

SHL Generative AI Internship Assignment – Approach Summary

To tackle SHL's challenge of building an intelligent assessment recommendation system, I developed a full-stack semantic search application that takes in job descriptions or text queries and returns the most relevant SHL assessments. The project includes data extraction, embedding-based semantic matching, an API service, an interactive UI, and evaluation metrics, all built from scratch within a short time frame.

The first step involved **data scraping** from SHL's online catalog. I extracted structured details such as assessment names, descriptions, durations, and skills tested using Python (BeautifulSoup and Requests). This formed the base dataset for the recommender engine. Since the public SHL site lists only around 24 items per page, one of the ongoing challenges is automating pagination to extract a broader range of assessments. I plan to address this immediately after submission to scale the model further.

Next, I implemented **semantic similarity** using the SentenceTransformers library. I chose the all-MiniLM-L6-v2 model for its efficiency and accuracy. I generated vector embeddings for all catalog assessments and stored them in memory. For every input query, I computed its embedding and retrieved the top 3 most similar assessments based on cosine similarity.

I built a **FastAPI backend** with a /recommend endpoint that accepts a query string and returns a ranked list of matching assessments in JSON format. This API enables seamless integration with external systems or frontends. Alongside, I developed a simple yet effective **Streamlit-based frontend**, allowing recruiters or users to test the system interactively by entering queries and viewing recommendations instantly.

To assess performance, I created a test set with queries and expected outputs and implemented an **evaluation module** that calculates Recall@3 and MAP@3. Though the current scores are low due to limited catalog size and lack of semantic diversity, the system framework is solid and ready for scaling once the full data is integrated.

Despite time constraints, I successfully brought together **data processing, semantic search, API design, frontend deployment, and evaluation**. The biggest challenge I am currently resolving is increasing the number of assessments through deeper scraping and filtering to improve relevance. I am also working on better normalization of assessment names to handle duplication and improve model accuracy.

This project showcases my end-to-end capability in building AI-powered tools using modern NLP, full-stack development, and agile problem-solving — all of which I'm excited to bring to SHL.