

CareerPath AI: A Hybrid Machine Learning And Large Language Model Framework For Multi-level Career Recommendation With Real-Time Job Market Integration

Mrs. S.T. Ramya¹, T.Sri Vardhini², B.Sneha Reddy³, M.Naga Sai⁴

¹Assistant Professor, Department of Information Technology, Matrusri Engineering College.

^{2,3,4}Students of Department of Information Technology, Matrusri Engineering College.

vardhinhithuppada@gmail.com,snehareddybolla@gmail.com

Abstract—Career selection is a critical decision that significantly influences a student’s academic trajectory and long-term professional success. However, students at the 10th, 12th, and graduate levels often experience uncertainty due to limited awareness, evolving job market demands, and insufficient personalized guidance. Existing career recommendation systems largely focus on specific domains and rely on static datasets, limiting adaptability and holistic assessment.

This paper proposes a Hybrid AI-Based Career Recommendation System that integrates Machine Learning models (Decision Tree and XGBoost) with Large Language Models (LLMs) to deliver explainable, adaptive, and personalized career guidance across Arts, Commerce, Science, Humanities, and vocational streams. The system first classifies users based on educational level and evaluates academic performance, interests, aptitude, personality traits, and psychometric responses. Structured prediction models generate career probabilities, while LLMs provide contextual explanations, learning pathways, and actionable recommendations.

To ensure industry relevance, real-time job market trends and skill-demand data are incorporated through external APIs. A skill-gap analysis module identifies competency gaps and recommends targeted courses and certifications. A feedback-driven ranking mechanism further refines predictions and assigns a probability-of-success score to each recommendation. Experimental evaluation demonstrates improved predictive accuracy and personalization compared to conventional career guidance approaches. The proposed framework bridges the gap between education and employability through a scalable, data-driven, and future-ready decision support system.

Key Words—Career Recommendation System, Career Path Prediction, Hybrid Machine Learning, Decision Tree, XGBoost, Large Language Models (LLMs), Psychometric Assessment, Skill Gap Analysis, Educational Data Mining, Real-Time Labor Market Intelligence.

I. INTRODUCTION

Career selection is one of the most critical decisions in a student’s life, significantly influencing academic progression, professional development, and long-term personal satisfaction. Students at key educational milestones—such as the 10th grade, 12th grade, and graduation level—often face substantial

uncertainty when choosing a career path. This confusion arises due to limited exposure to emerging opportunities, inadequate personalized guidance, peer and parental influence, and the rapidly evolving nature of global job markets. As industries transform under technological advancements, new roles continuously emerge while traditional roles evolve, making informed career planning increasingly complex.

Conventional career counseling approaches primarily rely on manual consultations, aptitude assessments, and standardized questionnaires. Although these methods provide foundational support, they are often generalized, time-consuming, and unable to deliver highly personalized recommendations. Moreover, many existing digital career recommendation systems focus predominantly on graduate-level job matching or STEM-oriented domains, offering limited consideration for students from Arts, Commerce, Humanities, and vocational backgrounds. Additionally, integration of real-time labor market intelligence and adaptive feedback mechanisms remains insufficient in most current solutions.

Recent advancements in Artificial Intelligence (AI) and Machine Learning (ML) have enabled the development of intelligent recommendation systems capable of analyzing large volumes of structured and unstructured data. Machine learning algorithms can model complex relationships between academic performance, skills, interests, and aptitude indicators to predict suitable career paths. Furthermore, Large Language Models (LLMs) enhance system capabilities by generating contextual explanations, personalized career insights, and structured learning roadmaps, thereby improving interpretability and user engagement.

To address the limitations of existing systems, this study proposes a multi-level AI-based career recommendation framework that integrates Decision Tree and XGBoost algorithms with LLM-driven personalization and real-time job market analytics. The system first classifies users based on educational level and evaluates academic, psychometric, and skill-based parameters to generate ranked career recommen-

dations. In addition, a skill-gap analysis module identifies missing competencies and suggests targeted learning resources aligned with current industry demands. By combining predictive modeling, contextual AI guidance, and live labor market insights, the proposed framework aims to bridge the gap between education and employability through an adaptive, scalable, and data-driven decision support system.

