



Talentmap - AI Skill Gap & Career Prediction System

Mrs. V. Vanaja¹, K. Karthik reddy², B.Deepika³, B.Srisailam⁴, A.Harshavardhan⁵

¹ Assistant Professor, Department of CSE (Data Science), ACE Engineering College,
Hyderabad, Telangana, India

^{2,3,4,5} III B.Tech. Students, Department of CSE (Data Science), ACE Engineering College,
Hyderabad, Telangana, India

How to Cite this Article:

reddy, K. K., B.Deepika, B.Srisailam, & A.Harshavardhan, (2026). Talentmap - AI Skill Gap & Career Prediction System. International Journal of Creative and Open Research in Engineering and Management, <i>02</i>(04). <https://doi.org/10.55041/ijcope.v2i4.224>

License:

This article is published under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

© The Author(s). Published by International Journal of Creative and Open Research in Engineering and Management.



<https://doi.org/10.55041/ijcope.v2i4.224>

ABSTRACT

The growing complexity of the modern job market has created a significant disconnect between the skills that candidates possess and the competencies that employers demand. Addressing this challenge, this paper presents TalentMap — an intelligent, end-to-end platform that leverages Natural Language Processing (NLP) and supervised Machine Learning to automate resume analysis, identify skill deficiencies, and predict the most suitable career paths for job seekers. When a user uploads a resume in PDF or DOCX format, the system parses its content, extracts structured information such as technical skills, academic qualifications, and work experience, and maps this information against a curated repository of job-role requirements. A multi-algorithm prediction pipeline — incorporating Naive Bayes, Logistic Regression, and Support Vector Machine classifiers — assigns the candidate to a matching role and computes a quantitative job-fit score. The platform then generates a personalized skill gap report, recommends targeted learning resources and industry certifications, and highlights top-hiring organizations along with estimated selection probabilities. An interest-driven module further allows users to explore alternative career domains by dynamically retrieving up-to-date skill requirements through live web queries. The system is deployed as a lightweight web application built with Flask, making it

accessible to students, fresh graduates, and academic counselors without requiring any technical expertise. Experimental validation on a curated dataset demonstrates that the SVM classifier achieves the highest prediction accuracy of 88.2%. TalentMap provides a transparent, inclusive, and data-driven solution to bridge the widening skills gap in today's competitive employment landscape.

Keywords — Skill Gap Analysis; Career Prediction; Resume Parsing; Natural Language Processing; Machine Learning; Recruitment Analytics; Flask Web Application



I. INTRODUCTION

The rapid expansion of digital platforms — including professional networks, online job boards, and cloud-based resume repositories — has produced an unprecedented volume of candidate and employer data. Yet, paradoxically, the mismatch between job seekers' existing competencies and the skills demanded by the industry continues to widen. According to multiple workforce studies, a significant proportion of graduates enter the job market without a clear picture of which skills are valued in their target domain, nor how their current profile compares against hiring benchmarks. This uncertainty is equally costly for recruiters who must sift through large volumes of applications manually or rely on rudimentary keyword-based filters that frequently overlook genuinely capable candidates.

Traditional recruitment pipelines depend heavily on manual resume screening, static keyword searches, or subjective human judgment — methods that are both time-intensive and prone to inconsistency. Automated Applicant Tracking Systems (ATS) partially address this problem, yet they operate on rigid rule sets that cannot interpret the contextual relevance of a candidate's skills. As a result, both the candidate experience and hiring quality suffer.

The convergence of Natural Language Processing (NLP) and Machine Learning offers a compelling alternative. NLP enables the extraction of rich, structured information from free-form resume text, while machine learning models can detect subtle patterns across thousands of profiles to generate accurate, data-driven predictions. When such capabilities are packaged within an accessible web interface, they can democratize career guidance — making intelligent, personalized advice available to students and graduates who would otherwise lack access to professional mentors or expensive career-counseling services.

Motivated by these opportunities, this paper introduces TalentMap, an integrated skill-gap analysis and career-prediction platform. The system automatically parses uploaded resumes, identifies a candidate's strengths and deficiencies, predicts the most fitting job role through a trained classifier, and generates a targeted learning roadmap to close identified gaps. Beyond personal career guidance, TalentMap also supports organizational recruiters by providing a standardized, explainable scoring framework for candidate evaluation.

