International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 06 | June - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

STUDENTS CAREER GUIDANCE SYSTEM

Jeenath Laila N[1], Sathya M[2], Jumana Muneera M[3] , Mary Jasmine M[4]

[1]First Author Professor,Department of CSE & Government College of Engineering, Tirunelveli
[2]Second Author PG student, Department of CSE & Government College of Engineering, Tirunelveli
[3,4]Third and Fourth Author UG student Department of CSE & Government College of Engineering, Tirunelveli

---

Abstract - Most students across the world are always in confusion after completion of higher secondary and the stage where on choosing an appropriate career path. The students don't have adequate maturity to accurately know what an individual must follow to choose a congenial career path. Every student undergoes a series of doubts or thought processes on what to pursue after 12th which is the great confusion. Then comes the next agony is have the essential skills for chosen stream.Career counseling system is used to predict the suitable department for an individual based on their skills assessed from the information about the student's interests in various fields and the personal traits . Also, the students who are pursuing their graduation are in dilemma about the career opportunities after their graduation,the system gets the information about their interests by getting input from the students and recommend suitable career domain by using machine learning algorithm.A Supervised machine learning algorithm K-NearestNeighbours(KNN)algorithm classifies the students depends upon their interest , academics results and personality into various fields and recommend the various career domain that helps the students in choosing right career path

Key Words: optics, photonics, light, lasers, templates, journals

1.INTRODUCTION

The right career direction boosts up the performance of students and increases their motivation level.The career counseling-based systems are most important to judge the skill of students and accordingly, finding right career role to them .Career counseling is required to solve the problems which come into the mind of students for where the career guidance will play an important role to guide students to opt for the right course. Various resources, and services are encompassed within this system, aimed at guiding students through the process of career exploration, goal setting, skill development, and job search. By providing valuable information and support, students are provided by right career guidance to navigate the complex world of work and make choices that align with their interests, skills, and aspirations.The primary goal of a career guidance system is to empower individuals to make correct decisions about their careers. It is recognized that choosing a career involves more than just selecting a random course stream and an available job, it requires an understanding of one's own abilities, interests, values, and goals. In this system students are provided with the tools by a career guidance system to assess themselves and explore a wide range of career options. One key aspect of a career guidance system is self-assessment. Assessments that allow students to identify their strengths, interests, and values are provided by the system.A Questionnaire that contains questions related to various fields of interest are provided by the system in which the students has to attend[1]The personality traits, aptitudes, interested course fields and preferred work environments are gained through these assessments. By understanding themselves better,choices about potential career paths that align with personal attributes and preferences of the students could be done.Resources and information about different course stream ,careers, industries, and job roles are provided to students by the system. This includes information about different course stream ,educational requirements,job descriptions, salary ranges, , and labor market trends.A clear understanding of the skills and qualifications needed for different careers, as Well as the potential job opportunities available, can be gained[5].Academics percentage, personality traits and area of interest that are given as an input to the questionnaire are used to recommend suitable career to the students by the system.Recommendation of suitable career is done by using different machine learning techniques.Various machine learning algorithms are used for this process.Each algorithm has its own pros and cons.

Machine Learning Approach

Machine learning is a subfield of artificial intelligence that focuses on the development of algorithms and models that allow computers to learn and make predictions or decisions without being explicitly programmed. It involves training a computer system with data, enabling patterns to be identified, inferences to be made, and performance to be improved over time.Various machine learning algorithms have been used for recommending an opt career to the students.A Supervised machine learning algorithm known as K-Nearest Neighbours(KNN) algorithm is used for the recommendation of suitable career. KNN algorithm is used for both classification and regression tasks. It is an instance-based and non- parametric learning algorithm, with predictions being made based on the similarity of input data to known data points.

KNN Algorithm

In KNN algorithm, the training data comprises labeled examples, where each example has a set of features and a corresponding class or value. When a new, unlabeled data point is presented, the algorithm calculates the distance between that point and the existing data points in the training set by using any one of the methods such as Euclidean Manhattan and Minkowski distance. The value of "K" represents the number of nearest neighbors to consider and it various depends on the problem and dataset to be considered.To make a prediction for the new data point, the algorithm selects the K nearest neighbors based on distance, and a majority vote is taken among those neighbors. The data point is then assigned to the class that has majority of nearest neighbours.

