



AI Career Role Recommendation System with Skill Gap Analysis, Career Progression, and Salary Simulation Using Resume Semantic Analysis

Ms. Neelam Mary Vijaya Nirmala¹, G. Pallavi², B. Jhansi Lakshmi³, G. Ramya⁴,
J. Divya Dimple⁵

Assistant Professor, Department of Information Technology,
KKR & KSR Institute of Technology and Sciences, Guntur, India¹

Students, Department of Information Technology KKR & KSR Institute of Technology and Sciences, Guntur, India²⁻⁵

Abstract: Due to changing job market trends, differing skill sets, and the absence of personalized career counselling and guidance systems, it has become difficult for job seekers or students to choose their path. Therefore, an AI-Based Career Guidance and Role Recommendation System is proposed in this project, which performs smart resume semantic analysis to recommend career roles based on skills, educational qualifications, amount of experience, and professional interests. Resume parsing and Natural Language Processing (NLP) based techniques are used to convert tacit knowledge contained in a candidate's resume, which is an unstructured collection of words, into structured information about technical skills, educational qualifications, professional certifications, and work experience. TF-IDF (Term Frequency – Inverse Document Frequency) vectorization is used to convert a resume into a high-dimensional feature vector. Cosine similarity based semantic similarity algorithms are then applied to map these candidate's profiles to role descriptor templates and industry-specific competency frameworks of matching job roles. The system includes a Skill Gap Severity Index to identify missing and underserved skills for target roles, a Customized Learning Plan Generator to recommend upskilling pathways, a Career Progression Simulation to visualize different career paths, and a Salary Range and salary levels. Peer Benchmarking System compares a users and Resume Improve intelligence improves resumes by analyzing them to identify missing skills and errors instead of only giving a score. profile with similar candidates having the same degree and experience level. The proposed system is based upon semantic profiling, similarity modeling, predictive career analytics, and resume improvement intelligence, providing students, freshers, and job seekers with personalized, accurate, and actionable recommendations to make informed career decisions and achieve sustainable growth in their respective careers.

Index Terms: Career Role Recommendation, Resume parsing, Skill Extraction, TF-IDF, Semantic Similarity, Artificial Intelligence, Salary Range & Simulation, Natural Language Processing (NLP), Skill Gao Severity Index, Skill-Based Career Mapping.

I. INTRODUCTION

Industry needs change rapidly and school and work readiness skills no longer align. For students and job seekers, career choice is increasingly complex. Conventional career services and online resources often give generic career options without a personalized path forward or exploring how to gain in-demand skills which are required for that career. This leads individuals to struggle with finding careers that will suit their skills, interests, and long-term career development.

With improved Artificial Intelligence and Natural Language Processing, automation of extracting data from unstructured resume data had gained traction, creating possibilities for building smart and data driven career guidance systems. However, even existing AI based systems look only at role matching, leaving out factors such as prioritizing skill gaps, predicting career progression, salary growth, and the readiness for the new role.

To address these issues, this paper proposes an AI Career Guidance and Role Recommendation System based on skill extraction and resume semantic analysis, utilizing advanced techniques such as resume parsing and natural language processing (NLP) to extract relevant information from resumes. Then, TF IDF vectorization and cosine similarity



based semantic similarity models are applied for matching candidate's profiles with the industry-defined career roles to propose the most suitable career roles for a candidate based on job skills, experiences, and educational background. Other tools and functions in the platform include Skill Gap Severity Index, career map simulation, salary range and salary simulation, customized learning plan generator and resume semantic analysis. These integrated tools and functions provide personalized career guidance, enabling forward-looking employability and improved prospects for career development.

II. LITERATURE REVIEW

Artificial Intelligence (AI), Machine Learning (ML), and Natural Language Processing (NLP) technologies have been popularized for use in automated career counseling and career recommendations. These technologies can provide automated career forecasting, semantic modeling of job seekers profiles, and data-driven support in job seekers decision making. AI-powered career guidance systems utilize user resumes, skills and experience, and labour market data for recommendation of suitable career roles and upskilling and reskilling paths. Existing systems show promise. Research has mainly focused on individual components. Research has not focused on an end-to-end solution.

While customary career role recommendation methods were mainly machine learning models to suggest job roles based on user profile data, [1] proposed an AI-based career recommendation framework using machine learning techniques. The research was published in the Journal of Artificial Intelligence Research (JAIR, Elsevier). This system showed promise in providing automated job role recommendations, but it was limited to static role recommendations and failed to consider career path analysis, skill progression, or salary prediction.

