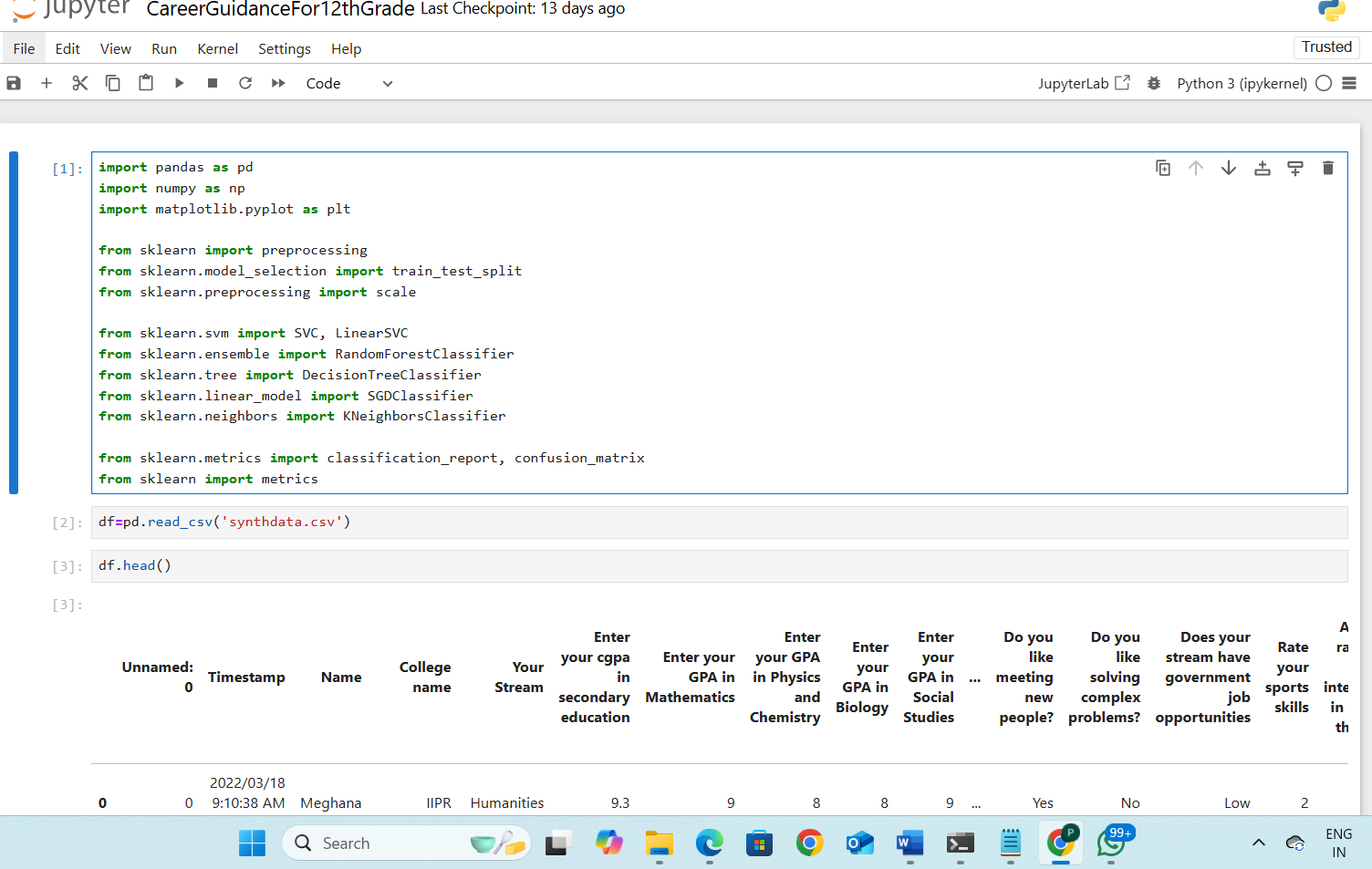
****

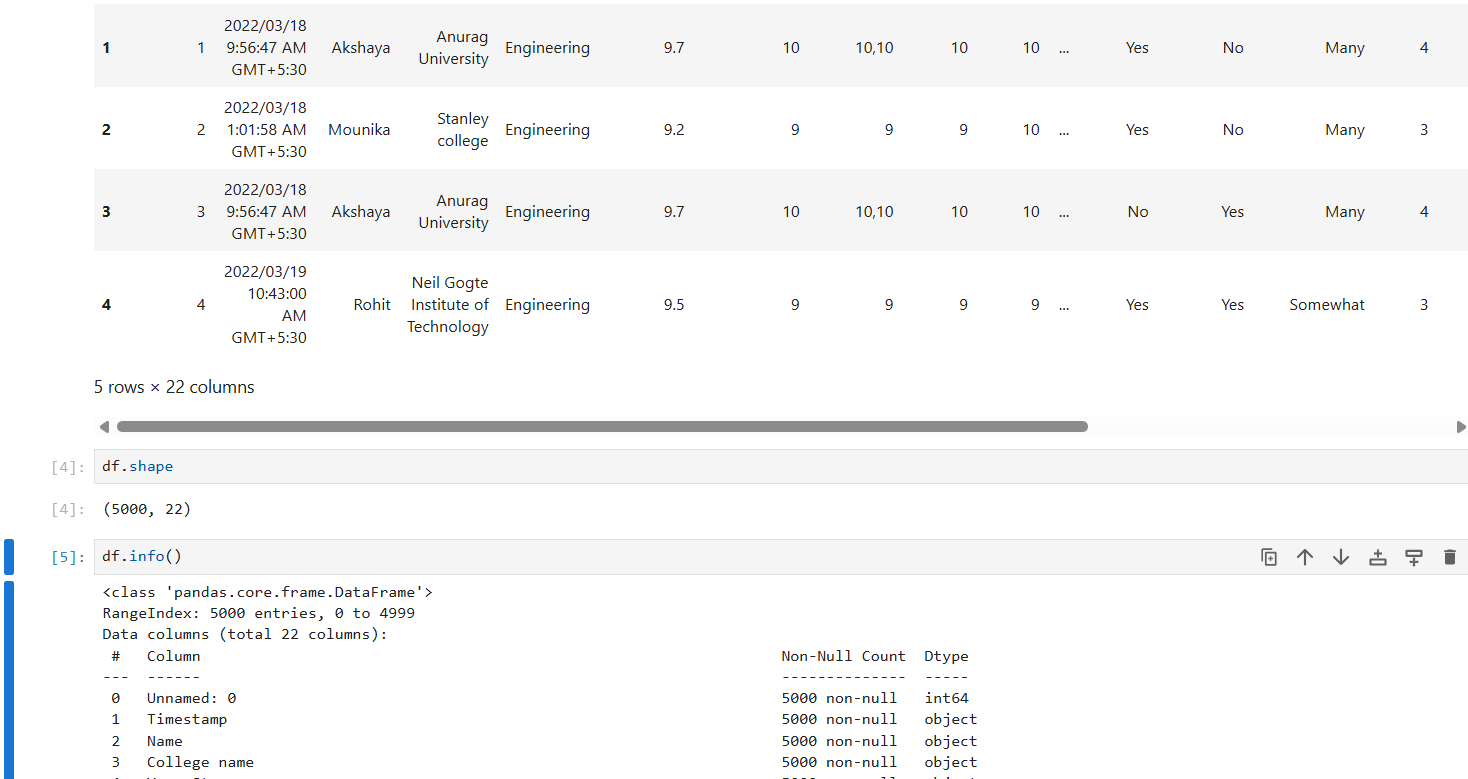
* + 1. **Class Diagram**

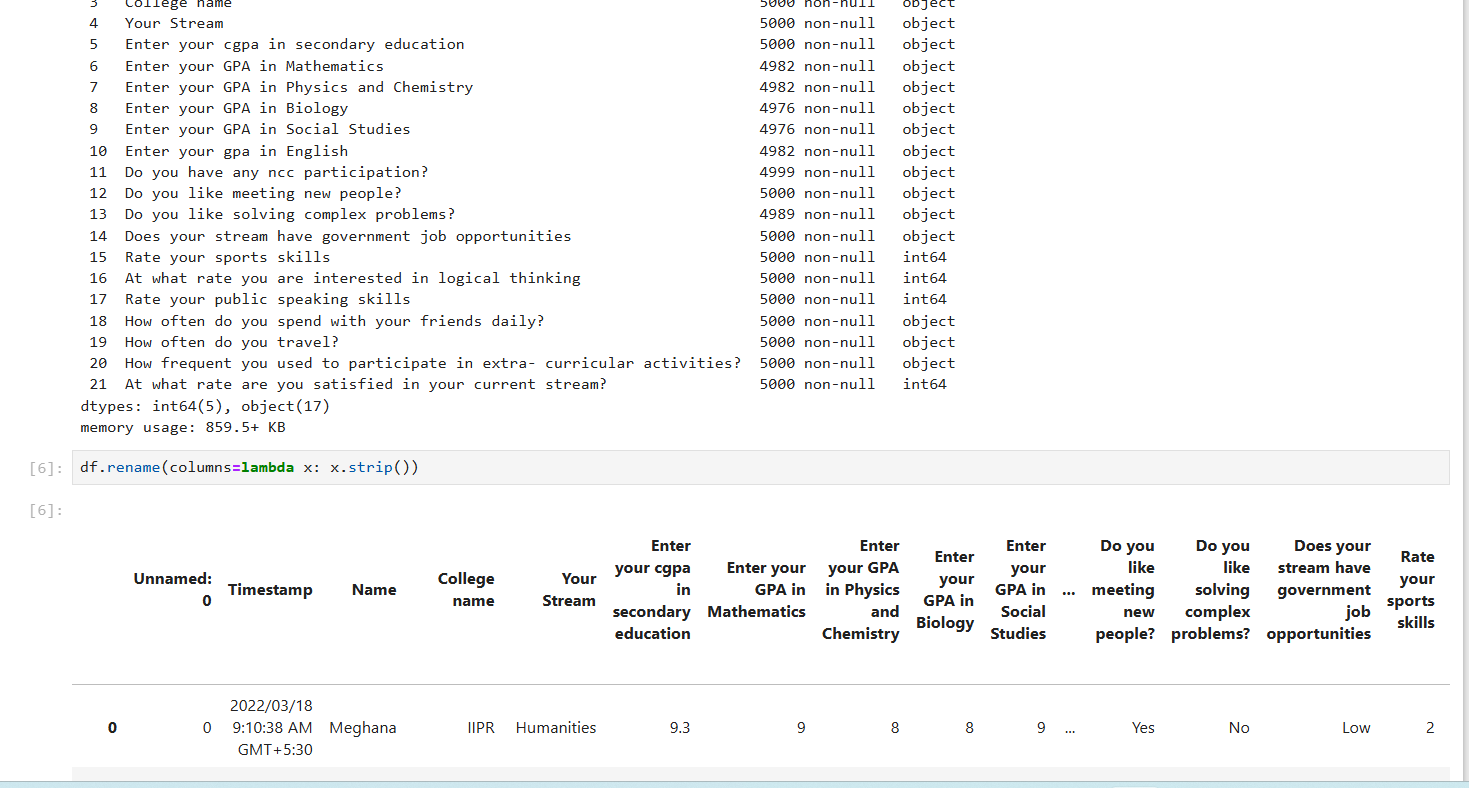
In software engineering, a class diagram in the Unified Modeling Language (UML) is a type of static structure diagram that describes the structure of a system by showing the system's classes, their attributes, operations (or methods), and the relationships among the classes. It explains which class contains information

