# Intelligent IT Career Advisor using Machine Learning

Hina Bediya, Dharmesh Mishra, Nikhita Mangaonkar

MCA Department, Sardar Patel Institute of Technology, India

Corresponding author: Dharmesh Mishra, Email: dharmesh.mishra23@spit.ac.in

The tremendous growth in the Information Technology (I.T.) field has made choosing a job more difficult than ever before, calling for intelligent, data-driven solutions to close the skills gap and meet business demands. This research study presents an innovative career advisory system that leverages the K-Nearest Neighbors (KNN) algorithm to recommend IT roles tailored to user skill profiles. Using proficiency ratings of skills across 17 IT domains like Cybersecurity, Database Management, Software Development, etc., in a dataset of 9178 entries, the model will provide accurate and personalized career advice. The model provides career predictions for positions like database administrator, software developer, business analyst, etc., by analysing both self-assessed skills and objective knowledge evaluations. This results in a classification accuracy of 98 percent, supported by robust cross-validation (0.83 ± 0.03) and high precision, recall, and F1-scores across all job roles. While addressing user input uncertainties, the dual-assessment system guarantees strong suggestions. The system, which targets IT students, fresh graduates, and professionals who are looking to make a career shift, demonstrates the power of machine learning in aligning individual competencies with the ever-changing demands of the job market. This research points out the importance of including intelligent systems in career planning in order to provide users with useful information and encourage well-informed choices in a constantly changing business environment.

**Keywords:** Machine Learning, career path prediction, IT careers, skill assessment, personalized guidance, k-nearest neighbors.

2025. In Puneet Kumar Gupta (ed.), *Computational Models for Intelligence and Automation*, 48–58. Computing & Intelligent Systems, SCRS, India. DOI: https://doi.org/10.56155/978-81-975670-1-8-5

# 1. Introduction

It might be difficult to choose a professional path, particularly in a field that is evolving quickly like Information Technology (I.T.). It might be challenging to determine which career path best suits one's interests and abilities with so many options to choose from, including database administration, software development, cybersecurity, and more. Many students, fresh graduates, and professionals considering a career change frequently struggle to decide which area of Information Technology to pursue. Conventional career guidance and decision-making resources can be restrictive because they frequently fail to offer individualized insights based on a person's true interests or abilities. We have created a machine learning (M.L.)-based system to help users choose the best I.T. career choices in order to solve this problem. By integrating two forms of input—users' self-reported degrees of proficiency in various I.T. domains and their responses to objective questions that assess their knowledge in certain fields—this system adopts a novel approach. The system may determine which I.T. professions such as database administrator, data scientist, business analyst, etc.—might be the most suitable for each user based on these inputs. This system's objective is to provide tailored, data-driven career advice. Through the analysis of industry-relevant data, the system lets users evaluate their current position in relation to the abilities needed for different I.T. professions. In this manner, the algorithm may identify areas where they might need more growth while also suggesting career paths that align with their existing interests and skill set. A wide range of people are targeted by this system, including professionals considering changing careers in I.T., recent graduates seeking to match their credentials with practical job openings, current I.T. students who aren't sure which specialization to pursue, and educational institutions seeking to offer career counseling to their students. In the end, the system is about closing the gap of knowledge between academic study and the changes in demand in the I.T. sector as a way of aiding people make well-informed decisions pertaining to their careers. This application helps users make the best career decisions, identify areas for skill improvement, and better position themselves for success in the ever-evolving I.T. industry.

Current career guidance systems, mainly in the IT industry, fail to provide personalized recommendations that are data-driven, and aligned with both individual skill sets and industry trends. Most current systems rely only on subjective self-assessments or static datasets that do not reflect the dynamic nature of requirements in careers. This paper introduces a new approach that improves accuracy through dual assessment, combining both self-reported proficiency with objective evaluations, to improve the accuracy of the career recommendation. The system achieves 98% accuracy by using the K-Nearest Neighbors algorithm, making it different it from the traditional methods. It bridges the gap between academic training and industry by assisting students, graduates, and professionals in making informed decisions while pursuing or changing careers within the IT sector.

