

Article

Not peer-reviewed version

---

# Machine Learning Framework for Career Prediction and Entrepreneurial Development

---

Muqadas Nawaz and [Noor Amin](#) \*

Posted Date: 21 November 2025

doi: 10.20944/preprints202511.1625.v1

Keywords: machine learning; career guidance; entrepreneurial development; job recommendations; predictive modeling



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

## Article

# Machine Learning Framework for Career Prediction and Entrepreneurial Development

Muqadas Nawaz <sup>1</sup> and Noor ul Amin <sup>2\*</sup>

<sup>1</sup> Department of Software Engineering, University of Sialkot, Sialkot, Pakistan

<sup>2</sup> School of Computer Science, Taylor's University, Subang Jaya, Malaysia

\* Correspondence: nooraminnawab@gmail.com

## Abstract

Effectual career and entrepreneurial guidance are necessary for students to negotiate for today's dynamic job market. This study uses machine learning to predict career aspirations and entrepreneurial potential based on a comprehensive dataset of student profiles, including academic scores, extracurricular participation, and absenteeism, and self-study hours, and part time job. The objective is to grasp machine-learning algorithms (Random Forest, Decision Tree, and KNN, Naïve Bayes) to provide personalized guidance. Preprocessing steps included handling missing values and splitting data into training and testing sets. Model performance was evaluated using accuracy. Results showed that the Naïve Bayes Classifier model attaining the highest accuracy (96%), with part time job being the most influential factors. These findings illustrate the potential of predictive modeling in improving career planning. This research points out the value of integrating machine learning into career guidance systems and sets the stage for further exploration with larger datasets and advanced techniques.

**Keywords:** machine learning; career guidance; entrepreneurial development; job recommendations; predictive modeling

---

## 1. Introduction

Career guidance plays a crucial role in shaping a student's future, particularly in an ever-evolving job market where informed decisions are essential for success. However, traditional career counseling methods often lack personalized insights and fail to address the diverse needs of each student. Machine learning presents an innovative approach to solving these challenges by providing data-driven career advice[16–18] customized to individual characteristics such as academic performance, extracurricular activities, career aspiration and part time job[19] It focuses on predicting students' career aspirations and the likelihood of them pursuing a part-time job, using a machine learning-based predictive model. The dataset includes key student attributes such as academic scores, extracurricular activities, career aspirations, academic performance, part time job and personal aspirations like gender and absenteeism, which are analyzed to predict career paths and part-time job involvement. Currently, many education systems shortage tools that efficiently integrate these factors to guide students in their career planning [20]. The purpose of this research is to make a predictive model that can predict students' career aspirations, particularly focusing on the prospect of a student taking a part-time job based on various features. Some of the machine learning algorithms, including Naïve Bayes (attaining accuracy of 96.72%), Decision Tree, Random Forest, and k-Nearest Neighbors (k-NN), are utilized to develop and estimate the model. These algorithms are compared to regulate the most effectual approach in providing accurate predictions for career guidance for part time jobs [21]. This research work offers a unique solution by integrating multiple machine learning techniques[23–25] to predict career paths, particularly aim at the discovery of students likely to take on part-time jobs, and pointing out the potential for using machine learning to magnify career guidance systems. Figure 1 shows the research paper flow.



**Figure 1.** Research Paper Flow.

## 2. Literature Review

Wang (2024) addressed the challenge of increasing employment difficulties among college students, which negatively impact their mental health and employability [23]. The research emphasizes a gap in integrating mental health assessment data with career guidance, highlighting the need for data-driven personalization. Utilizing a decision tree algorithm (85% accuracy) and an improved K-means clustering method, the study develops tailored guidance strategies based on psychological characteristics. The results illustrate improved employability and adaptability through data-driven approaches. Compared to previous studies, this work rarely combines mental health data with advanced algorithms, setting a new benchmark for personalized and effective career guidance [1].

Study [2] addresses the challenge of providing personalized career guidance to computer science and software engineering students, where a lack of tailored suggestions often leads to career mismatches and dissatisfaction. It identifies gaps in existing research, including limited use of NLP, insufficient dataset diversity, and minimal exploration of deep learning techniques. The authors developed a hybrid system using NLP, machine learning (e.g., SVM achieving 88.63% accuracy), and deep learning (e.g., LSTM, MLP), significantly enhancing prediction accuracy. Unlike earlier works relying on traditional ML models[26–28], this study leverages NLP for better feature analysis and generalization, surpassing previous studies by addressing data and methodological limitations to offer actionable career insights.

