Resume Evaluation Tool for Job Seekers

Emilija Milanovic, Changchen Yu

Job hunting in the tech and IT industries can be challenging. With numerous applicants competing for the same roles, recruiters often rely on software to quickly scan resumes and filter out candidates based on keywords that match the job description. This process means that even highly qualified candidates may be overlooked if their resumes are not tailored to the job posting. For this reason, we created a resume evaluation app to help job seekers assess how well their resumes align with a job description. The tool analyzes skills, verifies aspects such as education level, and suggests additional skills that could improve the resume's relevance.

We designed the app having as its primary target university students and recent graduates who are applying for entry-level positions like internships or new grad roles in the engineering and technology fields; we also believe that this tool can be useful for people looking to refine their resumes and, generally speaking, act as a flexible and scalable solution for resume optimization.

The project is implemented in Python and is divided into four main components: skill extraction, PDF parsing, skill matching, and an interactive interface. The user interface allows for the upload of the CV, it then parses the resume in PDF format and evaluates whether the education level meets job requirements. To accomplish this, we built a PDF parser that reads and organizes resume content into sections, such as education, skills, and experience.

Skill extraction is achieved through NLP-based models that we fine-tuned for a Named Entity Recognition (NER) task, which consists in identifying and classifying specific entities such as places, organizations, and in our case: skills, within unstructured text. During the project, we observed that combining the results of our fine-tuned model with the functionality provided by the SkillNer Python library, which uses a pre-defined database of skills and keyword matching techniques, greatly improved the coverage in identifying relevant skills.

For skill matching, we initially used FastText embeddings to measure semantic similarity between skills extracted from resumes and those in job descriptions. However, simply loading FastText proved to be computationally expensive and it also struggled with multi-word phrases, such as "machine learning" or "data analysis," treating them as separate words. To address these limitations, we switched to other SpaCy pre-trained model, en_core_web_lg, which generates pre-trained embeddings for both single words and phrases. Thanks to this change, we were able to both simplify implementation and improve accuracy and processing speed.

While implementing the skill matching algorithm, we encountered performance challenges related to cosine similarity calculations. Initially, we used nested loops to compare each skill in the resume with every skill from the job description, which turned out to be fairly inefficient

and time-consuming for larger sets of extracted skills. To improve the process of searching for similar vectors, we integrated FAISS (Facebook AI Similarity Search), which is a library designed to efficiently perform nearest-neighbour search on dense vectors and is optimized for, among many, high-dimensional data and batch processing of embedding. Thanks to this, we were able to significantly reduce computation time, and since the skill matching process is the most important part program, improving it was a significant step in resolving bottlenecks to speed and performance.

Finally, the interface is implemented in Streamlit and returns the matching skills and those missing from the CV from the job description.

Our project focuses on jobs in the tech field, but nothing prevents it from being applied to other sectors as well! Indeed, fine tuning the NER model is not too computationally expensive and can be tailored to extract skills specific to any industry, which makes the application versatile and applicable to a wide range of fields and support resume optimization and career development across various sectors in the future.