SMART CAREER GUIDANCE SYSTEM

KALIDINDI LAKSHMI CHATURYA[1], VATTI AVINASH REDDY[2], ESLAVATH GOPI[3], MALYALA GAYATRI[4], MRS. A. SWAPNA[5]

[1]*Department of CSE-AIML, Sreyas Institute of engineering and technology, Hyderabad, India*
[2]*Department of CSE-AIML, Sreyas Institute of engineering and technology, Hyderabad, India*
[3]*Department of CSE-AIML, Sreyas Institute of engineering and technology, Hyderabad, India*
[4]*Department of CSE-AIML, Sreyas Institute of engineering and technology, Hyderabad, India*
[5]*Assistant Professor, Department of CSE-AIML, Sreyas Institute of engineering and technology, Hyderabad, India*

Abstract

Choosing the right career is often confusing, when there are so many options available in these days. Many students and job seekers are not sure about which path is suitable to them the best because they are not completely aware of their skills and strengths. This project aims at solving this problem by developing a Career Guidance Recommendation System that gives personalized and suitable career suggestions.

The system takes simple inputs from the user such as level of education, interests, and basic skills like coding, communication, and logical thinking. These inputs are then analyzed using a machine learning model called Random Forest algorithm. The system compares the user's profile with real-world career data through a custom-made dataset and suggests the top three suitable career options along with confidence scores.

Along with recommendations, the system also highlights areas where the user can improve through a simple skill gap analysis. The application is designed as a web-based system, making it easy to use and accessible. Overall, it helps users make better and more informed career decisions. To provide a solution to this problem, the system proposed provides a more structured assessment and a data-driven system for providing career guidance.

Keywords: career recommendation, random forest, confidence scores, skill gap analysis, web-based application.

1. Introduction

Finding the right job in today's rapidly changing world is even tougher than ever before. With so many different options to choose from, many individuals have difficulty finding a job that matches both their abilities and long-term goals. However, most people do not use specific, detailed, up-to-date data to help them find a career to pursue, instead, they are usually influenced by generic advice from family, friends or online sources, all of which may not be very accurate, personalized, or reflect the most current career information available. As a result, many people make poor decisions about their careers by choosing jobs based upon their general skills, and the outcome is that they become disappointed and stagnant in their careers.

Most traditional career guidance techniques use simple questionnaires and/or manual guidance and assessment that provide limited depth of understanding and don't include all relevant details and data to provide an overall accurate assessment of a person. Many traditional assessment tools do not include all relevant components of a person, such as; the level of their technical skills, their communication style or ability, their ability to logically think through a solution and their personality in the decision-making process. Very rarely do traditional career assessment approaches provide actions and/or suggestions on how someone could develop his/her abilities to support the career they want to achieve which leads to the gulf of separation between stated career aspirations and actual ability.

To provide a solution to this problem, the system proposed provides a more structured assessment and a data-driven system for providing career guidance. To create a precise profile of each user, the system gathers detailed user information (for example, education level, personality type, areas of interest, and key skills) – coding ability, communication and logical reasoning – that demonstrate user competence and expertise. These data points are then used by the machine-learning model to develop an understanding of the user's overall characteristics.

The heart of the system is a machine-learning model that uses Random Forest as its base classifier. The Random Forest model is used to determine the relationship between user characteristics and various career options. Random Forest has been selected to check user data and career paths because it generates the most accurate and reliable predictions, through the aggregation of multiple decision trees, of all available machine-learning classifications. The model will be optimized.

Using the results of the analysis, the system produces a list of recommended career options with a confidence score for each option. The confidence score represents how closely the options presented in the result list match the user profile. Presenting multiple recommendations rather than a single recommendation creates an opportunity to explore different career options, ultimately providing users with additional opportunities to make better-informed decisions. The skill gap analysis also identifies the specific skills that need to be developed for users to succeed in these recommended careers.

The system is a web application that is available and usable by all individuals as it does not limit its purpose to a specific group of people. When a user enters their information into the system, the user receives their recommendations in real-time; thus, allowing any individual to receive his/her career options without the need for having a background in computer technology.

The design will continue to focus on being easy to use, but will also continue to have elements of persuasive elements and provide meaningful ways in which the system will help students, individuals looking to start a new career, and other persons change careers or find better career opportunities.

In general, the objective of this project is to provide a better, more valid, and a more personalized career guidance solution through the combination of user profile analysis with machine learning technologies. As a result of this project, the hope is that all individuals will make better decisions about their careers, and that they will move toward their careers with confidence and clarity.

