ADAMA SCIENCE AND TECHNOLOGY UNIVERSITY

SCHOOL OF ELECTRICAL ENGINEERING AND COMPUTING

DEPARTEMENT OF COMPUTER SCINECE AND ENGINEERING

NLP Project

Resume Scanner

Fidel Alemayehu ……UGR/17109/11

Saron Tibebe……...UGR /17615/11

Habtemariyam Sahelu ……UGR /17179/11

Dagemawi Demelash……A/UR15125/10

Biniyam Moges……..UGR/16897/11

Sub .To : Dr.Teshome

Sub. Date December 30,2021.

Resume Scanner

Resume scanning is a process where employers convert hardcopy resumes into electronic resumes. Basically, instead of a human reading the resume, the resume is first input into the company's computer database via a scanner. Then, if a resume corresponds to a job position, it's getting to an HR specialist.

What is resume scanner and how does it work?

Resume Scanner Compare your resume to your target job description. We'll give you a match score.

For our project we made a resume scanner using NLP techniques that accept resume and job description form the user using UI and tells how much percent the resume matches with the given Job description.

How we made resume scanner and the algorithms we used

We made the project with python

step1

Install docx2txt which convert documents into text

Step2

Importing Libraries

import docx2txt

Step3

Loading Data

we accept job description and resume using our UI from our computer

Step4

The docx2txt library converts the documents into text

Step5

We list the texts

text = {resume , Job_description}

Step6

import sklearn library

Scikit-learn (Sklearn) is the most useful and robust library for machine learning in Python. It provides a selection of efficient tools for machine learning and statistical modeling including classification, regression, clustering and dimensionality reduction via a consistence interface in Python

from sklearn import CountVectorizer

CountVectorizer is a great tool provided by the scikit-learn library in Python. It is used to transform a given text into a vector on the basis of the frequency (count) of each word that occurs in the entire text.

Step7

Count _matrix transform the numbers into matrix after CountVectorizer

Step8

import library from sklearn

from sklearn.metrics.pairwise import Cosine_similarity

sine similarity, or the cosine kernel, computes similarity between job description and resume scanner.

Step9

Then it will print similarity scores in 2*2 matrix

Step10

Get the match percentage

- **We assume that this resume scanner will help students and applicants a lot because it reduces the time they consume reading job description that doesn’t match their own.**

ATS is one of Resume scanner websites.

Companies often receive hundreds of resumes for every job posting they put out. They and their hiring managers don't have the time to go through each resume individually, so they use software, often called Applicant Tracking Systems (ATS), to filter through all the resumes they receive.

You need to make sure your resume is read correctly by these ATS and resume screening software. If it is not, your resume could get instantly rejected by these systems and never make its way to a recruiter or hiring manager.

Algorithms used by researchers

Step1

Natural Language Processing (NLP)

importing nltk packages

The Natural Language Toolkit (NLTK) is an open source Python library for Natural Language Processing

Step2

Python libraries/Python Packages

NLTK

beautifulsoup

Flask

Pandas

Sklearn

Step3

Predictive Analytics

Step4

Regular Expression/Rule Based Parsing

Step5

Named Entity Recognition (NER)

Named Entity Recognition (NER) is a standard NLP problem which involves spotting named entities (people, places, organizations etc.) from a chunk of text, and classifying them into a predefined set of categories. Some of the practical applications of NER include:

- Scanning news articles for the people, organizations and locations reported.
- Providing concise features for search optimization: instead of searching the entire content, one may simply search for the major entities involved.
- Quickly retrieving geographical locations talked about in Twitter posts.

Step6

Spacy’s NER

NER with spaCy

spaCy is regarded as the fastest NLP framework in Python, with single optimized functions for each of the NLP tasks it implements. Being easy to learn and use, one can easily perform simple tasks using a few lines of code.

Step7

BERT NER

pip install spacy

import spacy

nlp = spacy.load('en_core_web_sm')

Our way and the researchers’ way are a bit different but the outcome is almost similar and we hope our project will help peoples.