A
Mini Project Report on

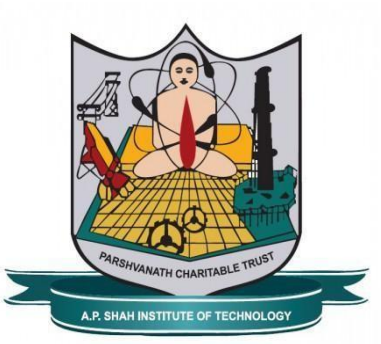
Serendipity-aware Career Recommendation System

Submitted in fulfillment of the requirements
for the degree of
BACHELOR OF ENGINEERING
IN
Computer Science & Engineering
Artificial Intelligence & Machine Learning

by

Atharva Patil (22106039)
Shraavani Salunkhe (22106031)
Brahmjot Singh (22106004)
Aniruddha Pawar (22106009)
Under the guidance of

Prof. Tejashri Kolhe



Department of Computer Science & Engineering
(Artificial Intelligence & Machine Learning)
A. P. Shah Institute of Technology
G. B. Road, Kasarvadavali, Thane (W)-400615
University Of Mumbai
2024-2025

A. P. SHAH INSTITUTE OF TECHNOLOGY

CERTIFICATE

This is to certify that the project entitled "**Serendipity-aware Career Recommendation System"** is a bonafide work of Atharva Patil (22106039), Shraavani Salunkhe (22106031), Brahmjot Singh (22106004), Aniruddha Pawar (22106009) submitted to the University of Mumbai in partial fulfillment of the requirement for the award of **Bachelor of Engineering** in **Computer Science & Engineering (Artificial Intelligence & Machine Learning).**

Prof. Tejashri Kolhe	Dr. Jaya Gupta
Mini Project Guide	Head of Department

A. P. SHAH INSTITUTE OF TECHNOLOGY

Project Report Approval

This Mini project report entitled "**Serendipity-aware Career Recommendation System"** by Atharva Patil, Shraavani Salunkhe, Brahmjot Singh, Aniruddha Pawar is approved for the degree of ***Bachelor of Engineering*** in ***Computer Science &Engineering***, (AI and ML)

2024-25.

External Examiner: _____

Internal Examiner: _____

Place: APSIT, Thane

Date:

Declaration

We declare that this written submission represents my ideas in my own words and whereothers' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission hasnot been taken when needed.

Atharva Patil (22106039)	Shraavani Salunkhe (22106031)	Brahmjot Singh (22106004)	Aniruddha Pawar (22106009)

ABSTRACT

The Serendipity-aware Career Recommendation System enhances career decision-making by combining predictive accuracy with serendipity, ensuring both relevant and novel career suggestions. Developed using Python, Streamlit, and Scikit-learn, it employs the Random Forest algorithm for robust career predictions based on user preferences, skills, and aptitude test results. To introduce unexpected yet meaningful career options, it utilizes TF-IDF and cosine similarity, expanding career possibilities beyond conventional choices. By integrating machine learning with serendipitous exploration, the system provides a personalized, insightful, and engaging approach to career guidance, helping users discover fulfilling career paths aligned with their potential.

Keywords: Career Recommendation, Serendipity, Random Forest, Machine Learning, TF-IDF, Cosine Similarity, Predictive Analytics, Career Guidance.

Index

Index			Page no.
Chapter-1			
	Introduction		1
Chapter-2			
	Literature Survey		4
	2.1	History	5
	2.1	Review	6
Chapter-3			
	Problem Statement		9
Chapter-4			
	Experimental Setup		11
	4.1	Hardware setup	12
	4.2	Software Setup	12
Chapter-5			
	Proposed system and Implementation		13
	5.1	Block Diagram of proposed system	14
	5.2	Description of Block diagram	14
	5.3	Implementation	16
Chapter-6			
	Conclusion		19
References			20

CHAPTER 1

INTRODUCTION

1. INTRODUCTION

Career selection is a pivotal decision shaped by an individual's interests, skills, and dynamic job market trends. Traditional career recommendation systems primarily focus on direct matches, often restricting users to predefined career paths. Our Serendipity-aware Career Recommendation System enhances career exploration by integrating machine learning-driven predictions with serendipitous suggestions, ensuring both accuracy and discovery. Developed using Python, Streamlit, and Scikit-learn, the system employs the Random Forest algorithm for reliable career classification while utilizing serendipity-based techniques to introduce unconventional yet meaningful career options. This approach fosters a personalized, insightful, and engaging career guidance experience, helping users uncover new and fulfilling career possibilities.

