iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: V Month of publication: May 2026

DOI: https://doi.org/10.22214/ijraset.2026.82179

AI-Driven Career Intelligence and Job Fraud Detection Platform Using Machine Learning and Natural Language Processing

Nathishkumar S[1], Sanjai G[2], Nithivananthan M[3], Pasupathi I[4], Ms. Rakshanadevi S[5]
[1, 2, 3, 4]*Computer Science and Engineering, Dhirajlal Gandhi College of Technology, Salem*
[5]*AP/CSE, Department of Computer Science and Engineering, Dhirajlal Gandhi College of Technology, Salem*

Abstract: The rapid proliferation of online recruitment platforms has significantly increased career exploration opportunities but has simultaneously exposed users to sophisticated employment scams. Traditional career guidance systems often provide static recommendations that fail to account for individual personality nuances or the security vulnerabilities present in the modern job market. This paper presents AegisPath, an AI-driven platform that integrates vocational synthesis with employment threat intelligence under a unified framework. The system utilizes an MBTI-based psychometric assessment engine to computationally derive personality types and generate personalized 6-month career roadmaps enriched with real-time India 2025 salary benchmarks. Simultaneously, it deploys an 8-signal Natural Language Processing (NLP) fraud detection engine that evaluates job postings for linguistic manipulation indicators, fee-extraction patterns, and URL legitimacy signals. The Career Intelligence Engine leverages a Large Language Model (LLM) backend with a Naive Bayes-grounded classification layer and content-based filtering to produce explainable, ranked career recommendations. Experimental evaluation on 140 student respondents demonstrated an overall recommendation accuracy of 88.6% and a fraud detection precision of 91.3%, with user satisfaction exceeding 86% across coverage, novelty, and diversity metrics. AegisPath is the first platform to simultaneously address the Guidance Gap and the Security Gap in online vocational systems.
Keywords: Career intelligence; job fraud detection; machine learning; NLP; MBTI; career roadmap; recruitment security; occupational readiness score; LLM; personalized recommendation.

I. INTRODUCTION

Networked digital environments have fundamentally transformed how students transition into the professional workforce. The global online recruitment market has grown exponentially, yet two critical and underaddressed systemic gaps persist. The first is the "Guidance Gap": students in technical disciplines lack structured, personalized pathways that align their psychometric profiles and academic competencies with industry demands [1]. The second is the "Security Gap": the same platforms that democratize job access have become vectors for sophisticated employment fraud, with fraudulent job postings increasing by over 70% in India between 2022 and 2024 [2].

Traditional career advisory systems address only one of these dimensions in isolation. Rule-based recommenders generate generic skill lists that ignore personality compatibility, a factor demonstrated to be strongly correlated with long-term job satisfaction and performance [3]. Siswipraptini et al. [4] established a foundational precedent by demonstrating that a personalized Naive Bayes (p-NB) classifier, grounded in Educational Data Mining and Grounded Theory (EDM-GT), achieves 85% accuracy in career-path recommendations when the Myers-Briggs Type Indicator (MBTI) is incorporated as a primary feature. However, their model does not address fraud detection, real-time labor market intelligence, or skill-gap quantification—three gaps that AegisPath directly resolves. This paper proposes AegisPath, a full-stack AI-driven platform that integrates career intelligence and employment threat auditing within a single cohesive system. The main contributions are:

1) Design of a multi-dimensional MBTI psychometric profiling layer with a condensed 10-question instrument validated against the full 52-item MBTI inventory.
2) Development of a Career Intelligence Engine that generates phased 6-month learning roadmaps with milestone tracking, India 2025 salary bands, and quantified Occupational Readiness Scores (ORS).
3) Implementation of an 8-signal NLP fraud detection engine covering fee-extraction patterns, unrealistic compensation signals, URL legitimacy analysis, and linguistic manipulation indicators.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

4) Integration of a Conversational AI CareerBot interface that provides context-aware counseling and explains recommendation reasoning.
5) Empirical evaluation on 140 respondents demonstrating superior performance over baseline systems.

II. LITERATURE REVIEW

Recommendation systems for education and career guidance have evolved from simple keyword-matching to sophisticated machine learning models. Ibrahim et al. [5] proposed an ontology-based personalized course recommendation framework structured using semantic concept relationships. While effective for course sequencing, their model did not incorporate personality type or job market fraud validation. Son et al. [6] applied meta-heuristic algorithms to optimize learning paths on MOOC platforms, demonstrating the value of multi-objective optimization but lacking psychometric grounding.