II. RELATED WORK

Early career recommendation systems were primarily based on rule-based expert systems and psychometric assessments, where predefined rules mapped user responses to specific career options. While these systems provided structured guidance, they lacked adaptability and personalization. With the advancement of Educational Data Mining (EDM) and Machine Learning (ML), researchers began applying classification algorithms such as Naïve Bayes, k-Nearest Neighbors (k-NN), Support Vector Machines (SVM), and Decision Trees to predict suitable career domains based on academic performance and aptitude scores. Decision Trees gained popularity due to their interpretability, while ensemble techniques like Random Forest and XGBoost improved prediction accuracy through advanced feature modeling and regularization. However, most existing ML-based systems rely on static datasets, focus primarily on STEM-oriented careers, and provide limited support for Arts, Commerce, and vocational streams.

Recent advancements in Artificial Intelligence have introduced Natural Language Processing (NLP) and Large Language Models (LLMs) to enhance personalization and user interaction in recommendation systems. LLMs enable contextual explanations, conversational guidance, and adaptive learning pathway suggestions. Additionally, some studies have incorporated real-time labor market analytics to align recommendations with industry skill demands. Despite these improvements, existing solutions often lack multi-level educational classification, integrated skill-gap analysis, and feedback-driven ranking mechanisms. To address these gaps, the proposed framework combines Decision Tree and XGBoost models with LLM-based personalization and real-time job market intelligence, delivering ranked, explainable, and adaptive career recommendations across diverse academic streams.

III. LITERATURE REVIEW

Recent research has demonstrated significant progress in the application of Artificial Intelligence (AI) and Machine Learning (ML) for career guidance systems. An AI-Powered Career Guidance System (IEEE, 2025) utilized the Big Five personality traits combined with ensemble classifiers such as Random Forest to recommend academic majors. The study employed data balancing techniques including SMOTE to enhance predictive performance and integrated a chatbot interface for improved interaction. Although the system achieved high classification accuracy in personality-driven major selection, its scope was confined to higher education students and lacked integration of real-time labor market intelligence and

multi-level educational classification.

Similarly, the AI-Driven Career Guidance for Graduates (IJSRED, 2025) focused on graduate-level career prediction by combining academic performance metrics with personality profiling. The framework improved personalization through supervised learning algorithms but remained limited to a single academic stage. The system did not incorporate advanced ensemble boosting methods or dynamic job trend data, restricting adaptability to evolving industry requirements. In comparison, the AI-Based Career Guidance System (IJCRT, 2025) explored skill-based career mapping using classification algorithms and NLP techniques. While it introduced feedback mechanisms for recommendation refinement, it relied primarily on static datasets and did not implement probabilistic career success estimation or structured educational-level segmentation.

Research has also expanded into employment-oriented systems. The Personalized Job Search with AI (IJERT, 2025) integrated Natural Language Processing (NLP), Large Language Models (LLMs), and external job APIs to provide real-time job matching based on resume analysis. Similarly, the Smart Resume Matcher (IJIRT, 2025) applied Named Entity Recognition (NER) and web scraping to extract skills and recommend job roles and upskilling courses. Although these systems effectively demonstrated real-time labor market integration and automated skill-gap identification, they primarily focused on employment matching rather than holistic academic-to-career mapping. Moreover, they lacked multi-stage educational classification, structured ensemble prediction frameworks, and comprehensive psychometric evaluation.

From the reviewed literature, it is evident that existing systems tend to address isolated components of career guidance—such as personality-based major selection, graduate-level prediction, or job matching—without offering an integrated, multi-level framework. The combined integration of structured ensemble machine learning models (Decision Tree and XGBoost), LLM-based contextual personalization, real-time labor market analytics, skill-gap assessment, and adaptive feedback learning across diverse academic streams remains underexplored. The proposed system aims to address these limitations by delivering a unified, scalable, and data-driven career recommendation framework spanning 10th, 12th, and graduate levels.

IV. SYSTEM ARCHITECTURE

The overall architecture of the proposed AI-Based Career Recommendation System is illustrated in Fig. 1. The system follows a layered and modular design integrating user interaction, backend processing, machine learning models, and real-time job market intelligence.

