International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 04 | April - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Career Guidance Launch Pad

Mr. Aditya Sabne
Department of **Artificial Intelligence & Data Science**
New Horizon Institute of Technology and Management
Thane(w) 400615, India
adityasabne216@nhitm.ac.in

Miss Gayatri Sarang
Department of **Artificial Intelligence & Data Science**
New Horizon Institute of Technology and Management
Thane(w) 400615, India
gayatrisarang216@nhitm.ac.in

Mr. Yash Nikam
Department of **Artificial Intelligence & Data Science**
New Horizon Institute of Technology and Management
Thane(w) 400615, India
yashnikam216@nhitm.ac.in

Prof. Shubhangi Soni
Department of **Artificial Intelligenc & Data Science**
New Horizon Institute of Technology and Management
Thane(w) 400615, India
Shubhangi.soni1210@gmail.com

Abstract— Choosing the right career is one of the most crucial decisions in one's life. This decision should not be influenced by anyone's pressure and should be mainly your own choice. Selecting the right career ensures that you gain financial growth over time which eventually leads to good well-being of an individual and hence a content society. At present in India, we have more than 250 career options available but according to a survey conducted by "India Today" in 2019, 93% of the student population of the country is familiar to hardly seven career options, as majority of the students and parents generally choose most conventional careers like Engineering, Medical, Law, Civil services, Management, Pharmacy etc. The reason for these appalling statistics is the lack of awareness and education of the varying professions available amongst the youth. Eventually pursuing a career which is not of your choice can lead to poor job satisfaction, lack of skills and an unsound mind. Therefore, taking the right decision at the right time is important. This is where Carrer Launch Pad comes to your rescue. Carrer Launch Pad is a web-based career guidance system which analyzes the profile of its user by conducting personality, technical tests and suggests diverse career options to its users. Our service does not stop at this point, we provide multiple informative material on the job opportunities available in the industry. Along with this we also display various skill building courses, projects, internships, job opportunities, startup options which can enhance your profile and increase your chances of getting employed. Our web application makes use of multiclass classification algorithms to predict the careers of an individual. After analyzing various multiclass algorithms like K nearest neighbor, SVM, Random Forest, One vs rest we implemented Random Forest because of being the most accurate algorithm. At present this system works for the SSC category of users and under-graduates (Information Technology and Computer Science domain). Our aim is to extend this system in all domains, make it accessible to all and become the one stop solution to all your career needs.
Keywords— career guidance system, multiclass classification,
personality test, Howard gardener test, unemployment.

I. INTRODUCTION

Due to the recent advancements in technology, the job sector requirements have changed drastically compared to the past 10 years. Many sectors have experienced consistent unemployment due to various factors such as overpopulation, illiteracy, automation, industrial conflict, seasonal nature of agriculture, decline of handicrafts and village industries. The root cause of this problem is a demand-supply gap in the employment sector. According to the recent CMIE Report, the unemployment rate in urban areas increased to 10.09% in December 2022 from 8.96% the previous month. India's unemployment rate has historically been high and it is predicted to rise in the next few years. According to the estimate, over the next four years, 10 million more people will join the ranks of the unemployed, bringing the total to 220 million. The government had also forecasted that there will be 1 million additional job vacancies over this time period, but the number of jobs would not be sufficient to counteract the population expansion in general. Some examples include low job availability in a certain sector due to oversaturation or a lack of skilled workers due to the majority of candidates not having the proper skills in a certain sector. Other factors include cyclical unemployment, which is caused by economic downturns or is related to changes in business conditions that affect the demand for workers. This leads to a vicious cycle where the fresher candidates with job experience have difficulties in getting employed and due to this unemployment, they are unable to gather sufficient work experience.

Our system aims to identify an individual's interests and natural instincts at an early stage and suggest a broad spectrum of career options to SSC and HSC students. This will make them aware of a wide variety of career options to choose from and also lead to the decentralization of the industry. The front end of our application is built using HTML, CSS, Bootstrap framework and uses Flask API. Flask helps to create a web server that helps to further run HTML templates. The users will be evaluated based on standard test procedures followed by Howard Gardner's Multiple Intelligence (Personality / Psychometric) Test and Skill Test containing technical questions from their respective domains. Interconnectivity between frontend and backend has been established easily with the help of the SQLite database, as it is the most popular used database which helps to store unstructured data (images, files, videos, etc.) and structured data together.

II. LITERATURE REVIEW

In May 2022, Ghimire et al. [1] published a paper titled "Introspection with Data: Recommendation of Academic Majors Based on Personality Traits [1]". The objective of their project was to help students find their most enjoyable academic path and aid guidance counsellors and parents with their recommendations. They used the Big Five Personality Trait index for classifying the user and better understanding their personality type. They surveyed over 500 people between 18 and 25 years old to capture personality traits and preference of college major and used that information to train a machine learning model to predict college major preference. They reduced the dimensions by half without decreasing the accuracy of the classifier by using two methods: Principal Component Analysis (PCA) and

the Social Science research on the Big-Five personality Traits (also known as OCEAN indices). By comparing different machine learning algorithms, they concluded that a decision tree model gave the most accurate classification with over 90% accuracy.

On 31 March 2022, Pandey and Maurya[2] published a paper on 'Career Prediction Classifiers based on Academic Performance and Skills using Machine Learning'[2]. Their

approach focuses on selecting the right courses for their career based on their academic interests and abilities. Various features such as student's mark percentage in 10th, 12th, B.Tech/Diploma, skills in communication etc., were taken into consideration, and the output variable was career options a student can choose. They proposed six machine learning algorithms and achieved the accuracy as follows: K - Nearest Neighbor > Gaussian Naive Bayes > Stochastic Gradient Descent > Support Vector Machine > Decision Tree > Logistic Regression > Neural Network > Random Forest.