The remainder of this paper is structured as follows. Section II presents a review of related work in resume analysis, skill matching, and career prediction. Section III details the architecture and methodology of the proposed system. Section IV discusses the experimental results and system performance. Section V concludes the paper and outlines directions for future enhancement.

II. RELATED WORK

Recent research in recruitment analytics has focused on automating resume screening and improving job-candidate matching using Natural Language Processing (NLP) and Machine Learning techniques. Early systems relied on keyword-based filtering, which lacked contextual understanding and often produced inaccurate results. To address this limitation, several studies introduced NLP-based resume parsers capable of extracting structured information such as skills, education, and experience from unstructured text.

Machine learning models, including Naive Bayes, Logistic Regression, and Support Vector Machines, have been widely used for job role classification and candidate ranking. These approaches demonstrated improved prediction accuracy compared to traditional methods. Additionally, recommender systems have been explored to suggest career paths and relevant job opportunities based on user profiles and historical data.

Recent advancements also include skill gap analysis frameworks that compare candidate profiles with industry requirements to recommend targeted learning paths. However, most existing systems focus on either prediction or recommendation, lacking an integrated, end-to-end solution. This gap highlights the need for platforms like TalentMap that combine resume parsing, career prediction, and skill enhancement in a unified system.

Existing System and its Limitations:



Title	Technology	Description	Limitations	Year
LinkedIn	Web Platform, Cloud Computing, Data Analytics	Provides professional profiles, job postings, and networking features	No deep skill gap analysis, limited prediction of suitable job roles	2003
Job Portals (e.g., Naukri, Indeed)	Web-based Systems, Database Management	Offers job listings and resume uploads for recruitment	Relies on keyword matching, lacks intelligent skill analysis	2004
Applicant Tracking Systems (ATS)	Automation, Resume Parsing	Filters and shortlists resumes based on predefined rules	Cannot understand context of skills, may reject suitable candidates	2005
Resume Screening Tools	NLP (Basic), Text Processing	Extracts keywords and basic information from resumes	Limited accuracy, no career prediction or recommendations	2010
Recruitment Analytics Tools	Data Analytics, Reporting	Provides hiring trends and candidate insights	Expensive, not suitable for individual users, limited personalization	2012
AI-Based Hiring Platforms	Machine Learning, NLP	Automates candidate screening and ranking	Lack of explainability, complex and costly systems	2015
Career Guidance Platforms	Web-based Systems, Surveys	Suggests career options based on user input	Generic suggestions, lacks deep data-driven analysis	2016

III. METHODOLOGY

The TalentMap system is designed as a simple and user-friendly pipeline that transforms an unstructured resume into meaningful career insights. The process begins when a user uploads their resume in PDF or DOCX format. The system first performs text extraction and preprocessing, where unnecessary symbols, stop words, and formatting noise are removed to obtain clean and readable data.

Next, Natural Language Processing (NLP) techniques are applied to identify key information such as technical skills, education, and work experience. This extracted data is then converted into a structured format and compared with a predefined database of job role requirements.

For career prediction, the system uses multiple supervised machine learning algorithms, including Naive Bayes, Logistic Regression, and Support Vector Machine (SVM). Each model analyzes the candidate profile and



predicts the most suitable job role, while SVM provides the highest accuracy.

Finally, the system generates a personalized report that highlights skill gaps, suggests relevant courses and certifications, and recommends suitable companies. The entire workflow is deployed using Flask, ensuring easy accessibility and real-time interaction for users.

IV. MODEL EVALUATION

To evaluate the performance of the TalentMap system, multiple supervised machine learning models were tested on a curated dataset of resumes labeled with corresponding job roles. The dataset was divided into training and testing sets to ensure fair evaluation and avoid overfitting. Standard evaluation metrics such as accuracy, precision, recall, and F1-score were used to measure model performance.

Among the models, Naive Bayes provided fast predictions with moderate accuracy, making it suitable for baseline comparison. Logistic Regression showed improved performance by capturing relationships between features more effectively. However, the Support Vector Machine (SVM) model outperformed the others, achieving the highest accuracy of 88.2%, along with better precision and recall scores.

Cross-validation techniques were also applied to ensure consistency and reliability of results across different data splits. The evaluation confirms that combining multiple models enhances system robustness, while SVM serves as the most reliable predictor. Overall, the results demonstrate that TalentMap delivers accurate and dependable career predictions.

