International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 03, March 2026)

AI-CareerMatch: A Data-Driven Model for Career Path Recommendations

Anandi P. Thakare[1], Shruti S. Borkar[2], Dr. Vinit A. Sinha[3]

[1,2]*MCA II yr Sem IV,* [3]*Assistant Professor, P.G. Dept of Computer Applications, PRMITR Badnera, City – Amravati country-India*

***Abstract*— The exponential growth of emerging technologies and the continuous evolution of job roles have significantly complicated career decision-making for students and early-career professionals. Conventional career counseling systems rely heavily on manual assessments, static aptitude tests, and generalized advice, which often fail to account for individual skills, interests, and rapidly changing industry demands. As a result, many individuals experience career mismatch, low job satisfaction, and underutilization of skills.**

This research proposes AI-CareerMatch, a comprehensive data-driven career recommendation system that leverages machine learning and data analytics to provide personalized career path suggestions. The model analyzes multiple parameters including academic background, technical skills, certifications, interests, and performance indicators. By applying supervised learning algorithms, the system predicts suitable career domains aligned with current market trends.

Experimental evaluation demonstrates that AI-CareerMatch improves recommendation accuracy and relevance compared to traditional approaches. The system aims to support educational institutions, career counselors, and individuals by offering scalable, intelligent, and adaptive career guidance solutions.

***Keywords*— Career Recommendation System, Machine Learning, Artificial Intelligence, Data Analytics, Career Guidance**

I. INTRODUCTION

Career selection is one of the most significant decisions in a student's life, influencing long-term professional growth, job satisfaction, and economic stability. In the modern digital era, the career landscape has expanded beyond traditional professions to include emerging roles in artificial intelligence, data science, cybersecurity, cloud computing, and business analytics. This expansion, while beneficial, has also increased confusion among students regarding suitable career choices.

Conventional career guidance approaches such as counseling sessions and aptitude tests often lack personalization and fail to incorporate real-time industry trends. Moreover, these methods are limited in their ability to analyze large volumes of student data efficiently.

As a result, students frequently make career decisions based on peer influence, societal pressure, or incomplete information.

Artificial Intelligence (AI) and Machine Learning (ML) provide powerful mechanisms for analyzing multidimensional data and generating personalized recommendations. AI-based recommendation systems have proven effective in domains such as e-commerce, healthcare, and education. Applying similar techniques to career guidance can significantly improve decision accuracy and reduce human bias.

This research proposes AI-CareerMatch, a machine learning-based career recommendation framework that analyzes student profiles and suggests suitable career paths based on data-driven insights.

AI-CareerMatch addresses these challenges by employing a machine learning-based recommendation framework that maps individual profiles to optimal career paths. The system is designed to adapt to changing skill requirements and industry trends, ensuring relevance and accuracy over time.

II. LITERATURE REVIEW

The rapid evolution of digital technologies has significantly increased the complexity of career selection for students. Traditional career guidance systems primarily rely on manual counseling methods, aptitude tests, and rule-based decision frameworks. Although these approaches provide basic recommendations, they often fail to consider multiple student attributes such as technical skills, personal interests, and changing industry demands. Early research in automated career guidance focused on rule-based expert systems. Sharma et al. [1] developed an intelligent career counseling model that provided recommendations based on academic performance and predefined rules. While the system offered structured guidance, its static rule-based design limited adaptability in dynamic job market environments.

With the advancement of data mining and machine learning techniques, researchers began exploring predictive models for career recommendation.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 03, March 2026)

Kumar et al. [2] proposed a career prediction system using classification algorithms such as Decision Trees and Naïve Bayes. Their model demonstrated improved prediction accuracy compared to traditional rule-based systems. However, the study mainly emphasized academic data and did not sufficiently integrate student interests or skill-based attributes.

Similarly, Patel et al. [3] investigated the application of artificial intelligence in career advisory systems. Their work highlighted the potential of intelligent systems in automating career counseling processes but also identified challenges related to personalization and scalability.

Data mining techniques have also been utilized to analyze student characteristics and career patterns. Patil et al. [4] applied clustering algorithms to group students based on skill similarities and academic performance. Although clustering helped identify patterns in student data, the unsupervised nature of the approach resulted in generalized recommendations rather than precise career predictions.

Recent research emphasizes the importance of integrating skill-based attributes with supervised learning techniques. Lee et al. [5] demonstrated that combining technical skill profiles with machine learning classification models significantly enhances the accuracy of career recommendation systems.

Despite these advancements, many existing systems still lack a comprehensive framework that simultaneously considers academic performance, technical skills, certifications, and personal interests. Furthermore, limited research focuses on comparing multiple machine learning algorithms for career prediction tasks.