In October 2021, Matheus and Takigawa[3] published a paper titled 'Evaluation of Machine Learning Algorithms for Multiclass Classification of Voice Calls from Power Systems Operations'[3]. In this paper the Brazilian government has used multiclass classification algorithms like Naive Bayes, Support Vector, K-nearest-neighbors, Multi-layer perceptron and Random Forest to categorize the communication between the operators into different classes which are used for analysis. They have manually labelled the input records into 12 classes such as proof of availability, time confirmation etc. The verbal conversations between the operators are recorded and preprocessing operations are performed. The multiclass algorithms are applied on the dataset and their accuracy is compared. Multi-Layer Perceptron provides the highest accuracy of 85% on the dataset. The authors have concluded by proposing that addition of more samples in the dataset can prevent overfitting and give more promising results.

In March 2021, Vignesh et al. [4] published a paper on "An Intelligent Career Guidance System Using Machine Learning"[4]. The research paper was based on a computerized career counselling system that was used to predict the suitable department for an individual based on their skills as assessed by an objective test. The target audience was classified into three types: 11th, 12th, and engineering students, in which they have focused their interest by recommending a specific department in engineering. For their targeted audience, the author has created three modules: the skill test assessment module, the prediction module, and the result analysis module. The skill test module was based on a combination of objective tests that included psychological and core skill-oriented questions. The prediction model was built with the help of KNN, SVM, Naive Bayes, and K-mean models; among them, KNN performed well with a 94.10 percent accuracy on the dataset they created using their methodologies. Moreover, the Kmean model was generally used to map each group of departments to the candidate's performance and to provide them with secondary and tertiary recommendations. The Result Analysis Module provides a detailed analysis of the candidate's performance and a suggested career in a systemic format to provide a quality understanding for the user.

In December 2020, Alsafy et al. [5] represented a paper on "Multiclass Classification Methods: A Review"[5]. In this paper the comparative study of different classification, Recognition, Diagnosis or Clustering algorithms and their limitations is given. They studied KNN, SVM, Decision tree, Naive bayes, C4.5, ANN, genetic algorithm, ANN, Fuzzy logic, k-mean, LDA, QDA. The data collected was both qualitative and quantitative. This result will provide features that are used to categorize, recognize, diagnose, or cluster data and perform features selected on it. They interpreted that different algorithms have their own advantages and disadvantages.

In December 2020, Hewage et al. [6] published a paper titled "Smart Human Resource Management System to Maximize Productivity" [6]. The end goal of their project was to build a smart human resource management system that reduces human judgment, time in the candidate selection process and predicts employee performance and attrition to motivate current employers to maximize productivity with minimal financial loss in the workplace environment. They used Object Character Recognition to extract raw data from employee resumes and applied unsupervised learning algorithms and natural language processing for the resume classification and skill assessment. They predicted the employee performance and attrition using supervised machine learning classification techniques and the result of the final performance was generated as a score for each employee.

In September 2019, Akkaya and Çolakoğlu [7] represented a paper on 'Comparison of Multi-class Classification Algorithms on Early Diagnosis of Heart Diseases'[7]. In this research paper their approach focuses on analyzing multi-class Classification Algorithms on Early Diagnosis of Heart Diseases. They proposed classification algorithms which are capable of to do multi-class classification like Logistic Regression, Gaussian Naïve Bayes, k-Nearest Neighbors, Support Vector Machines, Multilayer Perceptron, CART, Random Forest, Gradient Boosting Machine, Extreme Gradient Boosting. These algorithms were applied to solve many problems, including the diagnosis of heart diseases.

In August 2018, Anand et al. [8] presented a paper on "Automated Career Guidance Using Graphology, Aptitude Test, and Personality Test [8]," in which they helped users choose the appropriate profession based on their human behavior and analysis. To measure one's behavior and analysis, the author has developed an MCQ module, handwriting analysis, and result module. The MCQ module introduces aptitude and psychometric question sets to represent one's behavior. However, Myers-Briggs Type Indicator (MBTI) is used to classify psychometric traits. Thus, the result module displays careers obtained from handwriting analysis and the psychometric test results, which are mapped to the broad career domains obtained from the aptitude analysis. The accuracy obtained from the aptitude test result with the help of 120 testing sample datasets and 115 expected output datasets is a maximum of 95.83 percent. The psychometric test precision is 90.83 percent with the following 120 sample testing datasets and 109 expected output datasets. The handwriting analysis follows up to a 93.33 percentage score with the help of 530 sample testing datasets and 496 expected output datasets.

In April 2017, Alimam et al. [9] published a paper titled "Automated System for Matching Scientific Students to their appropriate Career Pathway based on Science Process Skill Model"[9]. The paper presents ideas on building a model for students in their high school and it is based on the Moroccan system of education (MES). They have classified the students on the basis of their grades and a skill test. In this system they have considered two situations: 1) where the student will have to study the same subjects during his next year and they have given the solution using RIASEC Based Model
Guidance. 2) Where the subjects will differ totally during the next year of studies and for this situation, they come with SPS (Science Process Skill) model. It allows the author's system to identify the real skills of students and represent the student's grades as a set of SPS vectors. Also, they have experimented the system on samples of 50 final year students of high school and they have calculated their affectation to choose the field on the basis of their grades and the field suggested by SPS model.