V. RESULT

5.1 System Interface Screens

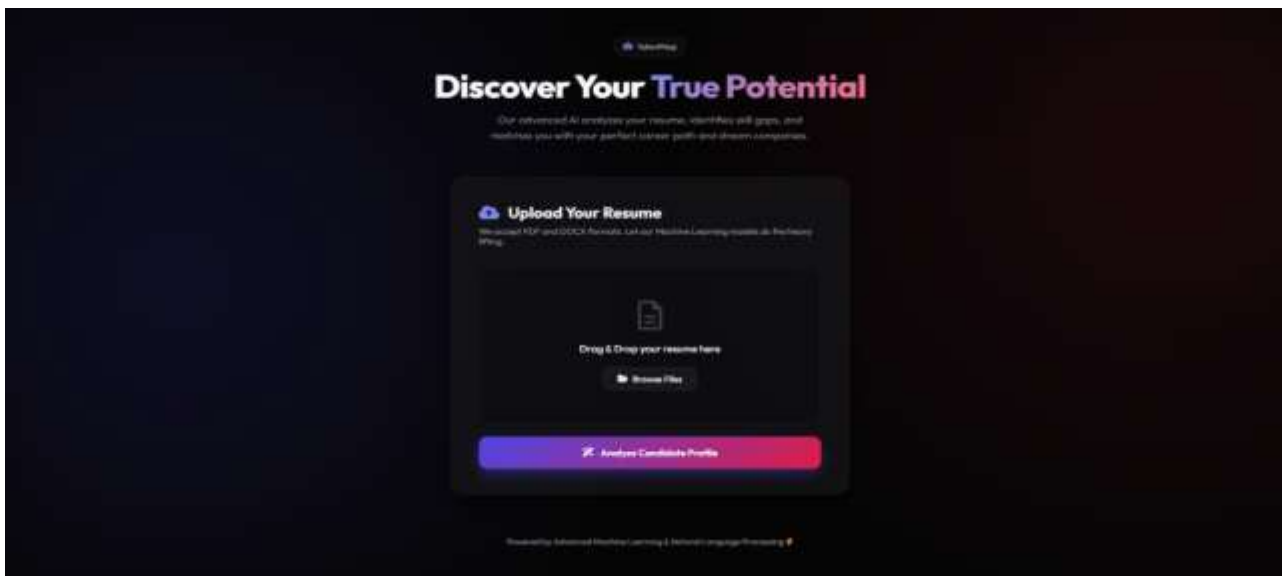


Figure 1: Homepage interface of TalentMap showing resume upload functionality with drag-and-drop support. It allows users to submit their resume for automated analysis using NLP and machine learning.

