**UNT AI/CS SUMMER RESEARCH PROGRAM**



**ATSIRE- Advanced Tracking System with Intelligent Resume Embeddings**

**Course: INFO 5900 Special Problem**

**Faculty Advisor: Dr. Ting Xiao**

**Student Advisor: Ravi Varma Kumar Bevara**

**Date of Submission: 07-28-2023**

**Team Members**

|  |  |  |
| --- | --- | --- |
| **Name** | **Department** | **Student ID** |
| Harshitha Pasem | College of Information | 11592106 |
| Sri Chandra Dronavalli | College of Information | 11585371 |
| Sai Tejesh Gonemadatala | College of Engineering | 11608671 |

**Project Abstract:**

During economic downturns, businesses face significant challenges in maintaining their operations and remaining profitable. In such times, it becomes even more crucial to hire highly qualified candidates who can help improve the company's operations and ensure its success. However, identifying the most suitable candidates is not always easy, and traditional recruitment methods may not be sufficient to evaluate a candidate's qualifications accurately. To address this issue, a recent study has proposed a unique method that uses embeddings to assess a candidate's qualifications. Embeddings are a type of deep-learning technique that can create comprehensive and multidimensional representations of data, such as resumes and job descriptions.

By creating embeddings for both resumes and job descriptions, the method enables more efficient comparison and matching of candidates, resulting in more accurate evaluations of their job compatibility. The approach proposed in the study has several advantages over traditional recruitment methods. First, it can handle both structured and unstructured data, which makes it more versatile and flexible. Second, it can capture the nuances of a candidate's qualifications better than traditional methods, which tend to rely on keywords and specific criteria. Third, it can scale to handle large volumes of data, which is essential when dealing with a high volume of applicants.

By using this method, businesses can significantly enhance their recruitment process and ensure that the most qualified candidates fill critical roles. This can improve overall efficiency and help mitigate risks associated with hiring underqualified or unsuitable candidates during economic uncertainty. Moreover, the method can be applied to various industries, making it a valuable tool for companies across different sectors.

**Data Specifications:**

**Resume Dataset:**

We came across the Resume dataset on Kaggle, which includes 962 records gathered two years ago spread across 25 distinct categories. It offers a compilation of job categories and resumes, helping us.

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

**Updated Resume:**

* Category: The various categories of the job
* Resume: The resume of the category.

**Job Description Dataset:**

We spent considerable time and effort in curating the job description to ensure that it is comprehensive and relevant. To achieve this, we have leveraged a range of resources, including web scraping techniques for LinkedIn, to gather data for each category. Through this process, we have masterfully crafted a job description that accurately reflects the requirements and expectations of the role, providing potential candidates with a clear understanding of the position and its responsibilities.

**Workflow:**

**A diagram of a diagram

Description automatically generated**

**Project Design:**

* Language: Python
* Tools:

Currently, our data is limited and manageable using Google Colab. However, as the data becomes larger and more complex, we can utilize TACC systems for better handling and processing of the data. TACC systems are designed to handle large amounts of data and complex computations, making them an ideal choice for data analysis tasks that require significant computing resources. By leveraging TACC systems, we can ensure that our data analysis is efficient and effective and can accommodate future growth in our data needs.

* Libraries: pandas, string, spacy, transformers, deep learning algorithms like Bert-base-nli-mean-tokens, Roberta-base-nli-mean-tokens, distil Bert-base-nli-mean-tokens, gpt2, t5,

**Importing the required libraries:**



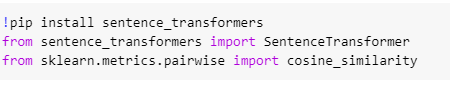
**Importing the required files:**



**Preprocessing the text:**



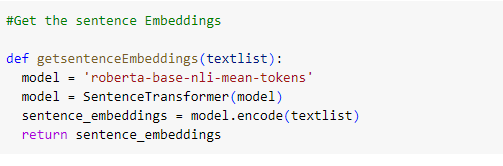
**Installing the Sentence Transformer:**