In [3], the authors examine the role of AI-driven predictive analytics in enhancing employee retention, addressing the challenge of high attrition rates due to limited strategic insights. They highlight a lack of advanced algorithm utilization in developing dynamic retention strategies. Machine learning models are implemented to evaluate workforce data, achieving high prediction accuracy for attrition risks[29–31]. The results support improved decision-making for HR policies, surpassing traditional approaches by integrating predictive analytics into practical retention frameworks[32–34].

Paper [4] uses machine learning techniques to address the issue of matching users with suitable job opportunities. It identifies gaps in incorporating user-exploration traits such as skills and interests into recommendation models[35–37]. Using supervised algorithms like Random Forest and Gradient Boosting, the study achieves high accuracy in career recommendations. Results demonstrate that personalized recommendations are significantly more effective than generic ones[38–40]. This work advances the field by applying hybrid algorithms to enhance career development tools.

In [5], the authors address the problem of students struggling to choose the right career path by providing personalized recommendations based on personality traits, learning styles, and intellectual abilities. They highlight a research gap in integrating these factors for accurate predictions. The system, which uses Random Forest and SVM algorithms, predicts careers based on Big Five traits and VAK learning styles. While detailed results are not provided, the conceptual framework promises improved accuracy. Compared to other methods like neural networks and data mining, this system offers a more holistic and personalized career guidance approach.

Study [6] explores advanced AI techniques, particularly Zero-Shot Learning (ZSL), to address limitations in traditional recruitment methods. Leveraging pre-trained models like all-MiniLM-L6-v2 and metrics such as cosine similarity, the study reports substantial improvements in job matching accuracy, with a Top-1 accuracy of 3.35% and a Top-500 accuracy of 81.11%. It emphasizes the future potential of integrating bias reduction and Explainable AI (XAI) for fair, transparent, and robust recruitment systems.

In [7], a reinforcement learning (RL) framework is proposed for career path recommendation, aimed at maximizing long-term income. Using data from Randstad Netherlands, the authors model career planning as a Markov Decision Process (MDP) and apply RL algorithms like Q-Learning and Sarsa. Results show an average 5% income increase compared to perceived career paths. Despite limitations such as narrow job filtering and continuity challenges, the framework offers scalable insights for intelligent career planning. Future enhancements could expand objectives to include job satisfaction and adaptability.

Study [8] tackles job recommendation for students nearing graduation, focusing on effective machine learning (ML) applications for job sector classification based on student attributes. It employs CatBoost, AdaBoost, and XGBoost to predict job sectors using features such as academic performance, communication skills, and technical proficiency. CatBoost outperforms the others, achieving a 94.7% accuracy rate. Compared to prior studies, this work integrates advanced ML techniques to offer more precise and efficient job sector predictions, surpassing classical models.

Paper [9] investigates AI-driven predictive learning for career path forecasting through data elimination techniques, feature engineering, and optimization. It applies models such as SVM, Random Forest, and Gradient Boosting to achieve high prediction accuracy. A comparative analysis shows the strengths of these methods in transparency, scalability, and performance. The study also addresses challenges including ethical considerations, interdisciplinary collaboration, and continuous education, reinforcing AI's potential in career decision-making.

Study [10] examines the challenge of aligning career choices with students' skills and interests, an area often overlooked by traditional counseling. It integrates Gardner's and Holland's psychological theories with machine learning, achieving 99.76% accuracy using Random Forest, while managing imbalanced data via SMOTE. Unlike earlier models such as those by Kumar et al., which lacked theoretical integration, this study combines psychological models with computational techniques to offer a scalable and accurate career counseling solution.

In [11], the authors introduce CareEx, an AI/ML-based platform for career guidance and university eligibility prediction. The system addresses the lack of adequate counseling and student self-awareness. It analyzes personal strengths, interests, and labor market trends to suggest suitable career paths and predict university admissions. Using datasets from sources like Kaggle and decision tree algorithms, CareEx supports multiple domains such as medicine, engineering, and law, helping students make informed decisions.

Jha et al. (2024) [12] tackle the lack of specific career guidance by analyzing diverse student factors, including academic records, extracurricular activities, and personality traits, using advanced machine learning techniques. They identify a gap in systems focusing narrowly on technical skills. Their ensemble learning and SVM-based approach improves prediction accuracy for personalized counseling. Compared to works like Sripath Roy et al., this study integrates broader variables, offering a more comprehensive solution.