2. Proposed

2.1 System Overview

The suggested solution implements the Random Forest method to identify the best three career options that a person could pursue according to his or her educational qualifications, personal attributes, and skills. The algorithm is capable of examining the basic skills such as programming and logical thinking to provide recommendations based on the analysis of an individual's preferences. The system will also recognize any deficiencies and suggest improvements.

2.2 System Architecture

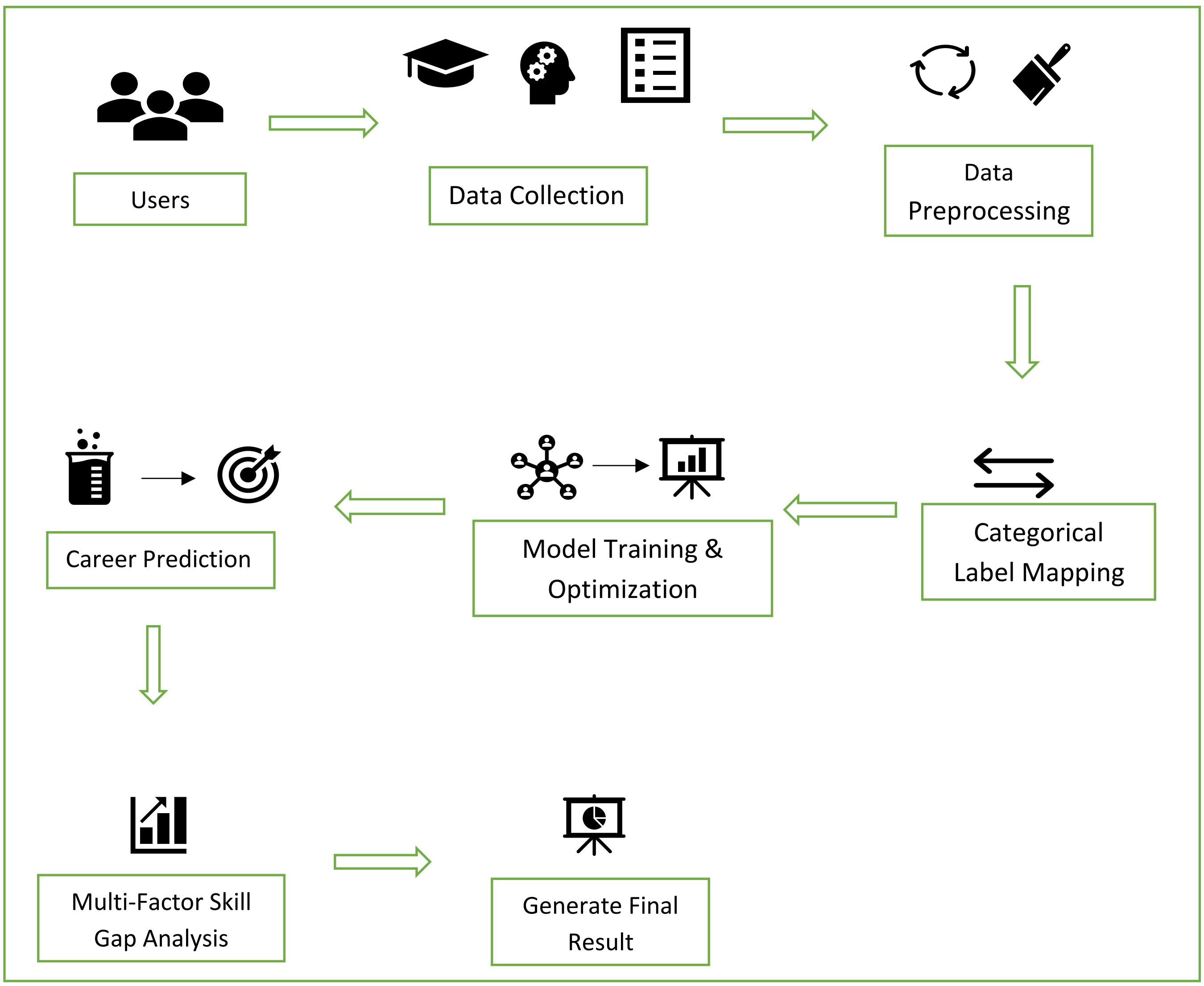

Figure 1: System Architecture for the Smart Career Guidance Platform, illustrating the end-to-end data pipeline from initial user input through Random Forest classification and subsequent skill gap analysis.

