

Developer: Ali Haider 2021-CS-38

Supervisor: Dr Usman Ghani

### Problem Statement

In the modern recruitment landscape, the growing volume of resumes and job descriptions has made it difficult to match the right candidates manually. Traditional methods are time-consuming, error-prone, and inefficient, often leading to suboptimal hiring decisions. With the rise of data-driven approaches and Natural Language Processing (NLP), there is an opportunity to automate and streamline this process

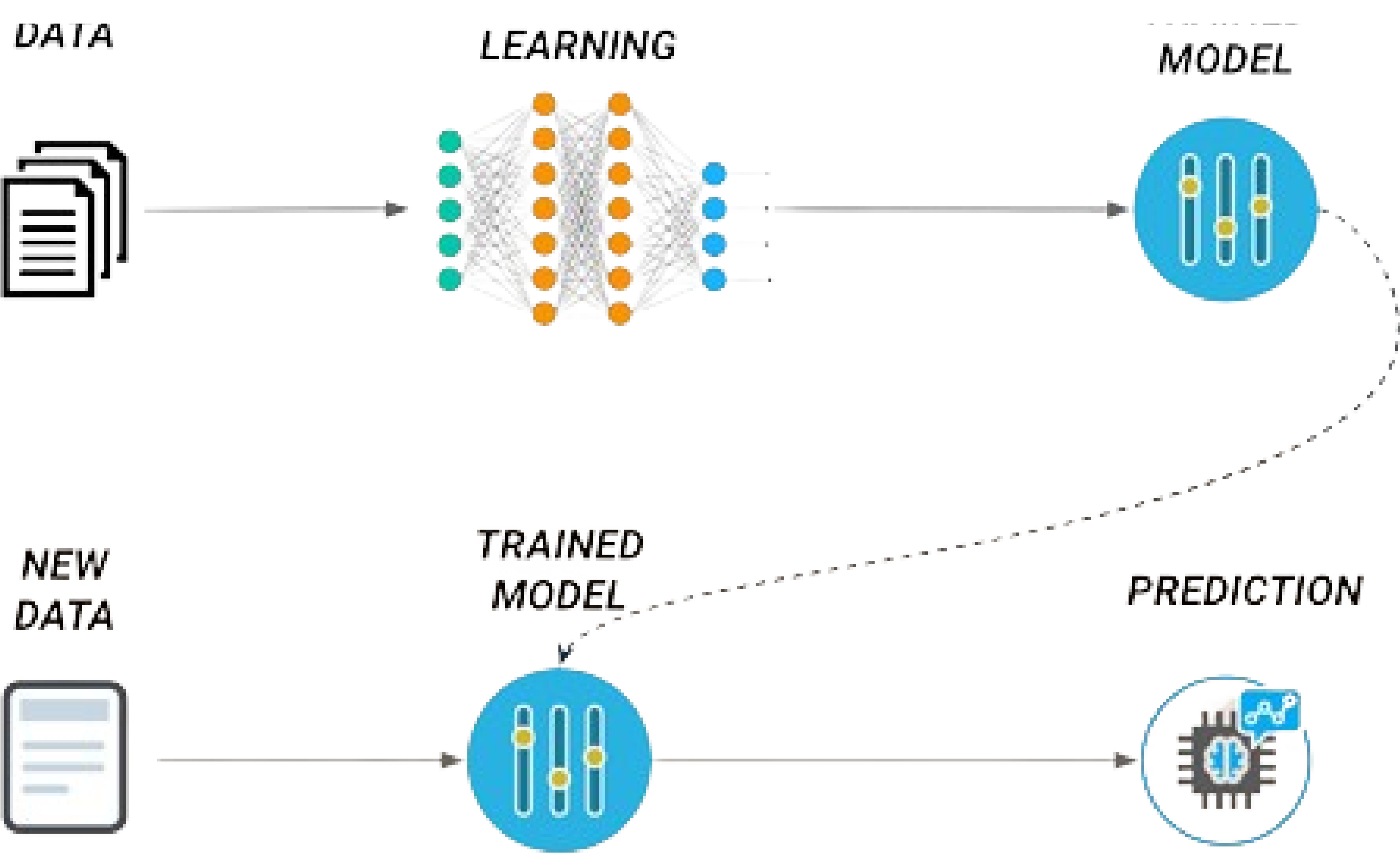
### Solution

This project utilizes cutting-edge NLP techniques to develop an automated Job Description and Resume Matching System. The system analyzes textual data from resumes and job descriptions, extracting meaningful features to calculate similarity scores. By leveraging models such as **DistilBERT** for semantic understanding and **Doc2Vec** for document similarity, the system can efficiently match candidates to job roles based on their qualifications, skills, and experiences.