In February 2017, Min NIE1 et al. [10] published a paper titled 'Advanced forecasting of career choices for college students based on campus big data' [10]. This paper presents a study of identifying career choices based on the behavioral choices of students during their university course. They have utilized approximately 4000 records of college students based on various features like My. They conducted a comparative study of 4 classification algorithms on the data of 6 semesters. Random Forest algorithm gave the highest accuracy. Following this the authors also calculated the importance of each feature on the prediction of the output (career choices) by calculating the Micro F1 value of each feature. The author also concluded that the precision of the system increases as students reach their final year as they become more sure of their choices.

III. PROPOSED METHODOLOGY

The career guidance systems which are currently available in the market conduct multiple tests and provide the candidate with a career report which generally consists of the career options available for the user. But what these systems usually lack is the inclusion of diverse career options and the future employment chances present in that career. To provide a robust solution to this problem we provide our users with all the possible options of employment like Start-ups, government schemes, Internships, Higher Education to ensure that each individual gets employed. Carrer Launch Pad is an interactive career guidance web application tool aimed at helping students find their strengths, weaknesses, skillsets and suggest them a flexible future career path. The target audience of this software is divided into 4 categories of users which are School Students (14-16 years age group), Junior College (1618 years age group), and Undergraduates (18-22 years age group) and Recruiters. The users will be evaluated based on standard test procedures: the Howard Gardner's Multiple Intelligence (Personality / Psychometric) Test and Skill Test containing technical questions from their respective domains. There is a recruiter portal also available in our system which will connect the recruiters with the prospective under-graduate students registered on our platform.

The system will consist of the following modules:

3.1 Test Modules

3.1.1. Personality Test Module

Personality tests are used to evaluate an individual's characteristics: 1. their strengths/weaknesses 2. areas of interests 3. preferences and style of interacting with people and situations etc. These unique traits of a person are identified based on their responses to a set of questions. These questions are usually standard in format and do not have a definite right answer. It normally consists of a range of choices from which the candidate has to select the answer with which they agree the most. Some of the common tests are: Myers-Briggs Type Indicator (MBTI), Howard Gardner's Multiple Intelligence Test, 16 Personality Factor Questionnaire, DISC Test, Big 5 personality test etc.

Our target audience ranges from secondary students to under-graduates therefore to keep the test simple, comprehensible yet effective we choose the Howard Gardner's Multiple Intelligence Test.

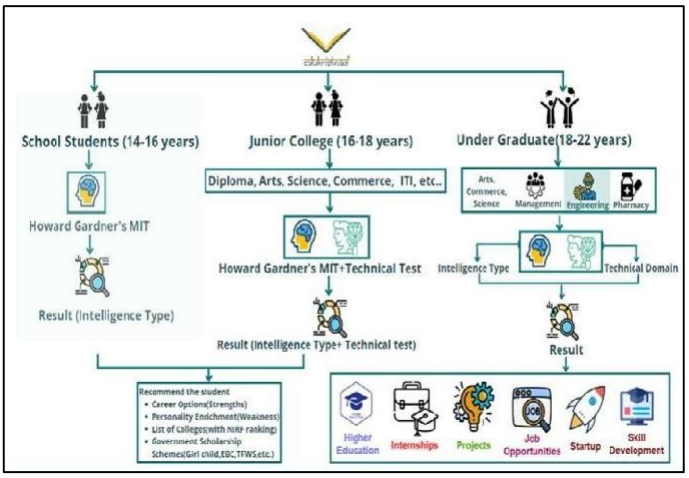


Fig 1: System Overview

Howard Gardner's Multiple Intelligence Test:

The theory of multiple intelligence was developed by Howard Gardner, an eminent psychologist, in 1983 in one of his books titled 'Frames of Mind'. Howard Gardner's theory of multiple intelligences proposes that humans are not born with a definite single type of intelligence but rather a person can be composed of multiple types of intelligences and each of them reflects a certain strength/liking the user has.

- Gardner introduced eight different types of intelligences consisting of: Linguistic, Logical/Mathematical, Spatial, BodilyKinesthetic, Musical, Interpersonal, Intrapersonal, and Naturalist.
- This theory also implies that an individual is not limited to a particular type of intelligence and

can possess multiple types of intellects. Each intelligence type can be associated with an occupation/career choice.

Referring to this standard model we have presented 5 questions of each intelligence to the user on our interface. The example below will illustrate the format of the questions: This use of this model for the different category of users is stated below:

A) SSC Users: The 8-10th grade students are unaware of the technical aspects of any career, hence to provide them with more generalized career options our system will test the students on their natural abilities through this personality test and suggest them with multiple career options based on their intelligence type. We have also ensured the resultant career option of a student is not restricted to only 1 intelligence. We have listed 36 sets in our system. Each set corresponds to a list of job options. Each job role has a specific intelligence type that a student needs to have as best (must needed), average and poor. If there are 5 job roles having the same group of intelligence types then they are clubbed into 1 set. In this manner we have 36 sets and 72 career options created.

Example: For example, a set has been created consisting of following career options - {Company Secretary, Lawyer, Financial Advisor, Business Manager}. All these job roles have their best intelligence as Linguistic; Average is Logical-Mathematical and Poor is Interpersonal. Hence, they are grouped together. The diagram below illustrates the concept stated above.

B) Other users: The result of the Personality test along with a technical test is used to assess and suggest flexible career options to this category of users.