2.2.1 Data collection

Data collection is the most important and crucial stage for the machine learning project. The consistency and quality of the input data ensure the accuracy and effectiveness of the algorithm. In order to collect the data for Smart Career Guidance System, we use multi-dimensional data gathering approach. Unlike the traditional systems which collects the, we collect the data to predicted an accurate career recommendation. A number of factors are taken into consideration when predicting the career path including educational level, personality, interest and basic skills like coding, communication, and logical thinking. All of them are taken into account in order to predict an accurate career path.

2.2.2 Data Preprocessing

Making the data informative is just as important as gathering the data. Data will be gathered from several sources, and there can be a lot of undesired and erroneous data. The main purpose of data preprocessing is to clean the data and replace them with the appropriate data and removing null and missing data. Complete trash values may even be present in the gathered data. The data might not be in the proper format or manner. To make data relevant and helpful for additional processing, all such instances must be confirmed and substituted with alternative values. Data has to be stored in a certain format. When pre-processing data, all of these factors are taken into account.

2.2.3 Categorical Label Mapping

Sine the machine learning algorithms operates on the numerical logic, we have to convert the categorical data such as personality into the machine understandable language. This helps the algorithm to understand the categorical data easily and ensures the accuracy of the system. One-Hot Encoding is used to convert the categorical data into numerical logic.

2.2.4 Model Training & Optimization

The model training and optimization is the main computational procedure in the machine learning project. Here the pre-processed data is used to teach the machine learning algorithm. Random Forest algorithm is used to analyse the pre-processed data. It greatly improves the system's capacity and it is chosen because of its capability of handling high-dimensional data and provides a non-biased probability distribution for career recommendation.

To normalize the inputs, Min-Max Scaling handles them fairly because, the inputs are different from each other.

$$\boxed{x' = \frac{x - \min(x)}{\max(x) - \min(x)}}$$

Where *x* is original input value and *x'* is normalized value between 0 and 1

2.2.5 Career Prediction

Career Prediction is where the system predicted the accurate career paths based on the inputs given by the user. Here Random Forest is used to predict the career paths. Random Forest algorithm is the collection of decision trees. To achieve this, Ensemble Prediction (Majority Voting) is used. Ensemble prediction formula is used to combine the opinions of the hundreds of decisions trees in random forest and gives the final predicted result.

$$\hat{Y} = mode\{T_1(x), T_2(x), \dots\dots, T_n(x)\}$$

Where $\hat{Y}$ is the final predicted result and $T_n(x)$ is nth tree's prediction in the forest.

Because the random forest is collection of multiple decision trees, it handles the data splitting. Gini Impurity is used to measure the purity of a node.

$$Gini(D) = 1 - \sum_{i=1}^{C} P_i^2$$

Where D is dataset at the specific node, C is total number of career classes, P_i is probability of user being classified into a career.

2.2.6 Multi Factor Skill Gap Analysis

A prescriptive diagnostic layer called Multi-Factor Skill Gap Analysis assesses the difference between user competencies and industry standards. The multi factor skill gap analysis uses weighted feature attribution to identify the most important areas for professional growth by combining academic, technical, and psychometric criteria. This gives users a focused approach to achieve the full-fledged career readiness. Here the system calculates the confidence scores for a specific career c as:

$$Confidence\ Scores = P(c) = \left(\frac{\sum_{i=1}^{n} I(T_i(x)=c)}{n}\right) \times 100$$

Where P(c) is the probability for career c , n is the total number of decision trees in your forest (defined in your code as n_ estimators =100) and T is the prediction made by the individual tree for input x.

After calculating the confidence scores for every predicted career, the top three career recommendations are selected to suggest for the user. Here the Skill Gap Analysis is also provided for the user for the top one predicted career path. It helps the user to know where they need the improvement to reach their goal.

2.2.7 Generate Final Result

The concluding stage of a system's processing is generating the final result. This module guarantees that the final product is a useful path for professional development rather than merely a prediction by bridging the gap between raw data science and user-centric counselling. Here the final result is displayed with top three career recommendations and the skill gap analysis is also provided to ensure the improvement in their professional career.

2.3 Methodology

A multi-stage computational pipeline is used in the Smart Career Guidance System's technique to map high-dimensional student data to professional positions. The Random Forest algorithm technique allows the system to transition from conventional rule-based logic to a probabilistic model that can handle non-linear correlations between technical capabilities, personality attributes, and academic experience.

