

Report

Approach :

- I started with making the pdf extractor, which took the pdf of cvs as input and then extracted all the information from it. I directly used the dataset from the kaggle.
- Next, I downloaded the job descriptions from the hugging face dataset, selecting the first 15 job descriptions for my task
- Then, I used the DistilBERT Tokenizer and Model for embedding of both job descriptions and resumes.
- Finally, I calculated the cosine similarity scores between each cv and the job description specified.

Challenges faced and solutions:

- The resumes didn't follow a particular format for skills and qualifications. It was tough to categorize each part. This has a possible solution of using LLMs for categorizing and summarizing skills and qualifications.
- The same issue mentioned above was with the job descriptions.
- Due to large amount of dataset it took a lot of computing time, limiting my results by using only few resumes.

Recommendations or insights from the matching process :

- As most of the companies with a number of applicants uses ATS (Applicant Tracking System), people should try to include keyword in their resume matching with the job description in which they want to apply.
- We can also use clustering for similar jobs and similar resumes to give better insights.

Top 5 candidates for each job description based on similarity scores :

Job Description 1:

Top Matched CVs:

- CV Detail 222 (Cosine Similarity Score: 0.94160)
- CV Detail 187 (Cosine Similarity Score: 0.94131)
- CV Detail 80 (Cosine Similarity Score: 0.94105)
- CV Detail 229 (Cosine Similarity Score: 0.93878)

- CV Detail 245 (Cosine Similarity Score: 0.93813)

Job Description 2:

Top Matched CVs:

- CV Detail 292 (Cosine Similarity Score: 0.91157)

- CV Detail 93 (Cosine Similarity Score: 0.91115)

- CV Detail 80 (Cosine Similarity Score: 0.91058)

- CV Detail 63 (Cosine Similarity Score: 0.90926)

- CV Detail 55 (Cosine Similarity Score: 0.90349)

Job Description 3:

Top Matched CVs:

- CV Detail 291 (Cosine Similarity Score: 0.94660)

- CV Detail 229 (Cosine Similarity Score: 0.94621)

- CV Detail 61 (Cosine Similarity Score: 0.94342)

- CV Detail 292 (Cosine Similarity Score: 0.94300)

- CV Detail 245 (Cosine Similarity Score: 0.94149)

Job Description 4:

Top Matched CVs:

- CV Detail 229 (Cosine Similarity Score: 0.94243)

- CV Detail 80 (Cosine Similarity Score: 0.93905)

- CV Detail 245 (Cosine Similarity Score: 0.93854)

- CV Detail 99 (Cosine Similarity Score: 0.93658)

- CV Detail 288 (Cosine Similarity Score: 0.93470)

Job Description 5:

Top Matched CVs:

- CV Detail 229 (Cosine Similarity Score: 0.94922)

- CV Detail 219 (Cosine Similarity Score: 0.94429)
 - CV Detail 213 (Cosine Similarity Score: 0.94176)
 - CV Detail 187 (Cosine Similarity Score: 0.94165)
 - CV Detail 124 (Cosine Similarity Score: 0.94118)
-

Job Description 6:

Top Matched CVs:

- CV Detail 14 (Cosine Similarity Score: 0.91570)
 - CV Detail 43 (Cosine Similarity Score: 0.91564)
 - CV Detail 100 (Cosine Similarity Score: 0.91482)
 - CV Detail 78 (Cosine Similarity Score: 0.91198)
 - CV Detail 28 (Cosine Similarity Score: 0.91185)
-

Job Description 7:

Top Matched CVs:

- CV Detail 229 (Cosine Similarity Score: 0.95471)
 - CV Detail 62 (Cosine Similarity Score: 0.95380)
 - CV Detail 280 (Cosine Similarity Score: 0.95333)
 - CV Detail 80 (Cosine Similarity Score: 0.95054)
 - CV Detail 15 (Cosine Similarity Score: 0.94965)
-

Job Description 8:

Top Matched CVs:

- CV Detail 229 (Cosine Similarity Score: 0.95621)
 - CV Detail 292 (Cosine Similarity Score: 0.95362)
 - CV Detail 80 (Cosine Similarity Score: 0.95222)
 - CV Detail 61 (Cosine Similarity Score: 0.95143)
 - CV Detail 291 (Cosine Similarity Score: 0.94811)
-

Job Description 9:

Top Matched CVs:

- CV Detail 229 (Cosine Similarity Score: 0.94849)
 - CV Detail 219 (Cosine Similarity Score: 0.94363)
 - CV Detail 80 (Cosine Similarity Score: 0.94356)
 - CV Detail 187 (Cosine Similarity Score: 0.94336)
 - CV Detail 15 (Cosine Similarity Score: 0.94336)
-

Job Description 10:

Top Matched CVs:

- CV Detail 282 (Cosine Similarity Score: 0.95457)
 - CV Detail 230 (Cosine Similarity Score: 0.95406)
 - CV Detail 186 (Cosine Similarity Score: 0.95396)
 - CV Detail 164 (Cosine Similarity Score: 0.95393)
 - CV Detail 299 (Cosine Similarity Score: 0.95113)
-

Job Description 11:

Top Matched CVs:

- CV Detail 229 (Cosine Similarity Score: 0.95394)
 - CV Detail 62 (Cosine Similarity Score: 0.95171)
 - CV Detail 17 (Cosine Similarity Score: 0.94941)
 - CV Detail 15 (Cosine Similarity Score: 0.94919)
 - CV Detail 280 (Cosine Similarity Score: 0.94899)
-

Job Description 12:

Top Matched CVs:

- CV Detail 229 (Cosine Similarity Score: 0.95166)
- CV Detail 80 (Cosine Similarity Score: 0.94854)
- CV Detail 63 (Cosine Similarity Score: 0.94726)

- CV Detail 245 (Cosine Similarity Score: 0.94550)
- CV Detail 219 (Cosine Similarity Score: 0.94493)

Job Description 13:

Top Matched CVs:

- CV Detail 80 (Cosine Similarity Score: 0.94128)
- CV Detail 229 (Cosine Similarity Score: 0.93852)
- CV Detail 245 (Cosine Similarity Score: 0.93621)
- CV Detail 15 (Cosine Similarity Score: 0.93522)
- CV Detail 56 (Cosine Similarity Score: 0.93443)

Job Description 14:

Top Matched CVs:

- CV Detail 192 (Cosine Similarity Score: 0.92840)
- CV Detail 321 (Cosine Similarity Score: 0.92397)
- CV Detail 231 (Cosine Similarity Score: 0.92192)
- CV Detail 124 (Cosine Similarity Score: 0.91557)
- CV Detail 161 (Cosine Similarity Score: 0.90898)

Job Description 15:

Top Matched CVs:

- CV Detail 229 (Cosine Similarity Score: 0.94768)
- CV Detail 15 (Cosine Similarity Score: 0.94375)
- CV Detail 222 (Cosine Similarity Score: 0.94355)
- CV Detail 80 (Cosine Similarity Score: 0.94341)
- CV Detail 167 (Cosine Similarity Score: 0.94116)