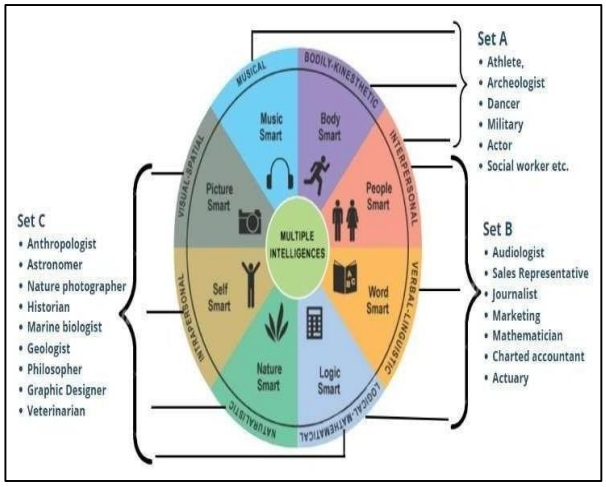


Fig 2. Creation of Sets

		Best	Average	Poor
Sr No	Job role	2	1	0
1	Astronomer	Logical - Mathematical, Naturalist,Spatial-Visualization, intrapersonal	Linguistic, Bodily, Interpersonal	Musical
2	Geologist	Logical - Mathematical, Naturalist, Bodily	Linguistic,Spatial-Visualization, Intrapersonal	Musical, Interpersonal
3	Marine Biologist	Naturalist,Logical - Mathematical	Bodily, Spatial-Visualization,Intrapersonal	Linguistic,Musical, Interpersonal
4	Veterinarian	Naturalist,Intrapersonal, Interpersonal	Linguistic, Logical - Mathematical,Bodily	Musical, Spatial-Visualization
5	Nature photographer	Naturalist, Intrapersonal,Bodily	Interpersonal, Linguistic	Logical - Mathematical,Musical,Spatial-Visualization
6	Recording engineer	Musical, Interpersonal, Logical - Mathematical	Linguistic, Intrapersonal	Bodily, Spatial-Visualization,Naturalist
7	Audiologist	Linguistic, Interpersonal, Logical - Mathematical	Intrapersonal, Naturalist,Musical	Bodily, Spatial-Visualization
8	Sound editor	Musical, Naturalist, Logical - Mathematical	Linguistic, Interpersonal, Intrapersonal	Bodily, Spatial-Visualization
9	Music teacher	Linguistic, Musical, Interpersonal	Bodily, Intrapersonal, naturalist	Logical - Mathematical, Spatial-Visualization
10	Actuary	Logical - Mathematical,Interpersonal, Intrapersonal,Musical	Linguistic, Bodily,	Naturalist, Spatial-Visualization,

3.1.2. Technical Test

The technical test is at present designed only for the Information Technology and Computer Science students. We have considered 6 major domains: Network and Infrastructure Management, Web Technology, AI ML, Data Science, IoT, Project Management, Security for designing the test. There are approximately 5-6 questions framed from each domain and these are shuffled and presented to the user on the UI of the website. The result of technical test (the highest scoring domain of the user) along with the personality test is combined and mapped to a job role. This test is utilized only for users pursuing/already completed an undergraduate course. The objective of a technical test is to assess a candidate's knowledge of a specific field or discipline. There will be different technical tests for evaluating comprehension of different domains. By incorporating personality test and technical test results, the candidate will be recommended with the suitable career option.

For example, if a student is technically proficient in Web Development and his/her personality trait is inclined towards a Naturalist then our system will suggest the user with various agricultural career/job opportunities that are available within this sector. The option of becoming an entrepreneur is also provided. The technical test is at present designed only for the Information Technology and Computer Science students. We have considered 6 major domains: Network and Infrastructure Management, Web Technology, AI ML, Data Science, IoT, Project Management, Security for designing the test. There are approximately 5-6 questions framed from each domain and these are shuffled and presented to the user on the UI of the website. The result of technical test (the highest scoring domain of the user) along with the personality test is combined and mapped to a job role.

3.2. Prediction Module

Machine learning algorithms, APIs, and datasets for training the machine learning model are only a few of the technologies used in the construction of this module. A machine learning algorithm operating at the web application's back-end performs the prediction in this module using the candidate's test results from the Test module.

3.2.1 Dataset:

The dataset used for the machine learning model is synthetic as there was no appropriate data available related to the core concept of this application. Since the system has three different categories of students, the dataset is different for each category and hence different machine learning models are built. Dataset for SSC students consist of 3400+ entries. There are eight different features available in the dataset which consist of scores of personalities such as linguistic, logical, spatial, bodily, musical, interpersonal, intrapersonal, and naturalist and there is one target label which gives the set as an output. Further, for the legitimacy of the dataset the mapping of eight different personalities and the job roles to be recommended are verified with the help of a career counselor.

Student	Linguistic	Musical	Bodily	Logical	Spatial	Interpersonal	Intrapersonal	Naturalist	sets
S1	11	6	14	16	19	11	18	19	set18
S2	12	6	13	16	19	11	18	19	set18
S3	13	5	12	16	20	14	18	19	set18
S4	14	5	11	20	17	14	18	19	set18
S5	11	5	14	20	17	11	19	19	set18
S6	12	6	13	18	17	12	19	19	set18
S7	13	7	12	20	17	12	19	19	set18
S8	14	7	11	20	20	12	19	19	set18
S9	11	7	14	16	20	13	19	19	set18
S10	12	7	13	16	16	13	19	19	set18
S11	13	6	12	16	16	13	19	19	set18
S12	14	6	11	16	17	13	19	19	set18
S13	12	6	14	17	16	12	19	19	set18

Fig 3. SSC Dataset

Likewise, for the UG category the following datasets are created:

- Third year: The dataset contains 1100+ entries which consists of scores of 8 personalities and 3 job domains - Network and Infrastructure Management, Web Technology, AI/ML and Data Science and these are the 11 features. There is one target label which gives the set as an output. A set is created by mapping dominant personality with strong technical domain and each set is mapped to a list of job roles. So basically, a set looks like {personality trait, best domain, list of job roles} and there are 12 such sets for this category. For example, {Logical, AI/ML and Data Science, (Data Scientist, Computer Vision Engineer, ML Engineer, Computer Vision and Machine Learning Researcher)} The dominant personality trait is Logical, strong technical domain is AI/ML and Data Science and job roles suggested are Data Scientist, Computer Vision Engineer, ML Engineer, Computer Vision and Machine Learning Researcher.