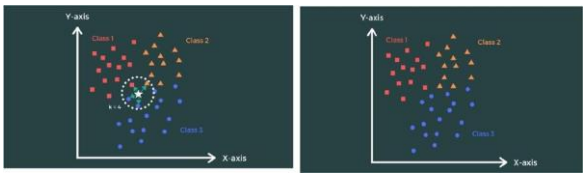

Fig -1 : KNN Classification

The performance of the algorithm is influenced by the choice of the value of K. maller value of K increases the sensitivity of the algorithm to local variations, potentially resulting in overfitting, while a larger value of K considers more neighbors, resulting in a smoother decision boundary but may overlook local patterns.KNN is considered a simple yet effective algorithm, particularly in cases where the data exhibits clear clustering or separation patterns. A career guidance system driven by the utilization of KNN algorithm is an innovative approach designed to assist students in making decisions about their career paths. The capabilities of the KNN algorithm are utilized by this system to provide personalized recommendations and guidance based on similarities between individuals and various career options.

2.LITERATURE REVIEW

TSaad Ahmed et al.[1] conducted a study on "The Online Career Guidance System" in 2018 ,where three main tests namely General Aptitude Test , Personal Interest Test , and Domain Test were conducted. Admin can add and remove questions in the database using MySQL.The advantage from this study is easy implementation and access of data,but predicted accuracy is low as it doesn't use any of the machine learning techniques.

Aniket Surve et al.[2] conducted a study on "Student Career Guidance System"in 2021 in which the self-assessment form is provided, the source for Massive Open Online Courses(MOOCs) relevant with the role predicted, Books and Certifications.The prediction is made using Logistic Regression model. Once a role is predicted,the frontend presents the user with the sources for the certifications and MOOCsby fetching it from the MongoDB database. The backend process started with data acquisition. Dataset was obtained from GitHub. By using Logistic Regression model output is provided in terms of probabilities. The advantage of using logistic regression is that it is very much easier to implement but this algorithm assumes linearity between the predicted variables and features that are used for prediction while these assumptions is not always true that makes this algorithm inefficient.

Sana Thamke et.al[3]conducted a study on "Web Based Scientific CareerCounselling System"developed using Naïve Bayes classification algorithm in which classification is done by calculating probability for each class label, which is the probability that a given data point belongs to a particular class based on the frequency of occurrence of that class in the training data. The predicted accuracy is low as it assumes that features of the input data are independent of each other, which may not always be true which makes this algorithm inefficient

K.Sai Vivek Reddy et.al[4] conducted a study on "Career Guidance System using Ensemble Learning" in 2022.Machine learning algorithms such as Decision Tree, Random Forest, Voting Classifier have been used.Dataset consists of numerical data of academic's marks, and characteristic data of a student. Dataset is taken from Kaggle, and it consists of 20,000 rows and shown in the form of input parameters.It has observed that the accuracy on those algorithms is 66%.

Maha Nawaz et.al[5] implemented a Web-based CareerCounseling "Automated Career Counseling System for Students" using case-based reasoning(CBR) and Decision Tree algorithm. Decision Tree is employed through Weka,a software collection of machine learning algorithms. The accuracy of the CBR algorithm is approximately 80% while the accuracy of the Decision Tree is 50-60%.The data used was collected from Graduate students of 5 different Universities of Lahore in the form of a questionnaire list.

3. EXISTING SYSTEM

Career guidance can help individuals identify careers that match their skills, interests, and personality traits. This process can help people make decisions about their career paths.The existing system described in the literature review have cerain limitations.The system was developed by using various machine learning algorithms such as Naïve bayes classification, Decision trees,Support vector machine(SVM). The disadvantage on Naïve Bayes algorithm is that,the probabilistic classification is done. Since the algorithm finds the conditional probability it assumes that the features are independent to each other. This assumption is not always true in different cases which makes this algorithm inefficient[3]. In decision trees when the size of the tree grows it leads to pruning i.e., Data compression techniques that causes overhead to the system.[4] The algorithmic complexity and memory requirements of SVM was very high.Choosing an optimal kernel function in case of non-linear data in SVM and this task is difficult.The problem in Logistic Regression is that it assumes a linear relationship between the features and the target variable.[2] It may not perform well when the relationship between the features and the target variable is non-linear which sometimes gave the ambiguous predictions.