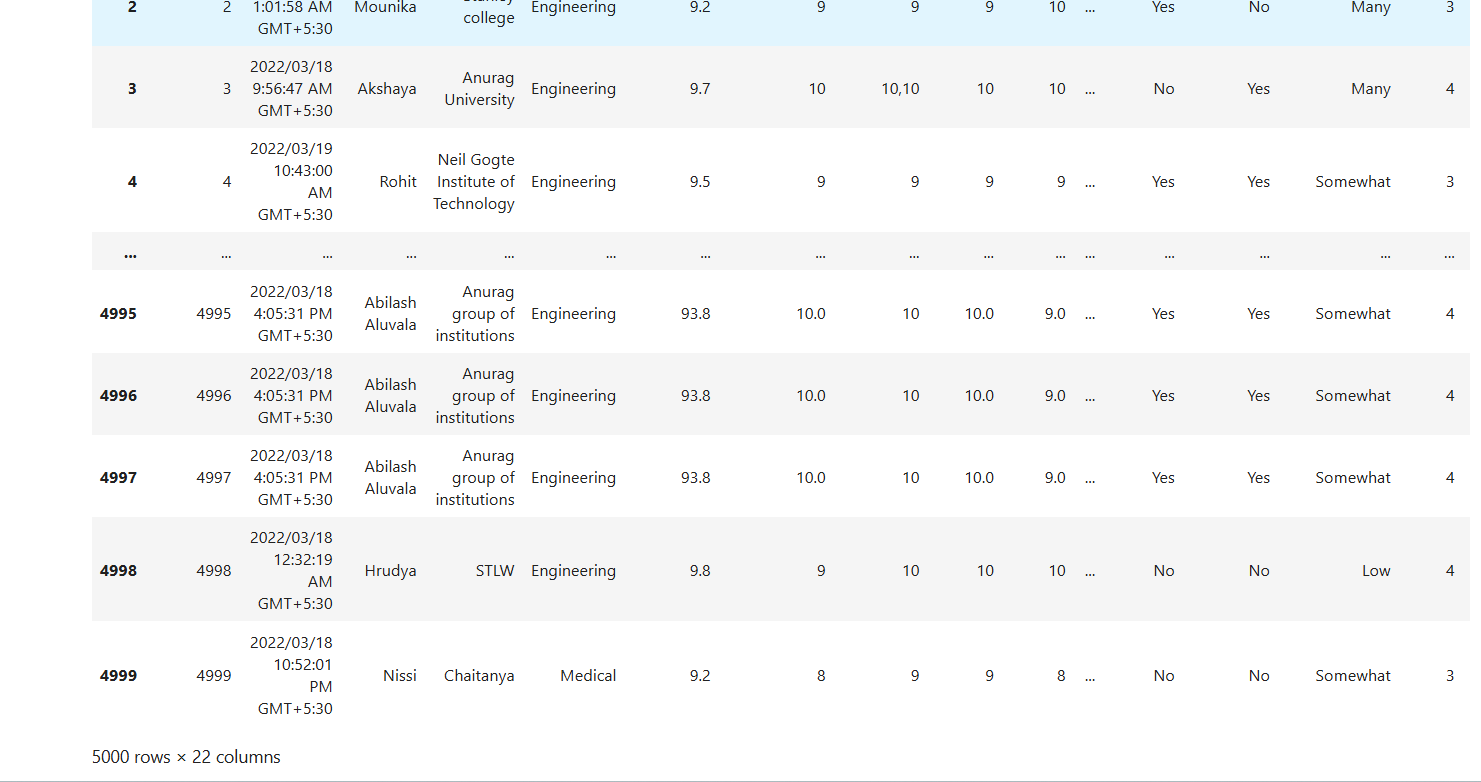
****

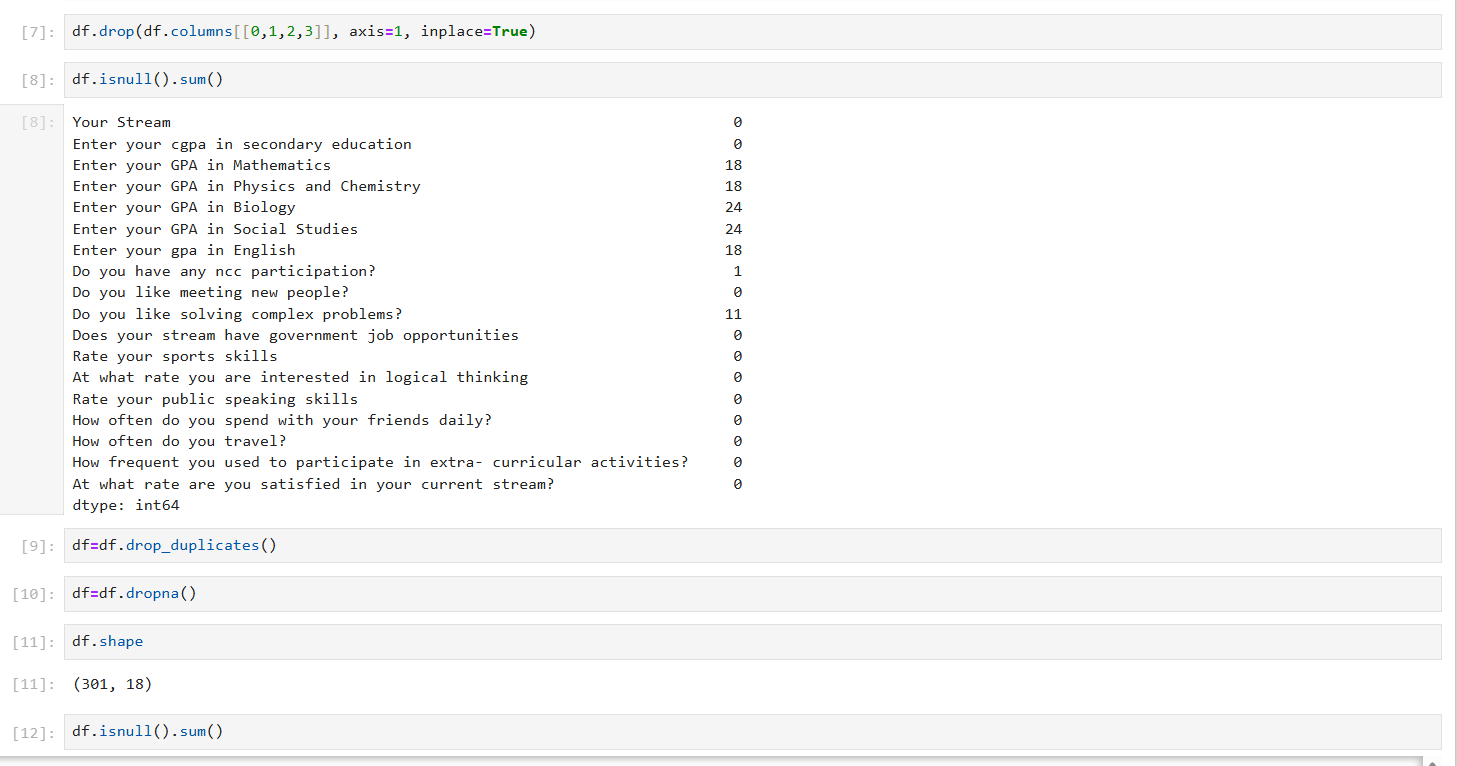
1. **CODING AND IMPLEMENTATION**

****

****

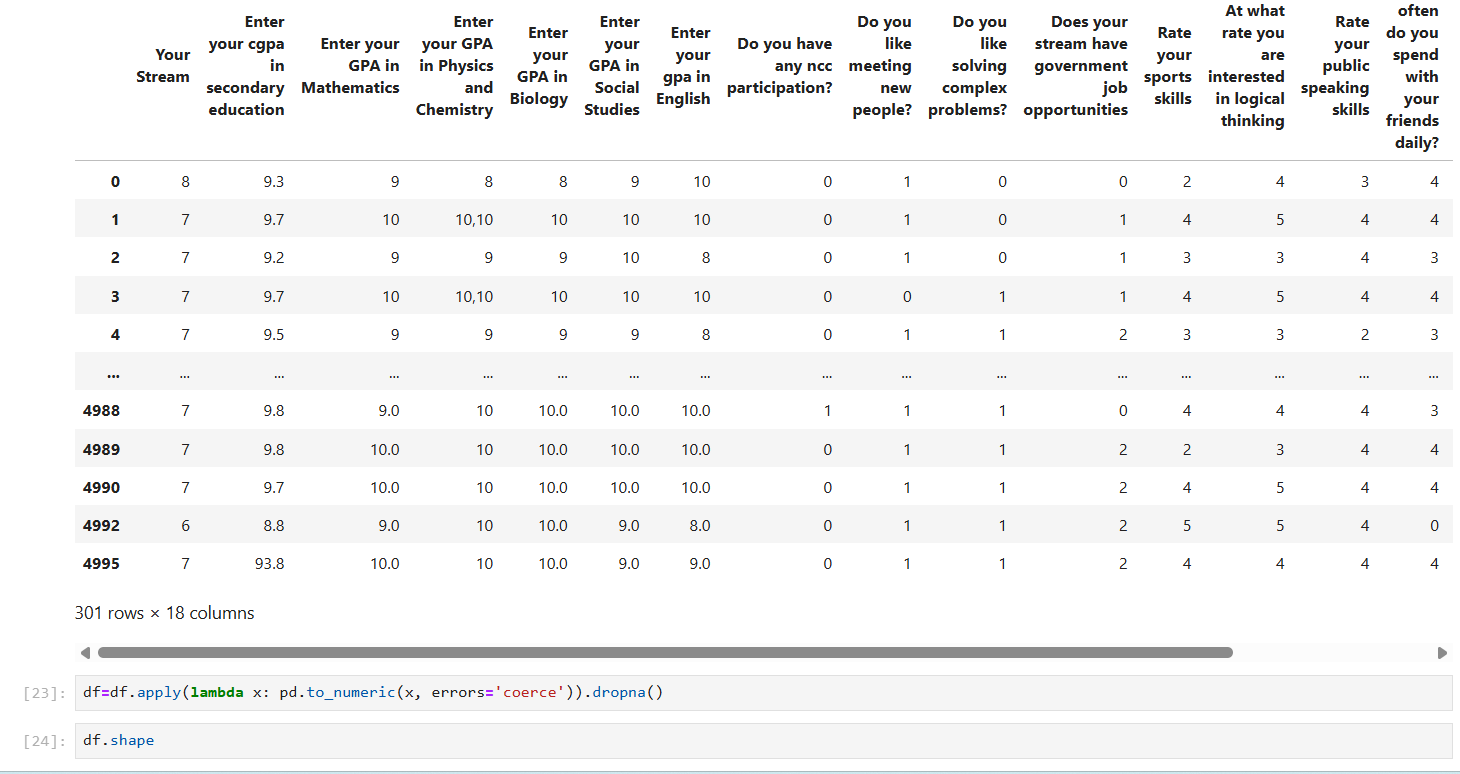


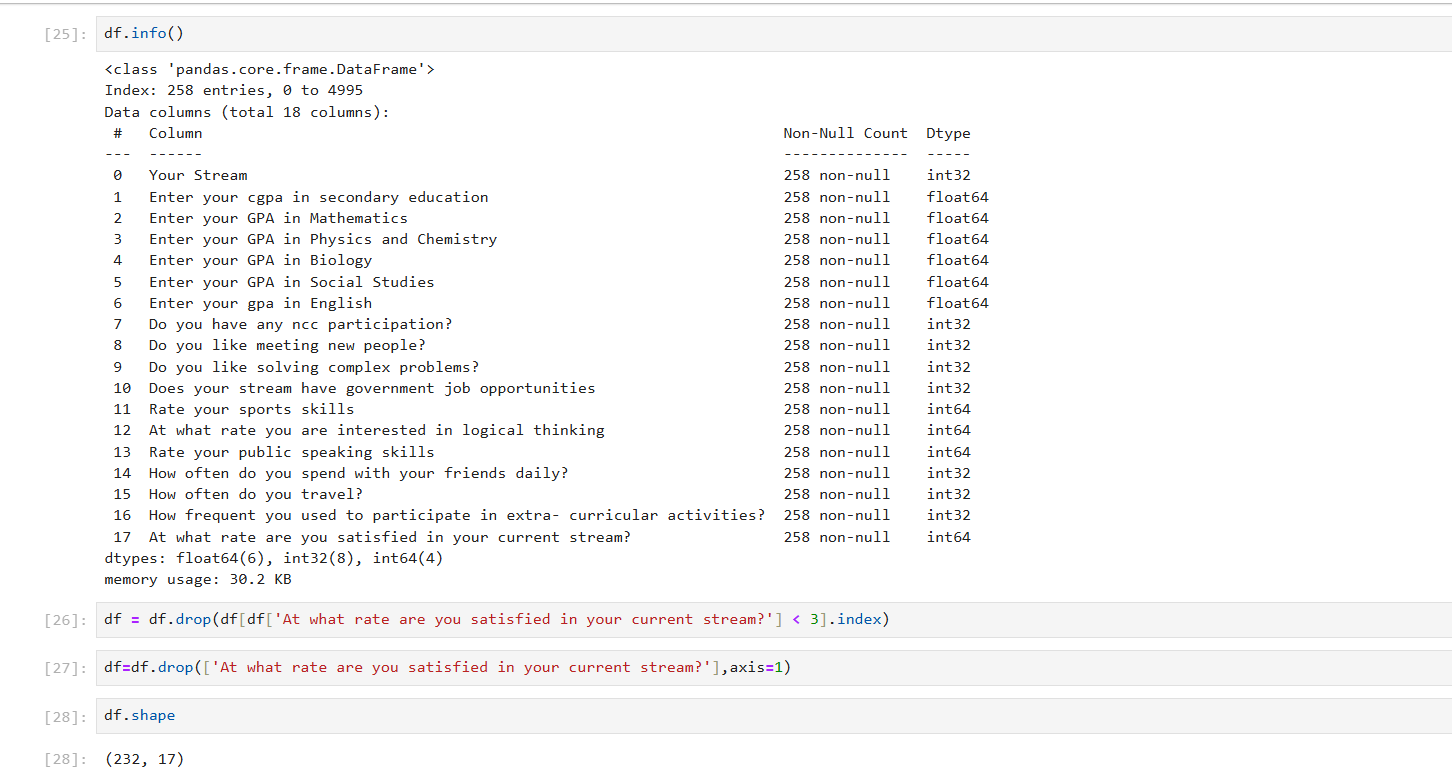
****

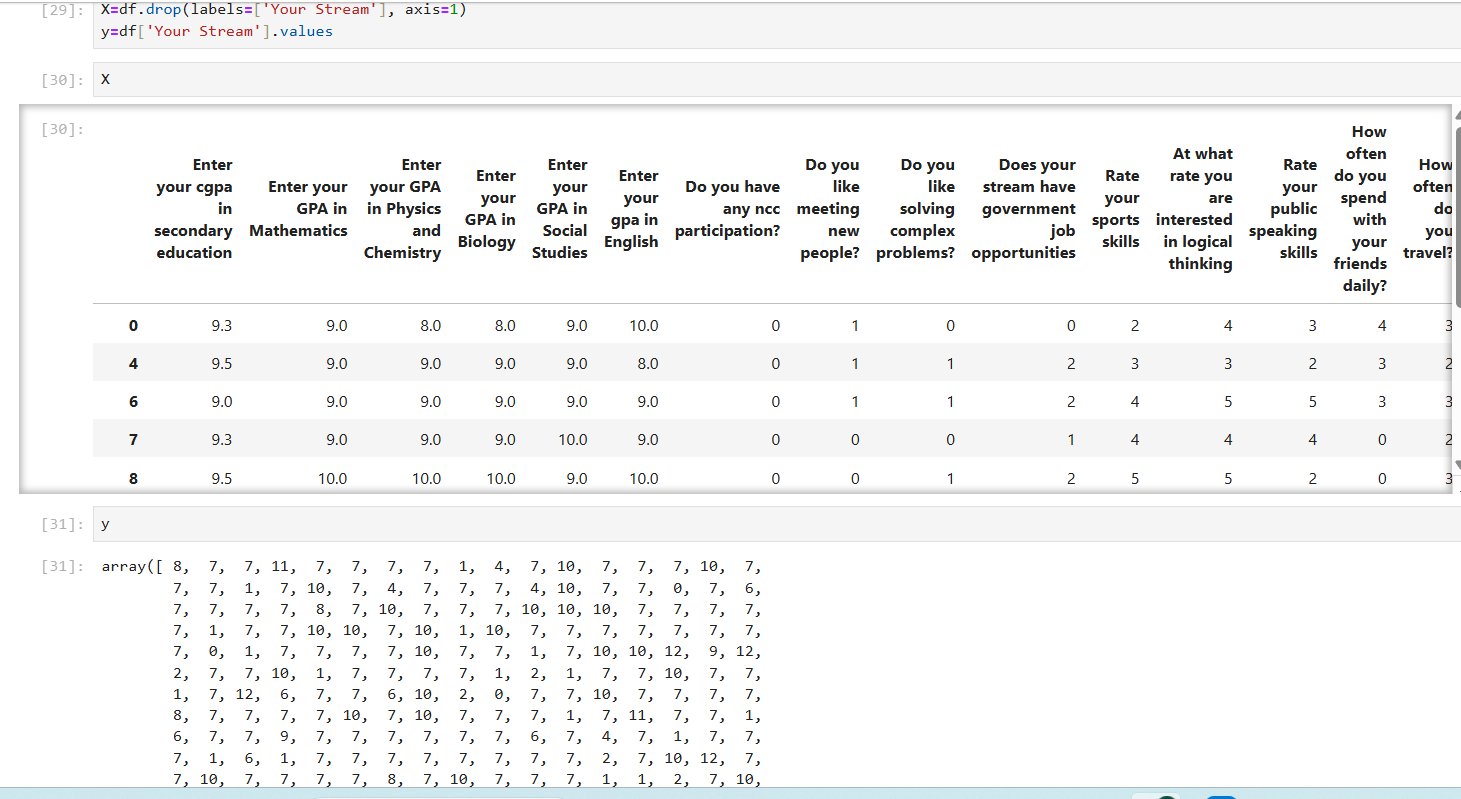
****

****

****

****

****

****

1. **TESTING**

**TYPES OF TESTS**

Test strategy and approach

Field testing will be performed manually and functional tests will be written in detail.

Test objectives

All field entries must work properly.

Pages must be activated from the identified link.

The entry screen, messages and responses must not be delayed.

Features to be tested

Verify that the entries are of the correct format

No duplicate entries should be allowedAll links should take the user to the correct page.

**Unit testing**

Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program inputs produce valid outputs. All decision branches and internal code flow should be validated. It is the testing of individual software units of the application .it is done after the completion of an individual unit before integration. This is a structural testing, that relies on knowledge of its construction and is invasive. Unit tests perform basic tests at component level and test a specific business process, application, and/or system configuration. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications and contains clearly defined inputs and expected results.

**Integration testing**

Integration tests are designed to test integrated software components to determine if they actually run as one program. Testing is event driven and is more concerned with the basic outcome of screens or fields. Integration tests demonstrate that although the components were individually satisfaction, as shown by successfully unit testing, the combination of components is correct and consistent. Integration testing is specifically aimed at exposing the problems that arise from the combination of components.