- Fourth Year: The dataset contains 1100+ entries which consists of scores of 8 personalities and 6 job domains - IoT, Security, Network and Infrastructure Management, Web Technology, AI/ML and Data Science and Project Management these are the 14 features. There is one target label which gives the set as an output.

3.2.2. Machine Learning Algorithms

The system used supervised machine learning to generate classifiers, because the dataset provided consists of both input and output or labeled dataset. We have used multiclass classification techniques while developing the machine learning models. Hence, there will be several class labels present in the dataset, where class labels are the sets. There are 36 class labels for the SSC category dataset and 12 class labels for undergraduate's dataset. So, the number of class labels may increase in the future, a multiclass classification technique is more appropriate for solving the problem.

The K-Nearest Neighbor algorithm is used to classify target values by estimating the distance between each neighbor using any equations such as Euclidean distance, Minkowsky, cosine similarity measure, chi square, and correlation. The KNearest Neighbor algorithm maintains all existing data and classifies new data points based on similarities. This means that when fresh data is generated, it may be quickly sorted into a suitable category using the K-NN method. There is no specific way to identify the best value for "K," therefore we must experiment with many values to get the best one.

A Naive Bayes classifier is a machine learning model that uses probabilistic learning to perform classification problems. The classifier's core is based on the Bayes theorem. In our system, it classifies whether the particular class label(set) is suitable for students, given the scores of his/her personality and technical test.

A Support Vector Machine primarily, it is used for Classification problems in Machine Learning. It creates the best line or decision boundary that can segregate ndimensional space into classes so that we can easily put the new data point in the correct category in the future. For multiclass classification, the same principle is utilized after breaking down the multi-classification problem into smaller

subproblems, all of which are binary classification problems.

The Random Forest is a classifier that contains several decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset. When using Random Forest for classification, each tree gives a classification or a "vote." The forest chooses the
classification with the majority of the "votes."

In One-vs-Rest(O-v-R) classification, for the N-class instances dataset, we have to generate the N-binary classifier models. The number of class labels present in the dataset and the number of generated binary classifiers must be the same. Since the datasets contain 36 class labels for SSC, 12 class labels for undergraduate's category that creates the multiclass problem. In our system we have used One-vs-Rest to solve multiclass problems, which will predict the set based on student performance. The set is mapped with the corresponding personality and job role using python programming. After implementing the above algorithms on the SSC dataset we have the performance shown by algorithms as follows : K-Nearest Neighbor - 96.21, Naïve Bayes - 91.72, Support Vector Machine - 94.34 , One-v-Rest
- 96.72 , Random Forest - 95.45. Considering the highest performing algorithms Knn, One-v-Rest and Random Forest we performed separate testing on these algorithms and concluded that Random Forest could classify the maximum number of samples correctly. Therefore we have proceeded with the implementation of the Random Forest algorithm in our system.

```
#Accuracy of classifiers
#KNN Algorithm
from sklearn.metrics import confusion_matrix
from sklearn.metrics import accuracy_score
knn_predictions = classifier.predict(X_test)
score=accuracy_score(Y_test,knn_predictions)
print("K nearest neighbor = ",score)
#Naive Bayes algorithm
naive_score=accuracy_score(Y_test,gnb_predictions)
print('Naive Bayes = ',naive_score*100)
#SVM Algorithm
from sklearn.metrics import accuracy_score
svm_score=accuracy_score(Y_test,svm_predictions)*100
print('SVM = ',svm_score)
#One vs Rest algorithm
o_score=accuracy_score(Y_test,prediction)*100
print('One vs Rest = ',o_score)
#Random Forest algorithm
score=accuracy_score(Y_test,y_pred)*100
print('Random Forest = ',score)
```

```
K nearest neighbor =  95.7124842370744
Naive Bayes =  91.7255297679112
SVM =  94.34914228052472
One vs Rest =  96.72544080604534
Random Forest =  95.7124842370744
```

Fig 4 : Accuracy output of all implemented algorithms

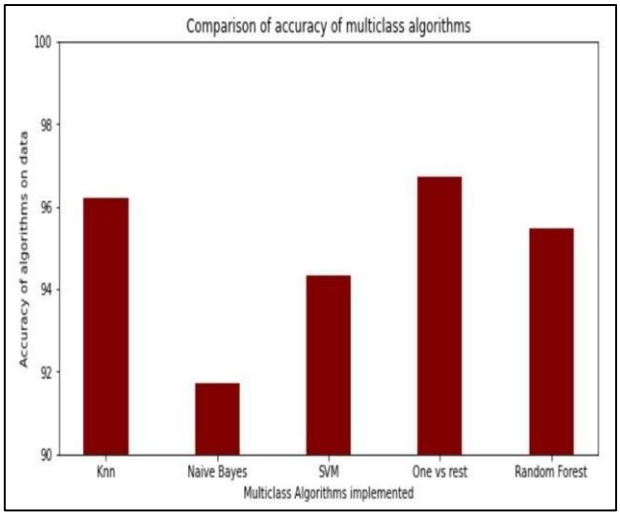


Fig 5. Bar graph to represent the accuracy each multiclass algorithms

3.2.3 Flask API

The system uses the Flask API (Application Programming Interface) to enable communication between the front end and back end. An application programming interface is used to send and receive requests and responses from the application. The machine learning model in our application receives the scores from the Test module via the flask API.