The proposed AI-Based Career Recommendation System follows a modular and layered architecture designed to ensure scalability, flexibility, and efficient integration of multiple intelligent components. The system is structured into five primary layers: User Interaction Layer, Application Processing Layer, Prediction and Intelligence Layer, External Integration Layer, and Data Storage Layer. Each layer performs a spe-

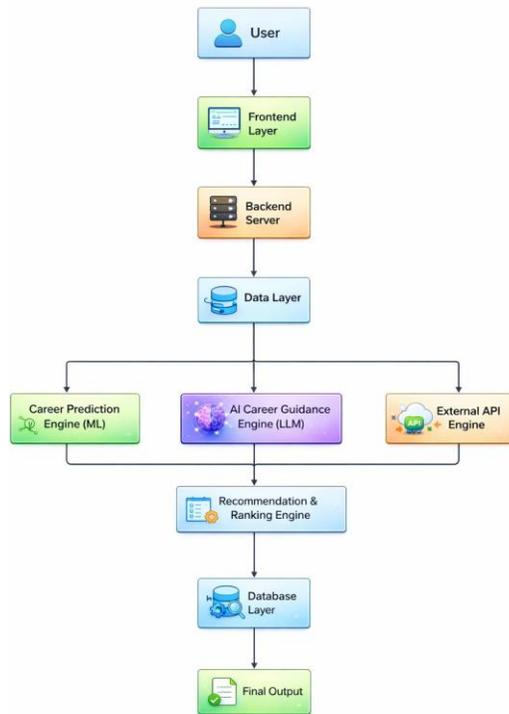


Fig. 1. Architecture of the Proposed AI-Based Career Recommendation System
 specific function while maintaining seamless communication with other components.

The User Interaction Layer serves as the entry point of the system, where users provide their educational level, academic performance details, interests, skills, aptitude responses, and psychometric test inputs through a structured interface. This layer ensures user-friendly interaction and secure data collection. The collected information is forwarded to the backend processing unit for further analysis.

The Application Processing Layer handles data validation, preprocessing, and feature engineering. In this stage, raw user inputs are cleaned, encoded, normalized, and transformed into structured feature vectors suitable for machine learning models. Educational-level classification is performed to segment users into 10th grade, 12th grade, or graduate categories, ensuring that recommendations are contextually relevant to the user's academic stage.

The Prediction and Intelligence Layer forms the core of the architecture. This layer consists of two primary components: the Machine Learning Prediction Engine and the LLM-Based Guidance Engine. The Machine Learning engine implements Decision Tree and XGBoost classifiers to generate probability-based career predictions across multiple domains. The outputs of these models are combined using a weighted hybrid scoring mechanism to improve accuracy and reliability. The LLM component processes prediction results and user context to generate personalized explanations, structured learning roadmaps, and alternative career suggestions. This integration

enhances interpretability and user engagement.

The External Integration Layer connects the system to real-time job market data sources through APIs. This layer retrieves information such as trending job roles, skill demand frequencies, salary insights, and emerging career paths. The retrieved data is utilized in the skill-gap analysis module, where user competencies are compared against industry requirements to identify missing skills and recommend targeted learning resources.

The Data Storage Layer consists of a centralized database that stores user profiles, psychometric responses, prediction outputs, skill-gap results, market trend data, and user feedback. The feedback mechanism allows continuous system improvement by refining model weights and updating ranking parameters based on user interaction outcomes. This adaptive loop ensures that the system evolves with changing industry conditions and user preferences.

The final output module integrates results from all layers and presents users with a ranked list of career recommendations, probability of success indicators, identified skill gaps, and personalized learning pathways. The modular design of the architecture ensures extensibility, allowing future enhancements such as reinforcement learning, multilingual support, and integration with academic institutions or recruitment platforms.

Overall, the system architecture enables efficient coordination between structured machine learning models, contextual AI reasoning, real-time labor market intelligence, and adaptive feedback mechanisms, resulting in a comprehensive and intelligent career recommendation framework.

The proposed AI-Based Career Recommendation System follows a structured, hybrid, and modular methodology designed to provide accurate, interpretable, and adaptive career guidance across multiple educational levels. The framework integrates supervised machine learning models, Large Language Models (LLMs), real-time labor market intelligence, skill-gap analytics, and a feedback-driven refinement mechanism to ensure a comprehensive and scalable solution.

