5th International Conference on Electronics and Sustainable Communication Systems (ICESC 2024)
IEEE XPlore Part Number: CFP24V66-ART; ISBN: 979-8-3503-7994-5

Carrer Compass:A Career Path Recommender using Machine Learning

Umesh Guru Sai Guntupalli
Dept. of Information Technology
Velagapudi Ramakrishna Siddhartha Engineering College
Kanuru, India
umeshgurusai@gmail.com

Madhavi Latha Pandala
Dept. of Information Technology
Velagapudi Ramakrishna Siddhartha Engineering College
Kanuru, India
madhavi.p@vrsiddhartha.ac.in

Dharma Teja Veeranki
Dept. of Information Technology
Velagapudi Ramakrishna Siddhartha Engineering College
Kanuru, India
vdharmatejaa@gmail.com

Pramodhini Kumbha
Dept. of Information Technology
Velagapudi Ramakrishna Siddhartha Engineering College
Kanuru, India
pramodhini@gmail.com

2024 5th International Conference on Electronics and Sustainable Communication Systems (ICESC) | 979-8-3503-7994-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICESC60852.2024.10689897

***Abstract*—In the ever-expanding corporate realm, making an appropriate career choice is crucial for building a successful career. Students often face confusion due to increased choices or lack of guidance, leading to missteps in career selection. Such decisions might result in individuals working in fields unsuited to their strengths and interests, reducing overall efficiency. The proposed project highlights the need for a personalized career prediction system. Traditional methods often fall short in capturing the interests and strengths of individuals while solely depending on standard tests. This research study focuses on building a personalized career recommendation system based on the skills and interests of an individual. The project promotes the idea of dynamic and tailor-made tests for identifying suitable careers and an interactive, user-friendly mobile application using the Flutter framework. Machine Learning models like Random Forest and Support Vector Machine are employed for predicting and recommending careers. By evaluating personality traits, interests, and skills, the application will recommend a career. Continuous improvement is fostered through user feedback, ensuring the model's adaptability to changing job markets. This approach provides a comprehensive solution to challenges faced by individuals while choosing a career.**

***Keywords*—Machine learning, Random Forest, Support Vector Machine, dynamic questionnaire, Flutter framework, career guidance, user feedback, career prediction system.**

I. Introduction

The necessity for a more personalized approach to career decision-making arises from the increasing number of available career choices and the potential for individuals to choose incompatible paths. The demand for solutions that guide individuals based on personality and skills is evident. Organizations often employ a strategy where new hires are trained and assigned roles based on performance during the training process. This can lead to issues if individuals are unable to perform due to incompatible career choices or may result in long-term dissatisfaction if they are promoted to unsuitable positions.

Traditional methods of career recommendation, which rely heavily on academic grades and marks, fail to capture the interests and talents of individuals. This lack of insight can hamper the recommendation process, eliminating the chance for individuals to build a potential career suited to their interests and talents. As a result, individuals may never identify their true interests or potential in a particular field.

To address this issue, a mobile application is proposed to assist students in identifying compatible career paths. This application offers tailor-made personalized tests based on each individual's skills and interests, with a dynamic flow of questions that adapts to the user's current responses. Machine learning models, including Random Forest and Support Vector Machine, are utilized to analyze responses and predict potential career paths. The questionnaire is designed to dynamically fetch the next question based on the user's response to the current question, ensuring a more personalized and accurate assessment.

II. Review of Literature

Zhang explored an advanced career guidance system that integrates big data technology. This approach focuses on analyzing students' current situations and applying big data to career planning education. The goal is to enhance career guidance by leveraging data analytics to provide more informed and effective career advice [1].

Madhan Mohan developed a "Career Prediction System" specifically for students in the computer science and engineering fields. This system utilizes decision tree and XGBoost models to evaluate students' skills in various areas. However, it is limited to computer science roles and does not account for individual interests [3].

Rane et al. introduced an interactive web-based career prediction system that uses machine learning techniques, including decision tree and K-Nearest Neighbors (KNN) algorithms. This system aims to assist students in choosing a career path after high school, addressing the challenges faced by younger individuals in making informed career decisions [4].