In this case, the request is to be assigned to a suitable personality trait and career option based on the candidate's performance. The response in this case is to give the projected personality trait and career option with knowledge about the candidate's performance. Python programming is used to implement this portion of the application programming interface.

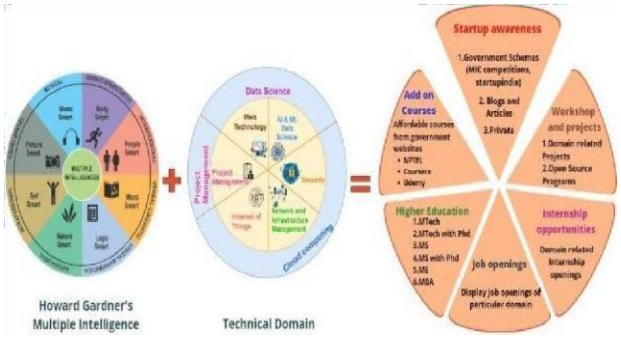
3.3. Results Module

Once the user completes the required tests on our platform they will be redirected to the results page. The distinctive factor of our system is the inclusion of multiple employment/skill building options that we provide to our users. The following features are provided to the users:

- Personality report: Based on the results of the Howard Gardner's Personality test the dominant personality traits of the user are highlighted along with their strengths and weaknesses, characteristics of personality etc.
- Career option: According to the result of the personality test (for SSC students) and personality + technical test(for other users) career options are predicted and presented to the user using a machine learning model.
- College Recommendation: The top-tier colleges for predicted career paths are displayed to the user, this data is collected from verified websites and colleges are arranged according to their NIRF ranking.
- Higher Education: For a under-graduate student, higher education options available in their particular major and the entrance exams information required to qualify are provided.
- Roadmaps: Along with the career option predicted for the user, information resources, roadmaps of job roles are also provided. For example the screenshot of our system's frontend below explains the roadmap of becoming a data scientist.
- Add on courses: In order to get employed the candidate should be equipped with appropriate skills. These skills can be developed by learning through courses, projects, internships etc. Trending courses hosted on nptel, swayam are linked on our platform using web scraping so that the user can navigate easily.
- Start-ups and government schemes: According to the survey conducted by Times of India the number of recognised start-ups have increased from 471 in 2016 to 72,993 as of June 2022. Start-ups are a crucial factor contributing towards self-employment as well as employment of other individuals. But the education/awareness regarding the start-up culture is quite less amidst the youth. Therefore our system enlightens on this point by providing information regarding start-up's and mentors on the website.
- Internships and Job opportunities: Job openings posted on sites like Internshala, Naukri, Indeed etc are scraped and displayed o the user.

Resume Builder : The users are provided with an option of building their resume on our website. This resume can be downloaded in pdf format and uploaded for the various job postings. The information which our system has collected from the user is alre ady displayed in the resume along with the personality strengths and technical robustness generated from their test result.

Fig 6. Result Module recruiting companies and the talent



acquisition teams will be able to access student data provides by Carrer Launch Pad.

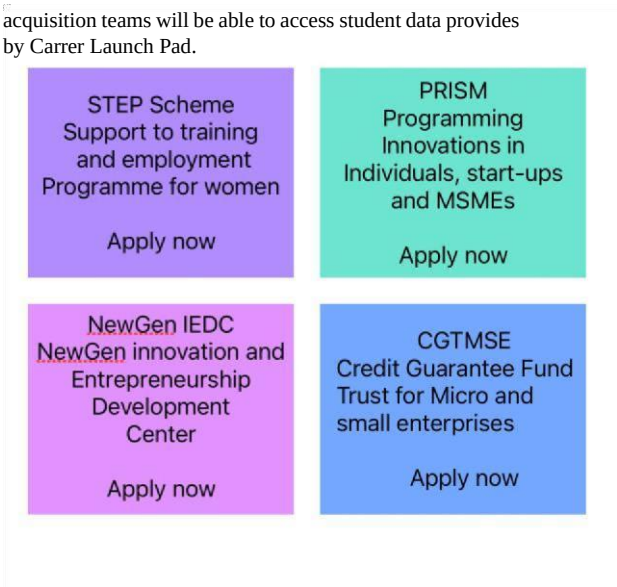


Fig 7. Screenshot of Start-up and government schemes.