The system initiates to collect the data from the users. This includes educational level, personality, interests, coding, communication and logical thinking. It then undergoes in the Data Preprocessing to ensure the collected data is clear and has no null values. As mentioned in system architecture (2.2.4) to calculate the normalized input values, we use the min-max scaling.

One-Hot Encoding is used to categorical inputs in order to bridge the gap between human characteristics and machine logic. This creates binary vectors from qualitative labels. In order for the Random Forest to generate accurate node splits without losing the category detail necessary for career matching, this transformation is essential. It is an essential process because, machine learning algorithms operates on the numerical logics.

A Random Forest Classifier is used by the main engine because of its capacity to lower variance and avoid overfitting. An ensemble of independent decision trees is built in order to train the model. The system determines the most pertinent characteristic (e.g., Coding vs. Logic) for a particular career path by calculating the Gini Impurity at each node split inside the trees. As mentioned in system architecture (2.2.5) Gini Impurity is used to check the purity of the node of a individual decision tree.

The system uses Ensemble Prediction (Majority Voting) to accomplish robust categorisation. As mentioned in system architecture (2.2.5) the final predicted result is calculated by using Ensemble Prediction (Majority Voting). It is used to combine multiple decision trees outcomes present in the forest into final prediction. It ensures the most accurate career prediction.Here all the multi factors such as confidence scores and skill gap analysis are calculated by the system. To calculate the confidence scores for each predicted career path, as mentioned in system architecture (2.2.6), we use that confidence scores formula. It uses it to calculate the confidence scores and based on the confidence score. The user also gets the skill gap analysis for the top one career recommendation. This ensures the user to understand where they really need improvement in order to reach their desired career. This is the final stage of the system to perform data decryption, which transforms the statistical probabilities into user-centric assistance. The system turns a raw machine learning prediction into a useful career roadmap by ranking the Top Three career suggestions and skill gap analysis.

2.4 System Analysis & Performance

Both the conceptual framework and the effectiveness of the Random Forest model (Performance) must be examined in order to assess the Smart Career Guidance System**.** These helps the system in consistently work to generate the accurate career recommendations.

2.4.1 System Analysis

The analysis of the Smart Career Guidance System puts special emphasis on exactly how the system processes information to resolve career confusion. Essentially, this is like having many expert opinions rather than just one. The Random Forest algorithm uses the student's grades, technical skills, and personality traits as input and analyses them through hundreds of separate decision paths, thereby allowing for an accurate recommendation.

2.4.2 System Performance

The evaluation of a system's effectiveness is based on how accurate its output is as well as how useful it is in generating positive results for the user. The ability of the model to be accurate than traditional career tests is due to the use of trees as ensemble models that are combined together to generate a response. As with evaluating traditional career assessments, the model clearly demonstrates that it has a superior level of accuracy by dismissing noise or irrelevant information as well as focusing only on particular skills that truly matter in relation to a position.

Also, an important aspect of the evaluation is the value for money; or in other words, the match percentage. The Match Percentage is used to represent the Confidence Score, of the system; the higher the percentage the greater the level of confidence that the outcome matches the results and corresponding Skill Gap Analysis. While providing the user's name is valuable, the primary purpose of the system is to identify the Skills that the user possesses compared to the Skills necessary to perform the Requirements of the Position. As a result, the product performs well for individual employment assessments based on the user's requirements**.**

3.Literature Survey

The application of machine learning techniques in career guidance and academic recommendation systems has gained significant research attention in recent years. Previous studies have explored various approaches such as fuzzy logic-based recommendation systems, K-Nearest Neighbor (KNN), Support Vector Machines (SVM), Decision Trees, Random Forest, and hybrid machine learning models for predicting suitable career paths and academic courses for students.

A detailed review of the existing literature shows that while many approaches have improved the accuracy of career prediction and course recommendation. The following table summarizes the representative prior works relevant to the proposed system, highlighting their methods, key findings, and limitations. Table 3.1 presents the previous research carried out in this field.