4. PROPOSED METHODOLOGY

The Students Career Guidance is a system that recommends career options to the student by getting the information about their interest in various course stream personality traits and academics percentage as input then the suitable career will be recommended. This helps the students to identify their interests among various fields and then leads to right career domain. It aims to overcome the limitations of the existing system by using an efficient machine learning algorithm KNN that helps in recommending the suitable career for the students. KNN algorithm is a non-parametric algorithm that does not make any assumptions about the distribution of the data .It is a simple and intuitive algorithm and recommendation of suitable domain is unambiguous which makes the system efficient. Information about the recommended career includes the skills required,availability of career options are also provided by the system.Figure 2 depicts the workflow of career guidance system.Dataset containing the students academic results ,interest and suitable career options are collected. Preprocessing of the data is done. A model is built using KNN algorithm.KNN model is trained and tested with training and testing data.There are 16 different career options and 6 major course stream the model used for classification The system takes input from the students and based on interests ,the system classifies the students into the respective domain.It also recommends suitable career role and also provided with the information about the course stream which includes the skills required for the particular course etcRecommendation of suitable career based on interest of the students helps to identify

the necessary steps such as acquiring specific skills or pursuing further education to reach their desired career destinations.

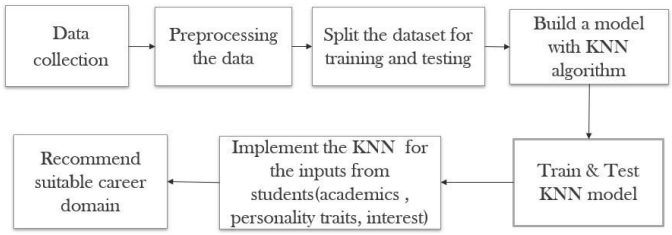

Fig -2 : Block Diagram

Data Collection

The datasets that contains the Student's academics, personal traits and their interests are collected from Kaggle website and Github. The dataset contains the attributes such as Student's name ,email-id ,phone number. Based on the student's interest and academics, the suitable career domain will be recommended.The dataset contains the different coursestream and various career paths.It contains different domain values includes diverse coursestream like Science, Arts, Medicine, Commerce Law, Engineering.Datasets also contains career domain as Database administrator AI ML specialist ,Graphics designer ,software developer,Technical writer etc.,

Data Preprocessing

The data was preprocessed to eliminate inconsistencies and error.It is essential to improve the performance.Preprocessing helps to make the data consistent eliminating any duplicate,and irregularities in the data and improve the accuracy of the results.The goal of data preprocessing is to transform the raw data into a clean, consistent, and suitable format that can be effectively utilized by machine learning algorithms All the null values, punctuation in the student's datasets are removed. Major procedures in data preprocessing

Label Encoding : Label encoding replaces each category with a numerical label.The label values in the student's dataset are encoded such as Not Interested- 0,Beginner-1 Average -2, Intermediate-3,Excellent-4,Professional-5.Depends on the encoded value the distance between each datapoints is calculated and suitable career domain will be recommended

Feature Selection

Feature selection is the process of selecting a subset of relevant features (variables, attributes) from the original set of features to improve the performance of a machine learning model. By selecting the most informative and discriminative features, feature selection can offer several benefits such as Improved model performance, Faster training and inference, Enhanced interpretability.

Dataset that has been used in this project contains student's attributes such as name,email id,phone number along with features about their interest,personal traits and academics that helps to classify the students into their respective domain.The attributes such as name,email-id doesn't contribute to the classification process.Hence those features were discarded and only the necessary features were selected.

Training And Testing Data

The dataset contains 9100 data about the student's interest that is split into training and testing datasets as 70% for training and 30% for testing. The training data is used totrain the KNN model, and the testing data is used to evaluate the performance of the model.

The scikit-learn library in python provides an implementation of train-test split evaluation procedure via the train_test_split() function.

X_train, X_test, y_train, y_test=train_test_split(X,y,test_size=0.3,random_state = 524

KNN Algorithm

KNN algorithm is one of the supervised machine learning algorithm.It is used to predict the career paths of students in this project. This algorithm calculates the distance between the new data points and the already existing datapoints in the training dataset by Euclidean distance formula.

Step-1.Data are collected and preprocessed .The preprocessed data is split into Training and testing data.The training data is used for training the K
NN algorithm.

Step-2.KNN algorithm is trained with different clusters of datapoints i.e.,
the different domain values.The KNN model learns from the labelled data that which datapoints belongs to the respective career domain options.

Step-3. The system takes the input from the students such as their academics Percentage ,Personality traits and Interests on different areas etc.,

Step-4.KNN takes the new data input and calculate the distance between each datapoints in the dataset .K value is fixed and K Nearest datapoints are selected.

Step-5.The domain values of the K nearest neighbours are used to predict the career domain of the new data input that are given by the students.