To address these limitations, the proposed AI-CareerMatch system introduces a data-driven framework that analyzes multidimensional student attributes and evaluates the performance of multiple supervised machine learning models, including Decision Tree, K-Nearest Neighbors, and Random Forest. This approach aims to provide more accurate and personalized career recommendations aligned with modern industry requirements.

III. Dataset Description And Feature Selection

A structured dataset of approximately 300 student profiles was created using a synthetic and sample dataset designed to simulate real-world academic and skill- based student information. The dataset includes attributes such as academic performance, technical skills, certifications, and career interests. These simulated records were used to evaluate the effectiveness of the proposed AI-CareerMatch system in predicting suitable career paths.

To ensure comprehensive career prediction, the dataset includes multiple categories of attributes. Academic attributes consist of overall grade point average (GPA), subject-wise performance, and academic specialization. Technical skill attributes represent proficiency levels in programming languages such as Java, Python, SQL, and web technologies. Certification-related attributes capture participation in industry-recognized certification programs and online courses. Additionally, interest- based attributes are incorporated to reflect individual career preferences, domain inclination, and long-term professional goals.

The target variable of the dataset corresponds to the predicted career domain, which includes categories such as Software Developer, Data Analyst, Data Scientist, Cybersecurity Analyst, Cloud Engineer, and Business Analyst. These career labels are defined based on current industry trends and job market requirements.

Feature selection is performed to identify the most influential attributes contributing to accurate career prediction. Numerical features such as GPA and skill proficiency scores are retained in continuous form, while categorical attributes, including interests and certifications, are transformed using label encoding and one-hot encoding techniques. Appropriate feature encoding ensures compatibility with machine learning algorithms and enhances model performance.

This well-structured dataset enables the AI-CareerMatch system to effectively learn meaningful patterns between student attributes and suitable career paths, thereby improving the accuracy and relevance of career recommendations.

IV. Machine Learning Algorithm Used

To predict suitable career paths, the AI-CareerMatch system employs multiple supervised machine learning algorithms. These algorithms are selected based on their effectiveness in classification tasks and their ability to handle multidimensional educational data.

A. Decision Tree Algorithm

The Decision Tree algorithm is a tree-structured classification technique that recursively splits the dataset based on feature values to generate decision rules. Each internal node represents a decision condition, while leaf nodes correspond to predicted career domains. Decision Trees are easy to interpret and provide transparent decision-making paths, which is beneficial for career guidance applications. However, they are prone to overfitting when trained on complex or noisy datasets, which may limit their generalization capability.

B. K- Nearest Neighbors (KNN) Algorithm

K-Nearest Neighbors is a distance-based learning algorithm that classifies a new instance by analyzing the labels of its nearest neighbors in the feature space. The algorithm relies on similarity measures such as Euclidean distance to determine closeness between student profiles. While KNN is simple to implement and effective for small datasets, its performance is sensitive to feature scaling and high dimensionality, which can affect prediction accuracy in large educational datasets.

C. Random Forest Algorithm

Random Forest is an ensemble learning algorithm that constructs multiple decision trees and combines their predictions to generate a final output. By aggregating results from multiple trees, Random Forest reduces variance and minimizes overfitting. This algorithm is particularly effective for career recommendation tasks due to its ability to capture complex relationships between academic performance, skills, and interests. The robustness, scalability, and high predictive accuracy of Random Forest make it well suited for intelligent career guidance systems.

V. MACHINE LEARNING FRAMEWORK FOR AI-CAREERMATCH SYSTEM

Figure 5.1 illustrates the workflow of the proposed AI-CareerMatch career recommendation system designed to provide personalized career guidance to students using machine learning techniques. The framework follows a structured pipeline beginning with raw data collection and ending with result visualization.

Initially, raw student data is collected, which includes academic performance, technical skills, certifications, interests, and career preferences. This raw data is subjected to data preprocessing, where missing values are handled, categorical attributes are encoded, and data normalization is performed to improve data quality.

After preprocessing, the system performs model selection, where appropriate learning techniques are chosen based on data characteristics. Both machine learning and deep learning approaches are considered at this stage. In the proposed implementation, traditional machine learning algorithms are primarily used for prediction, while deep learning models are explored as an extension.

The selected model undergoes hyperparameter tuning to optimize performance and avoid overfitting. The trained model is then evaluated using standard performance metrics. Finally, result visualization is performed to present career recommendations and performance analysis in an understandable format.

