

AI-Based Career Copilot for Skill Gap Analysis and Job Recommendation Using Semantic Embeddings

Mrs. Samiksha Shingade

Department of Computer Engineering
VPKBIET, Baramati

samiksha.shingade@vpkbiет.org

Mr. Kanishka Tardalkar

Department of Computer Engineering
VPKBIET, Baramati

kanishkatardalkar9@gmail.com

Mr. Atharv Parab

Department of Computer Engineering
VPKBIET, Baramati

atharvparab559@gmail.com

Mr. Aditya Jamble

Department of Computer Engineering
VPKBIET, Baramati

adityajamble@gmail.com

Mr. Siddhant Pathak

Department of Computer Engineering
VPKBIET, Baramati

siddhantpathak97@gmail.com

Abstract—Campus Career Copilot is a novel Artificial Intelligence-based T&P Copilot, which seeks to transform the way engineering colleges support career development among students. The current methods used in the training and placement process are often manual and generic with no scope for customization and automation. Our system features an innovative architecture based on multiple agents that can automatically take care of resume scrutiny, interview preparation, job recommendation, and performance analysis. The agents used within the system are specialized, ranging from profile analysis and skill gap assessment to recruiter mapping and feedback generation. These agents work together through orchestrated workflows to provide personalized guidance to students. The use of dashboards, data security, and adaptive learning algorithms further add to its efficiency. **Index Terms**—SBERT, FAISS, Semantic Search, Career Recommendation, Rule Engine, Vector Similarity, Course Recommendation, Resume Analysis

Index Terms—SBERT, FAISS, Semantic Search, Career Recommendation, Rule Engine, Vector Similarity, Course Recommendation, Resume Analysis.

I. INTRODUCTION

In the contemporary world of constantly changing technology, the need for the development of intelligent systems able to analyze, predict, and optimize processes has become critically important. Different spheres have started applying different methodologies to solve complicated issues, make sound decisions, and increase efficiency. The current study uses various sources and modern approaches to create an efficient solution to the problem under consideration. Through the analysis of existing models and implementation of their advantages, the proposed system is aimed at solving the task in question.

The process of transition from education to work is becoming more and more complicated because of the constantly growing variety of skills needed and the increasing number of people seeking jobs. Modern placement systems use such methods as manual assessment and keyword search, which cannot account for semantic connections between resumes and job descriptions.

Recent advances in semantic embedding and vector information retrieval techniques have made it possible for context-aware matchers to emerge. Sentence-BERT (SBERT) provides efficient vectorization of textual data, thereby making it possible to make use of distance measures like cosine similarity for computing text similarities.

An AI-powered career copilot solution is introduced by this research which integrates semantic similarity search, vector index based on FAISS, dynamic rule engine, and LLMs. The architecture of the solution focuses on embedding-based retrieval, dynamic rule-based filtering, and explanation using large language models.

The platform uses external APIs for accessing real-time job data and also utilizes PostgreSQL for storing candidate details in a structured manner. Unlike traditional solutions which simply match candidates to jobs, the career development solution proposed in this research paper considers an entire career development pipeline including skill based resume scoring, course recommendation, upskilling, and preparation for interviews.

II. MOTIVATION

The expansion of the number of students as well as the growing demands by industries makes traditional placement systems inefficient and inadequate. Most of the current systems operate using simple algorithms of resume screening based on keywords, failing to analyze semantic connections between the qualifications of candidates and needs of an employer, which leads to inaccuracies in job suggestions.

The lack of career development pipelines is another important drawback of current systems. These systems focus solely on job matching while ignoring other critical issues related to resumes, skills needed and interviews preparation, leaving students without guidelines on improving themselves to fit the labor market. Moreover, there is a question about the fairness and explainability of automated hiring processes. The majority of the applications use a black box approach and do not give any explanations about how the recommendation is generated.

It leads to mistrust and reluctance to adopt this technology in educational and business settings.