Furthermore,Sahid and Pratama proposed a model based on psychological aptitude tests, utilizing decision tree and naïve Bayes algorithms. While this model offers insights into psychological aptitudes, it lacks an interactive platform that could

maximize user engagement and provide more comprehensive career recommendations [5].

Lastly,Kumar et al. developed a platform that facilitates goal-setting and self-assessment, allowing students to identify their strengths and weaknesses. Despite its focus on personal development, this platform does not offer career prediction or recommendations based on individual interests [9].

III. Proposed Architecture and Methodology

A. Dataset

The dataset obtained from Kaggle includes comprehensive data relevant to career paths, encompassing a wide range of industry skills and corresponding job roles. The dataset also integrates the Big Five personality questionnaire to assess individual characteristics and align these traits with suitable career interests.

B. Design Methodology

The design methodology for this career recommendation system is structured into several key stages to ensure an effective implementation. The process begins with data collection and preprocessing. Initially, relevant career and skills data is gathered and cleaned to address any inconsistencies or missing values. Feature selection is then performed to identify and retain the most pertinent data attributes for training the machine learning models.

The next stage involves model development. A dynamic questionnaire is crafted, based on the Big Five personality traits and skills assessment, to gather detailed insights into each user's preferences and competencies. This questionnaire is designed to adapt based on user responses, ensuring a tailored approach. Machine learning models, specifically Random Forest and Support Vector Machine (SVM), are trained using this preprocessed data. Random Forest builds multiple decision trees and aggregates their predictions to enhance accuracy, while SVM identifies the optimal hyperplane that best separates different career categories based on user inputs.

Following model development, the focus shifts to application development. A mobile application is built using the Flutter framework to provide an interactive and user-friendly interface for the questionnaire and career recommendations. Concurrently, a Flask-based server is set up to handle backend operations, process user inputs through the trained models, and deliver career recommendations back to the application.

The subsequent stages involve integration and testing. The Flutter front-end is connected to the Flask back-end to facilitate seamless data flow and accurate predictions. Rigorous testing is conducted to validate the performance of the career recommendations and ensure the application functions correctly.

Finally, user feedback is collected to continuously refine the questionnaire and improve model accuracy. This feedback loop allows for adjustments to meet user needs effectively.

C. Algorithms

Random Forest

Input: Dataset D containing N features and a sample test input T

Output: Random forest tree [1] Define the number of trees to be considered in the forest Define the maximum depth of each tree Set the minimum samples that are required to split a node Select the features to be considered at each split Create an empty list to store the individual decision trees Divide the features and apply the algorithm of decision tree for each subset each tree in number of trees Randomly sample the data from the training dataset with replacement Construct a decision tree using the sampled data and add the tree to the forest Combine multiple decision trees to form a Random Forest ensemble model

Support Vector Machine (SVM)

Input: Dataset D containing N features and a sample test input T

Output: SVM model [1] Calculate the Loss function for every feature by using Eq. eq:loss1 and Eq. eq:loss2

$$L(y, f(x)) = \max(0, 1 - y \cdot f(x)) \tag{1}$$

$$SVMLoss = \frac{1}{n} \sum_{i=1} L(y_i, f(x_i)) + \lambda \| w \|^2 \tag{2}$$

Choose the feature with the minimum Loss function that also has the maximum margin. Include regularization parameter to the model Optimize the weights by calculating Gradients. Update gradients using regularization parameters if there is misclassification Repeat the process until classification is correctly done

D. Architecture Diagram

This chapter includes the proposed work's architecture diagram we used.

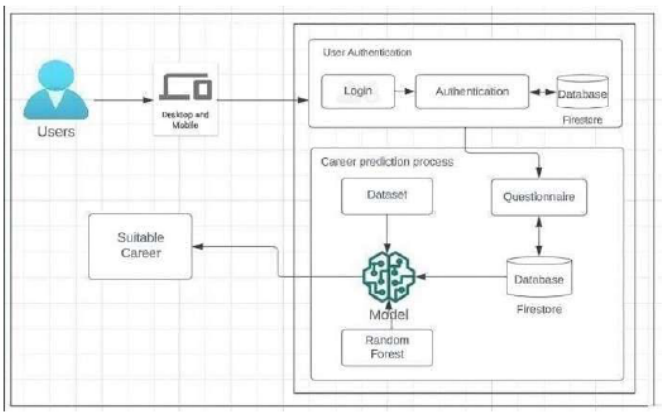

Fig. 1. System Architecture

Fig.1 outlines a comprehensive career recommendation system, where users engage through desktop or mobile devices. The process begins with user authentication, ensuring secure access via login and authentication mechanisms. Once

authenticated, user information and responses to a dynamic questionnaire, which assesses skills, interests, and personality traits, are stored in a Firestore database. This initial stage ensures that user data is securely managed and personalized for subsequent steps.