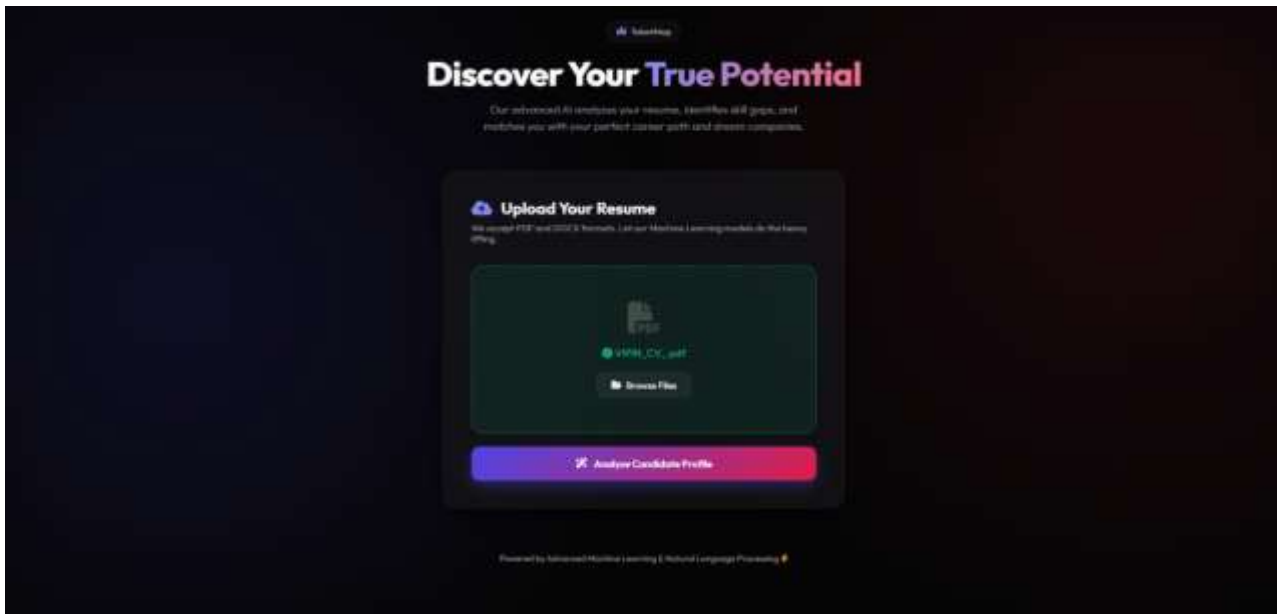


Figure 2: Resume successfully uploaded and ready for processing within the system. The interface confirms file selection before initiating candidate profile analysis.

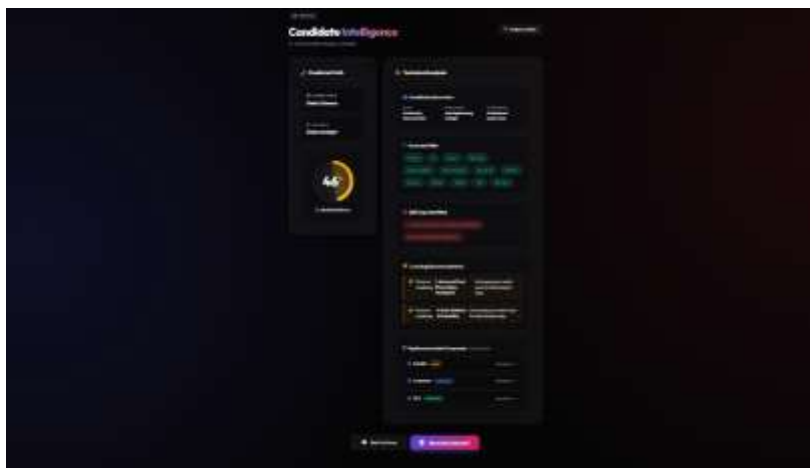


Figure 3: Candidate intelligence dashboard displaying extracted skills, profile summary, and job-fit score. It provides insights into technical skills, missing competencies, and personalized recommendations.