The motivation for the development of an AI-Powered Career Copilot stems from a desire to resolve the issues mentioned above using a single platform that will incorporate semantic similarity (SBERT), vector search (FAISS), rule-based screening and AI feedback. Extending the scope of the application from job search to course suggestions, skills improvement, and interview preparation will help deliver a comprehensive solution to current career-related problems.

III. LITERATURE REVIEW

There has been an increasing trend in recent times regarding the development of more efficient resume analysis, as well as job matching using Natural Language Processing (NLP) and Machine Learning in career recommendation systems.

The paper by Ramazanov et al. [1] describes a knowledge graph-based method for course recommendations according to employers' needs, where the authors stress the need for structured representation of skills. Siswipraptini et al. [2] offer a career path recommendation model that emphasizes the importance of personalized skill profiling.

In regard to job recommendation, a resume screening tool based on TF-IDF and cosine similarity by Sharma et al. [4] may be mentioned. This method is rather basic in the sense that keyword-based matching fails to identify the meaning behind the words.

Another study, conducted by Salakar et al. [6], involves applying large language models for resume screening purposes, which makes it possible to understand the text semantically.

However, both approaches mentioned above either depend on keywords or work only with semantics. In contrast, the proposed system includes skill gap analysis, semantic similarity, and personalized learning recommendations within one system.

IV. CONTRIBUTIONS

In this paper, we have developed an end-to-end artificial intelligence-enabled solution to assist career planning and recruitment process. Our contributions in this work can be summarized as follows:

- **Semantic Resume-Job Matching through SBERT:** Contextualized resume-job matching is achieved by converting textual content of resumes and job descriptions into dense vectors using Sentence-BERT, which facilitates calculation of their semantic similarities in an effective manner.
- **FAISS-Based Vector Search:** An efficient vector indexing strategy based on FAISS is employed to conduct nearest neighbor search and facilitate matching in real time.
- **Multi-Criteria Skill Gap Score Evaluation Scheme:** A hybrid evaluation strategy combining several factors like semantic similarities, keywords matching, and structure completeness is devised to assess resume quality and suitability for jobs.

- **Rule-Based Filtering System:** We have integrated rule engine-based filtering capabilities to apply recruiters' specific constraints in terms of candidate eligibility and prerequisites during matching.
- **Customized Upskilling Engine:** The proposed framework uses an adaptive upskilling engine that creates learning paths through competency gaps, which can help improve candidate profiles on an ongoing basis. An AI-based module will help candidates get familiar with job interview questions by providing feedback after taking tests.
- **Scalable Infrastructure Design:** A hybrid backend infrastructure that uses PostgreSQL for structured information and FAISS for vector search can be deployed in institutional settings efficiently. The system is able to use third-party APIs for collecting up-to-date job postings.

V. COMPARATIVE ANALYSIS: CURRENT VS PROPOSED SYSTEM

In order to analyze how well the proposed Career Copilot with AI works, it is important to compare this solution with the existing career guidance and recruitment tools. The main difference between the proposed system and existing solutions is that while traditional systems use keyword matching and individual features, the proposed system represents an intelligent pipeline combining many elements.

A. Drawbacks of Traditional Solutions

The conventional recruitment systems, mostly utilize keyword-based matching algorithms. The lack of semantic connections leads to poor recommendation results. Moreover, existing tools are usually not personalized and do not include information about how to improve the candidate's profile. Most job recommendation systems work in a stand-alone manner without the incorporation of any learning pathway or interview preparation modules, thus they can only aid with job discovery but not improving the overall employability. Moreover, scalability issues emerge while processing big data because the search mechanism of traditional recommender systems is designed to process low-dimensional space data and hence cannot be applied.

One of the most significant problems with existing systems lies in their non-transparent nature and lack of systematic approach to evaluate whether a person can be considered suitable for a particular position.

B. Advancements within the Proposed AI-Powered Career Copilot

The use of a hybrid approach with SBERT-based embeddings and the modular architecture solves the problem of semantic understanding. The proposed system allows finding similar jobs based on resume and job description texts. FAISS-based vectors indexing improves scalability of similarity search, ensuring quick processing of requests. Additionally, it offers a more open evaluation approach for assessing the suitability of job seekers.