The core of the system involves the career prediction process, utilizing a Random Forest machine learning model. The stored user responses, combined with a pre-existing dataset detailing various career paths and required skills, are fed into the model. This model processes the inputs to generate tailored career recommendations. The recommended career paths are then communicated back to users through the application interface.

IV. Results and Analysis

The primary objective of the career recommendation project is to recommend a career based on the skills and interests of an individual. Machine learning models were trained using Random Forest and Support Vector Machine (SVM) algorithms.The results and observations are as follows.

A. SVM Model

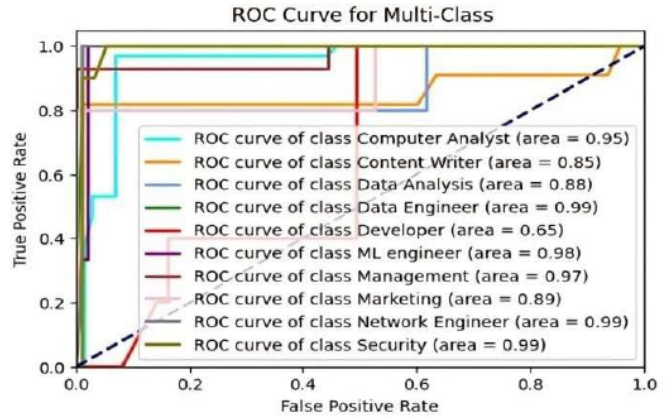

Fig. 2. ROC curve for SVM model

The ROC curve for the SVM model, shown in Figure 4.1, plots the true positive rate (sensitivity) against the false positive rate (1-specificity) at various threshold settings. This curve helps us evaluate the trade-off between the true positive rate and false positive rate across different decision thresholds. The area under the ROC curve (AUC) provides a single metric that summarizes the model's ability to discriminate between positive and negative classes.

Figure 4.2 presents the confusion matrix for the SVM model, which is a table used to describe the performance of a classification model on a set of test data for which the true values are known.

B. Random Forest Model

The ROC curve for the Random Forest model, depicted in Figure 4.3, also illustrates the model's performance across different threshold settings.