**Functional test**

Functional tests provide systematic demonstrations that functions tested are available as specified by the business and technical requirements, system documentation, and user manuals.

Functional testing is centered on the following items:

Valid Input : identified classes of valid input must be accepted.

Invalid Input : identified classes of invalid input must be rejected.

Functions : identified functions must be exercised.

Output : identified classes of application outputs must be exercised.

Systems/Procedures : interfacing systems or procedures must be invoked.

Organization and preparation of functional tests is focused on requirements, key functions, or special test cases. In addition, systematic coverage pertaining to identify Business process flows; data fields, predefined processes, and successive processes must be considered for testing. Before functional testing is complete, additional tests are identified and the effective value of current tests is determined.

**System Test**

System testing ensures that the entire integrated software system meets requirements. It tests a configuration to ensure known and predictable results. An example of system testing is the configuration-oriented system integration test. System testing is based on process descriptions and flows, emphasizing pre-driven process links and integration points.

**White Box Testing**

White Box Testing is a testing in which in which the software tester has knowledge of the inner workings, structure and language of the software, or at least its purpose. It is purpose. It is used to test areas that cannot be reached from a black box level.

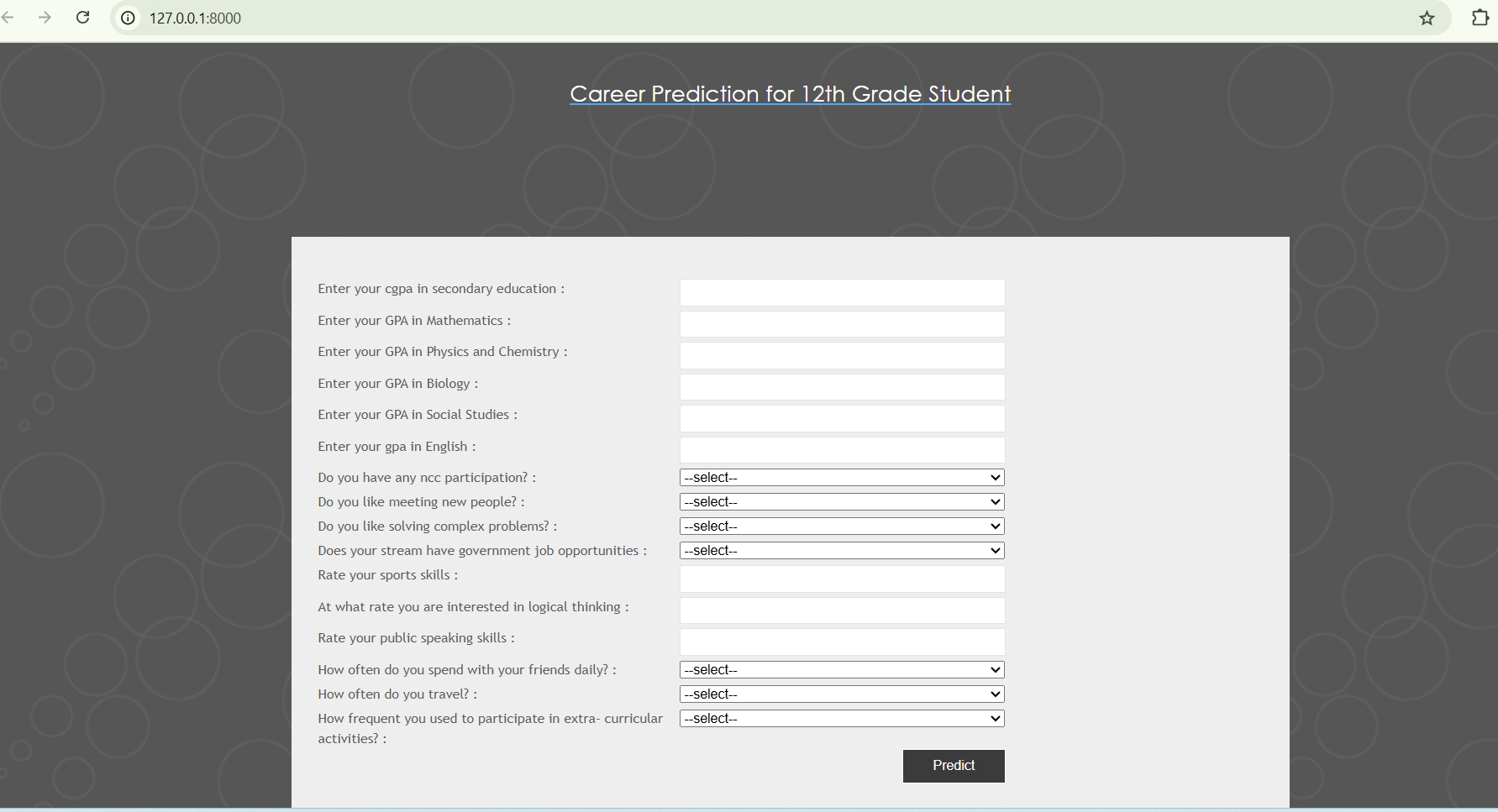
**Black Box Testing**

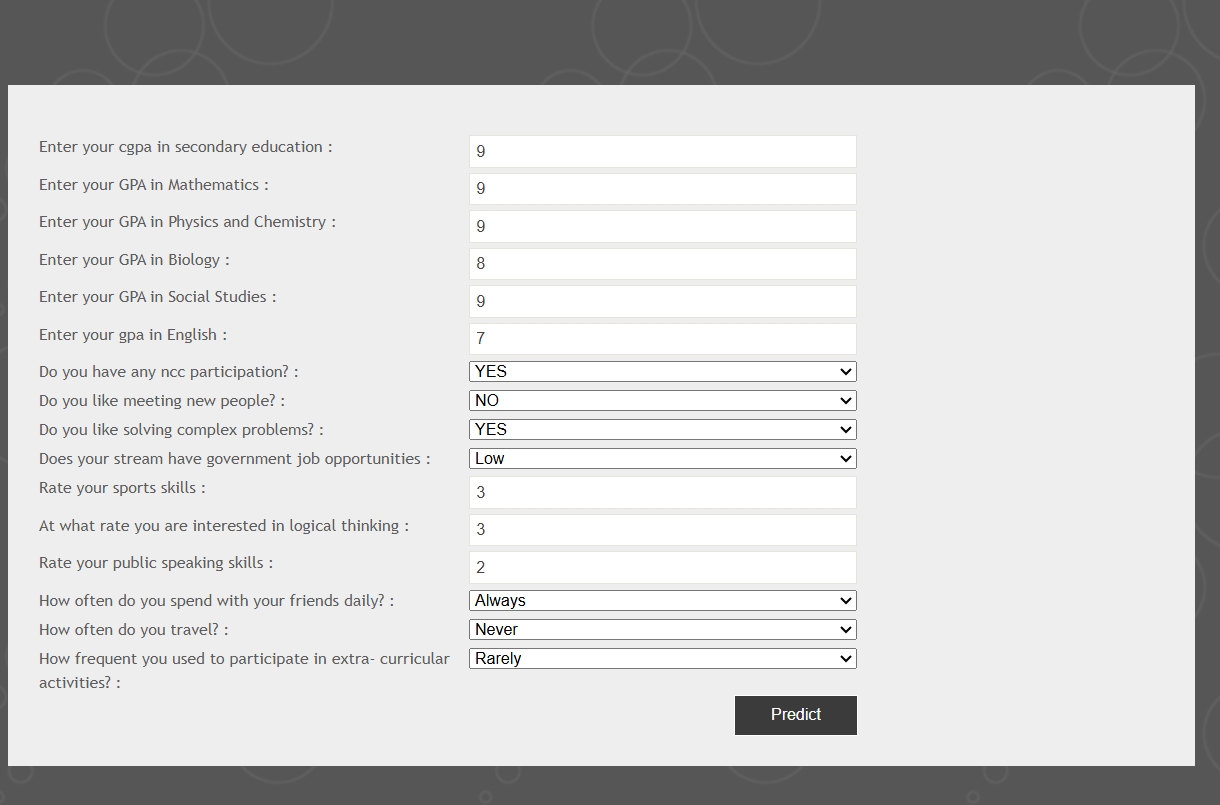
Black Box Testing is testing the software without any knowledge of the inner workings, structure or language of the module being tested. Black box tests, as most other kinds of tests, must be written from a definitive source document, such as specification or requirements document, such as specification or requirements document.

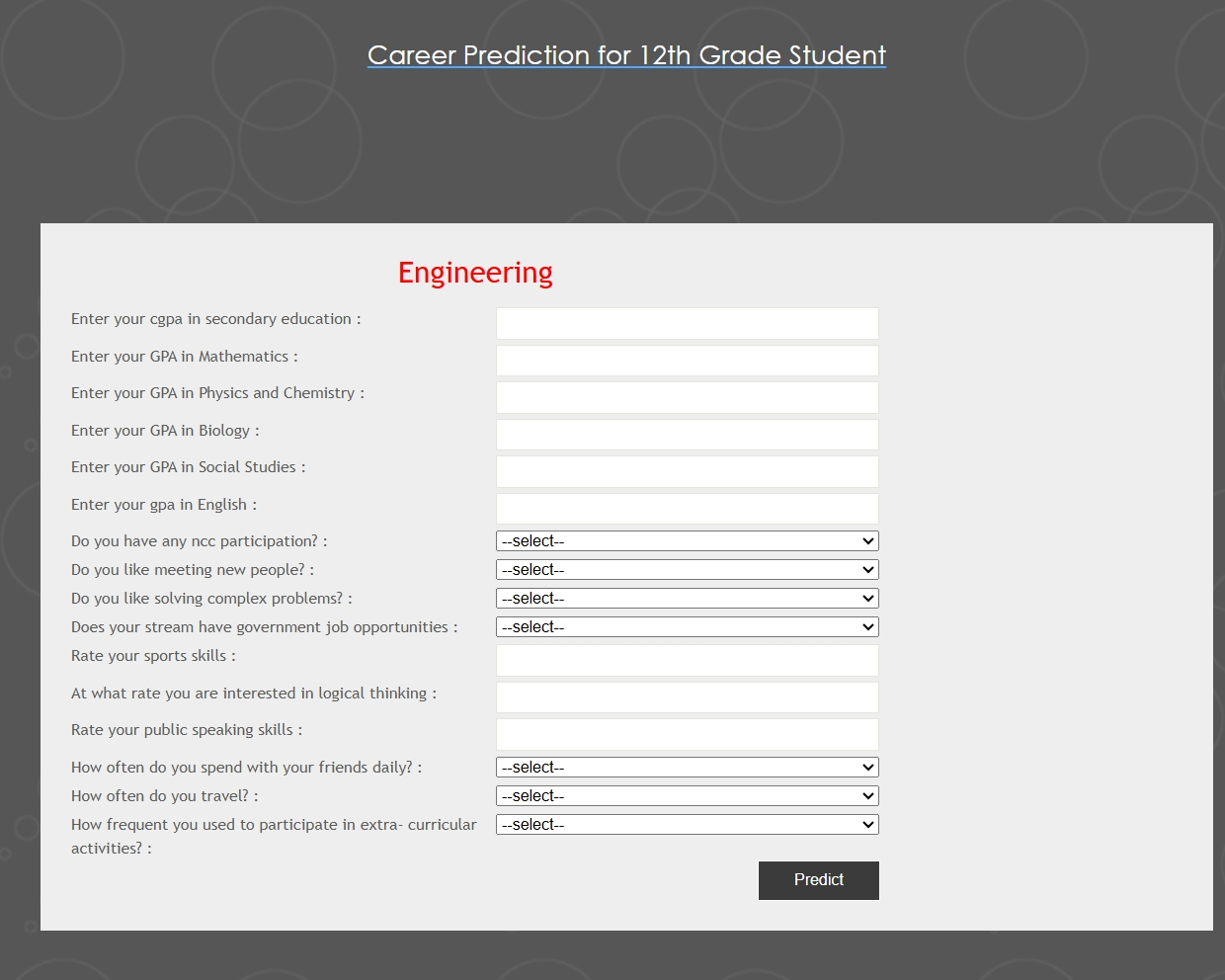
**Test cases**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Tested** | **Test name** | **Inputs** | **Expected output** | **Actual Output** | **status** |
| 1 | Load Dataset | Csv file | Read dataset | Load dataset | success |
| 2 | Split dataset | Train80% and test20% | Divide the training set and Testing set | Split train and Test | success |
|  | Train Model | Train dataset, random value, predicted class | Train with best accuracy | Train with best accuracy | success |
| 4 | Validate Model | No .of Epochs | Validate the Model with best fit | Model Generated | success |
| 5 | Predict accuracy and Error Rate | Accuracy | Plot expected accuracy and predicted accuracy | Plot expected predicted accuracy | success |
| 6 | Test Data | Test column | Predicted accuracy | Predicted accuracy | success |