# 2. Literature Review

For many years, career counselling has been an essential component of professional and educational growth, assisting people in navigating the challenges associated with making career decisions. As the Information Technology (I.T.) industry continues to develop at a rapid pace, there is an increasing need for data-driven, customized career guidance tools. This literature review examines important developments in the area of intelligent career guidance systems, especially those that use machine learning (M.L.) algorithms to provide tailored career advice.

## 2.1 Intelligent Career Guidance Systems

Conventional career counseling techniques frequently fail to offer customized advice based on people's unique skill sets and the changing needs of the job market. Roy (2020) underlined the significance of incorporating life skills and flexibility into career guidance [1]. Career counseling should be dynamic, student-centered, and relevant to each person's goals and abilities, according to the study. Shreyas et al. (2024) further developed this argument with the help of AI by discussing recent breakthroughs in machine learning algorithms in predicting careers and further establishing that these systems will give better and real-time advice [2].

In the context of I.T. careers, systems that provide individualized career recommendations based on the assessments of skills have lately been gaining popularity. A machine learning-based career advising system was proposed by Ajay et al. (2022) that uses a user's self-reported skill levels and preferences to predict appropriate career options [3]. Similar to this study, their methodology relies on algorithms such as K-Nearest Neighbors (KNN), in order to suggest content adapted to the user's profile. This study showed how such systems could be used to overcome the gap between academic qualification and the needs of practical employment, but it was not concerned with objective reviews of users' technical skills.

The use of ensemble learning algorithms in career predictions was investigated by Pandey and Maurya (2021), who showed how various machine learning models can be integrated to improve the accuracy of career forecasts [4]. This is in line with recent advances in this area, where hybrid approaches are being used to improve the system. Rutina and Soika (2020) discussed career guidance for first-year students in work-based learning and emphasized the early practical experience in career choices [5]. It thus shows how to guide such freshmen in their beginning academic and professional careers.

## 2.2 Machine Learningin Career Guidance

Over time, machine learning has developed into a useful instrument for providing data-driven career guidance. Westman et al. (2021) explored the application of artificial intelligence in career counseling and its capacity to improve decision-making [6]. They concentrated on using AI to assess soft and technical abilities as a comprehensive way to provide career guidance. Their study, however, was somewhat more conceptual and concentrated more on potential future developments and applications. Its AI career prediction models could not be empirically validated. In their discussion of intelligent career advice systems, Kanathur et al. (2023) suggested using a unidirectional model to apply a variety of datasets, such as individual preferences and academic and industrial trends [7]. Their approach addressed the multi-faceted nature of career planning but did not particularly focus on I.T. careers. It was therefore a gap in understanding how machine learning could be applied to this rapidly evolving field.

The objective of Sinha et al. (2023) was to employ machine learning algorithms to predict the career of the students based on their academic success and interests by aligning it with recent advancements in the individualized systems of career predictions [8]. Yadav et al. (2023) illustrated that ANN and M.L.P. classifiers are accurately used to predict the career, using information such as the academic background and personal preferences [9]. Bilon-Piórko and Thomsen (2022) summed up the modern concepts of career guidance, urging a holistic approach [10]. Ritonga and Wangid (2022) stressed the need to contextualize career guidance for I.T. students [11]. Ajgaonkar et al. (2023) presented EduKrishnaa, a web application using a multi-class classification algorithm for personalized career guidance [12].

## 2.3 Skill Mapping and Assessment

The next crucial element of the career guidance system is skill assessment and mapping. Our model includes skill proficiency, which ranges from "Not Interested" to "Professional," to make predictions with finer resolution based on user self-reporting. Furthermore, the I.T. domain proficiency assessments often comprise both subjective self-reports and objective tests, as presented in the dataset used for this study. This dual-assessment mechanism would make a more robust as well as correct representation of capability with the user as opposed to that of using one kind of input mechanism alone.