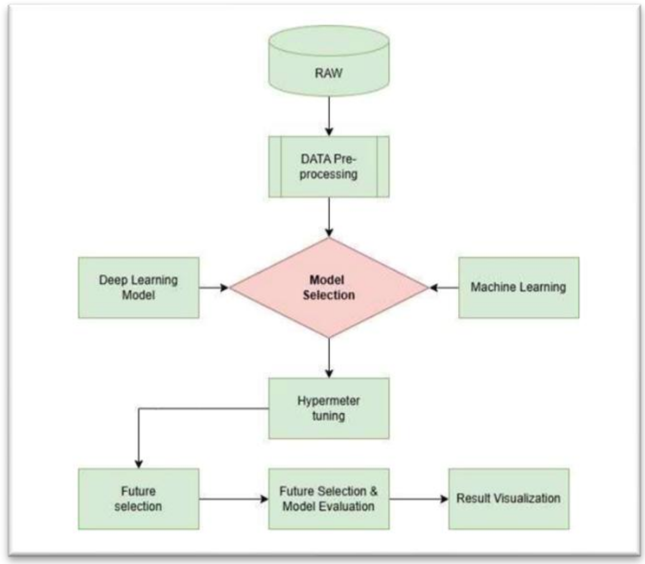

Fig.5.1. Workflow of AI-Driven Carrer Recommendation System

VI. EVALUATION METRICS

The performance of the proposed AI-CareerMatch system is evaluated using standard classification metrics, namely Accuracy, Precision, and Recall, as summarized in Table 6.1. These metrics are widely adopted in educational data mining and recommendation system research to assess prediction effectiveness and reliability.

Accuracy measures the overall correctness of the career recommendation model by calculating the proportion of correctly predicted instances among the total predictions. It is defined as:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Formula 6.1. Accuracy

Precision evaluates the relevance of the recommended career paths by determining the proportion of correctly predicted positive outcomes among all positive predictions:

$$\text{Precision} = \frac{TP}{TP + FP}$$

Formula 6.2 Precision

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 03, March 2026)

Recall represents the model's ability to correctly identify suitable career options for students by measuring the proportion of actual positive cases that are correctly predicted:

$$\text{Recall} = \frac{TP}{TP + FN}$$

Formula 6.3. Recall

These evaluation metrics are particularly suitable for career recommendation systems, as they ensure not only overall prediction accuracy but also the relevance and completeness of the recommendations. High precision minimizes the likelihood of suggesting unsuitable career paths, while high recall ensures that potential career opportunities are not overlooked.

Table 6.1 :
Description Of Evaluation Parameters

Parameter	Description
Accuracy	Measures the overall correctness of the model
Precision	Indicates the relevance of recommended career paths
Recall	Represents the ability to identify suitable career options

III. Result And Discussion

The proposed system is evaluated using a labeled dataset consisting of student profiles and corresponding career outcomes. Three supervised machine learning algorithms—Decision Tree, K-Nearest Neighbors (KNN), and Random Forest—are used for performance comparison. The experimental results are summarized in Table 7.1.

The results indicate that the Random Forest algorithm achieves the highest accuracy of 93%, along with superior precision and recall values. This improved performance is attributed to its ensemble learning approach, which effectively handles complex feature interactions and reduces overfitting. Decision Tree and KNN also show reasonable performance but are comparatively less accurate.

The experimental analysis confirms that ensemble-based models are more suitable for personalized career recommendation tasks.

Table 7.1:
Evaluation Of Proposed Techniques

Algorithm	Accuracy	Precision	Recall
Decision Tree	88%	86%	87%
KNN	85%	84%	83%
Random Forest	93%	92%	91%

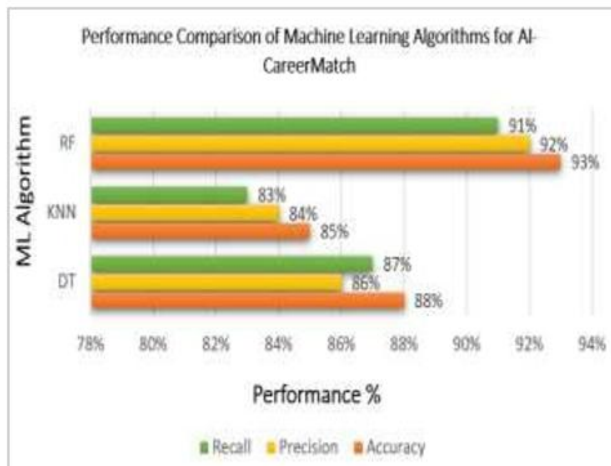

Fig.7.2. Performance comparison of Decision Tree, KNN, and Random Forest algorithms using Accuracy, Precision, and Recall metrics

Fig.7.2. In the proposed AI-CareerMatch system, the experimental results demonstrate the comparative performance of the machine learning algorithms employed in the AI-CareerMatch system using three standard evaluation metrics: accuracy, precision, and recall. The experimental results demonstrate that the Random Forest algorithm consistently outperforms the Decision Tree and K-Nearest Neighbors (KNN) models across all performance parameters. Random Forest achieves the highest accuracy of 93%, along with superior precision (92%) and recall (91%), indicating its strong capability to generate reliable and relevant career recommendations.