Table 3.1: The table shows prior work in smart career guidance system

Authors (Year)	Method/Approach	Key Findings	Limitations
Prof. Priyanka Shahane, Prutha Rinke, Taniksha Datar, Soham Badjate (2022)	Machine Learning using SVM, Random Forest, Decision Tree	Predicts suitable career interests based on academic and non-academic performance	Mainly focused on student interest prediction; limited domain scope
Balqis Albreiki, Nazar Zaki, Hany Alashwal (2021)	Student performance prediction using ML techniques (RF, KNN, ANN, DT)	Helps identify weak students early and supports academic intervention	More focused on performance prediction than direct career guidance
Yinkai Wang, Aowei Ding, Kaiyi Guan, Shixi Wu, Yuanqi Du (2021)	Graph-based Ensemble Machine Learning	Improves prediction stability and increases accuracy by up to 14.8%	Computational complexity is comparatively high
Byung-Hak Kim, Ethan Vizitei, Varun Ganapathi (2018)	Deep Learning (BLSTM / GritNet)	Provides accurate early-stage prediction of student outcomes	Mainly designed for online learning environments
Mack Sweeney, Huzefa Rangwala, Jaime Lester, Aditya Johri (2016)	Recommender Systems + Random Forest + Factorization Machines	Predicts next-term student performance and supports academic advising	More academic-course focused than career-focused
Kanishk Mandrelia, Rahul Chauhan, Ananta Pandey, Anam Khan (2025)	Random Forest, Decision Tree, SVM	Predicts student placement and job roles based on skills and academics	Mainly focused on CS students and placement roles

4. Experimental Setup

The design of the experimental setup for the Career Guidance System is a controlled and scalable environment to test machine learning (ML) against real time web Interactivity. The hardware architecture consists of a standard computing environment using an Intel i3 processor or higher with a minimum memory capacity of 4GB. This allows efficient handling of multi-threaded Python operations and model training cycles. The system is compatible with all four version of Windows (10/11), Linux and macOS - approximately 500MB of free storage will be available for project files, virtual environments and the main data set.

From the software perspective, the project is built off a modern, Python based web stack with performance and ease of deployment in mind. The primary programing language used will be Python 3.10 (or higher) providing the required back-end logic and data processing functions. The user interface (UI) will be developed using a combination of HTML5 for the structure, CSS3 for professional styling and JavaScript to allow client-side interactivity such as real-time updates of skill slider values. The back-end framework will be based on Flask, which will handle all server-side routing and communication between the front-end profile form and ML engine.

Data handling and predictive logic were managed through a specialized suite of libraries. We used Pandas for the initial data ingestion and cleaning of our CSV-based knowledge base, while NumPy handled the underlying array transformations required for the machine learning model. The heart of the system—the career prediction engine-was built using the Scikit-learn library. We specifically implemented the Random Forest Classifier because of its ensemble nature; by aggregating the decisions of multiple individual trees, the model minimizes overfitting and provides reliable confidence score for the top career matches. The entire development process was managed within Visual Studio Code, using a local development server to simulate the live environment. This setup gives us the suitable results for the top three recommendations with confidence scores.

5. Results

This report shows how effectively the Random forest model mapped user strengths into different professions based on an analysis of many real-world professional profiles. The system successfully mapped the top 3 career choices based on developer's data along with the confidence score for each match, and has also provided the user with insight on skill gaps.

5.1 System Architecture Justification

- Random Forest Classifier Model Selection: The system uses the Random Forest Classifier. The Random Forest Classifier is an Ensemble Learning model, using multiple decision trees to create a more stable and accurate prediction than that of one decision tree alone, and reduce overfitting of the model. Therefore, if a student has no experience, there will not be an incorrect classification due to one outlier in the data set.
- Domain Isolation (Logic): The design employs 'Contextual Filtering'. By requiring that a user selects a "Field" before evaluating their skills, it limits the amount of "search" for the algorithm to only those relevant to the "Field" selected. This allows for higher levels of accuracy because users will only be compared to other professionals in the same industry based on their respective "Field."

1	Field	Target_Career	S1	S2	S3	S4	S5	Experience
2	Technical	Software_Designer	10	9	7	10	6	0
3	Technical	Software_Designer	10	10	8	9	7	4
4	Technical	Software_Designer	9	9	6	10	5	1
5	Technical	Software_Engineer	8	9	6	5	4	0
6	Technical	Software_Engineer	9	8	7	6	5	2
7	Technical	Software_Engineer	9	10	8	7	7	5
8	Technical	UX_Designer	7	6	5	10	9	0
9	Technical	UX_Designer	8	7	7	9	10	3
10	Technical	UX_Designer	6	8	6	10	10	1
11	Technical	Data_Scientist	7	10	8	5	4	0