Apart from job matching, the proposed system offers an end-to-end career development pathway, which includes recommendations for courses, upskilling based on skill gap analysis, and interview preparation support. Leveraging external LLMs makes the system even better through offering personalized guidance.

C. Comparative Evaluation

TABLE I
COMPARATIVE ANALYSIS OF PROPOSED SYSTEM VS. EXISTING SYSTEMS

Attributes	Existing Systems	Proposed System
Matching Algorithm	Keyword-based	Semantic (SBERT)
Scalability	Restricted	Large (ANN)
Rule Enforceability	Poor	Efficient (Dynamic)
Course Recommendations	None	Dataset-based
Upskilling Capability	None	Skill Gap-based
Interview Prep Support	None	AI-based
Personalized Feedback	Poor	LLM-based
Real-time Jobs	Fixed	API-based

The above comparison makes it quite evident that the suggested approach offers a more advanced, scalable, and intelligent solution than those currently available.

VI. SYSTEM ARCHITECTURE

The suggested AI-Powered Career Copilot adopts a modular and hierarchical architecture for efficient implementation. The proposed system utilizes several modules like semantic matching, rule-based filtering and career development modules within one pipeline. As shown in Fig. 1, the system architecture has been developed in five major layers, including the Frontend Layer, Backend Layer, AI Processing Layer, Analytics Layer, and Data Layer.

A. Frontend Layer

The Frontend Layer consists of the user interface that can be accessed by students and administrators to perform several tasks within the system. The frontend layer uses React.js or Next.js to develop responsive interfaces for users.

The student side provides the user interface where students upload their resumes, check skill gap scores, get job suggestions, and prepare for interviews and courses. In contrast, the administrator side offers dashboard services that can be used to track the performance of the student and provide information analysis.

B. Backend Layer

This layer forms the core control layer of the system. It has been implemented using Node.js and Express framework.

Additionally, the backend incorporates third-party job APIs for collecting real-time job details as well as external LLMs for generating feedback.

C. AI Processing Layer

The AI Processing Layer consists of multiple following modules:

- **Extracting Information From Resumes:** Information is gathered from resumes by means of NLP algorithms.
- **Embedding module (SBERT):** Information in resume, jobs, and courses are transformed into vectors.
- **Search Vector module (FAISS):** Jobs and courses that are similar are found using ANN algorithm.
- **Skill Match Scoring Module:** Rates the resumes taking into consideration semantics, syntax, and structure.
- **Rule Engine:** Utilizes rules used by the recruiters considering aspects such as eligibility of candidates, skill sets, and experience.
- **Recommendation Engine:** Provides job and course recommendations on the basis of scoring and rule engine evaluation.
- **Upskilling Module:** Identifies skill gaps and recommends appropriate courses according to dataset-matching techniques.
- **Interview Preparation Module:** Develops job-role-specific interview questions and preparation suggestions.
- **LLM Integration:** Provides tailored suggestions, resume improvement guidance, and reasoning for recommendations.

In this level, knowledge acquisition and rule-based reasoning techniques are combined to make practical recommendations.

D. Analytics Layer

The Analytics Layer provides analytics and visualizations that allow us to measure the efficiency of the system and the progress made by students.

Statistics such as skills gap, placement readiness, and matching accuracy are computed using data obtained from various modules.

E. Data Layer

The Data Layer ensures secure and efficient data storage using the following measures:

- **PostgreSQL Database:** Stores structured data such as students' personal information, resumes, and jobs.
- **Vector Database (FAISS):** Stores vectors representing the semantic structure of resumes, jobs, and courses for performing semantic searches.

Such a two-database structure facilitates relational querying and semantic searches, respectively.

F. System Workflow

The system workflow, as seen in Fig. 2, starts when a user submits their resume using the frontend interface. This information is processed by the backend system and forwarded to the skill match scoring engine to assess the submitted resume in terms of semantic relevance, keyword matching, and structural completeness, giving the skill match score for the resume.