**Models used:**



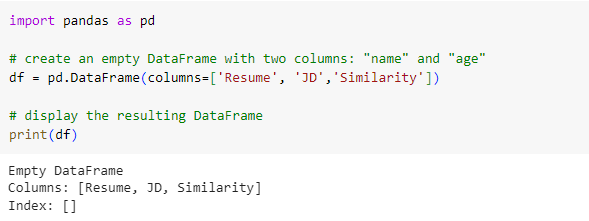
**Defining Sentence Embeddings:**



**Calling Sentence Embedding Function:**



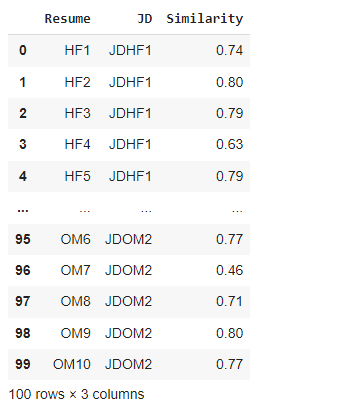
**Creating a dataframe for cosine similarity:**



**Finding cosine similarity:**



**Results of cosine similarities:**



**Visualizing the embeddings:**



A diagram of a graph

Description automatically generated

**Results:**

The research suggests that advanced ATS (Applicant Tracking System) outperforms both crowd outsourcing and traditional ATS methods in terms of recruitment and selection processes. The study's results were unambiguous, showing that the advanced ATS method produced outcomes more closely aligned with the advanced ATS method than the other two. One reason the advanced ATS method outperformed crowd outsourcing and traditional ATS methods is that it uses the latest deep-learning techniques to create comprehensive and multidimensional embeddings of both resumes and job descriptions. This approach enables more efficient comparison and matching of candidates, resulting in more precise evaluations of their job compatibility. Compared to traditional ATS methods that rely on keywords and specific criteria, the advanced ATS method can capture the nuances of a candidate's qualifications more accurately. Moreover, the advanced ATS method can handle both structured and unstructured data, making it more versatile and flexible than traditional ATS methods. It can also scale to handle large volumes of data, which is essential when dealing with a high volume of applicants. This scalability ensures that the recruitment and selection process can be carried out efficiently, even when there are many candidates to evaluate.

Overall, they show that advanced ATS is the most dependable and efficient approach for recruitment and selection processes. By using the power of deep-learning techniques, organizations can enhance their recruitment processes and achieve better outcomes. This can lead to significant benefits, such as increased efficiency, reduced costs, and improved business outcomes. Therefore, organizations looking to improve their recruitment and selection processes should consider adopting the advanced ATS method to achieve their goals.

**Future work:**

Improving the recruitment process is essential to ensure that organizations hire the most qualified and suitable candidates for their open positions. To achieve this, open communication and collaboration with recruitment teams are crucial. This approach can provide a better understanding of the positions being recruited for and the qualifications necessary for success, which can help to refine the recruitment process. To attract the most qualified candidates, it is also important to obtain more reliable resumes and job descriptions. This can be achieved by working closely with recruitment teams to ensure that job descriptions accurately reflect the skills, experience, and qualifications necessary for the position. Additionally, providing clear and concise job descriptions can help to attract candidates who are a good match for the role.

To further improve the recruitment process, organizations can explore the use of deep learning models, such as embeddings. These models can create comprehensive and multidimensional representations of both resumes and job descriptions, enabling a more efficient comparison and matching of candidates. By leveraging the power of advanced technologies, organizations can streamline the recruitment process and ensure that they are making the best hiring decisions for their organization.

Moreover, the use of deep learning models can help to identify candidates who may have been overlooked in the traditional recruitment process. These models can analyze a wide range of data, including educational background, work experience, and personal characteristics, to identify the most suitable candidates for a particular role. This can lead to better hiring decisions and improved business outcomes.

**References:**

* Parasurama, P., & Sedoc, J. (2021). Degendering Resumes for Fair Algorithmic Resume Screening. arXiv preprint arXiv:2112.08910.
* Bondielli, A., & Marcelloni, F. (2021). On the use of summarization and transformer architectures for profiling résumés. Expert Systems with Applications, 184, 115521