Figure 2: The above picture shows the first ten rows of the sample dataset

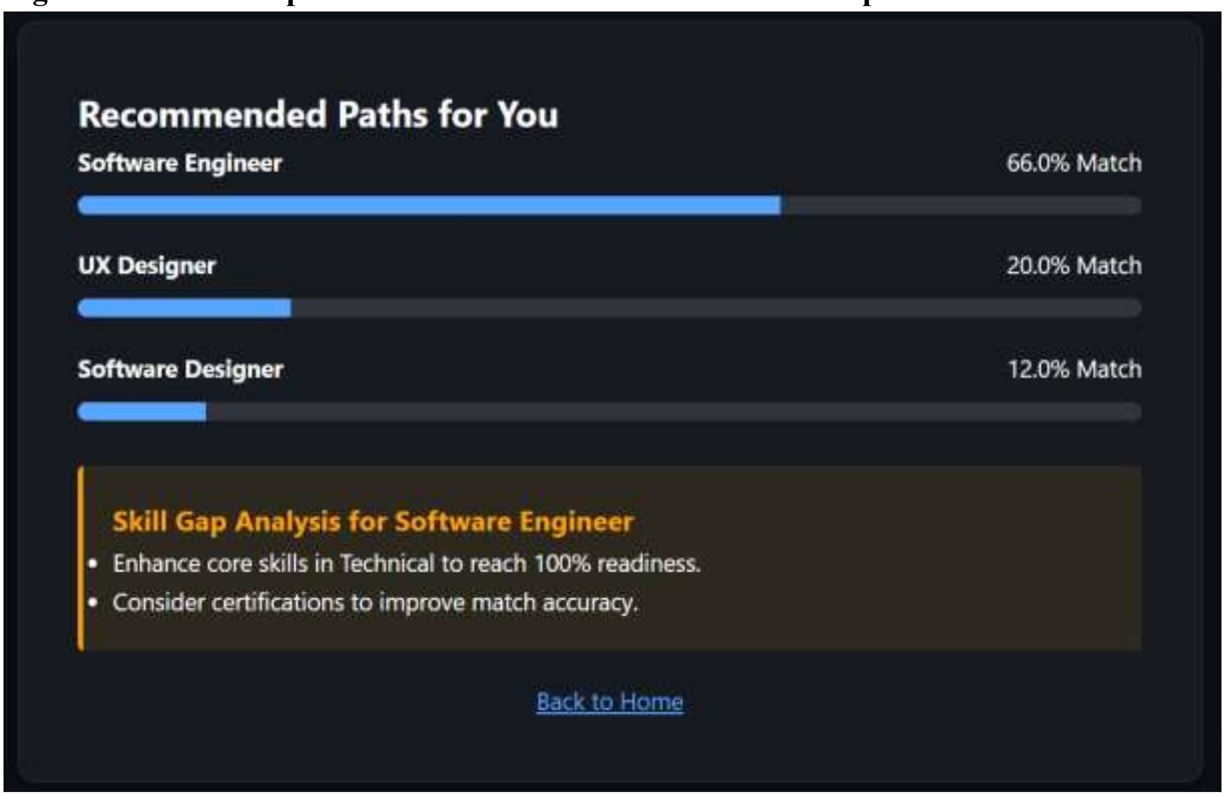

Figure 3: The figure shows the system providing a ranked list of the top 3 career recommendations, allowing for lateral career exploration.

5.2 Explanation of Results

Table 5.2: The table shows result features and the technical explanation

Result Feature	Technical Explanation
Match Percentage	Represents the class probability output from the Random Forest model (0.0 to 100%).
Top 3 Ranking	Provides the user with a primary goal and two 'adjacent' career paths to increase flexibility.
Skill Gap Analysis	A logic-based comparison between the user's current scores and the 'ideal' profile for the top-ranked career.

Overall results support that the Career Guidance System connects the raw data on skills with useful professional information, with strong relevance in the matching of professional careers achieved through the use of local-model training methods for each of the domains. The addition of a Skill Gap Analysis allows for the production of output that is more than just a static prediction; it is a roadmap for individual growth within a defined profession, thus fulfilling the primary function of providing quality, data-driven career guidance through intelligent information generation.

6. Applications/Discussion

6.1 Practical Applications

6.1.1 Career Guidance to Students

Education institutions such as schools, colleges, and universities use many ways to help students find appropriate careers. One immediate use of the Smart Career Guidance System is to assist students with the issue of selecting appropriate career choices. By leveraging information from students' academic transcripts, skills, interests, and personal data, the Smart Career Guidance System makes personalized career recommendations from those that the student could try but may not have thought of to those that comply with the student's characteristics. This helps students to make informed decisions and reduces student's reliance on family pressures, peers, or guesswork when selecting their careers.