Project Motivation

Most existing career recommendation systems prioritize precision but lack diversity in suggestions. Career decisions, however, often involve unexpected discoveries that align with latent interests and transferable skills. By incorporating serendipity through a randomization mechanism, our project aims to help users explore alternative career options they might not have initially considered. This encourages career flexibility and innovation, fostering a broader perspective in decision-making.

Project Overview

1. **User Input Collection**
    - Users take a Career Assessment Test and Aptitude Test to evaluate their skills, interests, and strengths.
    - Additional inputs, such as educational background and job preferences, are recorded to personalize recommendations.
2. **Data Preprocessing & Feature Engineering**
    - Handling missing values, encoding categorical data, and structuring features for better model performance.
    - Label Encoding is used to convert user interests and skills into numerical representations.

3. **Model Training & Career Prediction**
    - The Random Forest algorithm is trained on user responses to predict suitable career paths.
    - The model learns patterns and maps user profiles to optimal career choices based on their input.
4. **Serendipity Integration**
    - A serendipity-based randomizer introduces unexpected yet relevant career suggestions beyond conventional choices.
    - These alternative career options are identified based on transferable skills, emerging industries, and potential growth opportunities to broaden career exploration while maintaining relevance.
5. **Career Descriptions & Insights**
    - Provides detailed job descriptions, required skills, salary expectations, and industry trends for recommended careers.
    - Skill gap analysis helps users identify areas for improvement and potential career transitions.
6. **User Interface & Visualization**
    - A Streamlit-based UI presents career matches and serendipitous suggestions interactively.
    - Users receive a graphical overview of their test results, career recommendations, and serendipitous career options.

CHAPTER 2

LITERATURE SURVEY

2. LITERATURE SURVEY

2.1-HISTORY

Career recommendation systems have evolved from manual methods to AI-driven solutions. Initially, career counseling relied on psychometric assessments like the Holland Code (RIASEC) model, which categorizes individuals into six personality types (Realistic, Investigative, Artistic, Social, Enterprising, and Conventional) based on work preferences, and MBTI (Myers-Briggs Type Indicator), which classifies people into 16 personality types to align with potential careers, offering static career choices with limited personalization.

1. Rule-Based Systems (2000s)

Early digital systems used rule-based algorithms, relying on decision trees and predefined rules. However, they lacked adaptability and real-time labor market integration.

2. Machine Learning Integration (2010s – Present)

Advancements in machine learning (ML) introduced Random Forest, Decision Trees, and Neural Networks for career prediction, enhancing accuracy and personalization. Content-based filtering and collaborative filtering improved recommendations based on user data.

3. Serendipity in Career Guidance (Recent Advances)

Modern AI-driven systems integrate serendipity-aware approaches, introducing diversity-enhancing mechanisms to suggest unexpected but relevant career options, reducing the filter bubble effect.

2.2-LITERATURE REVIEW

Career recommendation systems have evolved significantly, leveraging advancements in artificial intelligence, data mining, and psychometric analysis. The literature highlights various methodologies used over time, from traditional counselling techniques to machine learning-based models.

[1] **Serendipity-Oriented Recommendations for Improving User Satisfaction**

Published in: Serendipity-Oriented Recommendations for Improving User Satisfaction WANG, N. (Author). 29 Jun 2023

Online recommender systems assist users in processing vast internet information but often suffer from overspecialization, focusing solely on relevance. Serendipity, balancing relevance with unexpectedness, enhances user satisfaction. To study this, a large-scale user survey was conducted to gather feedback on multiple recommendation objectives. Analyzing datasets from movies and e-commerce, key factors influencing perceived serendipity were identified. Using these insights, the Temporal Unexpected Recommendation (TUR) approach was developed, incorporating implicit and explicit distances with a time-aware GRU module. Experiments showed TUR improves unexpectedness while maintaining accuracy. Additionally, the study highlights biases in serendipity-oriented recommendations, paving the way for further research.

[2] **Research on Career Counselling Platform Based on Collaborative Filtering Recommendation Algorithm**

Published in: 2023 International Conference on Integrated Intelligence and Communication Systems (ICIICS)

The vast amount of online information has led to the rise of personalized recommendation systems, shifting users from active searching to automated suggestions. Collaborative filtering is a widely used technique, but it faces challenges like data sparsity and cold start problems. Cluster-based collaborative filtering helps overcome these issues, making it effective for career consulting platforms. This approach improves career planning for college students, enhancing employment rates and future career development. The method achieved 97.80% accuracy in recommendations.