Jemini-Gashi et al. (2023) found whether professional counseling sessions affect the outcome expectations, confidence of career-related abilities, and the goal of career [13]. Their study indicated the need for interventions to improve the level of self-confidence, coupled with the proper establishment of goals in influencing psychological factors that impact career choice. The consequences of their findings point to the larger context of career counseling: self-assurance and well-defined objectives, even though they are not necessarily directly related to machine learning. An AI-driven job advisor that incorporates both emotional and cognitive variables may benefit from this dimension's ability to improve individualized advice.

Despite its promise, machine learning-based career counseling systems still face a number of difficulties. The accuracy of the datasets required to train machine learning models is one of the main obstacles. The majority of current systems, like those covered by Kanathur et al. (2023), are based on static datasets that might not accurately represent the most recent employment trends [7]. In order to keep the career recommendations current and relevant, such work will eventually incorporate real-time job market data and industry needs.

### 2.4 State-of-the-Art Comparison

New developments in machine learning-based career counseling systems have looked into creative ways to increase precision and customization. Ajay et al. (2022) presented a KNN-based career advising model that relied on user-reported preferences but was limited in its comprehensiveness by the absence of an objective assessment of users' technical skills [3]. Similar to the current study, Mathur et al. (2022) used a dual-assessment mechanism to measure users' subjective and objective skills. In contrast to KNN, their usage of less complex algorithms, such as decision trees, produced predictions with a lower accuracy [14].

The ability of ensemble learning techniques to improve the accuracy of career predictions was shown by Pandey (2020) [15]. However, these techniques are less appropriate for real-time applications due to their computational cost. Although they provided only conceptual insights without empirical confirmation in the IT sphere, Westman et al. (2021) underlined the significance of integrating non-technical abilities into AI-driven career advising [6].

The KNN-based model used in this work, on the other hand, outperforms current techniques with a 98% classification accuracy and real-time predictions. By integrating subjective self-assessments with objective evaluations, the dual-assessment process offers a more thorough assessment of user competencies. By providing individualized, data-driven career recommendations that are accurate and effective, this method closes the gap between academic credentials and changing market expectations.
In conclusion, the dynamic nature of IT careers necessitates tailored solutions, despite the fact that the promise of AI and machine learning in career guidance systems has been the subject of countless research studies. By combining dual evaluations with industry-aligned techniques, the current study builds on earlier research by enabling users to make well-informed career decisions. In a constantly changing environment, this study helps to build a strong framework for career counseling by bridging the gap between individual skills and market expectations.

## 3. System Design and Methodology

### 3.1 System Overview

Based on user input, the suggested system makes recommendations for I.T. career routes using the K-Nearest Neighbors (KNN) algorithm. By examining preferences and skills, it makes choosing a career easier. Below is an overview of the application's key features:

### 3.1.1 Skill Assessment and Prediction

In order to assess their knowledge, users can either take an objective quiz or enter their level of expertise in a variety of I.T. domains (such as database management, cybersecurity, and more) on a scale from "Not Interested" to "Professional." The KNN model compares the user's data with previous skill-role mappings in the training dataset to forecast the most appropriate career roles based on the input that was chosen.

### 3.1.2 Career Path Prediction

The system forecasts which I.T.-related job roles would be best for the user based on their input. Database administrator, software developer, data scientist, and other positions could be among the possibilities. The system provides a list of possible job alternatives by matching the user's talents with the professional roles in the dataset using the KNN model. The method helps users make decisions more quickly by providing an instant profession recommendation once they enter their talent ratings.

### 3.1.3 Machine Learning Model Deployment

This system's primary machine learning algorithm is the KNN model. It predicts a career path by comparing a user's input (skills) with the dataset's historical data and identifying the most similar cases. The KNN model learns to forecast appropriate jobs based on user input thanks to the training data, which includes a variety of I.T. roles and related skill sets.

