

Task: Resume Matching with Job Descriptions Using PDF CVs

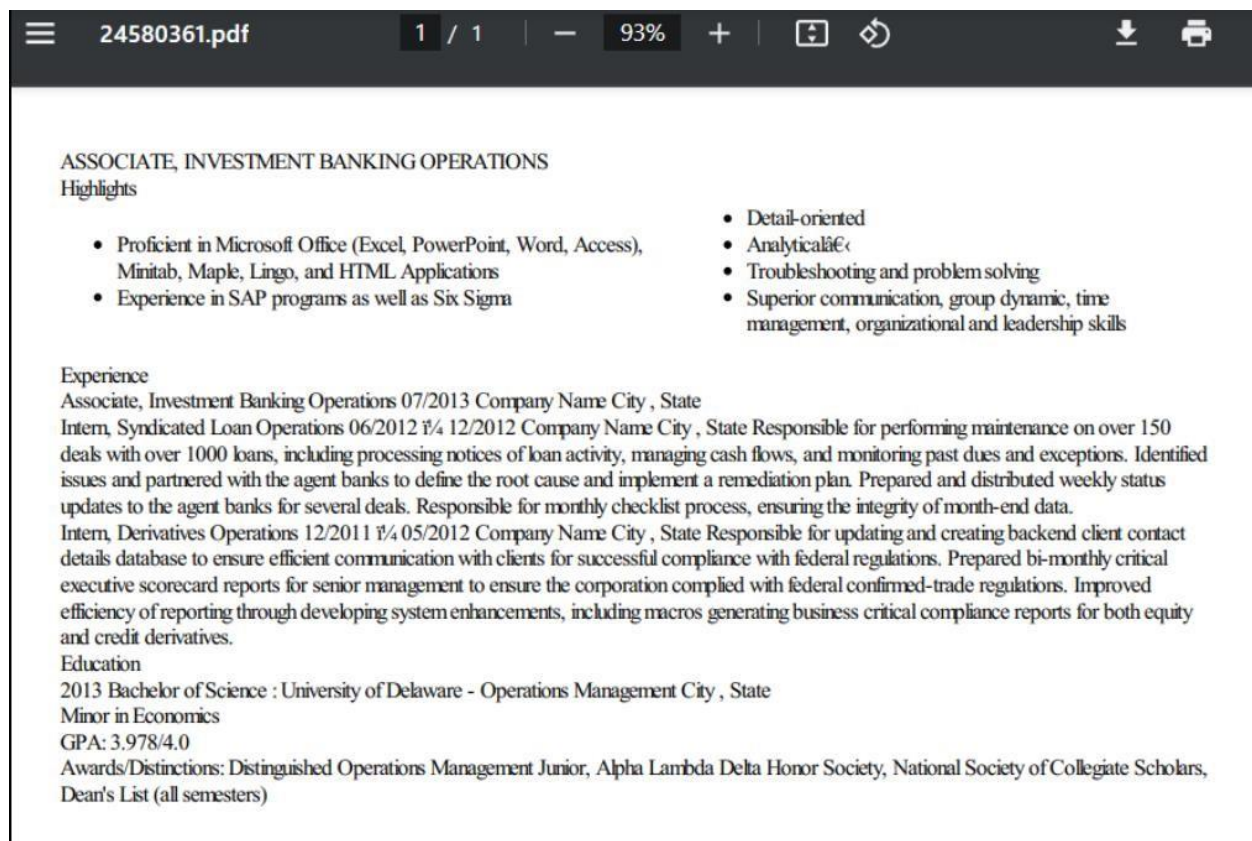
Challenges

1.extracting skills

2.using bert for word vectors

1.Extracting skills:-

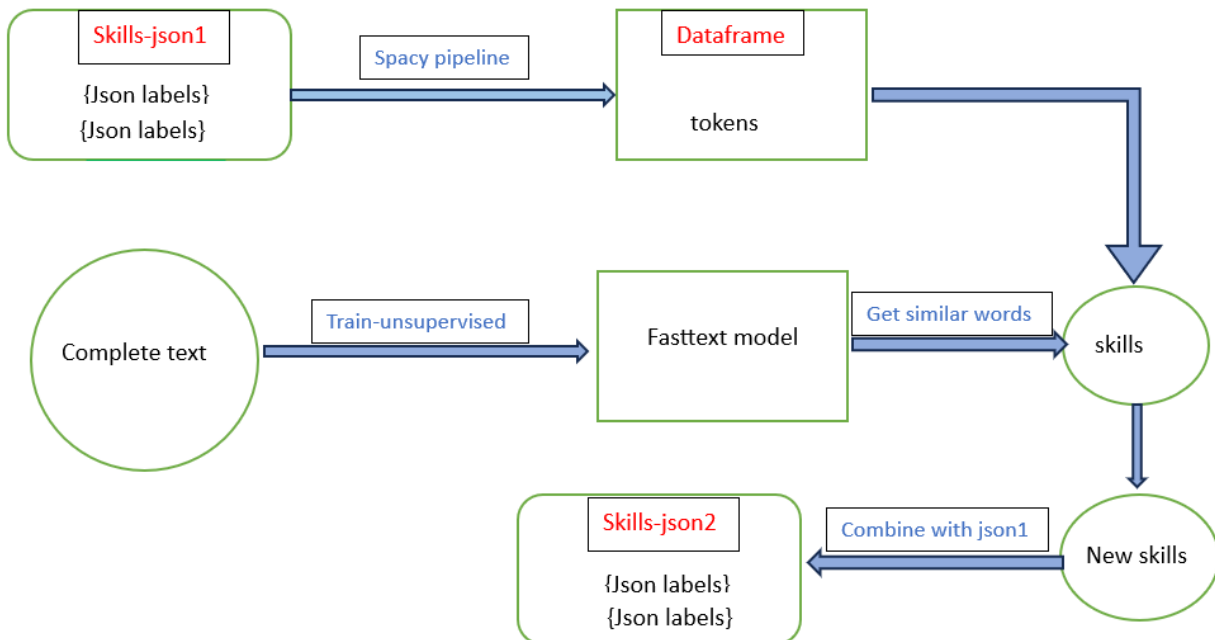
Each pdf's is in different format and layout. Where some pdf's didn't have skills section.



It's not easy to get skills from each resume, then using a labeled json data adding it to spacy pipeline.

Got some skills, using these skills generated new skills using fasttext model, which trained on complete resume text in an unsupervised way.

New skills can get using `get_nearest_neighbors()`, which score and word-based in score I added it new skills.



2.using bert for word vectors

We need complete sentence to get word vectors from bert, but we have tokens from entire text.

Training bert model with entire text is very time taking and computationally expensive.

Then took vectors using spacy large model(en_core_web_lg) from resumes and job description.

Aluru Hema sundher