[3] **Design and implementation of career planning system based on machine learning**

Published in: 2023 IEEE 3rd International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA)

To better cater to the needs of job seekers and provide enhanced career development directions, this study proposes a novel employment guidance and career planning system integrating machine learning algorithms, along with the creation of a career matchmaking framework. Firstly, leveraging the RippleNet algorithm model, we construct a knowledge graph-based

career planning recommendation system. The system-building process involves investigating key technologies in data processing and recommendation algorithms. Subsequently, through the implementation of database security design, we ensure the information security of users within the system, thereby maintaining the efficient operation of the system. The results indicate that the system is well-functioning, exhibits stable performance, and is capable of meeting the fundamental requirements of a career planning system.

[4] **Framework Design of Intelligent Career Recommendation System for Undergraduates**

Published in: 2022 International Conference on Intelligent Education and Intelligent Research (IEIR)

The development of artificial intelligence technology has proudly enhanced the quality of life and education of students. The outbreak of COVID-19 in early 2020 dealt a huge blow to the world economy and workplace environment, therefore planning a career path before graduation is a primary and core task for undergraduate students to succeed in this era. This paper introduces the framework design of an intelligent career recommendation system, which is based on the analysis of the required career ability and students' individual ability to achieve accurate career recommendations.

[5] **Job and Course Recommendation System using Collaborative Filtering and Naive Bayes algorithms**

Published in: 2023 2nd International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA)

With the rise of online learning, students face challenges in choosing courses aligned with their career goals. This study proposes a hybrid course and job recommendation system using Naïve Bayes and Collaborative Filtering to provide personalized suggestions based on student interests, skills, and aspirations. The system consists of two main components: course recommendation, which suggests relevant courses, and job recommendation, which identifies suitable job opportunities. By integrating additional data sources like student feedback, industry trends, and job market demands, the system enhances recommendation accuracy and helps students make informed career decisions.

[6] **Student Career Recommendation System Using Content-Based Filtering Method**

Published in: 2022 3rd International Conference on Artificial Intelligence and Data Sciences (AiDAS))

Graduates often face challenges in finding suitable career opportunities, particularly in the field of computer science. This project aims to develop a career recommendation system that helps students identify relevant job roles based on their interests and skills. The system utilizes content-based filtering to compare job listings, with data extracted -from JobStreet through web scraping. It features a user-friendly interface, allowing students to receive personalized recommendations effortlessly. The system has been tested using a functionality tester to ensure accuracy. Future improvements include integrating with career vacancy websites for real-time job updates and enhancing the filtering mechanism to provide more precise career matches.

CHAPTER 3

Problem Statement

3.Problem Statement

Students often face challenges in identifying career paths that align with their skills, interests, and aspirations due to limited guidance and an overwhelming number of choices. Traditional career counseling methods are often generic and lack personalization, making it difficult for individuals to explore diverse opportunities.

This project aims to develop a Serendipity-aware Career Recommendation System that utilizes Machine Learning (ML) techniques, particularly Random Forest, to analyze user preferences, skills, and aptitude test results. By integrating serendipity-based suggestions using Natural Language Processing (NLP) techniques like TF-IDF and Cosine Similarity, the system introduces unexpected yet relevant career options. This approach ensures a more comprehensive and personalized career exploration experience while maintaining accuracy and relevance.

Keywords: Serendipity-aware Career Recommendation, Machine Learning, Random Forest, Natural Language Processing, TF-IDF, Cosine Similarity.