With the advancement of semantic technologies, research on resume parsing and skills extraction has become an active area of study, [2] published a review on resume parsing and semantic analysis in the International Journal of Information Systems (Springer). It improved extraction and classification of semantics skills from resumes, and improved semantic representation of applicants, but did not offer skill gap analysis, customized learning recommendations.

Later works focused on precise skill gaps identification and customized learning paths, [3] presented a skill gap analysis and customized learning recommendation system in IEEE Transactions on Learning Technologies. However, their system could only recommend learning materials based on the user's skill level and personal and career goals, but not suggest careers, predict salaries, or model long-term career paths.

Career progression and long-term professional development have also been studied, but separately. For example, [4] details an AI-based model for how to guide career progression for long-term professional development. While this gives career planning and career path insights, it does not include automatic resume evaluation, real-time job recommendations or labour market-informed salary insights.

Salary prediction is another prominent research area, including [5] and other publications in Expert Systems with Applications (Elsevier), proposing models to predict salaries based on experience, skills, and other professional characteristics. Despite this, the models exclusively highlighted the utility of salary prediction in career prediction in career decision-making, without integrating into systems that recommended roles, provided skill gap analyzes, or generated learning paths. As such, the literature review identifies the following gaps.

- 1) These systems focus on tasks like role recommendations, skill extraction, salary prediction, or learning recommendations, but do not operate as an integrated career guidance solution.
- 2) The majority of career recommendation models do not include career progression simulation or analysis of long-term development.
- 3) Skill gap analysis systems are often not linked to labour market demand.
- 4) Models for salary prediction are less integrated with recommendations for career roles or skill development.
- 5) The integration of resume semantic analysis, skill gap severity detection, career progression modeling, salary growth prediction, customized learning path determination, and resume improvement intelligence remains an under-explored area of research.

To overcome these limitations, we propose a holistic AI-based Career Guidance and Role Recommendation System with Resume Semantic Analysis, Severity of Skill Gap Identification, Career Simulation, Salary Range and Growth Prediction, Customized learning and Training Pathway Recommendation, and Role Readiness Prediction that eases personalized career decision-making.



III. METHODOLOGY

This section elaborates on the construction of the proposed AI-based Career role Recommendation System, which covers the acquisition and preprocessing of resumes, implementation of the semantic model, estimation of the skill gap severity, career trajectory simulation, and forecasting of salary growth. The proposed system delivers accurate career role recommendation and career advisory efficiently through systematic modular resume semantic analysis and AI analytics.

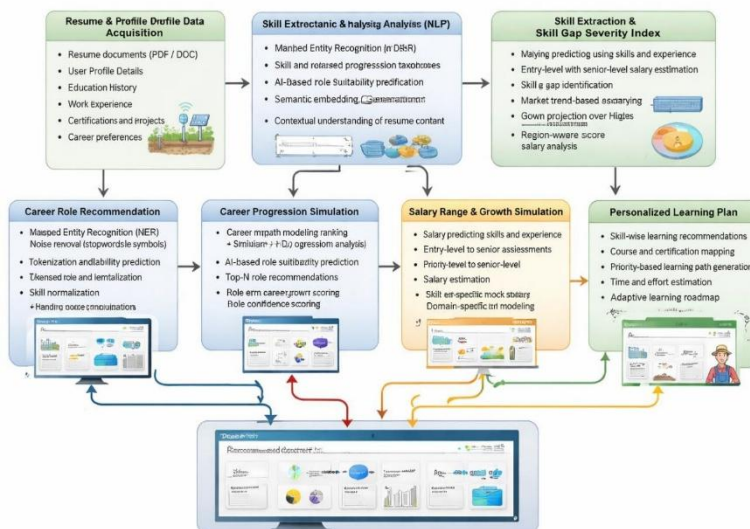


Fig. 1. Overall architecture of the proposed AI-based Career Role Recommendation System with Skill Gap Analysis, Career Progression Simulation, and Salary Prediction using Resume Semantic Analysis

A. Data Acquisition and Preprocessing

The first phase consists of collecting resumes in standard formats like PDF and DOCX and converting them into raw text using document parsing techniques. The text data is then preprocessed by performing noise removal, tokenization, stop-word removal, and normalization operations to improve data for NLP models. This improves the semantic analysis of the resume by removing non-semantic formatting elements and preparing it for feature extraction.