V. METHODOLOGY

The methodology begins with user segmentation based on educational level, categorizing individuals into 10th grade, 12th grade, or graduate-level groups. This classification is essential because career opportunities, eligibility criteria, and academic pathways vary significantly across these stages. By narrowing the recommendation space according to educational level, the system improves contextual relevance and predictive precision. For example, students in the 10th grade may receive



Fig. 2. Methodology Framework of the Proposed System stream selection guidance (Arts, Commerce, Science, Vocational), while graduates may receive domain-specific career role recommendations.

After classification, the system performs structured data acquisition through detailed questionnaires, psychometric assessments, and skill self-evaluation modules. The collected data includes academic records, subject strengths, interest preferences, aptitude indicators, personality traits, technical and non-technical skills, and extracurricular inclinations. Psychometric assessments help measure behavioral tendencies and cognitive attributes, while interest mapping captures domain alignment. The system converts this raw information into structured features through feature engineering processes. Academic scores are transformed into subject proficiency indices, interests are mapped into domain-based vectors, and skill ratings are standardized into measurable competency levels. This transformation ensures that diverse forms of user input can be computationally processed effectively.

To enhance model robustness, data preprocessing techniques are applied. Missing values are addressed through appropriate imputation methods to prevent information loss. Categorical data, such as stream preference or personality categories, are encoded into numerical representations. Continuous variables are normalized to maintain uniform scale and prevent bias toward features with larger magnitudes. Additionally, outlier detection techniques are implemented to remove anomalous entries that could negatively impact predictive performance. Where necessary, class imbalance handling techniques are applied to ensure that underrepresented career domains receive adequate representation during model training.

For career prediction, the system employs two supervised learning algorithms: Decision Tree and Extreme Gradient Boosting (XGBoost). The Decision Tree model is selected for its interpretability and structured hierarchical decision-making capability. It learns feature-based splits that classify users into career domains based on patterns observed in training data. Although highly interpretable,

Decision Trees may be prone to overfitting in complex datasets. To address this limitation, XGBoost is incorporated as a complementary model. XGBoost uses gradient boosting techniques to sequentially minimize prediction errors and enhance generalization. It is particularly effective in capturing nonlinear relationships and complex feature interactions. Both models independently generate probability-based predictions across predefined career categories.

To improve reliability and stability, a hybrid ensemble strategy is implemented. The outputs of both models are combined through a weighted aggregation mechanism that balances interpretability and predictive strength. The weighting strategy is determined through validation experiments to ensure optimal performance. This hybrid approach reduces model bias, enhances generalization capability, and increases overall confidence in career predictions.

Beyond statistical prediction, the system integrates Large Language Models to enhance contextual understanding and personalization. The LLM receives structured model outputs along with user-specific contextual information and generates human-readable explanations of recommended career paths. It also produces alternative suggestions, detailed career descriptions, and structured learning roadmaps outlining necessary qualifications, certifications, and skill development pathways. This component ensures that the system does not merely classify users but provides actionable and interpretable guidance, thereby improving user trust and engagement.

The methodology further incorporates a skill-gap analysis module that compares user competencies with industry-required skills obtained from real-time job market APIs. External data sources provide insights into trending roles, required qualifications, emerging skill demands, and labor market shifts. By aligning user profiles with industry benchmarks, the system identifies missing competencies and recommends targeted courses, certifications, and training programs. This mechanism bridges the gap between academic preparation and employability by aligning recommendations with current and evolving market requirements.

A ranking mechanism is then applied to prioritize recommended career options. The ranking process considers multiple factors, including model prediction confidence, market demand intensity, user interest alignment, and skill compatibility. This multi-factor evaluation produces a ranked list of career recommendations accompanied by a probability-of-success indicator. The probability score reflects how well the user's academic background, interests, and skills align with the selected career path.

Finally, the system incorporates an adaptive feedback mech-

anism to enable continuous improvement. After receiving recommendations, users can provide feedback regarding relevance, clarity, and usefulness. This feedback is stored in a centralized database and analyzed to refine model weights, adjust ranking parameters, and enhance future predictions. Over time, this adaptive learning loop improves system accuracy, responsiveness, and personalization capability.

Through the integration of structured machine learning models, contextual AI reasoning, dynamic labor market intelligence, skill-gap analytics, and feedback-driven optimization, the proposed methodology establishes a comprehensive, scalable, and future-ready career recommendation framework capable of supporting informed decision-making across diverse academic streams.

