# Resume-Matching-with-Job-Descriptions-Using-PDF-CVs

1. Approach to the Task:

Our approach to the candidate-job matching task involves the following key steps:

Step 1: Data Extraction and Preprocessing:

- We started by extracting text data from PDF resumes using the pdfplumber library.

- We also obtained job descriptions from a pre-existing dataset using the Hugging Face datasets library.

- Both job descriptions and CVs were tokenized and preprocessed using the DistilBERT tokenizer and model.

- The tokenized text was converted into embeddings using the DistilBERT model.

Step 2: Cosine Similarity Calculation:

- We computed cosine similarity between the job descriptions and CV embeddings. Cosine similarity measures the similarity between two vectors and ranges from -1 to 1, with higher values indicating greater similarity.

- Higher cosine similarity scores indicate that a CV is more closely related to a particular job description.

Step 3: Ranking and Selection:

- For each job description, we ranked the CVs based on their similarity scores in descending order.

- Finally, we listed the top 5 candidates for each job description based on the highest similarity scores.

2. Challenges Faced and Solutions:

- Data Extraction: Extracting relevant information from PDF resumes can be challenging due to varying formats. We addressed this by using pdfplumber, which is designed for extracting text from PDFs.

- Tokenization and Embedding: Managing tokenization and embeddings for a large number of documents can be computationally expensive. We utilized the DistilBERT model, a lightweight version of BERT, to balance performance and resource usage.

- Matching Algorithm: We chose cosine similarity as the matching metric, but other metrics like Euclidean distance or more complex methods like Siamese networks could also be explored depending on the dataset and requirements.

3. Top 5 Candidates for Each Job Description:

Based on our matching process, here are the top 5 candidates for each job description along with their similarity scores:

Here are the 15 Job Descriptions

Job Description 1:

* CV 1: DIGITAL MARKETING ASSOCIATE (Similarity Score: 0.9237)
* CV 2: PRODUCT AND WEB DESIGNER (Similarity Score: 0.9157)
* CV 3: SOCIAL MEDIA MANAGER (Similarity Score: 0.9120)
* CV 4: SOCIAL ENGAGE SALES (Similarity Score: 0.9115)
* CV 5: SENIOR ASSOCIATE BUSINESS DEVELOPMENT (Similarity Score: 0.9046)

Job Description 2:

* CV 1: DIGITAL MARKETING ASSOCIATE (Similarity Score: 0.9553)
* CV 2: SOCIAL ENGAGE SALES (Similarity Score: 0.9543)
* CV 3: DIRECTOR (Similarity Score: 0.9471)
* CV 4: SENIOR DIRECTOR OF BUSINESS OPERATIONS - HEALTHCARE (Similarity Score: 0.9422)
* CV 5: PRODUCT DESIGNER (Similarity Score: 0.9419)

Job Description 3:

* CV 1: BUSINESS DEVELOPMENT (Similarity Score: 0.9497)
* CV 2: DIGITAL MARKETING ASSOCIATE (Similarity Score: 0.9484)
* CV 3: SOCIAL ENGAGE SALES (Similarity Score: 0.9473)
* CV 4: SET DESIGNER (Similarity Score: 0.9430)
* CV 5: LICENSED UNITED HEALTHCARE INSURANCE SALES AGENT (Similarity Score: 0.9418)

Job Description 4:

* CV 1: DIGITAL MARKETING ASSOCIATE (Similarity Score: 0.9487)
* CV 2: BUSINESS DEVELOPMENT (Similarity Score: 0.9416)
* CV 3: DIGITAL MEDIA INTERN (Similarity Score: 0.9415)
* CV 4: PRODUCT DESIGNER (Similarity Score: 0.9414)
* CV 5: BUSINESS DEVELOPMENT REPRESENTATIVE (Similarity Score: 0.9411)

Job Description 5:

* CV 1: DIGITAL MARKETING ASSOCIATE (Similarity Score: 0.9487)
* CV 2: BUSINESS DEVELOPMENT (Similarity Score: 0.9416)
* CV 3: DIGITAL MEDIA INTERN (Similarity Score: 0.9415) CV
* 4: PRODUCT DESIGNER (Similarity Score: 0.9414)
* CV 5: BUSINESS DEVELOPMENT REPRESENTATIVE (Similarity Score: 0.9411)

Job Description 6:

* CV 1: PRODUCT DESIGNER (Similarity Score: 0.9203)
* CV 2: SUBSTITUTE TEACHER (Similarity Score: 0.9193)
* CV 3: PRODUCT AND WEB DESIGNER (Similarity Score: 0.9175)
* CV 4: SOCIAL MEDIA MANAGER (Similarity Score: 0.9151)
* CV 5: DIGITAL MEDIA INTERN (Similarity Score: 0.9128)

Job Description 7:

* CV 1: BUSINESS DEVELOPMENT (Similarity Score: 0.9680)
* CV 2: SET DESIGNER (Similarity Score: 0.9609)
* CV 3: DIGITAL MARKETING ASSOCIATE (Similarity Score: 0.9529)
* CV 4: SENIOR DIRECTOR OF BUSINESS OPERATIONS - HEALTHCARE (Similarity Score: 0.9505)
* CV 5: SOCIAL ENGAGE SALES (Similarity Score: 0.9502)

Job Description 8:

* CV 1: DIGITAL MARKETING ASSOCIATE (Similarity Score: 0.9656)
* CV 2: SOCIAL ENGAGE SALES (Similarity Score: 0.9590)
* CV 3: LICENSED UNITED HEALTHCARE INSURANCE SALES AGENT (Similarity Score: 0.9529)
* CV 4: SENIOR DIRECTOR OF BUSINESS OPERATIONS - HEALTHCARE (Similarity Score: 0.9525)
* CV 5: BUSINESS DEVELOPMENT (Similarity Score: 0.9523)

Job Description 9:

* CV 1: DIGITAL MARKETING ASSOCIATE (Similarity Score: 0.9555)
* CV 2: BUSINESS DEVELOPMENT (Similarity Score: 0.9546)
* CV 3: SOCIAL ENGAGE SALES (Similarity Score: 0.9505)
* CV 4: PRODUCT DESIGNER (Similarity Score: 0.9493)
* CV 5: DIRECTOR (Similarity Score: 0.9489)

Job Description 10:

* CV 1: ENGINEERING MANAGER (Similarity Score: 0.9543)
* CV 2: CONSULTANT (Similarity Score: 0.9522)
* CV 3: ENGINEERING SERVICES MANAGER (Similarity Score: 0.9506)
* CV 4: PRODUCT DESIGNER (Similarity Score: 0.9505)
* CV 5: QA ENGINEERING TEAM LEAD (Similarity Score: 0.9500)

Job Description 11:

* CV 1: BUSINESS DEVELOPMENT (Similarity Score: 0.9623)
* CV 2: DIGITAL MARKETING ASSOCIATE (Similarity Score: 0.9570)
* CV 3: SET DESIGNER (Similarity Score: 0.9519)
* CV 4: BUSINESS DEVELOPMENT SPECIALIST/AE (Similarity Score: 0.9514)
* CV 5: DIRECTOR (Similarity Score: 0.9508)

Job Description 12:

* CV 1: DIGITAL MARKETING ASSOCIATE (Similarity Score: 0.9564)
* CV 2: SOCIAL ENGAGE SALES (Similarity Score: 0.9524)
* CV 3: BUSINESS DEVELOPMENT (Similarity Score: 0.9506)
* CV 4: SENIOR DIRECTOR OF BUSINESS OPERATIONS - HEALTHCARE (Similarity Score: 0.9483)
* CV 5: BUSINESS DEVELOPMENT REPRESENTATIVE (Similarity Score: 0.9479)

Job Description 13:

* CV 1: DIGITAL MARKETING ASSOCIATE (Similarity Score: 0.9555)
* CV 2: SOCIAL ENGAGE SALES (Similarity Score: 0.9490)
* CV 3: PRODUCT DESIGNER (Similarity Score: 0.9476)
* CV 4: BUSINESS DEVELOPMENT (Similarity Score: 0.9443)
* CV 5: SOCIAL MEDIA MANAGER (Similarity Score: 0.9433)

Job Description 14:

* CV 1: QA ENGINEERING TEAM LEAD (Similarity Score: 0.9218)
* CV 2: DATABASE PROGRAMMER/ANALYST (.NET DEVELOPER) (Similarity Score: 0.9182)
* CV 3: ASP.NET WEB DEVELOPER (Similarity Score: 0.9182)
* CV 4: CONSULTANT (Similarity Score: 0.9156)
* CV 5: E-LEARNING DESIGNER (Similarity Score: 0.9132)

Job Description 15:

* CV 1: DIGITAL MARKETING ASSOCIATE (Similarity Score: 0.9547)
* CV 2: SOCIAL ENGAGE SALES (Similarity Score: 0.9545)
* CV 3: PRODUCT DESIGNER (Similarity Score: 0.9531)
* CV 4: BUSINESS DEVELOPMENT (Similarity Score: 0.9475)
* CV 5: SET DESIGNER (Similarity Score: 0.9474)

Recommendations and Insights:

- The matching process successfully identifies candidates most closely aligned with job descriptions based on skills and education.

- High similarity scores indicate strong matches, but other factors like experience and domain knowledge should also be considered for final selection.

- Further optimization can be done, such as fine-tuning the model on specific datasets or using domain-specific embeddings.

- Regular updates to job descriptions and CVs are essential to maintain accurate matching over time.

- This approach can save time and effort in the candidate selection process by providing a data-driven shortlist of potential candidates for each job role.

Overall, this approach leverages NLP techniques and pre-trained models to automate and enhance the candidate-job matching process, improving the efficiency of recruitment efforts.