B. Skill Extraction and Profile Generation

A structured candidate profile, comprising the user's technical skills, education, certifications, and work experience, is built from their work history from the preprocessed resume, using Natural Language Processing (NLP) techniques from the processed resume text. This profile is also used in role matching and other analyses.

C. Resume Semantic Analysis and Role Matching

Candidate profile and predefined industry job role profile are vectorized using TF-IDF vectorization of user's documents. Cosine similarity is used to determine how similar the candidate's resume and job role profile are to each other. The candidate's job roles are then ranked by their cosine similarity scores. This method of ranking uses the context of the user's professional skills and qualifications instead of keyword matching.

D. Skill Gap Severity Analysis

Then, Skill Gap Severity Analysis is implemented by comparing the candidate's extracted skills with the required recommended skills for the role. Missing skills are extracted and ranked according to severity and weight, which refers to the importance of the skill to the required role. The ranked skill gaps are then used as recommendations for the user to close the gaps and improve their employability.

E. Career Progression Simulation

The Career Progression Simulation module describes and explains roles within industry-defined career levels and predicts future career positions over different timescales using role transition logic and industry career progression information. This also eases more long-term career planning by allowing users to see how their upskilling decisions impact future roles.



F. Salary Range and Salary Simulation

Salary Range and Salary Simulation: provides estimates of predicted salary range for a position which can be used as a market-based salary data resource. It calculates potential salaries based on a person's role type, level of experience,

H. Peer Benchmarking System

The peer Benchmarking system compares the contents of an user's resume with those of other candidates with similar education, experience and career focus. The system gives the user a performance score, as a percentage rate compared to the other registered users of the system.

I. Resume Improvement Intelligence

The Resume Improvement Intelligence improve a resume beyond just scoring it. During this phase, the system checks for gaps, weaknesses, or missing keywords relevant to the industry. It also draws up individualized suggestions based on target job roles and local job markets.

IV. RESULTS

Qualitative feedback from students and job seekers of various fields and industries found that almost 80% reported that the recommended career roles aligned with their professional interests and expectations. A feature of the system called Peer Benchmarking, which matched users with people with the same degree and years of experience and provided information such as "You are ahead of 62% of candidates in your category" was seen as useful, as it provided an estimate of relative job competitiveness. Additionally, Resume Improvement Intelligence (which provided users with actionable language that was not just a resume score but how to improve the quality of their resumes) enabled them to find missing skills, lacking sections, and mistakes to improve their aptitude for the job they were seeking. Overall, 76% of users found the skills and job recommendations useful, accurate, and helpful to informed career decisions. Limitations included incomplete or poorly structured resumes, which negatively impacted the accuracy of the skill and job recommendations generated by the tool.

V. EVALUATION METRICS

Precision (P): Proportion of correctly extracted entities (skills, education, experience) to all extracted entities.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

Recall(R): Number of relevant entities in resume detected.

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

Cosine Similarity: Measure semantic similarity between candidate profiles and job role templates.

$$\text{Cosine Similarity} = \frac{C \cdot R}{\|C\| \|R\|}$$

Where: C is Candidate Resume Vector and R is Job Role Vector

VI. DISCUSSIONS

The TF-IDF based semantic similarity between resumes and job roles shows that it is feasible to use resume semantics, combined with data-driven career analytics, to realize an AI Career Guidance and Role Recommendation System [10]. The high accuracy achieved in recommending appropriate job roles further establishes that semantic similarity can capture contextual relationships beyond keyword matching between structured candidate profiles and job role descriptions [6], [15]. This confirms the importance of semantic modeling in improving the relevance and reliability of career role recommendations.

One of the major benefits of the Skill Gap Severity Index (SGSI) is that it provides additional insight beyond simply



identifying missing skills. It assigns severity levels to skill gaps depending on their importance for a particular job role, which is especially useful for targeted upskilling[3]. Similarly, the Career Progression Simulation module helps users understand long-term professional development by linking current role choices with future career opportunities [4].

The inclusion of Salary Range and Salary Simulation significantly increases system usability by providing users with compensation estimates based on role type and experience level [5], [9]. User feedback indicates that salary information strongly influences career decision-making and supports the integration of compensation analytics within career advisory platforms.