CHAPTER 4

Experimental Setup

4.Experimental Setup

Hardware Requirements

1. Processor: Intel Core i5– Minimum 2.5 GHz
2. RAM: 16GB
3. Internet Connection: Required for dependencies installation

Software Requirements

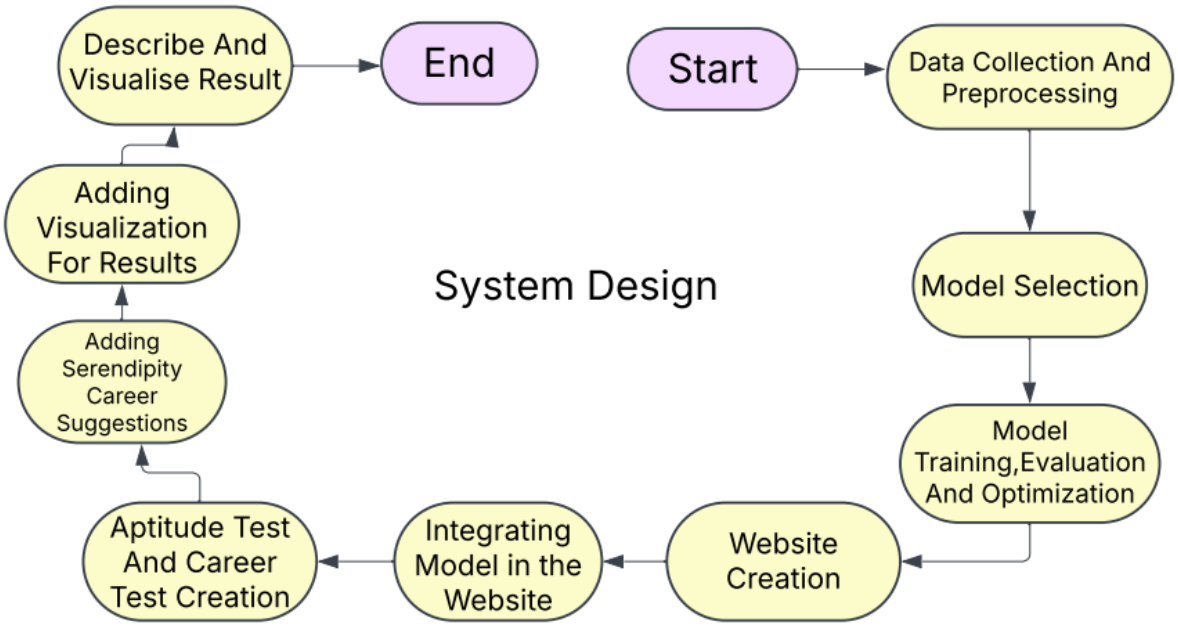
1. Programming Language: Python 3.8+
2. Frameworks & Libraries:
    - Streamlit (for UI)
    - Pandas (for data handling)
    - NumPy (for numerical operations)
    - scikit-learn (for ML model & serendipity logic)
    - SQLite (for database)
    - joblib (for model storage)
    - Matplotlib (for visualizations)
    - TfidfVectorizer (from scikit-learn, for text feature extraction)
    - Cosine Similarity (from scikit-learn, for similarity calculations in serendipity)
3. Development Environment: PyCharm / VS Code / Jupyter Notebook
4. Database: SQLite3

CHAPTER 5

Proposed system and Implementation

5. Proposed system and Implementation

5.1 Block diagram of proposed system



5.2 Description of block diagram

1. Understanding Problem Statement

Defining the **objective** of the system—building a career recommendation model that incorporates user preferences, aptitude, and serendipity. Identifying key challenges and user needs.

2. Data Collection and Preprocessing

Gathering relevant data from various sources, including career descriptions, skills, aptitude tests, and user responses. Preprocessing steps include cleaning data, feature extraction, and normalization for better model performance.

3. Model Selection

Selecting an appropriate machine learning model. The Random Forest algorithm is chosen for classification due to its efficiency in structured career recommendation tasks.

4. Model Training, Evaluation, and Optimization

Training the model on the preprocessed dataset and evaluating it using accuracy, precision, recall, and F1-score. Optimizing performance through hyperparameter tuning.

5. Model Deployment on Website

Integrating the trained model into a Streamlit-based web application, allowing users to interact with the system, take career tests, and receive recommendations dynamically.

6. Website Creation

Developing a Streamlit-based web application where users can register, take tests, view career recommendations, and track past results.

7. Aptitude Test and Career Test Creation

Designing two key assessments:

1.Career Test – Evaluates interests and skills to provide a career recommendation.

2.Aptitude Test – Measures cognitive ability and problem-solving skills, contributing to career suitability evaluation.

8. Adding Visualization for Results

Using graphical representations such as bar charts, pie charts, and tables to help users understand their aptitude test performance, skill mapping, and career matches.

9. Adding Serendipity Career Suggestions

Introducing unexpected but relevant career suggestions using TF-IDF vectorization and Cosine Similarity. This enhances user experience by suggesting alternative career paths similar to the primary recommendation.