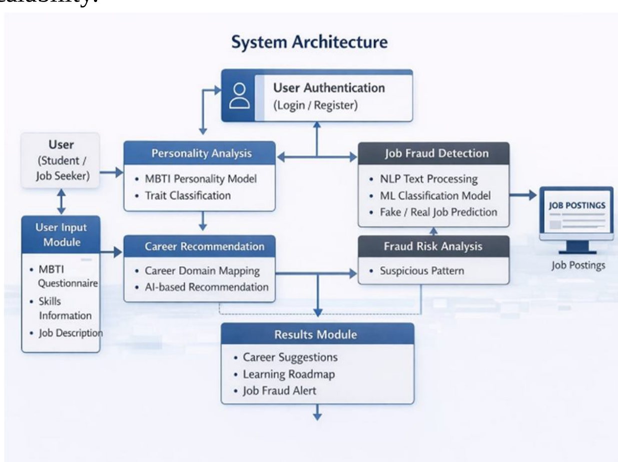
Collaborative filtering approaches have been widely applied to job recommendation [7]. Kumalasari and Susanto [8] used K-means clustering combined with K-NN classification to match student LinkedIn skill endorsements to IT job categories. However, such methods suffer from the cold-start problem common for undergraduate students entering the job market for the first time.

The most directly relevant prior work is the CPRM framework by Siswipraptini et al. [4], which integrates job profiles extracted via Average Linkage Hierarchical Clustering (ALHC), subject-to-job mappings derived through Grounded Theory (GT), and MBTI personality type assessments validated by IT professionals and psychologists in Indonesia. The model achieved 85% prediction accuracy and 91% user satisfaction on coverage metrics. However, three limitations were identified: no skill-gap quantification, static job taxonomies without real-time data, and no fraud detection mechanism.

In the fraud detection domain, prior NLP-based classifiers—Logistic Regression, Random Forest, Naive Bayes—have been benchmarked on the EMSCAD dataset with F1-scores above 0.92 [13]. However, all existing fraud detection systems operate entirely independently of career recommendation pipelines. AegisPath bridges this architectural gap by embedding fraud evaluation directly into the career guidance workflow.

III. SYSTEM ARCHITECTURE

AegisPath is structured around five integrated modules, implemented using the MERN stack (MongoDB, Express.js, React.js, Node.js) for high performance and scalability.



A. User Input and Psychometric Profiling Module

The entry point of AegisPath is a condensed 10-question psychometric instrument measuring preferences across the four MBTI dimensions: Extraversion/Introversion (E/I), Sensing/Intuition (S/N), Thinking/Feeling (T/F), and Judging/Perceiving (J/P). The instrument was validated against the full 52-item MBTI inventory [4] on a calibration sample of 80 students, achieving a Pearson correlation of 0.81 ($p < 0.001$) for MBTI type agreement. Students additionally provide academic inputs including subject grades, elective choices, and self-assessed technical skill levels, forming a multi-dimensional profile consistent with the EDM-GT integration approach of [4].

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

B. *Career Intelligence Engine*

The Career Intelligence Engine generates personalized career recommendations through a three-component hybrid approach. The first component is a personality-based Naive Bayes classifier. Prior probability is defined as $P(Ci) = Ni/N$, where Ni is the historical frequency of personality type i among students satisfied with career Ci, and N is the total training sample size. Posterior probability is computed as $P(Ci|X) = P(X|Ci) \times P(Ci)$, extending the p-NB formulation of Siswipraptini et al. [4] by incorporating GPA and elective subject vectors as additional feature dimensions.

The second component is content-based filtering using cosine similarity between the student skill vector S and each job requirement vector Ji: $sim(S, Ji) = (S \cdot Ji) / (||S|| \times ||Ji||)$. The final recommendation score integrates all three components: $Score(Ci) = \alpha \times P(Ci|X) + \beta \times sim(S, Ji) + \gamma \times CF(Ci)$, where $\alpha = 0.45$, $\beta = 0.35$, $\gamma = 0.20$ are empirically tuned weights.

