**Code Documentation: Job Description Matching with Resumes**

This code performs a matching process between job descriptions and resumes using cosine similarity, aiming to find the best-matched job description for each resume. The process involves using DistilBERT embeddings and cosine similarity to quantify the similarity between the embeddings of job descriptions and resumes.

**Libraries and Packages**

The code utilizes several Python libraries and packages:

* **os**: Operating system functions.
* **fitz** from PyMuPDF: Used for PDF parsing.
* **DistilBertTokenizer** and **DistilBertModel** from Hugging Face Transformers: For tokenization and DistilBERT model.
* **cosine\_similarity** from scikit-learn: Utilized for calculating cosine similarity.
* **pandas**: Used for data manipulation.
* **load\_dataset** from Hugging Face datasets: To load the job descriptions dataset.
* **numpy**: For numerical operations.

**Dataset Loading and Preparation**

* The code loads the Kaggle Resume Dataset using Hugging Face's **load\_dataset** and extracts job descriptions, company names, and job positions.
* Initializes a DistilBERT tokenizer and model from the Hugging Face Transformers library.

**Job Description Embeddings**

* Iterates through the job descriptions, tokenizes them using the DistilBERT tokenizer, and computes their embeddings using the DistilBERT model.
* These embeddings are stored for further use in the cosine similarity computation.

**Resume Processing**

* Reads resumes from a CSV file and extracts relevant information, including resume text, education, and skills.
* Tokenizes the resume text and calculates its embeddings using DistilBERT.

**Cosine Similarity Calculation**

* Utilizes cosine similarity to calculate the similarity scores between the embeddings of each resume and job descriptions.
* Identifies the top-matched job description for each resume based on the highest similarity score.

**Results Display**

* Prints the top match for each resume, including resume text, matched company, job role, a portion of the job description, and the similarity score.

**Testing with another resume**

To test the code with another resume:

1. Prepare the New Resume:

Make sure the new resume is in a format that the code can process. If the code is designed to read from a CSV file, ensure the new resume is added to the CSV file in the appropriate column.

1. Update the Input Data:

Update the code to use the new resume you want to test. Replace the relevant lines that load and process the resumes with the new resume's information.

1. Run the Code:

Execute the updated code to process the new resume and match it with the job descriptions.