### Dataset

The dataset used in this project comprises a collection of over 3000+ resumes gathered from various public sources, including Kaggle. It contains detailed information about candidates' professional experience, skills, education, and other key factors relevant for job matching.

### Methodology



### Implementation

#### NLP-based Approach (DistilBERT):

- Utilizes DistilBERT for tokenizing and embedding job descriptions and resumes.
- Compares embeddings to identify the most relevant resumes.
- Implements DistilBertTokenizer and DistilBertModel from the Hugging Face Transformers library.

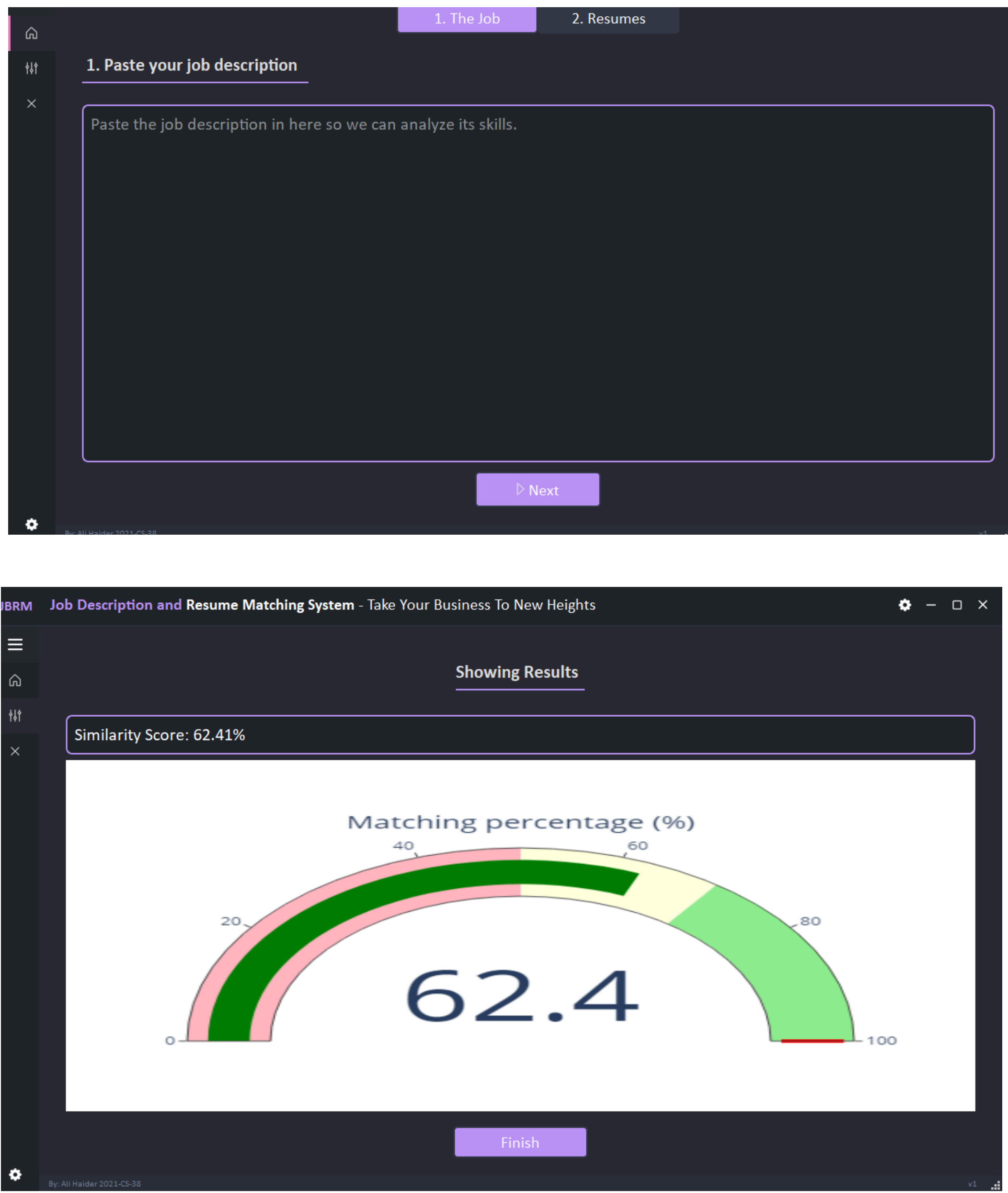
#### Vector Similarity Approach:

- Uses Doc2Vec to generate vector representations of job descriptions and resumes.
- Calculates cosine similarity to assess match quality.

#### Functional Application:

- Developed a functional desktop application for matching job descriptions to resumes.
- Supports analysis of multiple resumes or single resume for personalized job matching.

### Results



### Technologies