The following is the flowchart of system functionalities, which describes the fundamental steps from user input up to career recommendations (Figure 1).

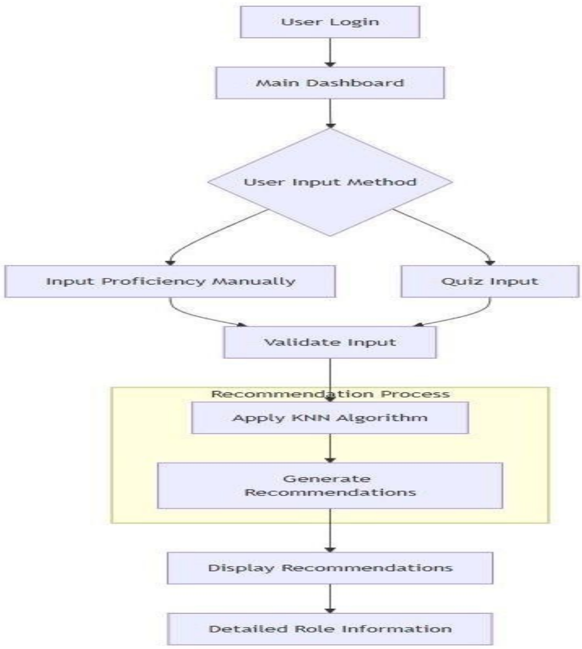


**Figure 1.** System Overview Flowchart

# 4. Methodology

## 4.1 Dataset Preparation

The dataset, which came from Kaggle, correlated skill levels in 17 different I.T. domains with 17 different I.T. occupations, including technical writer, database administrator, and A.I./M.L. specialist. Database basics, networking, programming, cyber security, data science, and other areas were among the proficiency areas.

### 3.1.4 Data Preprocessing

- **Proficiency Mapping:** To translate categorical proficiency levels into numerical numbers, a unique scale was created.
- ”Not Interested” was mappedto1, ”Poor”to2, ”Beginner”to3, ”Average”to5, ”Intermediate”to6, ”Excellent”to7, and ”Professional”to9.
- Higher precision for experienced users was given priority by this non-linear mapping, which made sure the model better represented differences in advanced skill levels than a conventional linear scale (e.g.,0–6).
- **Handling Missing Data:** Mandatory inputs ensured complete live data for predictions, while rows with missing values were eliminated because of the data set’s low null value presence.
- **Feature Encoding:** Label Encoder was used to convert career responsibilities into numerical labels.
- **Feature and Target Separation:** The dataset was divided into the target variable (y), which represents I.T. careers, and features (X), which reflect competency levels.

## 4.2 Model Selection

The K-Nearest Neighbors (KNN)algorithm was chosen because it works well for multi-labelpredictionsandmulti-classcategorisation.ImportantAlgorithmFeatures:

- **Flexibility for Multi-Label Outputs:** KNN allows users whose skill sets match a variety of roles to be recommended for numerous careers. To produce these predictions, the model makes use of a probability distribution that is derived from the nearest neighbors.
- **Hyper parameter Tuning:** To balance the trade-off between underfitting (high k) and overfitting (low k), the number of neighbors (k) was adjusted to 5. Cross-validation performance was used to determine this figure.

## 4.3 Model Comparison

Before settling on the KNN model, several machine learning algorithms are trained and tested to determine the best-performing algorithm for career trajectory. The models compared are:

**K-Nearest Neighbors (KNN):** KNN is a Simple Learning Machine algorithm that involves working on the Supervised Learning approach. It will classify new data points, depending on the similarity of the characteristics and the categories of old ones. The KNN algorithms save all available data and classify it as new based on the similarity obtained between the new data and previously saved data. KNN assigns the new data to the most appropriate category based on its proximity to other data points. In the training phase, KNN simply stores the dataset and upon receiving new data, it classifies it by looking at the 'k' closest examples in the feature space. The category with the majority among those nearest neighbors is assigned to the data point. KNN is intuitive, easy to implement, and effective in scenarios where explicit relationships between features are absent but patterns can be identified through proximity.