10. Describe and Visualize Result

Presenting predicted career recommendations in an interactive format with explanations. Using data visualization tools like Matplotlib and Streamlit charts to display skill matches and career insights.

5.3 Implementation

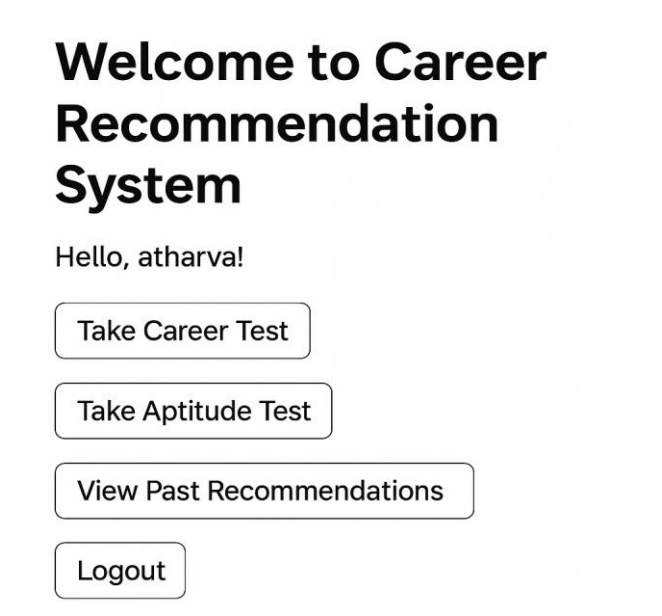


Fig-5.3.1-Home Page

Fig-5.3.2-Career Test

Fig-5.3.3-Career Test Result

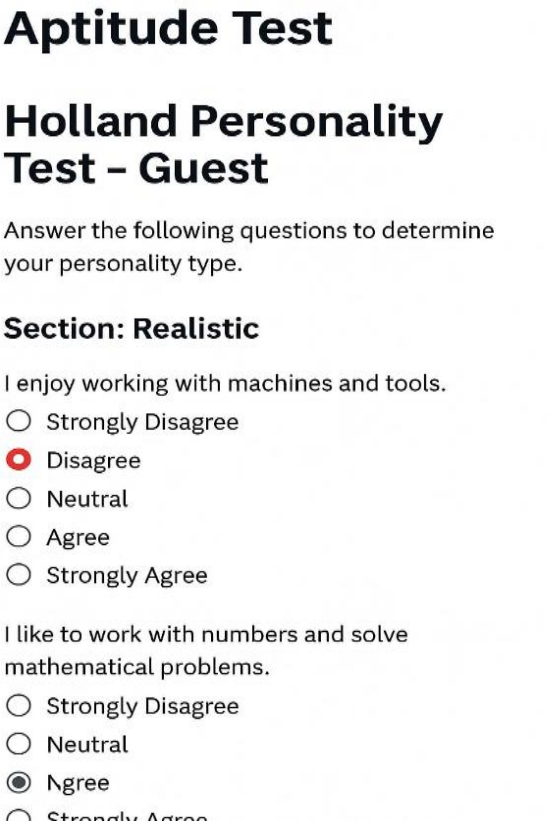


Fig-5.3.4-Aptitude Test

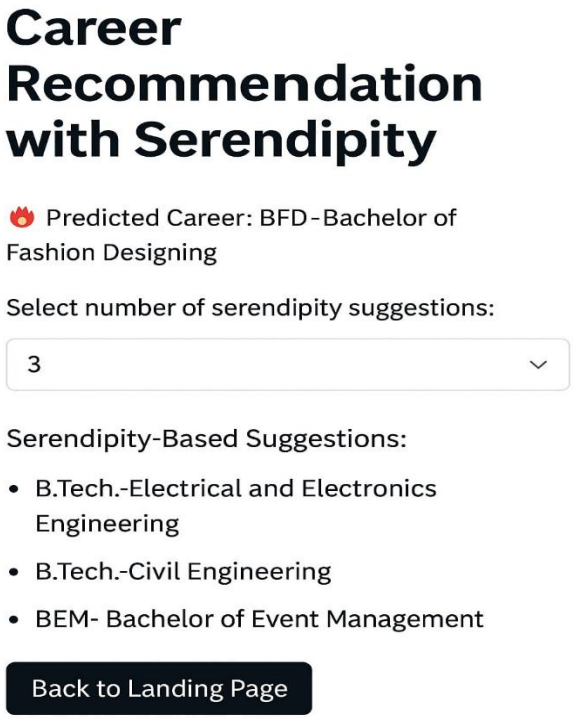


Fig-5.3.5-Recommendation with serendipity

Fig-5.3.6-Recommended Results

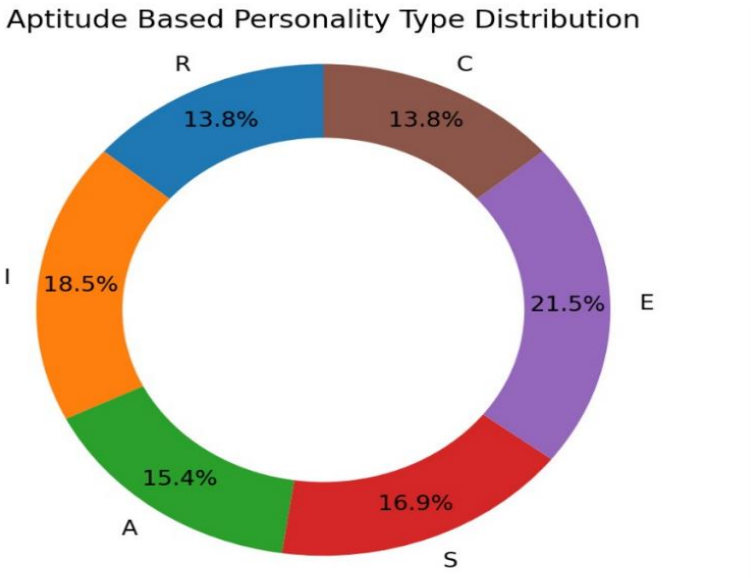


Fig-5.3.7-Personality Distribution Chart

CHAPTER 6

Conclusion

6. Conclusion

The Serendipity-aware Career Recommendation System successfully provides personalized career suggestions by leveraging machine learning techniques such as Random Forest and Natural Language Processing (NLP). By integrating aptitude tests, career assessment surveys, and serendipity-based suggestions, the system ensures a well-rounded and dynamic career exploration experience. Unlike traditional career counseling methods, this approach enhances user engagement by introducing unexpected yet relevant career options based on cosine similarity and TF-IDF-based analysis.

The implementation of Streamlit for an interactive user interface and SQLite for data storage ensures ease of use and accessibility. The system not only improves the accuracy of career predictions but also expands students' awareness of potential career opportunities. Future improvements can include real-time labor market analysis, deep learning-based recommendation models, and integration with industry mentors for guidance.