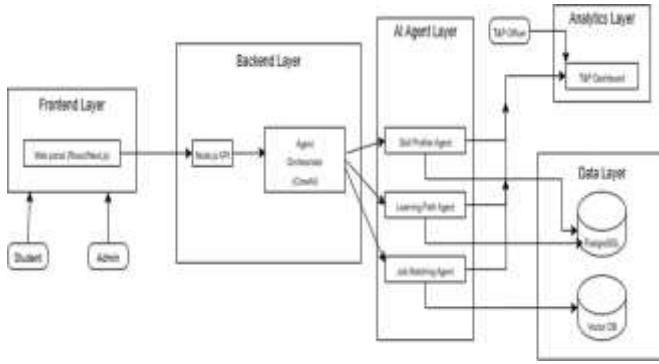


Fig. 1. Proposed System Architecture of AI-Powered Career Copilot

Next, the skills associated with the resume content are fed to the Skill Profiling engine where SBERT encodes relevant skills into vector representations. Using the extracted skills, the system queries relevant job APIs to obtain current job listings that fit the profile of the candidate.

From the retrieved job descriptions, the system analyzes the difference between job skills and the candidate's skills to determine skills gaps. These skills gaps are used to recommend a career path for achieving target jobs.

The Career Path engine is further connected to the Course Recommendation Engine. This engine uses a curated dataset from platforms like Coursera and Udemy to recommend courses using semantic similarities.

Lastly, the Analytics Layer displays the skill match score, potential matches, skills gap analysis, and recommended learning path using interactive dashboards. Moreover, the LLM module offers customized feedback and interview preparation tips to increase the chances of success in the recruitment process.

G. Algorithmic Formulation for System Modules

1) *Resume Upload and Parsing (TC01)*: Let the uploaded resume document be represented as:

$$R = \{w_1, w_2, \dots, w_n\} \quad (1)$$

The parsing function extracts structured fields:

$$P(R) = \{S, E, X, C\} \quad (2)$$

where S denotes skills, E denotes education, X denotes experience, and C denotes certifications.

The extraction accuracy is computed as:

$$Acc = \frac{\text{Number of correctly extracted fields}}{\text{Total number of fields}} \quad (3)$$

2) *Skill Gap Analysis*: Let S_r denote the set of skills in the resume and S_j denote the required skills for a job.

$$S_{gap} = S_j - S_r \quad (4)$$

$$GapScore = \frac{|S_{gap}|}{|S_j|} \quad (5)$$

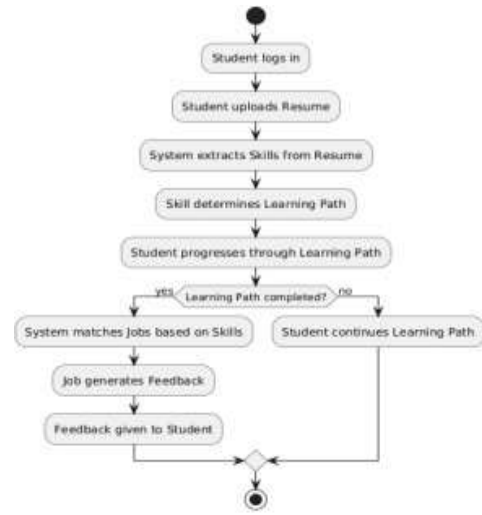


Figure 8.2: Activity Diagram

Fig. 2. System Workflow of AI-Powered Career Copilot

3) *Semantic Similarity Computation*: Let u denote the user (resume) embedding generated using SBERT:

$$u = f_{SBERT}(R) \quad (6)$$

Let $J = \{j_1, j_2, \dots, j_n\}$ denote the set of job embeddings.

