

SkillSense — AI-Driven Skill and Career Recommendation System

Overview

SkillSense is an intelligent recommendation engine that helps learners and professionals find the most relevant courses and jobs based on their existing skills and interests.

It combines machine learning-based text embeddings with semantic similarity search to match users to opportunities.

Core Components

1. Data Sources

- Coursera, edX, Skillshare, and Udemy course data.
- Job listings from multiple domains.
- Each dataset includes attributes like course title, institution, skills, duration, and level.

2. Preprocessing

- Course data is cleaned and unified into a common structure.
- Skills are extracted, formatted, and standardized.
- Missing attributes such as duration or rating are filled with defaults.

3. Embedding Generation

- Uses **SentenceTransformer: all-MiniLM-L6-v2** to convert job and course descriptions into numerical vectors.
- Embeddings capture semantic meaning, enabling similarity-based matching.

4. Vector Indexing

- FAISS (Facebook AI Similarity Search) indexes the embeddings.
 - Separate indices are created for jobs and courses.
 - Enables millisecond-level retrieval of top matches.
-

Recommendation Logic

When a user enters a skill or interest (e.g. “I know Networks”):

1. The input is embedded using the same model.
2. FAISS searches for top-N most similar jobs and courses.
3. The system returns ranked lists of:
 - **Recommended Jobs** — matched by required skills and description similarity.
 - **Recommended Courses** — matched by course content and skill tags.

Example Output

Input:

“I know Networks”

Recommended Jobs:

1. Network Engineer — Skills: Networking, Cisco, Linux
2. Systems Administrator — Skills: Networks, Infrastructure, Security
3. Cloud Support Associate — Skills: AWS, Networking, Virtualization

Recommended Courses:

1. “Computer Networks” (Coursera) — Skills: Networking, TCP/IP, Routing
2. “Introduction to Network Security” (edX) — Skills: Firewalls, VPN, Security Basics

Technologies Used

Component	Description
Pandas	Data loading and preprocessing
SentenceTransformers	Text embeddings using MiniLM
FAISS	Vector similarity search
Python	Core programming language
Google Colab	Development environment

Key Features

- Multi-platform course integration (Coursera, edX, Skillshare, Udemy)
 - Semantic job and course matching
 - Reusable and scalable architecture
 - Personalized recommendations
 - Lightweight and deployable on local systems
-

Future Scope

- Integration with LLMs for conversational recommendations
- Real-time resume parsing
- Adaptive learning path generation
- Expanded dataset integration