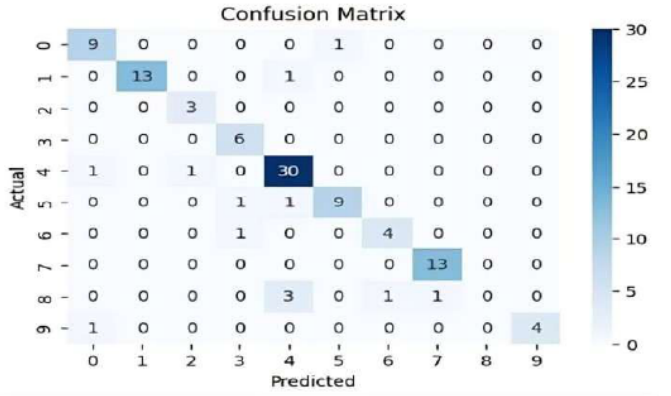

Fig. 3. Confusion matrix for SVM model

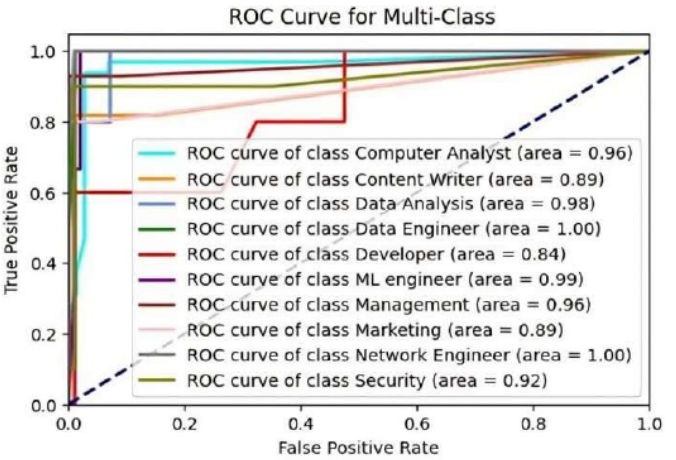

Fig. 4. ROC curve for Random Forest model

Figure 4.4 displays the confusion matrix for the Random Forest model. The matrix provides the same breakdown as the SVM confusion matrix, showing the counts of true positives, true negatives, false positives, and false negatives.By comparing these matrices, we can directly observe which model has fewer misclassifications and thus better performance.

The ROC curves illustrate that the Random Forest model has a higher true positive rate at each false positive rate compared to the SVM model, indicating better performance. The confusion matrices further confirm that the Random Forest model has fewer false positives and false negatives, leading to higher precision and recall. Overall, the visualizations and

TABLE I
Evaluation Metrics of Two Algorithms

	Precision	Recall	Accuracy
Random Forest	0.89	0.88	0.90
SVM	0.84	0.87	0.88

metrics demonstrate that the Random Forest model is more effective for our career recommendation system, which is why it was chosen for integration into the Flutter application.

The interface where the question is displayed to the user is

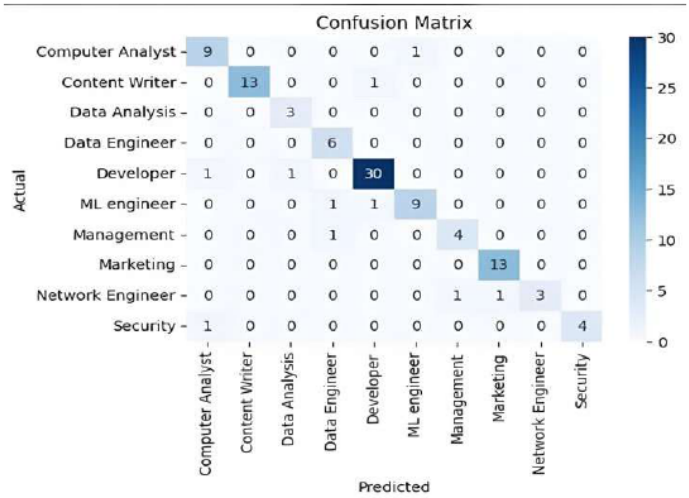

Fig. 5. Confusion matrix for Random Forest model

Fig. 7. Application output

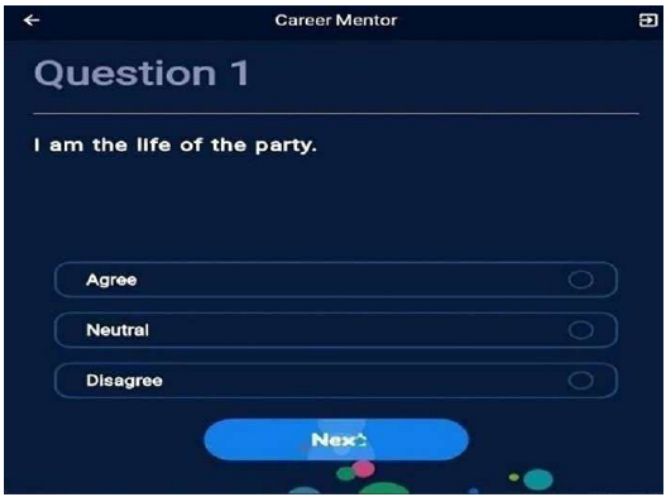

Fig. 6. Apllication Layout

illustrated in Figure 4.5. The selection of the next question is determined by the user's current response, and all responses are stored in a database.

Figure 4.6 illustrates the application output, which shows the recommended career to the user. This application is designed to consider the user's interests through a dynamic questionnaire and also takes into account the user's skills. Unlike many other applications, this approach ensures a more comprehensive assessment of the user. While the accuracy might be lower than other systems, this is due to the attempt to capture more complex relationships between interests and skills.

This study utilizes a questionnaire that assesses both personality and skills to recommend a career path to the user. The project is particularly beneficial for students who lack clarity about their career direction and struggle to make informed career choices. The approach addresses the common issue of students ending up in careers that do not align with their strengths and interests, thereby improving the likelihood of a satisfying and successful career.

