ABSTRACT

Choosing the right career is a challenging decision for students because a career is a significant period of one's life and with opportunities for progress. Students do not have sufficient knowledge to take the decision on their own which may lead to complications in future. In order to avoid future complications students should make the proper decision in selecting the highest benefit career for them. Due to the lack of guidance and unlimited online resources, students base their decisions on subjective perceptions of family and friends. This can potentially cause mismatch among interests, skills of students. To address the mentioned drawbacks, our project builds a recommender system based on student preferences in which it collects the student’s information was collected and recommends the better job domain based on the skills, interests and academic performance provided. In this project we have done comparative analysis among different algorithms like SVM, Decision tree, and Random Forest and have considered the model which has the highest precision, recall rate, F1-score and accuracy.

Keywords: Student preferences, Job domain, Recommendation system, SVM, Decision tree, Random Forest.

I. INTRODUCTION

Career is an essential aspect of a student's life as it defines their future prospects and plays a crucial role in their personal and professional growth. A good career choice can help a student attain financial stability, job satisfaction, and a sense of purpose in life. It can also provide opportunities for personal and professional development, as well as social recognition and prestige. Additionally, a chosen career can shape a student's future aspirations and goals, helping them focus their efforts towards achieving their desired outcomes. Overall, a well-planned and successful career can have a favorable influence on student's self-esteem, confidence, and overall quality of life. A recommendation system is a sort of data filtering system that recommends things or products to consumers based on their preferences, interests, and behaviors. E-commerce websites, social media platforms, music and video streaming services, and other online platforms all employ recommendation algorithms. They create suggestions for users by utilizing various algorithms and strategies such as collaborative filtering, content-based filtering, and hybrid filtering.

II. LITERATURE SURVEY

[1]. Job recommendation system based on skill sets by G. Mahalakshmi, A. Arun Kumar, (2022): This survey paper reviews the resume and job description using stop words removal and Porter’s stemming algorithm, followed by conversion of the text into a matrix using a tf-idf vectorizer. The system the similarity of the job descriptions and resumes, ranks the job descriptions based on their similarity scores, and identifies skills that the user may need to improve upon by comparing their resume with a skills dataset.

[2]. Suleiman Ali Alsaif’s (2022) Learning-Based Matched Representation System for Job Recommendation: This survey paper presents a content-based job recommendation system using Cosine or Jaccard similarities. Job skills are given higher weights than job domain while computing similarity scores. The system involves web scraping job offers from sa.indeed.com, pre-processing the data through tokenization and keyword extraction, and recommending jobs with high similarity scores for a job seeker's resume.

[3]. Trends and characteristics of career recommendation systems for fresh graduated students by Harco Leslie Hendric Spits Warnars (2022): This survey article uses a Systematic Literature Review (SLR) method to investigate the properties and applications of ITS. The research concerns covered include the approach to recommendation systems, AI techniques used, required characteristics, and assessment methodologies for career recommendation systems.

[4]. Student Career Recommendation System Using Content-Based Filtering Method by Adib Hakimi Abdul Rashid, Masurah Mohamad, Suraya Masrom, Ali Selamat (2022): Project aims to help computer science students find suitable careers using a system that recommends jobs scraped from Job Street website.
uses content-based filtering to match user preferences to job requirements and features a user-friendly interface.


Solution for career assistance that assists students in exploring IT job opportunities based on their education, talents, and experience. It employs a revolutionary similarity metric approach to match job roles to a variety of technical and non-technical talents, as well as job role network visualisation, to help students understand how their skills relate to emerging IT employment.

III. MODELING AND ANALYSIS

**Dataset**

The dataset contains 20 columns and 6901 rows. Each row corresponds to an individual and the columns represent various features related to their skills, capabilities, interests, and preferences. The columns include attributes like logical quotient rating, hackathons, coding skills rating, public speaking points, self-learning capability, additional courses completed, certifications, workshops, reading and writing skills, memory capability score, interested subjects, interested career area, type of company want to settle in, taken advice from seniors or elders, interested type of books, management or technical skills, hard/smart worker, ever worked in teams, introvert. The goal variable that we wish to forecast based on the known characteristics of an individual is the suggested job role. This dataset's goal is to suggest an appropriate employment role for a person based on their abilities, hobbies, and other important considerations. The proposed work role for a certain person was predicted using this dataset, which was used to train machine learning models.

