AI Resume Screening

Introduction

AI Resume Screening is a tool that uses artificial intelligence to automate the process of resume screening and shortlisting. The tool uses natural language processing and machine learning algorithms to analyze resumes and classify them to the job roles based on the words in their resume.

Our tool will be designed to make the hiring process easier for both HR teams and job seekers. Our idea is to develop a user-friendly web page that enables HR teams to select the specific job role they are recruiting for. Once the job role is chosen, a link is shared with potential candidates who can then upload their resumes.

Using pre-trained machine learning models, which are trained on thousands of resumes , our tool automatically filters out resumes that do not match the job requirements. Resumes are categorized based on specific keywords and phrases related to the job role ( for example , python developer – main words can be python , developed , project , computer science etc). If a candidate's resume matches the job description, they are immediately notified and their resume is uploaded to the company's database. Even if a candidate’s resume doesn’t match the description , it shows the potential role which is suitable for the uploaded resume.

Our tool saves HR teams valuable time and energy as they no longer need to manually sift through hundreds of resumes. Job seekers also benefit from receiving immediate feedback on their eligibility, streamlining their job search process.

Features

Automated resume screening: AI Resume Screening saves time and effort by automatically screening resumes based on job requirements and pre-defined criteria.

Improved accuracy: The tool uses advanced algorithms to analyze resumes, reducing the likelihood of human bias and improving the accuracy of the shortlisting process. We have achieved 90 percent accuracy

Efficient process: Each resume is categoried with the model in less than 5 seconds.

Detailed output : The HR gets a detailed output of the Name, Email, Location, Resume(in binary) and the Resume score based

Usage

The code consists of the following parts with its uses:

1.Upload\_Resume.py : This is the code which integrates the UI with the Trained model.

Use the command “streamlit run Upload\_Resume.py” which opens the web browser where you can upload your resume.

Alternatively the code can be hosted on github and the link can directly be sent to the candidate.

2.combineddata-ML.ipynb : It is the jupyter notebook which we used to train the final model on all algorithms. It also contains the accuracies of each algorithm we have used.

TF-IDF was used to extract the features from the pre-processed resume.

3.cv.pickle : It is the pickled file which contains the features of the model trained on the resumes by using TF-IDF. This pickle file is used to compare the features of the Uploaded resume with the model.

4.RF.joblib: It is the machine learning model (Random Forest) which had the highest accuracy. This is the model used to predict the category in which the resume fits.

5.SQL.txt : Contains the MYSQL queries to set up a database to store the details of the short listed candidates.

6.mail.py: It contains the Python code which automatically sends an email to the shortlisted candidates.

7.pages/Show\_Resumes.py: It displays the resumes of the shortlisted candidates by sorting them in descending order of the scores

8.HR.py: The page where HR enters the job role which is open for hiring, based on which shortlisting of candidates is done.

Other things to know:

The selected data will be stored in a MYSQL database.

The email feature to be enabled , the user has to configure the operating system’s environment variable to set the senders address , password and port number.

Datasets used for the model :

https://www.kaggle.com/datasets/gauravduttakiit/resume-dataset <https://www.kaggle.com/datasets/snehaanbhawal/resume-dataset>