The following diagram illustrates how the K-Nearest Neighbors algorithm classifies a new data point based on the proximity to its nearest neighbors (Figure 2).

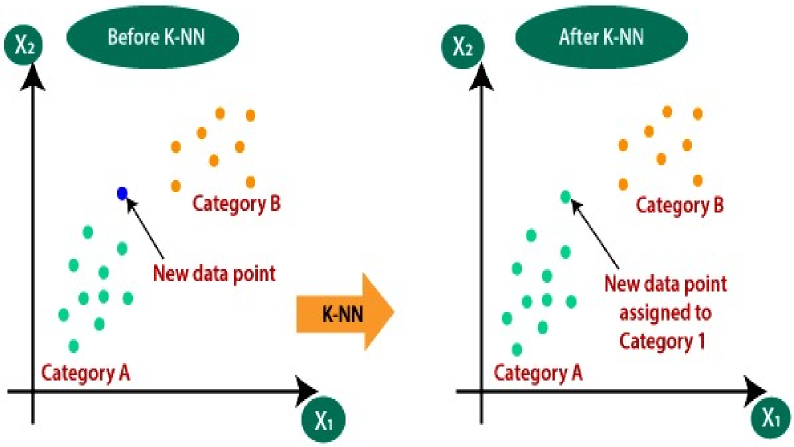


**Figure2.** K-Nearest Neighbors (KNN) Model

**Decision Tree:** A decision tree is fundamentally a hierarchical structure, where the internal nodes evaluate an attribute and the leaves represent final decisions. Building a decision tree is actually a process of dividing a set recursively with regard to different attributes' values. In each node, which happens to be the internal one, the algorithm decides which is the most suit able attribute that can be used for splitting the data according to some criteria. This process continues until a pre defined stopping condition is met, such as the maximum depth of the tree or ensuring that a minimum number of instances are in each leaf node.

The structure and working of the Decision Tree algorithm are represented visually in the diagram, highlighting how decisions are made at each node (Figure 3).

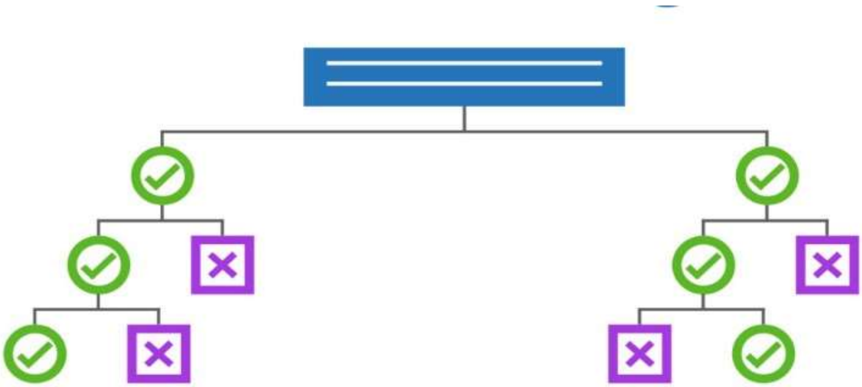

**Figure 3.** Decision Tree Model

**Support Vector Machine (SVM):** A Support Vector Machine (SVM) is a very powerful technique that is frequently used for regression, outlier detection, and a variety of classification tasks, both linear and nonlinear. Because of their remarkable adaptability, SVMs are used in a wide range of domains, including as spam filtering, handwriting recognition, face detection, gene expression analysis, text and image categorization, and anomaly detection. SVMs may categorize data that cannot be divided by a linear boundary by using kernel functions to convert data into higher-dimensional spaces.

A graphical representation of the Support Vector Machine algorithm demonstrates how it identifies the optimal hyper plane for classification (Figure 4).