V. Conclusion and Future Work

A. Conclusion

This project successfully develops a personalized career recommendation system using machine learning models, particularly Random Forest and Support Vector Machine (SVM), to offer tailored career advice based on an individual's skills and interests. The Flutter-based interactive mobile application ensures user-friendly engagement, dynamically adjusting the questionnaire based on user responses and storing this data for analysis. The results indicate that the Random Forest model outperforms SVM in accuracy, precision, and recall, making it the preferred model. By addressing the limitations of traditional career guidance methods that often overlook personal interests and focus solely on academic performance, this system aids users in making more informed and fulfilling career choices. Continuous user feedback allows the system to evolve and adapt to changing job markets and individual needs, providing a comprehensive career planning solution.

B. Future Work

Future work for this project includes integrating real-time job market data to provide users with insights into current industry trends and demand, recommending relevant courses and certifications to help users develop necessary skills, and creating structured learning paths for career progression. Enhancing the application's interactivity with features like chatbots and visual analytics, implementing a robust feedback mechanism for continuous improvement, and ensuring cross-platform compatibility are also priorities. Additionally, partnering with educational institutions to integrate the system into existing career guidance programs and incorporating psychometric testing to assess personality traits and cognitive abilities will further refine the recommendations and make the system more comprehensive.

VI. REFERENCES

[1] Alfred, Rayner, et al. "Towards paddy rice smart farming: a review on big data, machine learning, and rice production tasks." Ieee Access 9 (2021): 50358-50380.

[2] Ismail, M. A. F., et al. "E-PADI: An IoT-based paddy productivity monitoring and advisory system." Indones. J. Electr. Eng. Comput. Sci 14.2 (2019): 852.

[3] Lubis, H., et al. "Monitoring System of Rice Plant Growth Using Microcontroller Sensor." Journal of Physics: Conference Series. Vol. 1235. No. 1. IOP Publishing, 2019.

[4] Mekala, Mahammad Shareef, and Perumal Viswanathan. "CLAY-MIST: IoT-cloud enabled CMM index for smart agri- culture monitoring system." Measurement 134 (2019): 236-
244.

[5] Sreekantha, D. K., and A. M. Kavya. "Agricultural crop monitoring using IoT-a study." 2017 11th International conference on intelligent systems and control (ISCO). IEEE, 2017.

[6] Ghanshala, Kamal Kumar, Rahul Chauhan, and R. C. Joshi. "A novel framework for smart crop monitoring using internet of things (IoT)." 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC). IEEE, 2018.

[7] Wicaksono, Muhammad Galang Satrio, Erma Suryani, and Rully Agus Hendrawan. "Increasing productivity of rice plants based on IoT (Internet Of Things) to realize Smart Agri- culture using System Thinking approach." Procedia Computer Science 197 (2022).

[8] Sheng, Rodney Tai-Chu, Yu-Hsiang Huang, Pin-Cheng Chan, Showkat Ahmad Bhat, Yi-Chien Wu, and Nen-Fu Huang. "Rice growth stage classification via RF-based ma- chine learning and image processing." Agriculture 12, no. 12 (2022).

[9] James, Jerrin. "Plant growth monitoring system, with dy- namic user-interface." In 2016 IEEE Region 10 Humanitarian Technology Conference (R10-HTC), pp. 1-5. IEEE, 2016.

[10] Liqiang, Zhao, Yin Shouyi, Liu Leibo, Zhang Zhen, and Wei Shaojun. "A crop monitoring system based on wireless sensor network." Procedia Environmental Sciences 11 (2011).

[11] Lee, Meonghun, Jeonghwan Hwang, and Hyun Yoe. "Agricultural production system based on IoT." In 2013 IEEE 16Th international conference on computational science and engineering, pp. 833-837. IEEE, 2013.

[12] Keshtgari, Manijeh, and Amene Deljoo. "A wireless sensor network solution for precision agriculture based on zigbee technology." (2011).

[13] Lavanya, G., Chellasamy Rani, and Pugalendhi GaneshKumar. "An automated low cost IoT based Fertilizer Intimation System for smart agriculture." Sustainable Comput- ing: Informatics and Systems 28 (2020)