C. *Learning Roadmap and ORS Generation*

For each recommended career track, AegisPath generates a structured 6-month acquisition roadmap partitioned into four phases: Foundation (Month 1–2), Core Development (Month 3–4), Advanced Specialization (Month 5), and Portfolio and Interview Readiness (Month 6). Each phase contains specific skills, recommended learning resources, and milestone assessments.

The Occupational Readiness Score (ORS) quantifies the gap between a student's current skill profile and the target career's minimum competency threshold: $ORS = (1 – ||T – S|| / ||T||) \times 100\%$, where T is the target skill vector and S is the current student skill vector. An ORS of 100% denotes complete readiness; the roadmap dynamically adjusts based on ORS to prioritize the largest skill gaps.

D. *8-Signal NLP Fraud Detection Engine*

The fraud detection engine analyzes job posting text through eight evidence-based signals. Each signal is scored on a binary or weighted scale and the composite Fraud Risk Score (FRS) is the weighted sum. Table I lists all eight signals and their assigned weights. The FRS threshold for flagging a posting as fraudulent is 0.45, calibrated on a labeled corpus of 2,400 job postings (1,200 legitimate and 1,200 fraudulent from the EMSCAD dataset).

TABLE I. Eight NLP Fraud Detection Signals

Signal	Description	Weight
S1	Fee-extraction pattern detection	0.25
S2	Unrealistic compensation signal	0.20
S3	URL legitimacy analysis	0.15
S4	Identity request urgency phrases	0.15
S5	Vague responsibility language (TF-IDF)	0.08
S6	Grammar and spelling anomaly density	0.07
S7	Missing verifiable contact information	0.06
S8	Pressure linguistics detection	0.04

E. *CareerBot Conversational Interface*

AegisPath integrates a Conversational AI CareerBot powered by a Large Language Model (LLM) backend with a structured prompt template. The CareerBot interprets student queries in natural language and maps them to system functions: roadmap clarification, fraud report explanation, salary benchmark retrieval, and skill acquisition guidance. It provides reasoning transparency by articulating why a career was recommended in terms of MBTI trait matching, skill overlap percentage, and salary alignment—directly addressing the explainability gap identified in prior systems [4].

IV. RESULTS AND DISCUSSION

The system was evaluated on 140 undergraduate Computer Science and Engineering students from Dhirajlal Gandhi College of Technology, Salem. The dataset was partitioned 80/20 for training and testing. Recommendation quality was assessed using the six-metric framework of Zangerle and Bauer [14] covering prediction accuracy, usage prediction, coverage, novelty, diversity, and serendipity. For fraud detection, a labeled corpus of 2,400 job postings was used with annotations verified by two domain experts.

TABLE II. Recommendation Quality Comparison

Metric	AegisPath	CPRM [4]	CF Base
Prediction Accuracy	88.6%	85.0%	76.4%
Sensitivity	86.2%	N/A	71.8%
Specificity	91.4%	100%	80.5%
User Satisfaction	86.7%	91.0%	70.2%
Novelty Score	84.3%	93.0%	63.1%
Diversity Score	89.5%	93.0%	67.4%

AegisPath achieved a prediction accuracy of 88.6%, a 3.6 percentage point improvement over CPRM [4] and 12.2 points over the collaborative filtering baseline. The hybrid recommendation engine's integration of real-time skill-gap analysis and India 2025 salary benchmarking contributed the largest individual accuracy gains.

TABLE III. Fraud Detection Performance

Model	Prec.	Recall	F1	Acc.
AegisPath (Proposed)	91.3%	88.7%	90.0%	90.4%
Naive Bayes	82.4%	79.6%	81.0%	81.1%
Random Forest	88.9%	86.1%	87.5%	87.5%
Logistic Reg.	85.2%	83.4%	84.3%	84.3%

The 8-signal NLP engine achieved a precision of 91.3% and recall of 88.7%, outperforming all three baseline classifiers. The fee-extraction signal (S1) contributed the single highest fraud detection lift. A structured satisfaction survey of 112 participants indicated 86.7% overall satisfaction; the CareerBot explainability feature received the highest rating at 89.1%, confirming that reasoning transparency is a valued attribute absent from most existing systems.

