

Smart Career Path Recommender

Project Proposal for Information Storage and Retrieval

Name	NetID
Ubaid Khan Mohammed	mohammed.u
Manoj Siva Sai Deepank	deeps45
Yaswanth Yaradoddi	yaswanthreddy

October 18, 2025

Abstract

The **Smart Career Path Recommender** is an intelligent, data-driven system designed to bridge the gap between individual skill profiles, learning resources, and real-world job opportunities. Unlike existing learning or job platforms that function in isolation, our tool integrates both domains, providing personalized and actionable recommendations. By leveraging modern Information Retrieval (IR) and Natural Language Processing (NLP) techniques, it identifies relevant job roles, skill gaps, and courses that align with user aspirations. The project aims to demonstrate how semantic text analysis and embedding-based similarity models can power next-generation educational and career recommendation systems.

1. Introduction

In the modern digital economy, individuals are inundated with an overwhelming number of career choices and online learning options. However, existing recommendation systems, such as those found in Coursera, LinkedIn Learning, or Indeed, typically operate in silos. Learning platforms focus on suggesting popular courses, while job portals emphasize matching resumes to postings. This fragmented approach fails to provide holistic career guidance that connects education, skills, and professional goals.

Our proposed tool, **Smart Career Path Recommender**, addresses this gap by creating a unified recommendation framework that personalizes both learning and career pathways. By integrating open datasets from MOOCs, job descriptions, and skill taxonomies, and combining them with NLP-based semantic similarity models, we generate context-aware suggestions. The system thereby helps users identify which skills to develop next and which opportunities best align with their evolving profile.

2. Function of the Tool

The core function of the tool is to map users' skill sets and interests to relevant career trajectories, recommended courses, and potential job roles. Given user input such as educational background, skill list, and career goals, the system:

- Recommends matching job roles and corresponding industry requirements.
- Identifies specific skill gaps and suggests targeted learning resources.

- Provides examples of similar career transitions or real-world role models.

This functionality combines IR-based retrieval with semantic embeddings, enabling more intelligent and personalized recommendations than traditional keyword-based systems.

3. Need and Target Audience

Choosing the right career path is a significant challenge for students and professionals, especially with the rapid evolution of technology and job markets. The recommender tool is designed to serve:

- **Students** exploring suitable domains based on their interests and skill proficiency.
- **Professionals** seeking upskilling paths for a career switch (e.g., software engineering to data science).
- **Academic advisors and institutions** wanting to guide learners with objective, data-driven insights.

By aggregating and organizing massive information sources, the system provides a structured, transparent, and personalized learning-to-career pathway.

4. Existing Tools, Uniqueness, and Challenges

Current systems like LinkedIn Learning or Coursera suggest popular courses but lack an integrated understanding of user-specific goals and job market dynamics. Similarly, job platforms like Glassdoor or Indeed do not suggest how users can acquire missing skills.

Our tool stands apart by:

- Combining learning and employment data in a single recommendation framework.
- Using semantic text embeddings to understand contextual meaning rather than relying on keyword overlap.
- Generating interpretable outputs such as “career maps” and “learning bridges.”

Challenges include:

- Aligning heterogeneous data from different sources and formats.
- Ensuring semantic consistency across diverse terminology and job descriptions.
- Handling data sparsity for users with limited profiles or non-standard skills.

5. System Design and Algorithmic Approach

Data Sources

The system will use multiple open data repositories:

- **Kaggle job-skill datasets** containing role descriptions and required skills (kaggle-dataset).
- **MOOC metadata** (Coursera, edX) with course names, categories, and objectives. (coursera-dataset).
- **O*NET and Open Skills Project** skill taxonomies for standardization.

Core Algorithm

1. Convert text data (user profile, jobs, courses) into dense vector embeddings using a pre-trained transformer model (all-MiniLM-L6-v2).
2. Compute cosine similarity between user and resource embeddings.
3. Rank results and recommend top- k roles or learning paths.
4. Apply clustering or hybrid filtering to refine recommendations.

This approach mirrors IR principles of vector-space modeling and relevance ranking, enabling efficient and explainable retrieval.

6. Implementation Resources

Technologies and Tools:

- NLP libraries: HuggingFace, SentenceTransformers.
- Machine learning and data tools: scikit-learn, pandas, faiss.
- Web interface: Flask for visualization and live demo.

7. Evaluation and Demonstration

The prototype will be demonstrated using realistic test cases that simulate different user profiles. Its usefulness will be evaluated through:

- **Quantitative metrics:** Precision@K and Recall@K for retrieval accuracy.
- **Qualitative evaluation:** User satisfaction and relevance feedback.
- **Visualization:** Interactive dashboard showing ranked results, similarity graphs, and career trajectories.

8. Timeline and Milestones

Week	Milestone
1	Data collection and cleaning (jobs, skills, courses).
2	Text preprocessing and embedding generation.
3	Implementation of similarity and ranking pipeline.
4	Integration of recommender system and visualization.
5	Dashboard development and testing with case studies.
6	Final evaluation, documentation, and presentation.

Outcome: A fully functional, data-driven prototype showcasing how NLP and IR principles can transform career guidance into a personalized, intelligent, and scalable system.