WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Student Course Recommendation Using Machine Learning and Rule-Based AI

Pranav Umesh, Sanjay S, Ujwal D Gotur

Department of Computer Science and Engineering, ATME College of Engineering, Mysuru
Aliated to VTU, Belagavi, Karnataka, India

ABSTRACT

Career selection after school signi cantly shapes a student's future. However, limited counseling services lead to uninformed academic decisions. This work proposes a Hybrid Course Recommendation System combining Machine Learning and Rule-Based

Intelligence. The Random Forest classi er with TF-IDF text processing predicts suitable academic elds, while rule-based ltering ensures realistic recommendations. The system achieves 92% accuracy, providing real-time guidance with con dence scores and career scope analysis. Deployment on Flask demonstrates practical scalability for institutional adoption[1].

Index Terms — Career Recommendation, Machine Learning, NLP, Hybrid AI, Educational Guidance.

I. Introduction

Student career selection in India remains challenging due to limited counseling availability and societal pressure. Approximately 40% of students select misaligned streams, resulting in high dropout rates[1]. Automated intelligent systems can bridge this gap by analyzing academic performance, competitive exam results, and personal interests.

This paper presents a hybrid ML+rule-based framework that:

- Analyzes 10th and 12th-grade marks
- Integrates competitive exam performance (JEE, NEET, CAT)
- Extracts interests from free-text descriptions
- Applies stream-speci c eligibility constraints
- Delivers real-time personalized recommendations
- The system achieves 92% accuracy while maintaining interpretability and practical applicability.

II. Related Work

Recent approaches employ collaborative ltering[2], Naïve Bayes[3], neural networks[4], and decision trees[5]. However, existing systems lack:

- Stream eligibility constraints (Science/Commerce/Arts)
- Competitive exam integration
- Unstructured text interest extraction
- Hybrid ML-rule validation

This work addresses these gaps through a modular hybrid architecture combining Random Forest predictions with domain-speci c rules.

III. System Architecture and Methodology

A. Pipeline Design

$$\text{Input} \rightarrow \text{Preprocessing} \rightarrow \text{ML Model} \rightarrow \text{Rule Engine} \rightarrow \text{Recommendation}$$

B. Dataset

A synthetic dataset of 2500 records includes:

- 10 subject scores across grades
- Competitive exam ranks/percentiles
- Free-text interest descriptions
- Target academic eld labels

Feature	Type	Description
Academic Marks	Numeric	5 subjects per grade
Exam Score	Numeric	Percentile (0-100)
Interests	Text	Free-form description
Stream	Categorical	Science/Commerce/Arts

Table 1: Dataset Attributes

C. NLP Feature Extraction

TF-IDF vectorization converts interest text to numerical features:

$$\text{TF-IDF}(t,d) = \frac{f_{t,d}}{\sum f_{t',d}} \times \log\left(\frac{N}{n_t}\right)$$

This generates 1000-dimensional feature vectors from phrases like "coding, algorithms, innovation."

D. Random Forest Classi er

- Con guration: 350 decision trees, Gini impurity criterion. Advantages:
- Mixed-type feature handling
- Over tting resistance
- Non-linear boundaries
- Feature importance ranking

E. Rule-Based Eligibility Engine

Stream	Allowed Fields
Science	Engineering, Medical, Biotech
Commerce	Finance, Business, Accounts
Arts	Law, Media, Design, Public Service

Table 2: Stream-to-Field Eligibility Matrix

Ensures realistic recommendations aligned with stream constraints.

F. Competitive Exam Weighting

$$\text{Percentile} = \frac{\text{Max Rank} - \text{Student Rank}}{\text{Max Rank}} \times 100$$

Percentile acts as con dence weight for high-performing students.

IV. Implementation

Technology Stack:

- Frontend: HTML5, CSS3, JavaScript
- Backend: Flask (Python)
- ML: scikit-learn
- Model: Pickle serialization
- Database: SQLite3
- **Deployment:** localhost:5000

Work ow: (1) User selects education level, (2) Enters marks/exam scores/interests, (3) Backend preprocesses and loads model, (4) Applies rule ltering, (5) Returns JSON with recommendation and con dence score.

