SHL Assessment Recommendation Engine

Solution Overview

Introduction

Hiring requirements are typically expressed in unstructured natural language, while assessment catalogues are structured and domain-specific. This mismatch creates a semantic gap that makes it difficult to automatically map job descriptions to suitable assessments. Job roles may use abstract or informal terminology that does not directly correspond to assessment names or categories.

This project proposes a **Semantic Assessment Recommendation Engine** that recommends relevant SHL assessments for a given job description. The system leverages Natural Language Processing (NLP) and similarity-based retrieval to provide accurate, interpretable, and scalable recommendations aligned with hiring intent.

Problem Statement

The objective is to recommend the **top-K most relevant SHL assessments** for a given job description without relying on labeled training data.

Challenges

- Semantic mismatch between job descriptions and assessment metadata
- Ambiguity in skill terminology
- Need for efficient retrieval across a large catalogue

The system must remain explainable and suitable for real-world recruitment workflows.

Data Description

Two datasets are used:

- **Assessment Catalogue**: Assessment name, type (Aptitude, Technical, Behavioral), skills tested, description, and difficulty level.
- **Job Role Dataset**: Role title, role description, and required skills.

These datasets simulate real hiring scenarios and SHL product information.

System Architecture

The recommendation engine follows a semantic retrieval pipeline:

Job Description → Text Preprocessing → Feature Extraction → Similarity Scoring → Ranking → Top-K Recommendations

The modular architecture supports scalability and future model enhancements.

Methodology

Text from job descriptions and assessment metadata is preprocessed using lowercasing, stopword removal, and basic normalization. Identical preprocessing is applied across inputs to ensure consistency.

Feature extraction is performed using **TF-IDF vectorization**, which converts text into numerical vectors while capturing term importance. Both job descriptions and assessments are represented in a shared feature space.

Cosine similarity is used to compute semantic relevance between job descriptions and assessment vectors. Assessments are ranked by similarity score, and the top-K results are returned as recommendations.

Evaluation Strategy

To evaluate recommendation quality, a manually curated ground-truth mapping between job roles and relevant assessments is used.

Metrics

- **Precision@K**: Measures the relevance of recommended assessments
- **Recall@K**: Measures coverage of relevant assessment

Results show that semantic similarity-based retrieval provides meaningful recommendations even without supervised training data.

Technology Stack

- **Language**: Python
- **Libraries**: Pandas, NumPy, Scikit-learn
- **NLP Techniques**: TF-IDF, Cosine Similarity
- **Demo Interface**: Streamlit (optional)

The lightweight stack ensures reproducibility and ease of deployment.

Results, Limitations, and Future Work

The system effectively maps abstract job descriptions to relevant assessments and outperforms keyword-based matching. However, TF-IDF does not fully capture deep contextual semantics, and evaluation data is limited in size.

Future enhancements include transformer-based embeddings (Sentence-BERT), learning-to-rank models, and feedback-driven recommendation refinement.

Conclusion

This project demonstrates a practical and scalable **Assessment Recommendation Engine** aligned with SHL's hiring solutions. By leveraging NLP and similarity learning, the system bridges the semantic gap between job descriptions and assessment products, enabling accurate and interpretable assessment recommendations.