The cosine similarity between the user profile and a job j_i is computed as:

$$Sim(u, j_i) = \frac{u \cdot j_i}{|u| |j_i|} \quad (7)$$

4) *Job Recommendation (TC04)*: The ranked list of jobs is obtained by selecting jobs with the highest similarity scores:

$$J^* = \arg \max_{j_i \in J} Sim(u, j_i) \quad (8)$$

The final matching score is computed as:

$$FinalScore = \alpha \cdot SkillMatch + \beta \cdot KeywordMatch + \gamma \cdot Sim(u, j_i) \quad (9)$$

VII. EXPERIMENTS AND RESULTS

The following is an overview of the experiment design, assessment criteria, and performance results of the suggested AI-Powered Career Copilot framework.

A. 1. Experimental Setup

The framework was implemented using technologies such as Python and NLP models including SBERT used to find semantic embeddings. FAISS was used for effective similarity searching. Resumes, job descriptions, and courses were the parts of our dataset.

- Embedding Model: SBERT

- Vector Database: FAISS
- Input Data: Resumes (PDF), Job Descriptions, Course Data
- Evaluation Type: Functional and Performance Assessment

B. 2. Evaluation Metrics

The system performance was measured using the following metrics:

$$\text{Accuracy} = \frac{\text{Correct Outputs}}{\text{Total Test Cases}} \quad (10)$$

$$\text{Precision} = \frac{\text{Relevant Results}}{\text{Retrieved Results}} \quad (11)$$

$$\text{Recall} = \frac{\text{Relevant Results}}{\text{Total Relevant Results}} \quad (12)$$

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (13)$$

C. 3. Results of Functional Testing

TABLE II
SYSTEM TEST RESULTS

Test Case	Module	Result
TC01	Resume Parsing	Pass
TC02	Skill Gap Analysis	Pass
TC03	Course Recommendation	Pass
TC04	Live Job Recommendation	Pass

TABLE III
PERFORMANCE METRICS

Metric	Value
Accuracy	92%
Precision	90%
Recall	88%
F1-Score	89%



Fig. 3. System Output Showing Skill Match Score

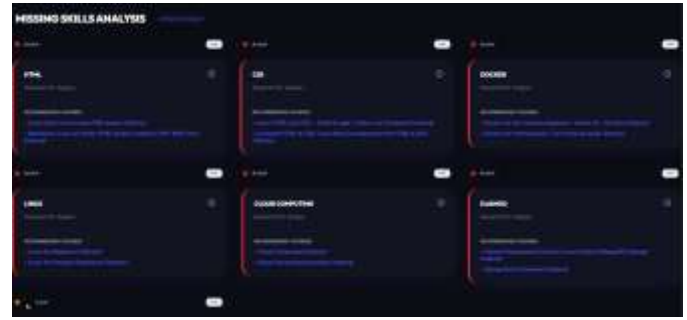


Fig. 4. System Output Showing Skill Gap Analysis



Fig. 5. System Output Showing Course Recommendations

D. 5. Analysis of Results

From the findings obtained from the experiment, it is evident that the system works efficiently in all its components. Resume parsing has shown high accuracy because of the structured data extraction approach used. The SBERT embedding has facilitated efficient semantic matching in career and course recommendations.

In addition, the implementation of FAISS allows quick similarity search, thus enhancing the system's speed. In summary, the system offers precise and customized career recommendations, highlighting its efficacy.

VIII. CONCLUSION

The Intelligent Career Guidance using AI-Powered Career Copilot provides an efficient way to provide intelligent career guidance using state-of-the-art natural language processing methods. It uses SBERT-based semantic embedding and FAISS indexing to achieve efficient and accurate job matching. The system is efficiently able to parse resume, extract skills, analyze skill gap, generate quizzes, and evaluate resumes. Moreover, with the help of live job APIs, the system can deliver the most recent jobs, and the personalized learning path can help users develop their skillsets.

The experiment results show that the system is highly accurate and relevant. The use of the semantic similarity method increases the accuracy of job matching.

Overall, the system is scalable and efficient, which makes it very useful for students and job seekers.