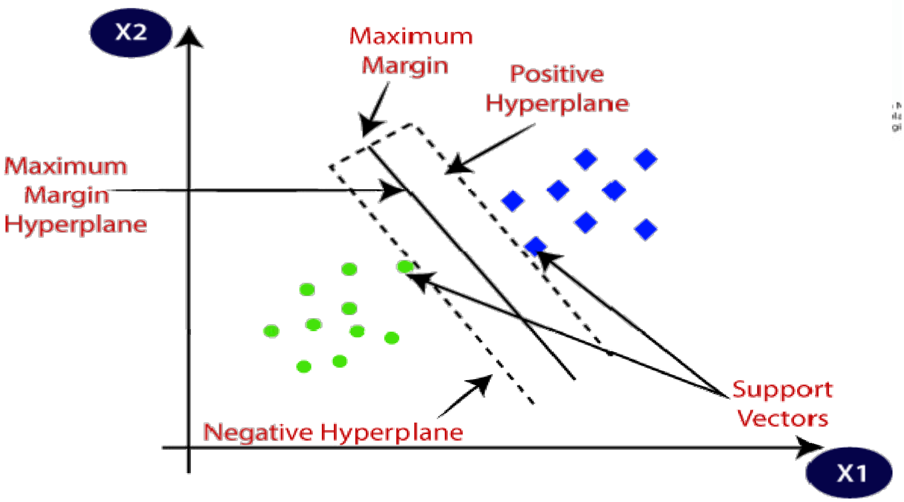


**Figure 4.** Support Vector Machine (SVM) Model

The same data set was used to train each model, and accuracy ratings were used to assess each model. The accuracy results for each model are shown below (Table 1):

**Table1.** MODELACCURACYCOMPARISON

| Model | Accuracy |
|---|---|
| K-Nearest Neighbors(KNN) | 0.98 |
| Decision Tree | 0.93 |
| Support Vector Machine(SVM) | 0.95 |

As seen, the KNN algorithm produced the highest accuracy score of 0.98, and thus it was the best model for this task. The Decision Tree and SVM models, though fairly good, did not better KNN in terms of classification accuracy. Based on this finding, the KNN model has been selected for the final career path prediction system.

## 4.4 Model Training and Testing

**Data Splitting:** To ensure class balance, the dataset was split using a stratified split in to training (80percentage) and testing (20percentage) subsets.

**Cross-Validation:** To evaluate the robustness of the model during training, a5-foldcross-validation technique was employed. Cross-validation accuracy was 83 percentage on average.
**Model Training:** To address any class imbalances, distance-based weighting was used to train the KNN classifier.

### 4.5 Prediction Workflow

- Ratings of users' proficiency are converted to numerical values.
- To determine the five closest neighbors, the system computes distances.
- The user is supplied with recommendations based on their skill sets once career opportunities have been compiled and ranked by probabilities.

## 5. Result Evaluation

A thorough confusion matrix, test set accuracy, and cross- validation accuracy were all used to assess the K-Nearest Neighbours (KNN) model's performance.

**Cross-Validation Accuracy:** $0.83 \pm 0.03$ In order to estimate the model's performance on unknown data, cross-validation was utilized to evaluate the model's stability and generalisation capabilities across several dataset subsets. Although there might be some fluctuation because of the structure of the dataset, the outcome shows reasonable generalisation.

**Test Set Accuracy:** 0.98 Witha test set accuracy of 0.98, the KNN model showed a great capacity to categorise unseen data, demonstrating its resilience and task-handling capabilities.

**Confusion Matrix:** The true class (actual labels) is represented by each row in the confusion matrix, whereas the forecast class (model predictions) is represented by each column. The number of accurate predictions for each class is shown by a high value on the diagonal, which denotes accurate classification. Misclassifications, in which the model predicts a class that differs from the actual label, are shown by off-diagonal values (Fig. 5).