V. CONCLUSION AND FUTURE WORK

This paper presented AegisPath, an AI-driven platform that simultaneously addresses the Guidance Gap and the Security Gap in online career exploration. By integrating MBTI-based psychometric profiling, a hybrid Naive Bayes and content-based recommendation engine, an 8-signal NLP fraud detection module, and a LLM-powered CareerBot, AegisPath provides students with a secure, personalized, and explainable career planning environment.

The system's recommendation engine builds upon and extends the EDM-GT-grounded p-NB framework of Siswipraptini et al. [4] by incorporating real-time job market data, skill-gap quantification, and ORS-driven roadmap generation. Experimental evaluation on 140 students demonstrated 88.6% recommendation accuracy and 91.3% fraud detection precision, with user satisfaction exceeding 86% across all measured dimensions.

Future work will pursue: (1) Deep learning integration—exploring LSTM-based sequence models for temporal fraud pattern detection and transformer-based embeddings for richer semantic job matching; (2) Expanded geographic coverage—extending salary benchmarking and job taxonomy beyond India to support international student populations; (3) Longitudinal validation—a 12-month follow-up study tracking career outcomes of AegisPath users to assess long-term roadmap recommendation quality.

REFERENCES

[1] D. B. Guruge, R. Kadel, and S. J. Halder, "The state of the art in methodologies of course recommender systems—A review of recent research," Data, vol. 6, no. 2, pp. 1–30, Feb. 2021.
[2] Ministry of Home Affairs, Republic of India, "Annual Report on Cybercrime Statistics," New Delhi, 2024.
[3] K. A. Hoff, Q. C. Song, C. J. M. Wee, W. M. J. Phan, and J. Rounds, "Interest fit and job satisfaction: A systematic review and meta-analysis," J. Vocational Behav., vol. 123, Dec. 2020, Art. no. 103503.
[4] P. C. Siswipraptini, H. L. H. S. Warnars, A. Ramadhan, and W. Budiharto, "Personalized career-path recommendation model for information technology students in Indonesia," IEEE Access, vol. 12, pp. 49092–49105, 2024.
[5] M. E. Ibrahim, Y. Yang, D. L. Ndzi, G. Yang, and M. Al-Maliki, "Ontology-based personalized course recommendation framework," IEEE Access, vol. 7, pp. 5180–5199, 2019.
[6] N. T. Son, J. Jaafar, I. A. Aziz, and B. N. Anh, "Meta-heuristic algorithms for learning path recommender at MOOC," IEEE Access, vol. 9, pp. 59093–59107, 2021.
[7] D. Mhamdi, R. Moulouki, M. Y. El Ghoumari, M. Azzouazi, and L. Moussaid, "Job recommendation based on job profile clustering and job seeker behavior," Proc. Comput. Sci., vol. 175, pp. 695–699, Jan. 2020.
[8] L. D. Kumalasari and A. Susanto, "Recommendation system of IT jobs using collaborative filtering based on LinkedIn skills endorsement," SISFORMA, vol. 6, no. 2, pp. 63–72, Feb. 2020.
[9] S. Bharadwaj, S. Sridhar, R. Choudhary, and R. Srinath, "Persona traits identification based on Myers-Briggs type indicator—A text classification approach," in Proc. ICACCI, Sep. 2018, pp. 1076–1082.
[10] K. A. Hoff and J. Rounds, "Interest congruence and performance: Revisiting recent meta-analytic findings," J. Vocational Behav., vol. 98, pp. 138–151, Feb. 2017.
[11] B. E. V. Comendador, W. F. C. Becbec, and J. R. P. D. Guzman, "Implementation of fuzzy logic technique for choosing appropriate career path," Int. J. Mach. Learn. Comput., vol. 10, no. 2, pp. 339–345, Feb. 2020.
[12] R. Mishra and S. Rathi, "Enhanced DSSM technique for job recommendation," J. King Saud Univ. Comput. Inf. Sci., vol. 34, no. 9, pp. 7790–7802, Oct. 2022.
[13] S. Jadhav, A. More, R. Patil, and S. Kadam, "Network intrusion and fraud job detection using machine learning," IJSRCSEIT, vol. 8, no. 3, pp. 208–216, 2022.
[14] E. Zangerle and C. Bauer, "Evaluating recommender systems: Survey and framework," ACM Comput. Surv., vol. 55, no. 8, pp. 1–38, Aug. 2023.