VI. RESULTS

The proposed AI-Based Career Recommendation System was evaluated using structured student profile data comprising academic performance, psychometric indicators, interests, and skill attributes. The dataset was divided into training and testing subsets to ensure reliable performance assessment. The Decision Tree classifier provided interpretable rule-based predictions with satisfactory accuracy, while the XGBoost model demonstrated superior predictive performance due to its boosting mechanism and enhanced generalization capability. The hybrid ensemble approach, combining both models, produced more stable and accurate career predictions compared to individual classifiers.

The integration of Large Language Models (LLMs) significantly enhanced qualitative performance by generating contextual explanations, structured guidance, and personalized career insights. Real-time job market integration ensured that recommendations were aligned with current industry demand trends. Additionally, the skill-gap analysis module effectively identified missing competencies by comparing user profiles with industry-required skills, enabling actionable upskilling recommendations.

The final system output is presented in the form of a structured Career Recommendation Report, which includes ranked career suggestions, probability of success indicators, identified skill gaps, and a personalized learning roadmap.

VII. DISCUSSION

The results indicate that the proposed hybrid framework improves predictive accuracy and recommendation reliability compared to single-model approaches. XGBoost demonstrated superior performance due to its boosting mechanism and ability to model complex feature interactions, while the Decision Tree provided better interpretability. The hybrid weighted model effectively balanced accuracy and explainability, leading to more stable and confident career predictions.

The integration of Large Language Models enhanced personalization by generating contextual explanations and structured learning roadmaps, addressing limitations of purely statistical models. Real-time job market integration further strengthened industry alignment, and the skill-gap analysis module provided actionable upskilling guidance. However, system performance depends on dataset quality and external API reliability. Future improvements may include larger datasets, reinforcement learning from user feedback, and advanced explainable AI techniques to enhance transparency and adaptability.

VIII. CONCLUSION

This study proposed a hybrid AI-Based Career Recommendation System to assist students at the 10th, 12th, and graduate levels in making informed career decisions. The framework integrates Decision Tree and XGBoost models with Large Language Model-based personalization and real-time job market analytics to generate accurate and ranked career recommendations.

The hybrid approach improved prediction reliability, while the LLM component enhanced interpretability through contextual explanations and learning roadmaps. The integration of skill-gap analysis and real-time industry trends ensured practical and market-aligned guidance. Overall, the system demonstrates the effectiveness of combining machine learning, contextual AI, and dynamic labor market intelligence to deliver a scalable and future-ready career recommendation solution.

IX. FUTURE SCOPE

Future developments may include multilingual support and mobile application deployment to improve accessibility and user reach. Incorporating advanced labor market forecasting techniques could further enhance the system's ability to anticipate emerging career trends. Integration with academic institutions and recruitment platforms would enable seamless academic-to-employment mapping, while collaboration with educational and industry partners could facilitate real-world validation and large-scale adoption of the framework.



Fig. 3. Career Recommendation Report Generated by the System

X. REFERENCES

- [1] A.I. N. El-Khalili, K. S. Dardese, N. T. Alotaibi, Y. Mubaydeen and I. Tahboub, "AI powered Career Guidance System," 2025 1st International Conference on Computational Intelligence Approaches and Applications (ICCIAA), Amman, Jordan, 2025, pp. 1-6, doi: 10.1109/IC-CIAA65327.2025.11013279.
- [2] B.Kharade, R. M., Desai, S. (2025). AI-driven career guidance for graduates. International Journal of Scientific Research and Engineering Development (IJSRED), 8(4), 1258–1265.
- [3] C.Lokam, G. V., Patil, P. S., Kukade, A. S., Dhule, A. N., Kothiwale, R. A. (2025). AI-based career guidance system. International Journal of Creative Research Thoughts (JCRT), 13(5), 843–847.
- [4] D.Babu, R. K., Sreerala, T., Kumar, R., Sumith, S. (2025). Personalized job search with AI: A recommendation system integrating real-time data and skill-based matching. International Journal of Engineering Research Technology (IJERT), 14(4), [page numbers].
- [5] E.Chetan, K. R. (2025). Smart resume matcher: AI-powered job and skill recommendation system. International Journal of Innovative Research in Technology (IJIRT), 11(9), 597–605.