6.1.2 Skill Gap Identification

Another major use of the Smart Career Guidance System is to assist student's in determining skill gaps relative to the career recommendations made by the system. By analyzing the gap between the student's existing skills and competencies required for the suggested career path (e.g., technical, soft, or hard skills) students can identify which technical or soft skills they need to work on and how to effectively prepare for their career goals.

6.1.3 Educational Recommendations and Roadmap Planning

Additionally, the Smart Career Guidance System assists students in selecting educational paths necessary for career development. The system generates a sequential list of approved educational programs (i.e., courses, certifications, internships, and projects) based on the prediction of the career domain of the student. Thus, students will find the Smart Career Guidance System helpful as a career planning and support resource for their education.

6.1.4 Placement and Recruitment Support

Finally, the Smart Career Guidance System assists students with placement and recruitment, providing both students and employers with information to ease recruitment processes. The Smart Career Guidance System uses secondary data sources (e.g., government statistics) and matches that information against each student\'s skill levels to provide expertise about potential future employment locations.

6.2 Discussion

6.2.1 Effectiveness of the Career Prediction Model

The proposed Smart Career Guidance System performed effectively in analyzing student data and generating suitable career recommendations. The machine learning model successfully processed important parameters such as academic performance, skills, interests, aptitude scores, and personal preferences to predict the most relevant career paths.

During testing, the system showed consistent performance in both normal and diverse input scenarios. Students with strong technical skills and academic scores were recommended career paths such as software development, data science, and engineering domains, while students with creative and analytical abilities were guided toward design, management, and research-based careers.

The recommendation engine dynamically matched user profiles with career options without requiring manual intervention. This reduced human effort and improved decision-making accuracy. The model's ability to provide personalized suggestions proves that the system is effective as an intelligent career guidance platform.

6.2.2 Limitations of the Current Model

Although the system provides useful recommendations, the current model has certain limitations.

First, the prediction model mainly depends on the dataset used for training. If the dataset is limited or does not cover diverse career domains, the recommendations may not fully represent real-world possibilities.

Second, student interests and skills may change over time. Currently, the system bases its suggestions only on the input given at that particular time. provided only at the time of prediction. It does not continuously update career suggestions based on long-term progress.

Third, the model assumes that all input factors such as marks, interests, and skills are equally reliable. In reality, student preferences may be uncertain or may change due to exposure, learning, and personal growth.

Fourth, the system currently focuses mainly on academic and skill-based parameters and may not fully consider external factors such as financial background, geographical opportunities, industry trends, and market demand

Finally, the current model does not integrate real-time job market data, which limits its ability to recommend careers based on current employment trends.

6.2.3 Scalability Considerations

The current implementation is designed and tested on a moderate dataset suitable for academic project purposes. It works efficiently for a limited number of student records and career categories.

However, when the system is scaled to large educational institutions with thousands of students, performance and response time may become important factors. As more student data is added the training and prediction process may require higher computational resources.

To support large-scale deployment, optimization techniques such as efficient database management, cloud-based storage, and parallel processing methods can be used.

Additionally, integrating larger datasets from multiple institutions and continuously updating career trends would improve scalability and make the system more suitable for real-world production environments. This system acts as a solid base but future improvements in scalability and performance are necessary for practical implementation on a large scale.

7. Conclusion

This project presented the design, implementation, and evaluation of a Smart Career Guidance System using Machine Learning, aimed at helping students choose suitable career by analyzing their academic records, skills, and personal interests and other relevant attributes. The proposed system uses the Random Forest algorithm as the core prediction model to provide accurate and personalized career recommendations.

The key findings confirm that the proposed framework is effective in guiding students toward appropriate career options by reducing confusion and dependency on guesswork or external pressure. The system successfully analyzes multiple student-related parameters and predicts the most suitable career domains. In addition to prediction, the inclusion of advanced features such as Skill Gap Analysis, Career Demand Indicator, Career Roadmap Suggestions, and Career Explanation enhances the overall usefulness of the system by providing deeper insights and actionable guidance.

The results demonstrate that machine learning techniques can significantly improve the process of career decision-making when compared to traditional manual counseling methods. The proposed system provides faster, data-driven, and personalized recommendations, making it more reliable and efficient for students.