1. **OUTPUT SCREENS**

****





1. **CONCLUSION**

In the system, we have designed and developed a web-based application for a career guidance system which provides suitable recommendations for a candidate in choosing an appropriate department. The recommendation provided in the proposed system is more accurate than the existing career guidance system. We have used the K-Nearest Neighbor algorithm to classify the skill sets of the candidate and predict a suitable department with respect to the performance of the candidate and we have also used K-Means Clustering algorithm and the clusters formed is by splitting the students' scores of the particular skill set and determining the rate of success for various departments in every cluster. The rate of success in each of the clusters is calculated and that will be used for department recommendation purposes where there will be a higher success rate and a low failure rate. In this project, the career guidance system has been researched thoroughly and then designed and developed a web-based application with expected outputs. In the near future, the framework's accuracy rate will be enhanced and additional features will be used for recommending a suitable department and also the outliers of the framework will be removed gradually.

1. **FUTURE ENHANCEMENTS**
2. **Real-time Mentorship Programs**

* **Smart Mentor-Student Matching** using AI to align students with mentors based on goals, interests, and communication style.
* **Instant Messaging & Live Sessions** within the platform for seamless communication.
* **Mentorship Analytics** to track interactions, goals achieved, and student engagement.
* **Virtual Internship Referrals** facilitated by mentors.
* **Multi-language Support** for mentors and students across regions.

1. **Career Path Roadmaps**

* **Dynamic Roadmaps** tailored to the student’s assessment results and updated as they progress.
* **Industry-Aligned Milestones** that reflect real-world expectations and emerging trends.
* **Interactive Timeline Interface** showing short-term and long-term career goals.
* **Integration with MOOCs** like Coursera/edX to recommend courses at each stage.
* **Alerts & Reminders** to keep students on track with their roadmap.

1. **Ai-Based Career Suggestion Engine**

* **Machine Learning-Driven Insights** from historical data to improve suggestions.
* **Real-Time Industry Demand Mapping** to align career suggestions with market trends.
* **Behavioral Pattern Learning** to refine suggestions over time based on user interaction.
* **Feedback Loop** for students to rate suggestions, enhancing algorithm accuracy.
* **Multi-Domain Compatibility** supporting STEM, arts, commerce, and emerging fields like AI, robotics, and design.

1. **REFERENCES**

[1]. A.M. El-Halee's, and M.M. Abu Tair, “Mining educational data to improve students’ performance: A case study,” International Journal of Information and Communication Technology Research, 2011, pp. 140-146.M. tech Er. Rimmy Chuchra. " Use of data mining techniques for the evaluation of student performance: a case study ". International Journal of Computer Science and Management Research, 1(3):425-433, October 2012.

[2]. Muhfuza Haque Md. Hedayetul Islam Shovon. " Prediction of student academic performance by an application of k- means clustering algorithm ". International Journal of Advanced Research in Computer Science and Software Engineering, 2(7):353-355, July2012.

[3]. S. Karthik M. Sukanya, S. Biruntha and T. Kalaikumaran, “Mining Data mining: Performance improvement in education sector using classification and clustering algorithm,” ICCCE In Proceedings of the International Conference on Computing and Control Engineering, 2012.

[4]. Brijesh Kumar Bhardwaj and Saurabh Pal. Data mining:" A prediction for performance improvement using classification & (IJCSIS) International Journal of Computer Science ^d\_ Information Security, 9(4), April,2011.

[5]. H. Anandakumar and K. Umamaheswari, “An Efficient Optimized Handover in Cognitive Radio Networks using Cooperative Spectrum Sensing,” Intelligent Automation & Soft Computing, pp. 1-8, Sep. 2017. doi: 10.1080/10798587.2017.1364931

[6]. Haldorai, A. Ramu, and S. Murugan, “Social Aware Cognitive Radio Networks,” Social Network Analytics for Contemporary Business Organizations, pp. 188-202. doi:10.4018/978-1-5225- 5097-6.ch010 [7]. H. Anandakumar and K. Nisha, “Enhanced multicast cluster-based routing protocol for delay tolerant mobile networks,” International Journal of Information and Communication Technology, vol. 7, no. 6, p. 676, 2015.

[8]. Md. Hedayetul Islam Shovon and Mahfuza Haque." An approach of improving student’s academic performance by using k-means clustering algorithm and decision tree". (IJACSA) International Journal of Advanced Computer Science and Applications, 3(8):146-149, August,2012.

[9]. T. M. Mitchell. Machine Learning. McGraw-Hill Companies, New York, USA, 1997, ISBN 0-07-042807-7.

[10]. Jhan and M. Kamber.Data Mining. Concepts and Techniques, Simon Fraser University, Morgan Kaufmann publishers, ISBN 1­ 55860-489-8,2001.

[11]. ShiNa, Shi Na, Liu Xumin, and Guan Yong. Research on k-means clustering algorithm: An improved k-means clustering algorithm. In Intelligent Information Technology and Security Informatics (IITSI), 2010 Third International Symposium on, pages 63-67, 2010.

[12]. Powes, D.M.W,"Evaluation: From Precision, Recall and F-measure To Roc, Informedness, Markedness & Correlation “, Journal of Machine Learning Technologies, ISSN: 2229-3981 & ISSN: 2229- 399X, Volume 2, Issue 1, pp-37- 63, 2011.

[13]. Holmes, Sr., O. W. (1858). The autocrat of the breakfast-table. Cambridge, MA: Houghton, Mifffflfflin.

[14]. Horan, J. J. (2010). The Virtual Counseling Center: Its niche, resources, and ongoing research and development activity. Journal of Career Assessment, 18, 328-335.

[15]. Hornyak, D. A. (2007). Utilizing cognitive information processing theory to assess the effectiveness of DISCOVER on college students’ career development. Dissertation Abstracts International Section A: Humanities and Social Sciences,68(6-A), 2319.Iaccarino, G. (2001).