References

[1] Y. Xu, E. Wang, Y. Yang and H. Xiong, "GS-RS: A Generative Approach for Alleviating Cold Start and Filter Bubbles in Recommender Systems," in IEEE Transactions on Knowledge and Data Engineering, vol. 36, no. 2, pp. 668-681, Feb. 2024, doi: 10.1109/TKDE.2023.3290140.

[2] Y. Dong, "Research on Career Counselling Platform Based on Collaborative Filtering Recommendation Algorithm," *2023 International Conference on Integrated Intelligence and Communication Systems (ICIICS)*, Kalaburagi, India, 2023, pp. 1-5, doi: 10.1109/ICIICS59993.2023.10421730.

[3] C. Sun, "Design and implementation of career planning system based on machine learning," *2023 IEEE 3rd International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA)*, Changchun, China, 2023, pp. 68-73, doi: 10.1109/EEBDA60612.2024.10485905.

[4] W. Liu, L. Niu and J. Zhao, "Framework Design of Intelligent Career Recommendation System for Undergraduates," *2022 International Conference on Intelligent Education and Intelligent Research (IEIR)*, Wuhan, China, 2022, pp. 256-261, doi: 10.1109/IEIR56323.2022.10050076.

[5] D. P. B, M. K, V. K and Y. D, "Job and Course Recommendation System using Collaborative Filtering and Naive Bayes algorithms," *2023 2nd International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA)*, Coimbatore, India, 2023, pp. 1-4, doi: 10.1109/ICAECA56562.2023.10200758.

[6] A. H. A. Rashid, M. Mohamad, S. Masrom and A. Selamat, "Student Career Recommendation System Using Content-Based Filtering Method," *2022 3rd International Conference on Artificial Intelligence and Data Sciences (AiDAS)*, IPOH, Malaysia, 2022, pp. 60-65, doi: 10.1109/AiDAS56890.2022.9918766.