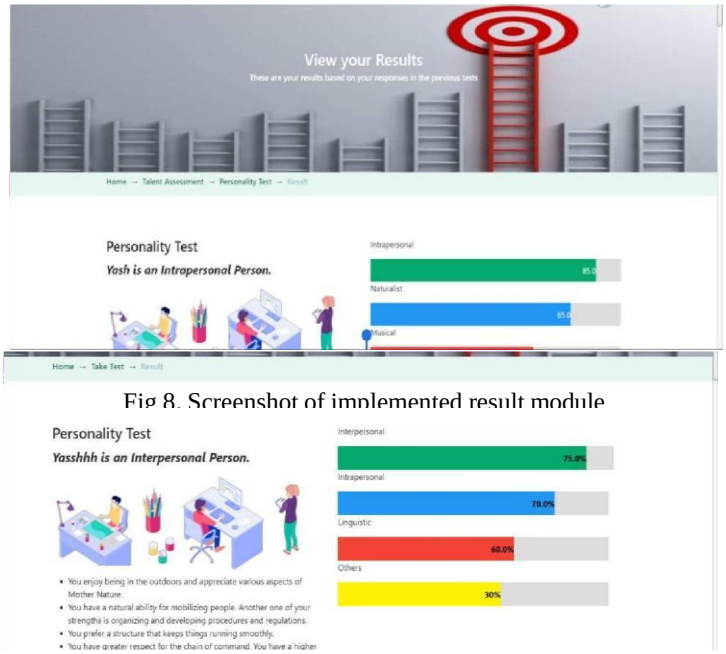
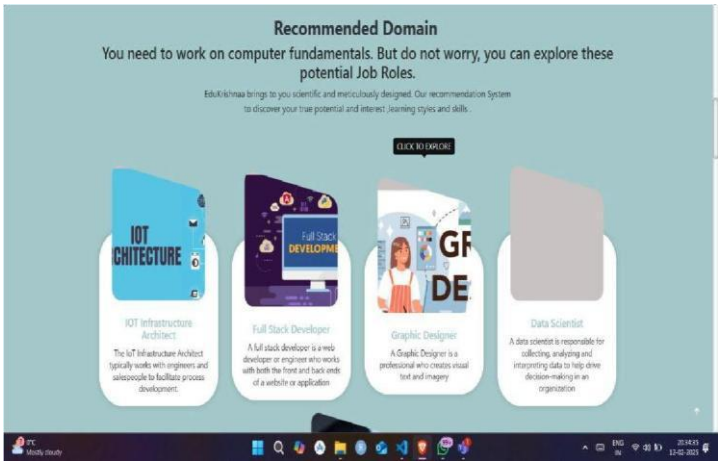
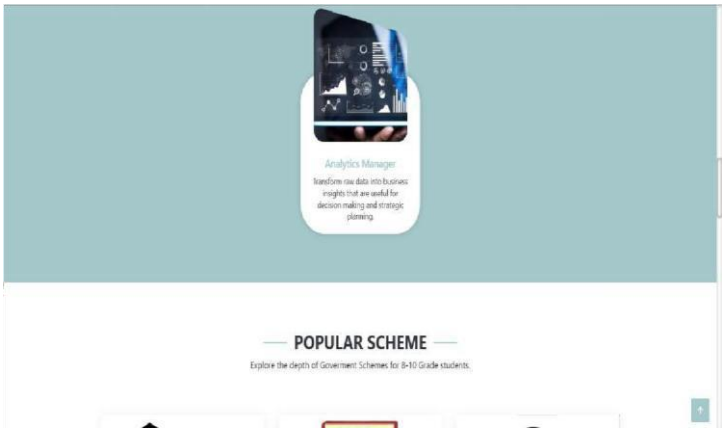


Fig 8. Screenshot of implemented result module

3.4. Recruiter Module

Our system connects the under -graduate students registered on our platform to the recruiters by allowing both the audience of users to see each other 's data meaning the users of carrer launch pad will be able to see the job openings posted by the recruiting companies and the talent acquisition teams will be able to access student data provided by Carrer Launch Pad . Using this method we are trying to link the talented candidates to the employing organizations.