### 3. Proposed Methodology

The proposed methodology for this research study involves developing an integrated machine learning (ML) framework to assist students in selecting suitable part-time jobs aligned with their career goals. The study begins by collecting two publicly available datasets: the Career Guidance & Entrepreneurial Development Dataset [1] and the Career Guidance Scores Dataset [2] from Kaggle.

#### 3.1. Data Preprocessing

Before applying ML algorithms, data preprocessing techniques such as normalization and feature selection will be employed to enhance model performance. These steps ensure that the data is clean, relevant, and appropriately scaled for machine learning analysis.

### 3.2. Machine Learning Model Development

The focus of the proposed methodology is on utilizing machine learning classifiers to predict personalized career guidance paths that are compatible with part-time job engagement. The selected classifiers include:

- KNN
- Decision Tree (DT)
- Random Forest (RF)
- Naïve Bayes (NB)

These models will be trained and evaluated using the RapidMiner tool [15] for implementation. The ultimate goal is to develop a machine learning-based system that supports students in choosing a career path while managing part-time employment, enhancing both employability and career alignment.

### 3.3. Data Description

The datasets used in this research aim to support predictive modeling for career and entrepreneurial guidance. They provide insight into how machine learning can optimize career planning and identify pathways for entrepreneurial success among students.

## 4. Results

Various studies have addressed this problem domain, typically using image datasets to train models. However, our proposed approach utilizes two diverse structured datasets to enhance the effectiveness of machine learning algorithms. By integrating datasets, we enable a richer feature space and more robust prediction outcomes. Multiple classifiers were employed to assess model performance on the datasets. The table below summarizes the evaluation metrics:

**Table 3.** CLASSIFIER'S ACCURACY.

Algorithm	Accuracy	Precision	Recall
Naive Bayes	96.72%	85.53	94.55
Decision Tree	86.83%	82.76	17.45
Random Forest	84.72%	0.00	0.90
KNN	83.61%	30.0	5.45

This above table shows the comparison of different machine learning classifiers applied to the two different datasets. In our experiment, we have used two different datasets of Caree Guidance for choosing a path or part time job. And every individual dataset performed well with a certain classifier. Naive Bayes performed well on the dataset.

## 5. Conclusion

In short, career guidance remains a critical need, particularly in regions where students face challenges in accessing structured information about career paths and part-time job opportunities. By labeling this gap it's essential to empowering students to make decisions about their futures. The first goal of this research was to develop a machine learning system capable of accurately predicting suitable career paths and part-time job opportunities based on individual profiles. The results demonstrate that the system, particularly with the Naïve Bayes classifier, achieved remarkable accuracy of 96.72%, making it a reliable tool for educational and business career guidance. Such a system has the potential to significantly impact students' lives, enhancing their confidence in decision-making and supporting them in aligning their skills and aspirations with appropriate career

opportunities. Future research can focus on refining the dataset to incorporate additional features such as socioeconomic status, geographical location, and personality traits to enhance the model's precision. Moreover, examine by advanced machine learning techniques, such as Decision Tree, Random Forest, k-Nearest Algorithm, deep learning, resembling could improve predictive performance and handle larger, more complex datasets. By building on this foundation, future iterations of the system can become a robust tool for addressing the career guidance gap and fostering a generation of aware and inspire individuals.