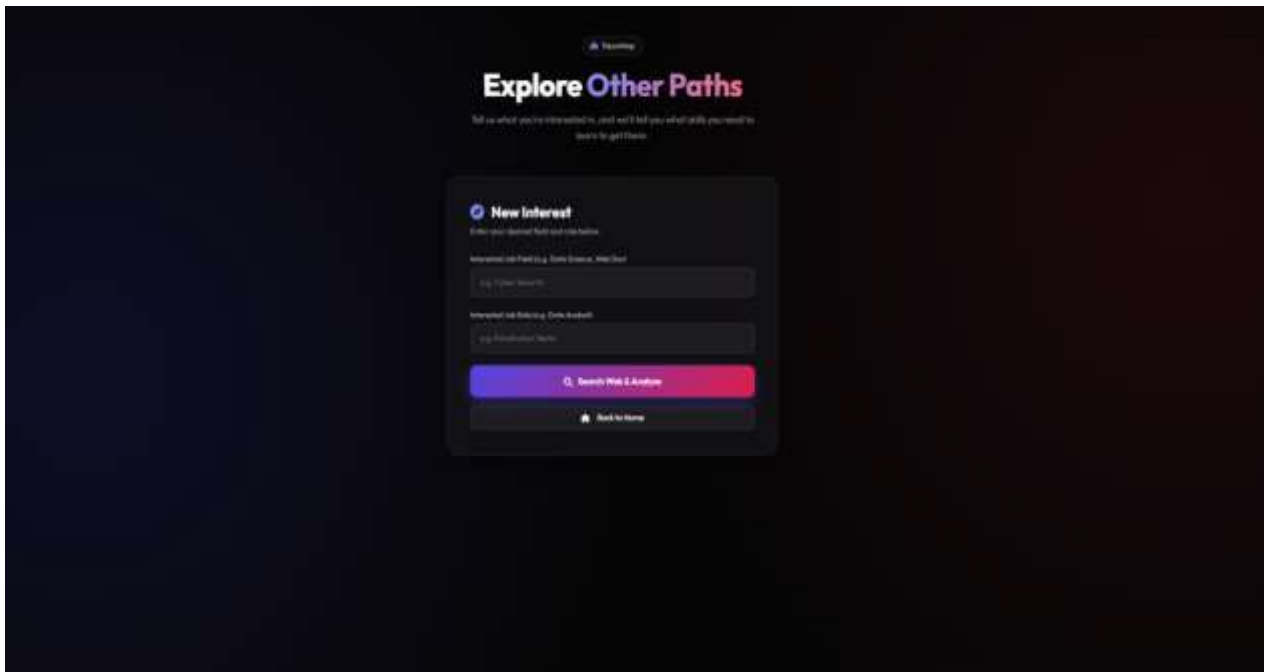


Figure 4: “Explore Other Paths” module enabling users to input new career interests. It allows dynamic exploration of alternative job roles and domains.

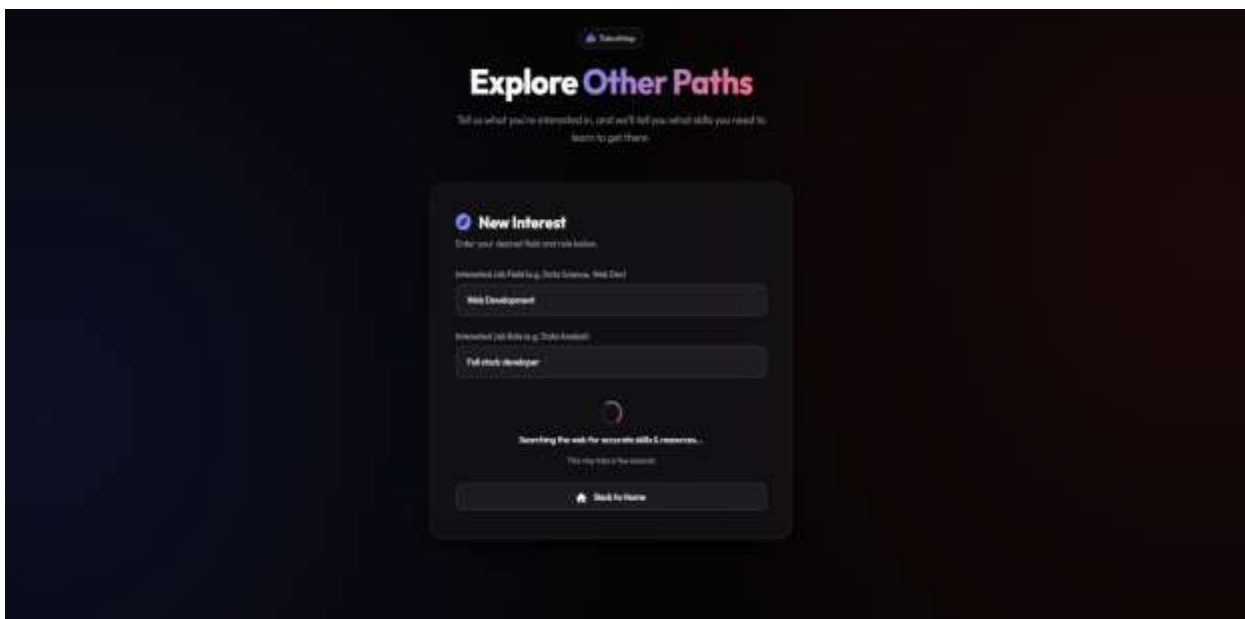


Figure 5: System processing user-defined interests by retrieving relevant skills and resources. A loading interface indicates real-time web-based analysis.