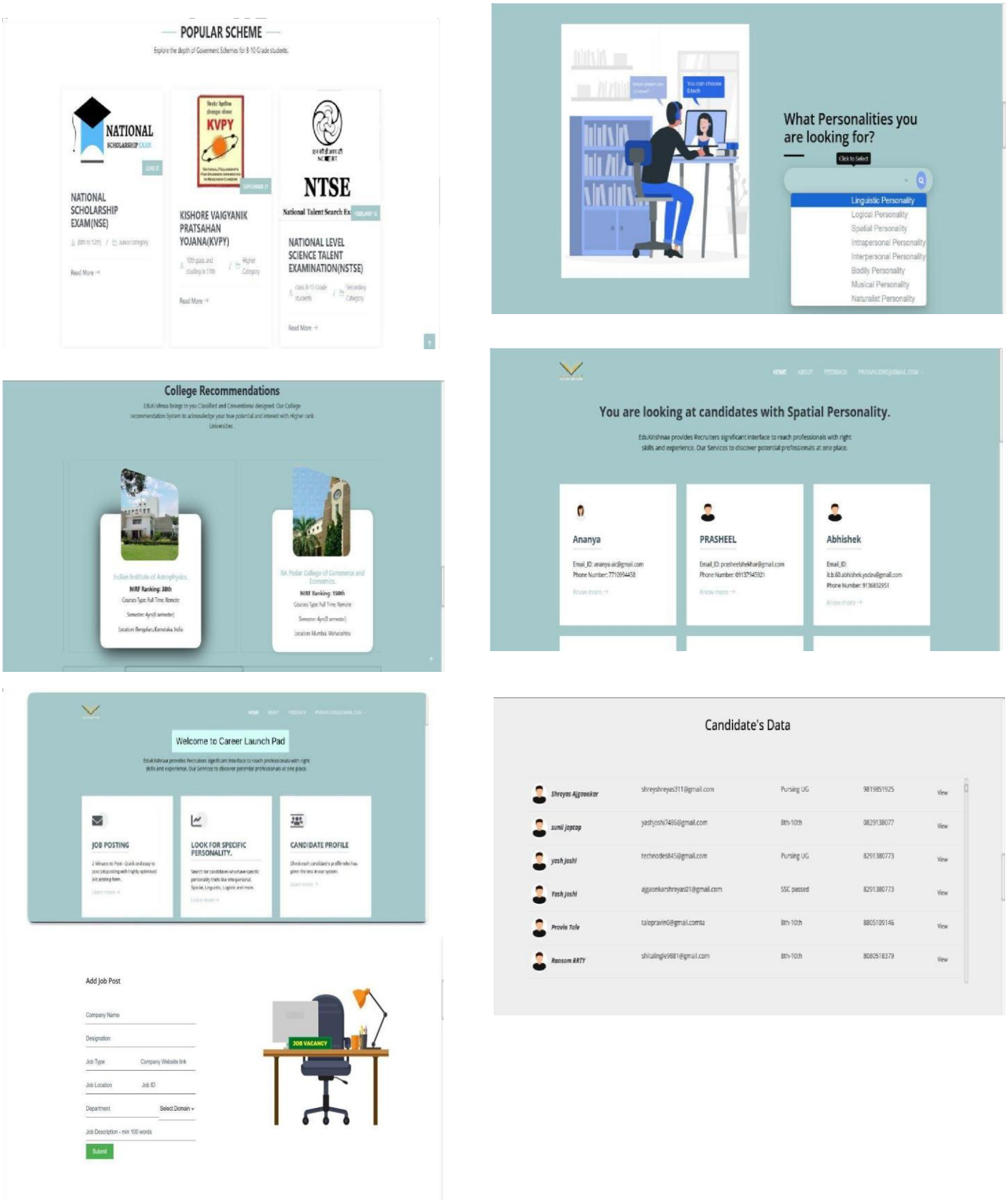


Fig 9. Result Screenshot

IV. Conclusion

In future the need for carrer guidance systems is going to increase with the exponential growth in carrer options and job opportunities. Educating our youth about the various options available along with their p roper information is of utmost importance, and this should be accessible to all. This project is still under implementation. At present our s ystem is working for the following category of users (14-16 years age group), Undergraduates (18-22 years age group) and Recruiters, and for the undergraduate category we have considered the Computer Science and Information Technology domain.

V. Future Work

Currently we have implemented this project on local machines. In the near future we will host this project on cloud and collect real time data from students, so that we can improve the accuracy of our system and alter the features/functionalities based on their feedback. Our aim is to include all the categories of career options available like Management, Science, Arts etc. For this we need expert guidance of career counselors on the type of technical questionnaires that should be designed. We also want to work with multiple recruiting companies who can post their job openings regularly on our platform. We want this platform to be accessible to all. This project has the capability of being converted into a business model with the integration of recruiters, startup mentors, career counselors and that is exactly our ultimate goal.

References

[1] Aashish Ghimire, Travis Dorsch, John Edwards, “Introspection with Data: Recommendation of Academic Majors Based on Personality Traits”, Intermountain Engineering, Technology and Computing (IETC), May 2022, ISBN:978-1-6654-8653-8

[2] Akanksha Pandey, L S Maurya, “Career Prediction Classifiers based on Academic Performance and Skills using Machine Learning” SSRG International Journal of Computer Science and Engineering, Volume 9 Issue 3, 5-20, March 2022, ISSN: 2348 – 8387.

[3] Matheus N. S. M. de Lima, Fabrício Y. K. Takigawa, “Evaluation of Machine Learning Algorithms for Multiclass Classification of Voice Calls from Power Systems Operations”, 7th International Conference on Engineering and Emerging Technologies (ICEET), 27-28 October 2021, Istanbul, Turkey.

[4] Vignesh S, Shivani Priyanka C, Shree Manju H, Mythili K, “An Intelligent Career Guidance System using Machine Learning”, International Conference on Advanced Computing & Communication Systems (ICACCS), March 2021, ISSN: 2575-7288.

[5] Baidaa M Alsafy, Zahoor M. Aydam, Wamidh K. Mutlag,“Multiclass Classification Methods: A Review” International Journal of Advanced Engineering Technology and Innovative Science (IJAETIS),Volume 5, Issue 3, December 2020, Page No: 01-10, ISSN:2455-1651.

[6] Hewage H.A.S.S, Hettiarachchi K.U, Jayarathna K.M.J.B, Hasintha K.P.C., A. N. Senarathne, J. Wijekoon, “Smart Human Resource Management System to Maximize Productivity”, International Computer Symposium (ICS), December 2020, ISBN:978-1-7281-9255-0.

[7] Berke Akkaya, Nurdan Çolakoğlu “Comparison of Multi-class Classification Algorithms on Early Diagnosis of Heart Diseases” SBIS Young Business and Industrial Statisticians Workshop on Recent Advances in Data Science and Business Analytics, September 2019, ISBN:978- 6055005-95-5.

[8] Apoorva Anand, Mrs. Dhanashri Patil, Shivani Bhagwat, Suyog Karanje, and Vihang Mangalvedhekar, “Automated Career Guidance Using Graphology, Aptitude Test and Personality Test”, International Conference on Computing Communication Control and Automation (ICCUBEA), August 2018, ISBN:978-1-5386-5257-2.

[9] Mohammed Abdellah Alimam, Hamid Seghiouer, Mohammed Amine Alimam, Mohammad Cherkaou “Automated System for matching Scientific Students to their appropriate Career Pathway based on Science Process Skill Model”, IEEE Global Engineering Education Conference (EDUCON), April 2017, ISSN: 2165-9567.

[10] Min NIE1, Lei YANG1, Jun SUN1, Han SU1, Hu XIA1, Defu LIAN1, Kai YAN “Advanced forecasting of career choices for college students based on campus big data”, February 2017.