

Name: Azizul Muktadir
Chowdhury
ID: 22341040
Section : 1
Course: CSE 440

Project Report: Resume Analyzer using Natural Language Processing

Introduction:

The purpose of this project was to develop a resume analyzer using natural language processing techniques. The analyzer is designed to take resumes in PDF format as input and perform various analyses on them. The goal is to extract relevant information from the resumes and provide insights to the user, such as identifying top skills and comparing candidate profiles.

Methodology:

The analyzer was built using the Python programming language and several natural language processing libraries, including spaCy, Gensim, and PyPDF2. The analyzer uses spaCy to parse resumes and extract relevant keywords. It then calculates the cosine similarity score to compare resumes and rank them accordingly. Additionally, the analyzer uses Gensim to perform Latent Dirichlet Allocation (LDA) analysis on a dataset of resumes, identifying the top topics and visualizing them using pyLDAvis.

Results:

The analyzer was able to successfully parse resumes in PDF format and extract relevant keywords using spaCy. It was also able to compare resumes using cosine similarity and rank them accordingly. The LDA analysis using Gensim was able to identify the top topics in a dataset of resumes and visualize them using pyLDAvis. These results provide valuable insights to users, enabling them to make better-informed decisions when assessing candidate resumes.

Conclusion:

The resume analyzer developed in this project is a valuable tool for companies and recruiters. It provides an efficient way to analyze and compare resumes, identify top skills and topics, and gain a better understanding of candidate profiles. Future work can include incorporating additional machine learning models to improve the accuracy of the analyzer and expanding its functionality to include additional features, such as sentiment analysis and entity recognition.

Recommendations:

Based on the results obtained from this project, it is recommended that companies and recruiters use the developed resume analyzer to streamline the recruitment process. The analyzer provides valuable insights and saves time by automating the analysis and comparison

of resumes. Companies can also consider expanding the functionality of the analyzer to include additional features, as mentioned above, to further improve its usefulness.

Acknowledgements:

We would like to express our sincere gratitude to the natural language processing community and the developers of spaCy, Gensim, and PyPDF2 for their contributions to this project. We also extend our appreciation to our team members who have worked tirelessly to develop this resume analyzer.

References:

Spacy: <https://spacy.io/>

Gensim: <https://radimrehurek.com/gensim/>

PyPDF2: <https://github.com/mstamy2/PyPDF2>

pyLDAvis: <https://github.com/bmabey/pyLDAvis>

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