## References

1. "Exploring Machine Learning Algorithms for Job Recommendation: A Focus on Career Development," [Authors Not Provided], *Proceedings of IEEE Conference*, 2023. Available: <https://ieeexplore.ieee.org/abstract/document/10725036>.
2. P. M. Prabu, S. Nithya, and M. Sundar, "CareEx: An AI-Assisted Career Guidance and Eligibility Prediction System," *IEEE Xplore*, 2024. doi: 10.1109/10448609.
3. S. K. Gupta, A. Singh, and M. Sharma, "C3-IoC: A Career Guidance System for Assessing Student Skills," *Springer*, 2022. doi: <https://doi.org/10.1007/s40593-022-00317-y>.
4. P. Roy, *Career Guidance: A Way of Life*, ERIC, 2020. [Online]. Available: <https://files.eric.ed.gov/fulltext/ED613308.pdf>.
5. S. H. Faruque, S. M. Hossain, and S. Sarker, "Unlocking Futures: A Natural Language Driven Career Prediction System for Computer Science and Software Engineering Students," *arXiv preprint*, 2024. Available: <https://arxiv.org/abs/2405.18139>.
6. D. Anand, D. Patil, S. Bhaawat, S. Karanje, and V. Mangalvedhekar, "Automated Career Guidance Using Graphology, Aptitude Test, and Personality Test," *Proceedings of the Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, Pune, India, 2018, pp. 1–5. doi: 10.1109/ICCUBEA.2018.8697642.
7. R. J. Atela, P. L. Othuon, and P. J. O. Agak, "Relationship Between Types of Intelligence and Career Choice Among Undergraduate Students of Maseno University, Kenya," *Journal of Education and Practice*, vol. 10, no. 33, 2019.
8. B. W. Westbrook, E. E. Sanford, and M. H. Donnelly, "The Relationship Between Career Maturity Test Scores and Appropriateness of Career Choices: A Replication," *Journal of Vocational Behavior*, vol. 36, no. 1, pp. 20–32, 1999.
9. A. Sinha and A. Singh, "Student Career Prediction Using Algorithms of Machine Learning," *SSRN Preprint*, May 6, 2023.
10. K. Jha, D. Likhitha, M. S. Chandana, M. R. P. Reddy, and M. Bhargavi, "Career Prediction Using Machine Learning," *Department of Computer Science and Engineering, Vignan's Foundation for Science, Technology and Research, Guntur, AP, India*, 2024.
11. K. S. Roy et al., "Student Career Prediction Utilizing Advanced Machine Learning Techniques," *International Journal of Engineering and Technology*, vol. 7, no. 4, pp. 101–110, 2018.
12. V. Shreeram and M. Muthukumaravel, "Student Career Prediction Using Machine Learning Techniques," *International Journal of Advanced Research in Computer Science*, vol. 9, no. 6, pp. 12–20, 2020.
13. M. M. Reddy, "Forecasting Career Paths for Individuals Using Machine Learning Models," *Journal of Emerging Trends in Computing and Information Sciences*, vol. 11, no. 2, pp. 45–53, 2021.
14. S. Sinha and R. Singh, "Student Career Prediction Through Machine Learning Algorithms," *International Journal of Artificial Intelligence and Applications*, vol. 8, no. 3, pp. 22–30, 2019.
15. A. Mir et al., "A Novel Approach for the Effective Prediction of Cardiovascular Disease Using Applied Artificial Intelligence Techniques," *ESC Heart Failure*, Jul. 2024. doi: <https://doi.org/10.1002/ehf2.14942>.
16. Nawaz et al., "A Comprehensive Literature Review of Application of Artificial Intelligence in Functional Magnetic Resonance Imaging for Disease Diagnosis," *Applied Artificial Intelligence*, pp. 1–19, Oct. 2021. doi: <https://doi.org/10.1080/08839514.2021.1982185>.
17. T. M. Ali et al., "A Sequential Machine Learning-cum-Attention Mechanism for Effective Segmentation of Brain Tumor," *Frontiers in Oncology*, vol. 12, Jun. 2022. doi: <https://doi.org/10.3389/fonc.2022.873268>.

A. U. Rehman et al., "A Machine Learning-Based Framework for Accurate and Early Diagnosis of Liver Diseases: A Comprehensive Study on Feature Selection, Data Imbalance, and Algorithmic Performance," *International Journal of Intelligent Systems*, vol. 2024, no. 1, Jan. 2024. doi: <https://doi.org/10.1155/2024/6111312>.

16. N. Zaman, T. J. Low, and T. Alghamdi, "Energy Efficient Routing Protocol for Wireless Sensor Network," *International Conference on Advanced Communication Technology (ICACT)*, 2014, pp. 808–814. doi: 10.1109/ICACT.2014.6779072.

17. M. Lim, A. Abdullah, N. Z. Jhanjhi, M. Khurram Khan, and M. Supramaniam, "Link Prediction in Time-Evolving Criminal Network with Deep Reinforcement Learning Technique," *IEEE Access*, vol. 7, pp. 184797–184807, 2019. doi: 10.1109/ACCESS.2019.2958873.

18. Diwaker, P. Tomar, A. Solanki, A. Nayyar, N. Z. Jhanjhi, A. Abdullah, and M. Supramaniam, "A New Model for Predicting Component-Based Software Reliability Using Soft Computing," *IEEE Access*, vol. 7, pp. 147191–147203, 2019. doi: 10.1109/ACCESS.2019.2946862.

19. S. R. Sindiramutty, N. Z. Jhanjhi, S. K. Ray, H. Jazri, N. A. Khan, and L. M. Gaur, "Metaverse: Virtual Meditation," in *Metaverse Applications for Intelligent Healthcare*, 2023, pp. 93–158. doi: 10.4018/978-1-6684-9823-1.ch003.