- **Number of records:** 6901
- **Number of attributes:** 20
- **Column names:** 'Logical quotient rating', 'hackathons', 'code skills rating', 'public speaking points', 'self-learning capability?', etc. 'Extra-courses did', 'certifications', 'workshops', 'reading and writing skills', 'memory capability score', 'interesting subjects', 'interested career area', 'kind of company wish to settle in?', and 'memory capability score' "Received input from elders or seniors," "Interesting Type of Books", 'Management or Technical', 'hard/smart worker', 'worked in teams ever?', 'Introvert', and 'Suggested Job Role'
- **Output label:** 'Suggested Job Role'

![Figure 1: Dataset](image)
We used Decision Tree, Random Forest, Support Vector Machine algorithms in this project which predicts the job role based on the student's details.

**Decision Tree:** A supervised learning method called a decision tree can be used to solve classification and regression problems, but it is typically favored for doing so. It is a tree-structured classifier, where internal nodes stand in for a dataset's features, branches for the decision-making process, and each leaf node for the classification result. The Decision Node and Leaf Node are the two nodes of a decision tree. While Leaf nodes are the results of decisions and do not have any more branches, Decision nodes are used to create decisions and have numerous branches. The given dataset's features are used to execute the test or make the decisions.

**Random Forest:**
In order to further increase prediction effectiveness, bagged decision trees, which produce a large number of decorrelated trees, have replaced random forests. They are a well-liked 'off-the-shelf' or 'off-the-box' learning algorithm with strong prediction performance and few hyperparameters. As a result of the individual tree regression, random forests produce an average predictive value. Random forests don't resolve well enough.

**SVM-Support Vector Machine:**
SVM is a flexible and effective machine learning technique used for regression and classification tasks. It is capable of doing both linear and non-linear classification and regression tasks on high-dimensional data. By maximizing the distance between the closest data points of various classes, SVM is intended to determine the optimum hyperplane that divides the data into multiple groups. SVM is a strong method that can handle noisy data since it has a regularizing parameter that can help prevent overfitting. SVM is utilized in many different applications despite the fact that it can be computationally expensive for very large datasets. This is because of its versatility and accuracy.
Figure 4: Activity Diagram

Figure 5: System Architecture

Figure 6: Technical Architecture
IV. RESULTS AND DISCUSSION

Precision, recall, and F1-Score are three standard measurements used to assess a classification model’s performance. They are calculated as follows:

Precision is the proportion of real positives among all examples predicted as positive by the model. Precision is calculated mathematically as: Precision = True Positives / (True Positives + False Positives)

Recall is the fraction of true positives among all positive examples is measured by recall. Recall is determined as follows: Recall = True Positives / (True Positives + False Negatives)

F1-score: The F1-score is a weighted average of precision and recall that balances both measurements to get a single score that represents the model’s overall performance. It is determined as follows: F1-score = 2 * ((Precision * Recall) / (Precision + Recall))

<table>
<thead>
<tr>
<th>Performance/Algorithm</th>
<th>Precision</th>
<th>Recall rate</th>
<th>F1-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision Tree</td>
<td>0.0875</td>
<td>0.0868</td>
<td>0.0870</td>
</tr>
<tr>
<td>SVM</td>
<td>0.0655</td>
<td>0.0688</td>
<td>0.0669</td>
</tr>
<tr>
<td>Random Forest</td>
<td>0.7886</td>
<td>0.0768</td>
<td>0.0734</td>
</tr>
</tbody>
</table>

Accuracy is a machine learning performance parameter that measures the proportion of correctly identified cases out of the total number of instances examined. In other words, accuracy measures how well a model can predict the correct outcome for a given set of inputs.

Figure 2: Accuracy Graph

V. CONCLUSION

In conclusion, a career recommendation system based on student preferences has the potential to become a valuable tool for students seeking career guidance. By analyzing a student’s interests, skills, and educational background, the system can suggest potential career paths that are in line with the student’s goals and aspirations. With the help of advanced technology and machine learning algorithms, these systems can provide more accurate and personalized career guidance to students, helping them make informed decisions about their future career paths.

VI. FUTURE SCOPE

A career recommendation system based on student preferences has the potential to become a valuable tool for students seeking career guidance. Such a system can be improved by incorporating new technology and machine learning algorithms to present students with more personalized and accurate recommendations. By analyzing a student’s interests, skills, and educational background, the system can suggest potential career paths that are in line with the student’s goals and aspirations. The system can also provide career exploration tools, such as job descriptions, salary data, and career path options. This can help students gain a better
understanding of potential careers and make more informed decisions about their future. In addition, the system can provide insights into the current job market and industry trends, enabling students to make informed decisions about their career path. This might assist them in keeping abreast of the most recent advances in the sector and in making smart decisions regarding their education and future. Ultimately, the future scope is vast and holds great potential for both students and career professionals. By incorporating advanced technology and machine learning algorithms, these systems can provide more personalized and accurate career guidance to students, helping them make informed decisions about their future.

VII. REFERENCES