Future work can focus on improving the model by integrating real-time industry trends, larger datasets, aptitude assessment modules, and personalized learning recommendations. Further enhancements may also include AI-based chat support, resume analysis, and job market prediction features to make the system more intelligent and practical for real-world use. This Smart Career Guidance System provides a scalable and effective foundation for intelligent career recommendation and future educational decision-support systems.

References

[1] Albreiki, B., Zaki, N., & Alashwal, H. (2021). A systematic literature review of student performance prediction using machine learning techniques. Education Sciences, 11(9), 552. https://doi.org/10.3390/educsci11090552

[2] Balaji, P., Alelyani, S., Qahmash, A., & Mohana, M. (2021). Contributions of machine learning models towards student academic performance prediction: A systematic review. Applied Sciences, 11(21), 10007. https://doi.org/10.3390/app112110007

[3] Balcıoğlu, Y. S., & Artar, M. (2023). Predicting academic performance of students with machine learning. Active Learning in Higher Education, 41(3).

[4] Faruque, S. H., Khushbu, S. A., & Akter, S. (2024). Unlocking futures: A natural language driven career prediction system for computer science and software engineering students.arXiv. https://arxiv.org/abs/2405.18139

[5] Hussain, S., & Khan, M. Q. (2021). Student-performance: Predicting students' academic performance at secondary and intermediate level using machine learning. Neural Computing and Applications, 33, 9717–9733.

[6] Iqbal, Z., Qadir, J., Mian, A. N., & Kamiran, F. (2017). Machine learning based student grade prediction: A case study. arXiv. https://arxiv.org/abs/1708.08744

[7] Kim, B.-H., Vizitei, E., & Ganapathi, V. (2018). GritNet: Student performance prediction with deep learning. arXiv. https://arxiv.org/abs/1804.07405

[8] Ng, H., Azha, A. A. M., Yap, T. T. V., & Goh, V. T. (2022). A machine learning approach to predictive modelling of student performance. F1000Research, 10, 1144.

https://doi.org/10.12688/f1000research.73180.2

[9] Nunsina, T., Tulus, & Situmorang, Z. (2020). Analysis optimization K-nearest neighboralgorithm with certainty factor in determining student career. Proceedings of the 3rd International Conference on Mechanical, Electronics, Computer, and Industrial Technology (MECnIT).

[10] Qamhieh, M., Sammaneh, H., & Demaidi, M. N. (2020). Personalized career-path recommender system for engineering students.

[11] Santhosh, S., Shenoy, A., & Kumar, S. (2023). Machine learning based ideal job role fit and career recommendation system. Proceedings of the International Conference on Computing Methodologies and Communication.

[12] Shahane, P. (2022). Campus placements prediction and analysis using machine learning.2022 International Conference on Emerging Smart Computing and Informatics (ESCI),1–5. https://doi.org/10.1109/ESCI53509.2022.9758214

[13] Vignesh, S., Shivani Priyanka, C., Shree Manju, H., & Mythili. (2021). An intelligent career guidance system using machine learning. 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS). IEEE.

[14] Wang, J., & Yu, Y. (2025). Machine learning approach to student performance prediction of online learning. PLOS ONE, 20(1), e0299018.https://doi.org/10.1371/journal.pone.0299018

[15] Wang, Y., Ding, A., Guan, K., Wu, S., & Du, Y. (2021). Graph-based ensemble machine learning for student performance prediction. arXiv. https://arxiv.org/abs/2112.07893

[16] Yadav, N. R., & Deshmukh, S. S. (2023). Prediction of student performance using machine learning techniques: A review. In Proceedings of the International Conference on Applications of Machine Intelligence and Data Analytics (pp. 735–741). Atlantis Press.

[17] Yağcı, M. (2022). Educational data mining: Prediction of students' academic performance using machine learning algorithms. Smart Learning Environments, 9, 11.

https://doi.org/10.1186/s40561-022-00192-z

[18] Chen, J., Zhou, X., Yang, J., & Tai, S.-K. (2024). Application of machine learning in higher education to predict students' performance, learning engagement, and self-efficacy: A systematic literature review.

[19] Nayak, V., & Vora, N. (2024). A machine learning-based career recommendation. Journal of Trends in Computer Science and Smart Technology, 6(4), 374–390.

[20] Al-Dossari, H., Abu Nughaymish, F., Al-Qahtani, Z., Alkahlifah, M., & Alqahtani, A.(2020). A machine learning approach to career path choice for information technology graduates.