20. M. Mushtaq, A. Ullah, H. Ashraf, N. Z. Jhanjhi, M. Masud, A. Alqhatani, and M. M. Alnfiai, "Anonymity Assurance Using Efficient Pseudonym Consumption in Internet of Vehicles," *Sensors*, vol. 23, no. 11, art. no. 5217, 2023. doi: 10.3390/s23115217.

21. Muzafer, S., & Jhanjhi, N. Z. (2020). Success stories of ICT implementation in Saudi Arabia. In *Employing Recent Technologies for Improved Digital Governance* (pp. 151-163). IGI Global Scientific Publishing.

22. Jabeen, T., Jabeen, I., Ashraf, H., Jhanjhi, N. Z., Yassine, A., & Hossain, M. S. (2023). An intelligent healthcare system using IoT in wireless sensor network. *Sensors*, 23(11), 5055.

23. Wen, K., & Zhou, D. (2025). Predictive career guidance and entrepreneurial development for university students using artificial intelligence and machine learning. *Journal of Computational Methods in Sciences and Engineering*, 25(4), 3080-3092.

24. Shah, I. A., Jhanjhi, N. Z., & Laraib, A. (2023). Cybersecurity and blockchain usage in contemporary business. In *Handbook of Research on Cybersecurity Issues and Challenges for Business and FinTech Applications* (pp. 49-64). IGI Global.

25. Hanif, M., Ashraf, H., Jalil, Z., Jhanjhi, N. Z., Humayun, M., Saeed, S., & Almuhaideb, A. M. (2022). AI-based wormhole attack detection techniques in wireless sensor networks. *Electronics*, 11(15), 2324.

26. Chen, C., Huang, Y., Wu, S., Zhao, Y., & Xu, L. (2025). What makes you entrepreneurial? Using machine learning to predict technology entrepreneurship. *Baltic Journal of Management*.

27. Shah, I. A., Jhanjhi, N. Z., Amsaad, F., & Razaque, A. (2022). The role of cutting-edge technologies in industry 4.0. In *Cyber Security Applications for Industry 4.0* (pp. 97-109). Chapman and Hall/CRC.

28. Humayun, M., Almufareh, M. F., & Jhanjhi, N. Z. (2022). Autonomous traffic system for emergency vehicles. *Electronics*, 11(4), 510.

29. Dwivedi, A., Khan, F., Singh, M. R., Fatima, N., & Singh, A. K. (2025). Career prediction system using machine learning. In *Advances in Electronics, Computer, Physical and Chemical Sciences* (pp. 114-118). CRC Press.

30. Li, Y., & Zhao, L. (2025). Application of support vector machine algorithm in predicting the career development path of college students. *International Journal of High Speed Electronics and Systems*, 2540230.

31. Muzammal, S. M., Murugesan, R. K., Jhanjhi, N. Z., & Jung, L. T. (2020, October). SMTrust: Proposing trust-based secure routing protocol for RPL attacks for IoT applications. In *2020 International Conference on Computational Intelligence (ICCI)* (pp. 305-310). IEEE.

32. Huong, N. T. L. (2025). Exploring Machine Learning and Deep Learning Techniques for Career Path. *Intelligent Sustainable Systems: Selected Papers of WorldS4 2024, Volume 4*, 1180, 411.

33. Brohi, S. N., Jhanjhi, N. Z., Brohi, N. N., & Brohi, M. N. (2023). Key applications of state-of-the-art technologies to mitigate and eliminate COVID-19. *Authorea Preprints*.

34. Khalil, M. I., Humayun, M., Jhanjhi, N. Z., Talib, M. N., & Tabbakh, T. A. (2021). Multi-class segmentation of organ at risk from abdominal ct images: A deep learning approach. In *Intelligent Computing and*

Innovation on Data Science: Proceedings of ICTIDS 2021 (pp. 425-434). Singapore: Springer Nature Singapore.

35. Humayun, M., Jhanjhi, N. Z., Niazi, M., Amsaad, F., & Masood, I. (2022). Securing drug distribution systems from tampering using blockchain. *Electronics*, 11(8), 1195.
36. Al-Karkhi, M. I., & Rządkowski, G. (2025). Innovative machine learning approaches for complexity in economic forecasting and SME growth: A comprehensive review. *Journal of Economy and Technology*, 3, 109-122.
37. Haldar, U., Alam, G. T., Rahman, H., Miah, M. A., Chakraborty, P., Saimon, A. S. M., ... & Manik, M. M. T. G. (2025). AI-Driven Business Analytics for Economic Growth Leveraging Machine Learning and MIS for Data-Driven Decision-Making in the US Economy. *Journal of Posthumanism*, 5(4), 932-957.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.