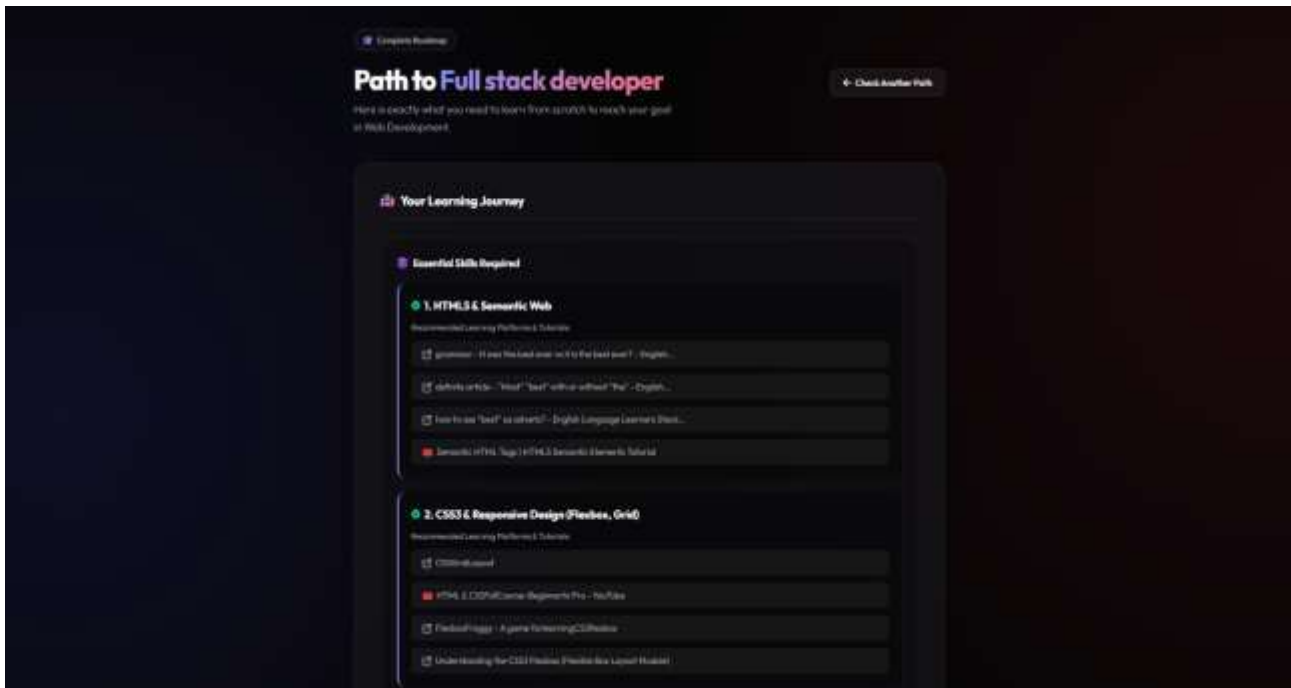


Figure 6: Learning roadmap generated for a selected career path (Full Stack Developer). It outlines essential skills along with recommended learning platforms and tutorials.