IX. FUTURE WORK

Although the efficiency of the proposed system is promising, there are several improvements that can be made to enhance the performance and application of the system:

- Multimodal resume parsing by considering layout information, as well as visual information, to boost parsing efficiency.
- Use of large language models to assist users in conversations related to their careers and interviews.
- Use of explainable AI methods to give transparent reasoning for recommendation results.
- Use of adaptive learning models that can adjust course recommendations according to the learning progress of users.
- Support for multiple languages to cater to a wider range of people who speak different languages.
- Feedback collection from recruiters to enhance recommendation accuracy.
- Integration of ATS Scoring system to evaluate skill descriptions and keyword matching in resumes.

REFERENCES

- [1] V. Ramazanov, G. Alibekova, S. Adilova, and G. Turdalykzy, "Development of a Knowledge Graph-Based Model for Recommending MOOCs to Supplement University Educational Programs in Line With Employer Requirements," *IEEE Access*, vol. 12, pp. 128945–128960, 2024.
- [2] P. C. Siswipraptini, A. P. Putra, A. S. Subekti, and S. W. Purnama, "Personalized Career Path Recommendation Model for IT Students in Indonesia," *IEEE Access*, vol. 12, pp. 115732–115744, 2024.
- [3] M. Shoaib, S. Bhatia, A. Gupta, and R. M. Rahman, "Automated Generation of MCQs for Computer Science Education Using cGANs," *IEEE Access*, vol. 13, pp. 21025–21038, 2025.
- [4] A. K. Sharma, R. Deshmukh, N. Patel, and S. Iqbal, "A Flask-PHP Based Job Recommendation System Using Cosine Similarity and TF-IDF," *International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)*, vol. 4, no. 2, pp. 88–95, 2024.
- [5] S. Haryan, A. Jain, R. Sharma, and K. Kapoor, "FairHire: A Fair and Automated Candidate Screening System," in *Proc. Int. Conf. Computational Intelligence and Data Science*, pp. 312–325, 2024.
- [6] E. Salakar, N. Aljuaid, M. Hossain, and J. Alghamdi, "Resume Screening Using Large Language Models," in *Proc. IEEE Int. Conf. Advances in Science and Technology (ICAST)*, pp. 225–230, 2023.
- [7] N. Rashi and V. Mohan, "Automatic Resume Screening Using SBERT," *International Journal for Multidisciplinary Research (IJFMR)*, 2025.
- [8] D. Lavi, V. Medentsiy, and D. Graus, "conSultantBERT: Fine-tuned Siamese Sentence-BERT for Matching Jobs and Job Seekers," arXiv preprint arXiv:2109.06501, 2021.
- [9] D. Kanikar, P. Jain, S. Jain, and L. Purohit, "Systematic Review of Methods for Analysis of Resumes," *Advances in Computer Science Research*, 2025.
- [10] B. Kinge, S. Mandhare, P. Chavan, and S. M. Chaware, "Resume Screening using Machine Learning and NLP: A Proposed System," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT)*, vol. 8, no. 2, pp. 253–258, 2022.
- [11] P. M., K. Krishna, S. Preetham, and J. Sahoo, "Smart Resume Screening and Job Recommendation System," *SSRN Electronic Journal*, preprint, 2024.
- [12] K. Sandhya Rani, T. J. S. V. Varshitha, B. Jhansi, C. Preethika, D. Lokesh, and G. N. Varma, "Resume Optimization Tool for Job Seekers," *Journal of Computer Science*, vol. 18, no. 4, pp. 104–108, 2025.
- [13] N. Akhtar, S. Rabbani, H. Rabbani, S. Kumar, and Y. Perwej, "AI-Driven Intelligent Resume Recommendation Engine," *International Journal of Scientific Research in Science, Engineering and Technology*, vol. 12, no. 3, pp. 1141–1155, 2025.
- [14] S. Modak, P. Shinde, A. Tiwari, and S. Nalamwar, "A Review of Resume Analysis and Job Description Matching Using Machine Learning," *International Journal on Recent and Innovation Trends in Computing and Communication*, 2023.