This enhanced performance can be attributed to the ensemble learning nature of Random Forest, which combines multiple decision trees to effectively capture complex relationships among student attributes while minimizing overfitting.

The Decision Tree model exhibits moderate performance with an accuracy of 88%, precision of 86%, and recall of 87%. While it provides interpretable decision rules, its single-model structure limits its generalization capability when handling diverse and multidimensional student profiles. In contrast, the KNN algorithm records comparatively lower performance, with accuracy, precision, and recall values of 85%, 84%, and 83% respectively. This reduced performance may be due to its sensitivity to feature scaling and dependence on distance measures, which can affect prediction reliability in high- dimensional datasets.

Overall, the results confirm that ensemble-based learning approaches are more effective for personalized career recommendation systems. The observed performance trends validate the effectiveness of the proposed AI-CareerMatch framework in aligning student profiles with suitable career paths using data- driven insights. These findings highlight the potential of machine learning techniques, particularly Random Forest, in enhancing the accuracy and relevance of intelligent career guidance systems.

Random Forest achieved the highest accuracy due to its ensemble learning approach and ability to handle complex feature interactions.

IV. CONCLUSION

This research successfully presented AI-CareerMatch, a machine learning-based career recommendation system that addresses the limitations of traditional career guidance methods by offering personalized and data- driven insights. By analyzing academic performance, technical skills, and individual interests, the system demonstrates the practical applicability of artificial intelligence in supporting informed career decision- making.

Experimental results confirm that the Random Forest algorithm outperforms Decision Tree and KNN models in terms of accuracy, precision, and recall, highlighting the effectiveness of ensemble learning techniques for career prediction tasks.

Overall, the proposed framework provides a scalable, reliable, and adaptable solution that can assist students, educational institutions, and career counselors in aligning individual capabilities with evolving industry requirements.

REFERENCES

[1] Müller, M. (2024). Advancements in software engineering through artificial intelligence. International Journal of Scientific Research & Engineering Trends, **10**(4). https://ijsret.com/wp-content/uploads/2024/07/IJSRET_V10_issue4_320.pdf

[2] Fiamma, P. (2018). International Journal of Scientific Research in Science, Engineering and Technology. International Journal of Scientific Research in Science Engineering and Technology. https://doi.org/10.32628/ijsrset

[3] Understanding machine learning's impact on software development through AI-driven code generation. (2025). IJSRD – International Journal for Scientific Research & Development, **13**(2), 214–215.

[4] Gokhe, R. (2025). The evolution of AI-powered code generation tools for web developers. IJNRD – International Journal of Novel Research and Development, **10**(11), b261–b263.

[5] Thirumaleswarlu, T., Tejasri, K., Mahender, G., & Thirumalesh, K. (2025). A comprehensive analysis of AI code generation techniques. International Journal of Progressive Research in Engineering Management and Science (IJPREMS), **5**(4), 936–940. https://www.ijprems.com

[6] Transformer-based code generation: Automating software development with AI. (2025). IARJSET – International Advanced Research Journal in Science, Engineering and Technology, **12**(3). https://doi.org/10.17148/iarjset.2025.12334

[7] Singh, M. (2024). The impact of artificial intelligence on software development. International Journal of Computing and Artificial Intelligence, **5**(2), 192–198. https://doi.org/10.33545/27076571.2024.v5.i2c.113

[8] Krishna, N. K., Thakur, N. D., & Meka, N. H. S. (2024). Enhancing software engineering practices with generative AI: A framework for automated code synthesis and refactoring. World Journal of Advanced Engineering Technology and Sciences, **13**(1), 672–681. https://doi.org/10.30574/wjaets.2024.13.1.0463

[9] Sarma, W., Tiwari, S., & Dey, S. (2024). Revolutionizing software engineering with generative AI and large language models: Strategies for innovation and efficiency. International Research Journal of Engineering and Technology (IRJET), **11**(12), 540. https://www.irjet.net

[10] Chafik, N., & Benchekroun, A. (2020). Integrating artificial intelligence in software engineering: Enhancements and challenges in the development lifecycle. IRE Journals, **3**(12), 253–254.

[11] Durrani, U. K., Akpinar, M., Bektas, H., & Saleh, M. (2025). Impact of artificial intelligence on software engineering phases and activities (2013–2024): A quantitative analysis using zero-truncated Poisson model. IEEE Access, **13**, 95535–95547. https://doi.org/10.1109/access.2025.3574462