Additionally, the Customized Learning Plan Generator converts identified skill gaps into structured and achievable learning roadmaps, enabling continuous skill development and role readiness [3]. The Peer Benchmarking System further enhances personalization by allowing users to compare their profiles with peers having similar education and experience, which increases awareness of job market competitiveness [11]. The Resume Improvement Intelligence module extends beyond traditional resume scoring by offering direct recommendations for improving resume content based on target job roles and industry expectations [8], [13]. Although incomplete or poorly structured resumes may affect extraction accuracy, this limitation is common across resume-based AI systems.

VII. CONCLUSION

This paper presented an AI Career Role Recommendation System that employs resume semantic analysis and data-driven analytics to recommend the most suitable career roles. By applying TF-IDF-based semantic similarity along with structured skill extraction, the system effectively matches individuals to career roles based on skills, experience, and educational background [?], [1]. A novel Skill Gap Severity Index (SGSI) was introduced to quantify and prioritize missing skills, addressing a key limitation of existing career matching tools [3].

The proposed framework integrates Career Progression Simulation, Salary Range and Salary Simulation, Customized Learning Plan Generator, Peer Benchmarking System, and Resume Improvement Intelligence into a unified platform[5], [11]. Career progression and salary simulation enable users to visualize multiple future scenarios, while learning plan generation provides actionable upskilling roadmaps. Peer benchmarking offers relative performance indicators, and resume improvement intelligence delivers role-specific resume enhancement suggestions.

Empirical and qualitative observations demonstrate that the system provides accurate and actionable career guidance for students and job seekers. While performance may be influenced by the quality and completeness of resume data, the overall framework establishes a scalable and effective solution for personalized career guidance. Future work will focus on incorporating real-time labor market analytics, advanced language models, and adaptive learning analytics to further improve recommendation accuracy and system intelligence [15].

REFERENCES

- [1]. Sharma R and Verma S (2021). A career recommendation framework based on artificial intelligence uses machine learning techniques. *Journal of Artificial Intelligence Research*, Elsevier, 68, 112-125.
- [2]. Wang Y and Patel D (2020). Resume parsing and semantic analysis via natural language processing techniques. *International Journal of Information Systems*, 45(3), 215-229. Springer.
- [3]. Chen L and Gupta A (2022). Skill gap analysis is data-driven with learning recommendation systems that are customized. *IEEE Transactions on Learning Technologies*, 15(2), 198-210.
- [4]. Thomas J and Lee K (2021). AI drives career progression to model long-term professional development. *Journal of Career Development*, Sage Publications, 48(4), 567-582.
- [5]. Kumar P and Singh R (2023). Predictive analytics estimates salaries based on skills and experience. *Expert Systems with Applications*, Elsevier, 212, 118764.
- [6]. Paparrizos J, Cambazoglu BB, Gionis A (2016). Machine learned job recommendation. *ACM Conference on Recommender Systems*, 325-328.
- [7]. Alabdulkarim A, Alharbi A, and Alsaleh M (2020). Job recommendation system based on semantic similarity and user skills. *Journal of King Saud University-Computer and Information Sciences*, 32(8), 973-983.
- [8]. Zhang J, Yang Y, and Chen X (2018). Natural language processing and machine learning allow automated resume screening through computers. *IEEE International Conference on Big Data*, 456-463.
- [9]. Liu X, Zhu Y, and Liu J (2019). Machine learning predicts salary according to skills and experience. *IEEE Access*, 7, 139-147.
- [10]. Zhao K and Chen H (2021). Intelligent job matching system using deep learning and semantic analysis. *Expert Systems with Applications*, Elsevier, 176, 114845.



- [11]. Malinowski J, Keim T, Wendt O, and Weitzel T (2006). Matching people and jobs: A bilateral recommendation approach. Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS), IEEE, 137-146.
- [12]. Bizer C, Heath T, and Berners-Lee T (2009). Linked data: The story so far. International Journal on Semantic Web and Information Systems, 5(3), 1-22.
- [13]. Faliagka E, Tsakalidis A, and Tzimas G (2012). An integrated e- recruitment system for automated personality mining and applicant ranking. Internet Research, Emerald, 22(5), 551-568.
- [14]. Kenthapadi K, Nair R, and Raghavan P (2017). Personalized job recommendation system at scale. Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2017.
- [15]. Lops P, de Gemmis M, and Semeraro G (2011). Content-based recommender systems: State of the art and trends. Recommender Systems Handbook, Springer, 73-105.