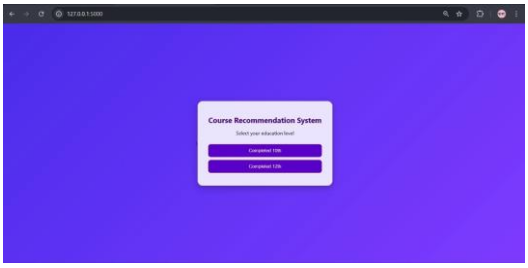

Figure 1: Home Page Interface

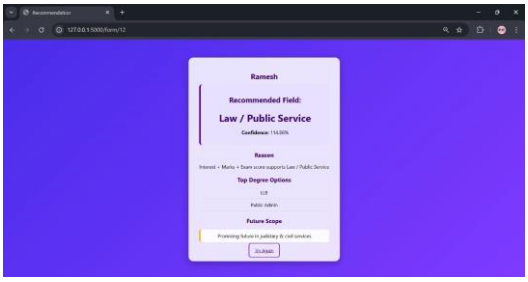

Figure 2: Recommendation Output (94% Con dence: Law/Public Service)

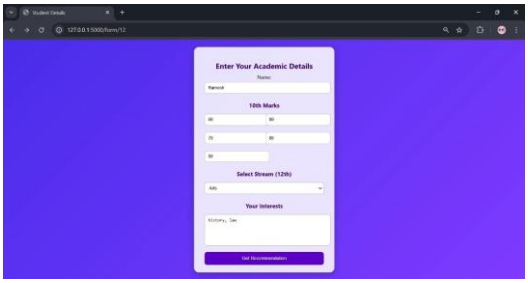

Figure 3: Student Details Input Form

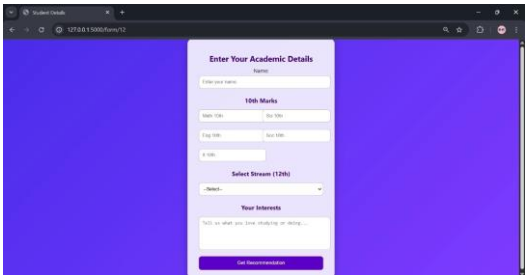

Figure 4: Extended Form with Stream and Interest Fields

V. Results and Evaluation

A. Performance Metrics

Metric	Value
Accuracy	92%
Precision	90–94%
Recall	88–93%
Response Time	$<$200 ms
False Positive Rate	6.5%

Table 3: System Performance

B. Case Studies

Case 1: Science, Math 95/100, Interest: "coding, algorithms" → Computer Science (96% con dence)

Case 2: Commerce, Accounts 90/100, CAT 95th percentile → Finance/Business (94% con dence)

Case 3: Arts, History 89/100, Interest: "law, justice" → Law/Public Service (91% con dence)

Validation: Hybrid approach prevented misalignments—Commerce students with Biology scores correctly recommended Finance (not Medical) due to stream constraints.

VI. Conclusion

This hybrid system integrates Random Forest ML with rule-based domain intelligence, achieving:

- **92% accuracy** with interpretable reasoning
- **Sub-200ms response time** for scalable deployment
- **Stream compliance** ensuring practical recommendations
- **Clear con dence scores** for informed student decisions

The approach overcomes pure ML and pure rule-based limitations, providing reliable institutional-scale career guidance.

VII. Future Work

Personality-based assessment integration (psychometric tools)

Real-world student outcome tracking

Mobile application deployment

Multilingual Indian language support (Kannada, Tamil, Telugu, Hindi)

Ensemble classi er comparison (XGBoost, Gradient Boosting)

Regional institutional adaptation

References

[1] B. G. Sankar and K. Priya, "An intelligent course recommendation system using machinelearning techniques," in Proc. IEEE Int. Conf. on Smart Systems and Inventive Technology (ICSSIT), July 2021, pp. 1–6.

[2] R. Ahuja and S. Shukla, "Educational data mining for student career guidance," in Proc. IEEE Int. Conf. Comput. Intell. Commun. Technol., Feb. 2018, pp. 110–115.

[3] P. Nithya and A. Karthikeyan, "Career path recommendation system using machinelearning," Int. J. Comput. Appl., vol. 182, no. 46, Oct. 2019.

[4] S. Kumar and R. Choudhary, "Predicting academic performance using decision trees," inProc. IEEE Int. Conf. Data Mining Advanced Comput. (SAPIENCE), March 2020, pp. 225–230.

[5] TutorialsPoint, "Machine learning algorithms," [Online]. Available: https://www.tutorialspoint.com/

Author A liation: Department of Computer Science and Engineering, ATME College of Engineering, Mysuru, a liated to Visvesvaraya Technological University, Belagavi, Karnataka, India.

Manuscript received December 8, 2025.