Step-6.The new datainput is assigned to the domain category that has the maximum of nearest neighbours among the K - Nearest Neighbours.

Thus the career options will be recommended by the system.

KNN algorithm is implemented by KNeighboursClassifier() class from scikit-learn Library in python

model = KNeighborsClassifier(n_neighbors=5)
model.fit(X_train.values,y_train.value

Finally, the model is tested on the data to get the predictions.
y_pred = model.predict(X_test.values)

RESULTS AND DISCUSSION

Figure 3 shows the HomePage of the web application that has the description ofsystem and a link where the students able to login

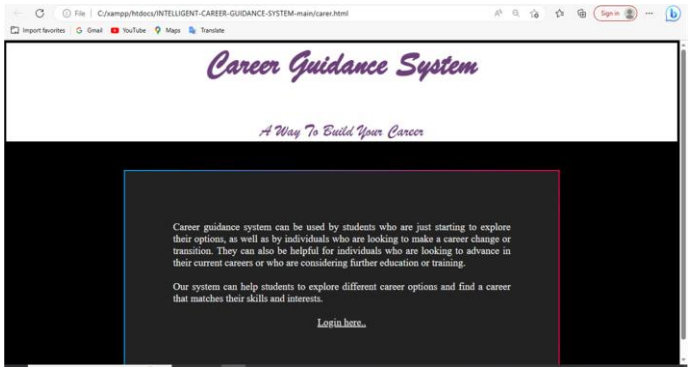

Fig -3 : Homepage for Career guidance system

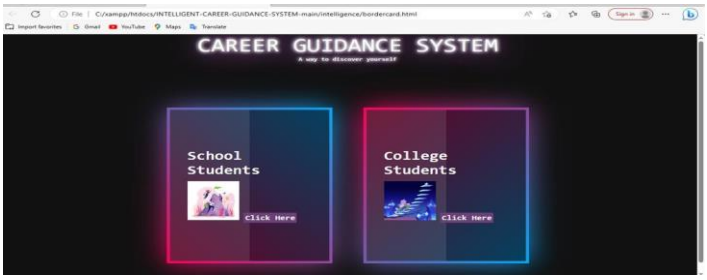

Fig -4 : DashBoard

Figure 4 shows a dashboard page in which both the higher secondary students and students who are in their graduation can login into the system by clicking the respective links

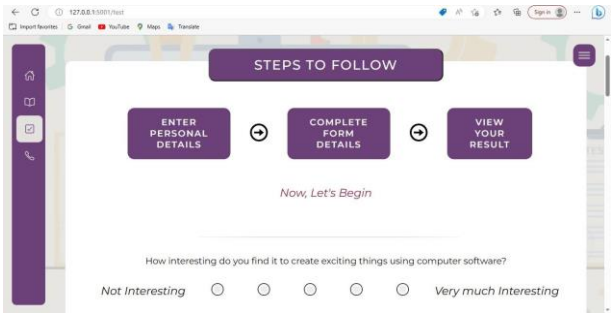

Fig -5 : Questionnaire for the students

Figure 5 depicts the page in which the students can give their interest from not interested to the most interesting to a set questions that are related to different fields of course stream based on the interest of them.

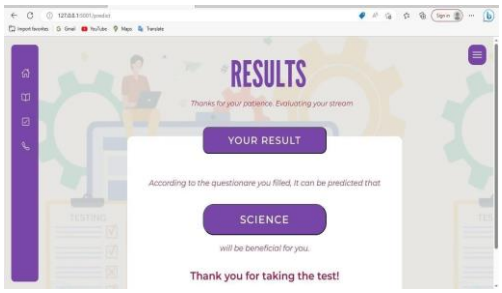

Fig -6 : Recommendation of suitable course stream

Fig 6 depicts that ,after the students have answered the questionnaire, the system recommends the suitable course stream based on the responses from the students.There are six major course stream hat the system recommends such as Arts,Science,Medicine,Commerce,Engineering,Law and the students get benefited from the system on choosing a right career after higher education. The career guidance system takes the input from the students that was given by students and suitable career options was recommended