**Classification Report:** The precision, recall, and F1-score for each class—essential metrics for evaluating the model's performance—are provided in the classification report (Table 2). The F1- score balances these two metrics to offer an overall effectiveness measure. Precision assesses the accuracy of positive predictions, while recall shows the model's capacity to recognize all pertinent instances. In every class, the model exhibits good recall and precision.

**Table 2.** Performance metrics by class

| Class | Precision | Recall | F1-Score | Support |
|---|---|---|---|---|
| AIML Specialist | 0.89 | 1.00 | 0.94 | 117 |
| API Specialist | 0.95 | 1.00 | 0.98 | 103 |
| Application Support Engineer | 0.93 | 0.97 | 0.95 | 115 |
| Business Analyst | 0.96 | 0.97 | 0.97 | 108 |
| Customer Service Executive | 0.97 | 0.99 | 0.98 | 99 |
| Cyber Security Specialist | 0.98 | 1.00 | 0.99 | 104 |
| Data Scientist | 0.97 | 0.99 | 0.98 | 113 |
| Database Administrator | 1.00 | 0.95 | 0.97 | 120 |
| Graphics Designer | 0.98 | 0.98 | 0.98 | 91 |
| Hardware Engineer | 1.00 | 0.98 | 0.99 | 102 |
| Helpdesk Engineer | 0.99 | 0.97 | 0.98 | 100 |
| Information Security Specialist | 1.00 | 0.94 | 0.97 | 105 |

| | | | | |
|---|---|---|---|---|
| Networking Engineer | 1.00 | 0.99 | 1.00 | 105 |
| Project Manager | 1.00 | 0.98 | 0.99 | 96 |
| Software Developer | 1.00 | 0.98 | 0.99 | 118 |
| Software Tester | 1.00 | 0.96 | 0.98 | 117 |
| Technical Writer | 1.00 | 0.96 | 0.98 | 123 |
| Accuracy | 0.98 | | | 1836 |
| Macroavg | 0.98 | 0.98 | 0.98 | 1836 |
| Weightedavg | 0.98 | 0.98 | 0.98 | 1836 |

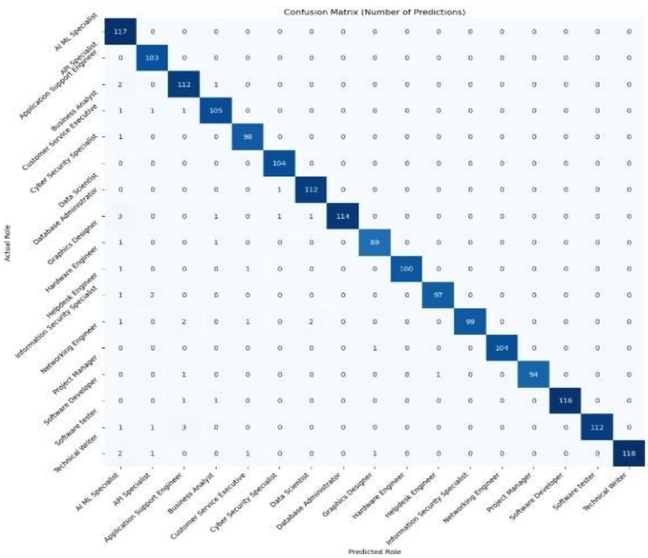


**Figure 5.** Confusion Matrix

In summary, the KNN model demonstrates high performance in classifying roles with minimal misclassification.

## 6. Future Enhancements

Real-time job market data may be added into the system in the future for upgrading and providing recommendations on what kind of career the users are fit for according to prevailing trends in the market. Further testing on personality traits and soft skills may be integrated additionally to give a better perspective on the strengths of each user. The technology may help users find the appropriate skills for the careers they desire by offering custom learning pathways and enabling connection with online courses. In addition, through tools such as an AI career advisor, multilingual support, and the ability to be accessed through mobile apps, the user experience would increase, helping them make wise decisions regarding career choices and keeping up-to- date about job openings

## 7. Conclusion

To conclude, the machine learning-based career path predictor will provide a data-driven and personalized methodology that will help people navigate the really complicated IT job landscape. Based on objective skill measurements as well as user self-assessment, the system then produces customized career guidance tailored to an individual's existing interests and abilities. The KNN algorithm ensures precise predictions, and the user-friendly interface makes the procedure easy to use for a wide range of users. As the system grows, it will continue improving its suggestions to help the users make informed decisions about their career paths. This project, by filling the gap between industry demands and education, hopes to enable users to pursue rewarding and well-suited professions in the rapidly evolving IT sector.