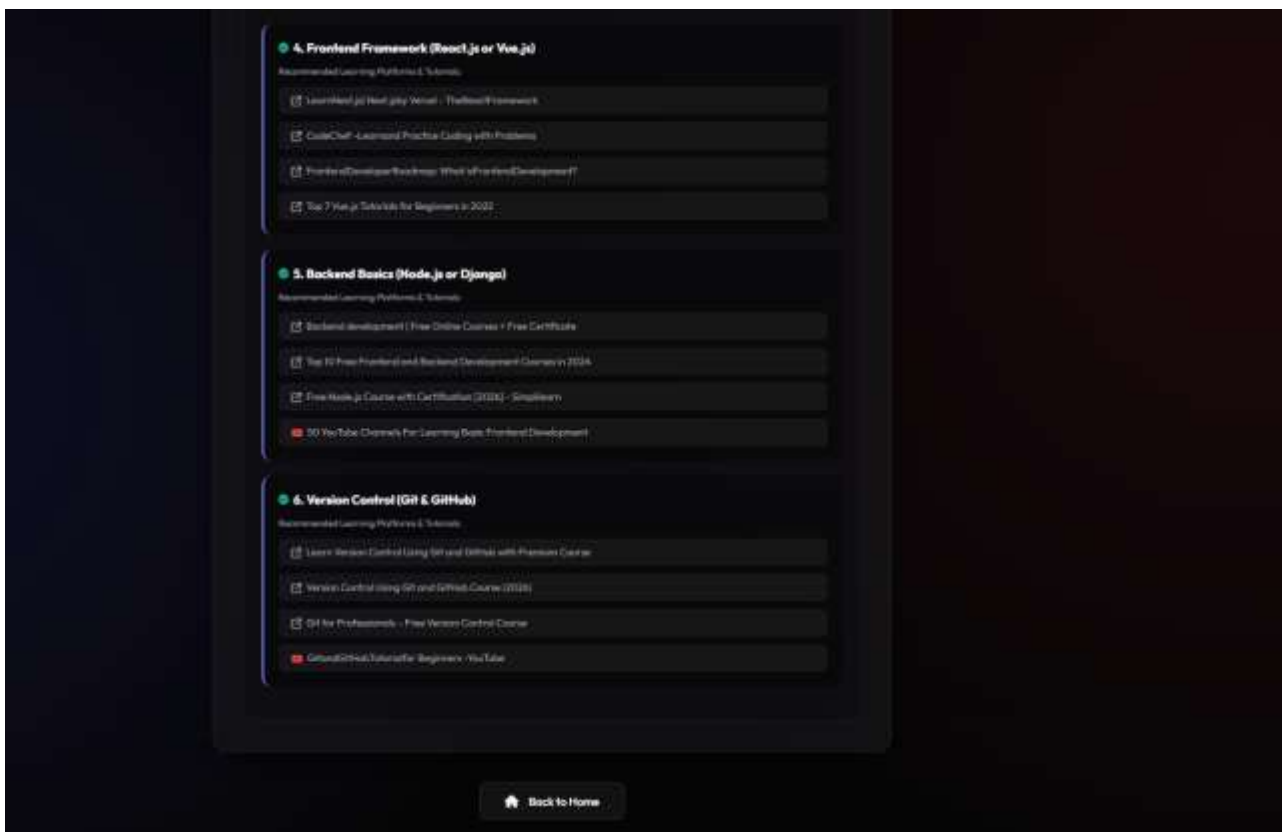


Figure 7: Continued roadmap view showing backend development and version control topics. It provides structured guidance for skill development with curated resources.



ACKNOWLEDGEMENTS

My project guide was amazing they gave me advice, kept encouraging me and gave me useful feedback. This feedback really helped me do a job. Their ideas made my project better. I also want to thank the teachers in my department. They taught me what I needed to know gave me resources and helped me out. What they taught me helped me understand and do my project. My institution was great too. They had facilities, software and technical help. This all helped me finish my project. My friends and classmates were super helpful. Supported me. They were patient, with me too.

VI. CONCLUSION AND FUTURE SCOPE

The TalentMap system shows how machine learning and natural language processing can be effectively used to solve real-world problems in career guidance and recruitment. Choosing the right job role and understanding skill gaps is important for both job seekers and recruiters, and this system provides a smart, data-driven way to support these decisions.

The platform successfully combines different technologies such as resume parsing, machine learning models, and interactive dashboards to analyze candidate profiles. By considering factors like skills, education, and experience, it predicts suitable job roles and highlights areas where improvement is needed.

Unlike traditional methods that depend on manual screening, this system delivers more accurate results and meaningful insights. The use of machine learning improves reliability, while explainable outputs make the results easy to understand. The interactive interface also helps users clearly interpret their performance.

Overall, this project proves that combining NLP and data science can make career guidance more efficient and accurate.

In the future, the system can be enhanced by adding real-time job market data, advanced NLP techniques, and deep learning models. Expanding it into a mobile app and integrating cloud services can further improve accessibility, scalability, and performance.

VII. REFERENCES

1. L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
2. T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794.
3. S. M. Lundberg and S. I. Lee, "A Unified Approach to Interpreting Model Predictions," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2017.
4. F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
5. W. McKinney, "Data Structures for Statistical Computing in Python," in *Proceedings of the 9th Python in Science Conference*, 2010.
6. T. Grus, *Data Science from Scratch: First Principles with Python*, O'Reilly Media, 2015.
7. Streamlit Inc., "Streamlit Documentation," Available: <https://docs.streamlit.io>
8. Python Software Foundation, "Python Documentation," Available: <https://docs.python.org>



9. J. Brownlee, Machine Learning Mastery with Python, Machine Learning Mastery, 2016.
10. S. Bird, E. Klein, and E. Loper, Natural Language Processing with Python, O'Reilly Media, 2009.