Fig -7 : Questionnaire for the students

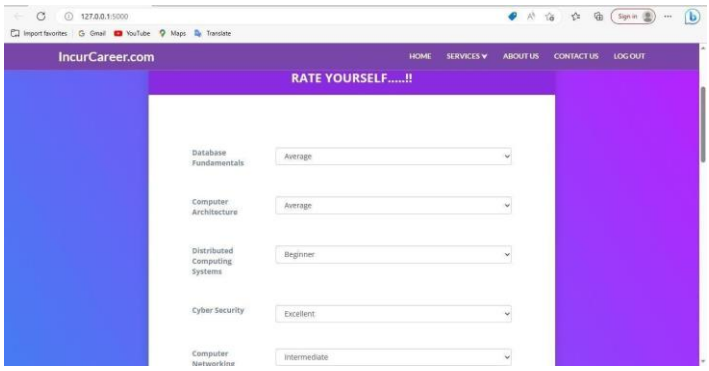

Figure 7 depicts the questionnaire that students has to fill the questionnaire on academics levels, personality traits & interest. After that suitable career domain based on their skills, area of interest will be recommended by the system

Fig -8 : Recommendation of Suitable career role

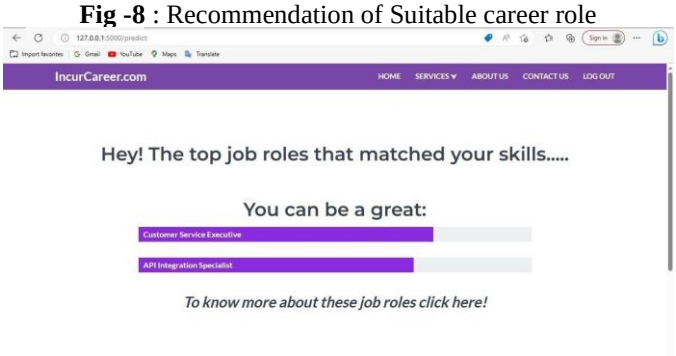

Figure 8 depicts the page in which the suitable career roles will be recommended by the system based on the questionnaire filled by the students and area of interest. It also Provides the description of the job roles, qualifications needed for the job roles.

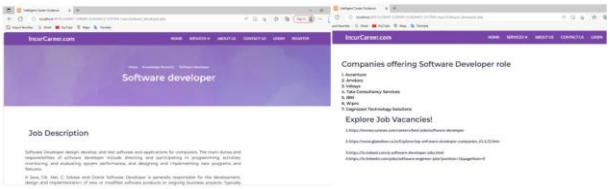

Fig -9 : Knowledge NetworkPage

Discussion:

1.The accuracy rate is nearly 95% .The system's ability to provide accurate and relevant recommendations are crucial for its effectiveness.

2..A well-designed and intuitive interface can enhance user engagement and satisfaction. An interactive and User-friendly interface is provided by the system.

Recommendation of career along with the information about different career paths, industries, educational programs, job market trends, and required skills. comprehensive information about each career option is provided by the system to make well-informed career decisions.
Career guidance system provides numerous advantages on recommending suitable career domain still it has limitations are only fewer input parameters are considered for classification process which may cause inconsistencies in recommendation process and the career domain values and the course stream that are considered for recommendation in the system are limited

CONCLUSIONS

Career guidance system is a web application based on a classification model designed to predict career options for both Higher Graduate Students and Undergraduate Students. One of the main problems students face is the lack of flexibility to pick the curriculum and job of their personal preference. The problem was successfully rectified by the system. K- nearest neighbor algorithm which is one of the efficient algorithms was used to classify the students based on their interest to the respective domain. The correct guidance,resources, and recommendations were provided that help students to explore various career options, align their interests and skills and make correct career choices.

REFERENCES

1. Aniket Surve, Amit Singh, Shivam Tiwari “Student Career Guidance System using Machine Learning” - International Research Journal of Engineering and Technology in 2021.

2. Alharbi Z, Cornford J, Dolder L and De La Iglesia B, “Using Data Mining Techniques to Predict Students at Risk of Poor Performance,” SAI Computing Conference (SAI), IEEE ,2016.

3. K.Sai Vivek Reddy , M.Akhila Reddy , Gagandeep Kaur “Career Guidance System using ensemble learning”- 2nd International Conference on Advancement in Electronics & Communication Engineering in 2022.

4. Ojekudo Nathaniel Akpofure,Tombari Tete “Design and Implementation of a Career Guide Information System for Secondary School Leavers An Object Oriented Approach”- International Journal of Science and Research in 2013

5. Roopkanth, K., & Bhavana, V, “Student Career Area Prediction using Machine Learning,” IEEE 2018.

6. Saad Ahmed, Suraj Prakash Maurya, Vaibhav Khalane “Online career guidance system”-International Research Journal of Engineering and Technology in 2018.