## References

[1] P. Roy, "Career Guidance: A Way of Life," Tathapi (UGC Care Journal), vol. 19, no. 39, pp. 22, 2020.

[2] Shreyas, R., Praveen, C., Shreyas, S., & Kiran, Y. C. (2024). A literature survey on AI-driven career path prediction. International Journal of Advanced Research in Science, Communication and Technology (IJARSCT), 4(1), 578.

[3] Ajay, D., Nilesh, S., Anirudh, D., Huzaif, S.,and P. V. Nagare, "An intelligent career guidance system using machine learning," International Research Journal of Modernization in Engineering Technology and Science, vol. 4, no. 3, pp. 101-108, 2022.

[4] Pandey, A.,& Maurya, L.S.(2021). Prediction of undergraduate students' career using various machine learning and ensemble learning algorithms. Webology,18(6), 3506.

[5] Rutina, K., & Soika, I. (2020). Career guidance for first-year students in work-based learning. Rural Environment. Education. Personality, 13, 441.

[6] S. Westman, J. Kauttonen, A. Klemetti, N. Korhonen, and others, "Artificial intelligence for career guidance – Current requirements and prospects for the future," IAFOR Journal of Education, vol. 9, no. 4, pp. 43-62, 2021.

[7] R. Kanathur, R. Akhilesh, A. S., and M. Kanathur, "Intelligent career guidance systems," International Journal of Innovative Science and Research Technology, vol. 8, no. 8, pp. 2117-2124, 2023.

[8] Sinha, A., Garima, & Singh, A. (2023, May 6). Student career prediction using algorithms of machine learning. Chandigarh University.

[9] Yadav, A. K., Dixit, A., Tripathi, A., Chowdhary, S. K., & Jangra, V. (2023). Career prediction system using ANN MLP classifier. In Proceedings of the 14th International Conference on Computing Communication and Networking Technologies (ICCCNT). IEEE.

[10]

[11] Bilon-Piórko, A.D., & Thomsen, R. (2022). A critical review of contemporary concepts of agency in the field of career guidance: Life- design counselling and a systems theory framework. British Journal of Guidance & Counselling, 50(4),533-546

[12] Ritonga, A. N., & Wangid, M. N. (2022). Career guidance services to take student career choice decision. European Journal of Education Studies, 9(6).

[13] Ajgaonkar, S., Tale, P., Joshi, Y., Jore, P., Jakate, M., Lavangare, S., & Kadam, D. (2023). EduKrishnaa: A career guidance web application based on multi-intelligence using multiclass classification algorithm. In Multi-disciplinary Trends in Artificial Intelligence (pp. 601–610). Springer.

[14] L. Jemini-Gashi, D. Be¨rxulli, J. Konjufca, and L. Cakolli, "Effectiveness of career guidance workshops on the career self-efficacy, outcome expectations, and career goals of adolescents: An intervention study," International Journal of Adolescence and Youth, vol. 28, no. 1, pp. 1- 14, 2023.

[15] T. Mathur, S. D. Pawar, M. Singh, and P. Kumar, "Intelligent career guidance system using machine learning," International Journal of Advanced Research in Computer Science and Engineering (IJARIIE), vol. 8, no. 6, pp. 647-655, 2022.

[16] Pandey, V. (2020). A review on career guidance and counselling needs for students. International Journal for Research in Applied Science and Engineering Technology, 8(5), 2500-2504.