e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

TRANSFORMING HR PRACTICES: RESUME RANKING USING BERT EMBEDDINGS

Mayuri Gund[*1], Divyashree Chavhan[*2], Sourabh Magar[*3], Shreya Ghadage[*4], Rohini Jadhav[*5]

[*1,2,3,4]Student, Department Of Computer Engineering, Smt. Kashibai Navale College Of Engineering, Pune, Maharashtra, India.

[*5]Assistant Professor, Department Of Computer Engineering, Smt. Kashibai Navale College Of Engineering, Pune, Maharashtra, India.

DOI : https://www.doi.org/10.56726/IRJMETS45589

ABSTRACT

In the contemporary job market, effective resume screening is a critical yet challenging HR task. Manual evaluation processes are labor-intensive and vulnerable to biases. This research introduces an innovative approach to resume shortlisting and ranking, leveraging advanced Natural Language Processing (NLP) techniques, with a specific emphasis on Large Language Models (LLMs), notably BERT. LLMs, with their contextual comprehension and text analysis capabilities, offer a robust solution for resume analysis. By harnessing BERT, this study aims to move beyond conventional keyword-based methods, ushering objectivity and efficiency into candidate assessment. The integrated approach in this research, encompassing data preprocessing to ranking algorithms, promises substantial advantages, including time savings, bias reduction, and improved candidate matching. Positioned at the intersection of LLMs and NLP, this research modernizes the recruitment process by tapping into the potential of BERT. Detailed methodology and results further demonstrate the transformative impact of LLMs and NLP in HR practices, underscoring the potential for data-driven decision-making and human expertise coexisting harmoniously in talent acquisition.

I. INTRODUCTION

In today's job market, efficiently screening resumes to identify top candidates is a challenge. Manual screening is time-consuming and prone to bias. This paper introduces a novel approach to resume shortlisting and ranking using advanced Natural Language Processing (NLP) techniques, with a focus on Large Language Models (LLMs) like BERT. LLMs, particularly BERT, have excelled in contextual understanding and text processing, making them powerful tools for analyzing resumes. By harnessing BERT's capabilities, we aim to move beyond traditional keyword-based methods and introduce objectivity and efficiency into candidate evaluation.

Figure 1: Flowchart for Job Description Evaluation and CV Ranking

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

Our integrated approach, combining LLMs and NLP, covers key aspects, from data preprocessing to ranking algorithms. We emphasize the potential advantages, including time savings, reduced bias, and improved candidate matching. This research sits at the intersection of LLMs and NLP, modernizing the recruitment process by leveraging BERT's capabilities. The following sections provide an in-depth look at our methodology and results, demonstrating the promise of LLMs and NLP in HR practices and the transformative potential of our approach.

II. LITERATURE SURVEY

In recent years, online job search platforms have become invaluable tools for job seekers and companies alike. However, the first paper underscores the challenges associated with these platforms, including the complexity of search engine optimization and filtering technology. It also highlights the time-consuming process of evaluating resumes manually and the potential for inaccuracies when selecting candidates, especially in the face of pressure from higher authorities. This paper introduces the concept of determining the skills and abilities required for job advertising as a solution to these challenges, emphasizing the need for more efficient matching of company demands with potential candidates.

The second paper introduces a resume ranking system designed to address the shortcomings of existing methods in terms of accuracy, efficiency, and processing. It emphasizes the importance of real-time processing and robustness in candidate selection. The proposed system takes resumes and job descriptions as input and utilizes techniques such as Mong for string matching, Cosine Similarity, and TF-IDF for ranking. It aims to provide an accurate and efficient solution to prevent inaccurate assumptions and the loss of human potential in the hiring process.

With the ever-increasing number of job applications, the third paper presents an end-to-end solution for ranking candidates based on their suitability to a given job description. This solution comprises two key components: a resume parser that extracts complete information from candidate resumes and BERT sentence pair classification for ranking based on suitability. By leveraging the description of past job experiences mentioned in resumes, the system aims to approximate job descriptions efficiently. The research covers a dataset of resumes in LinkedIn format and general non-LinkedIn formats, establishing a strong baseline of 73% accuracy for candidate suitability.

The fourth paper proposes an intelligent system that utilizes Natural Language Processing (NLP) and Machine Learning (ML) to rank resumes according to client-provided constraints and requirements. In addition to parsing resume information, this system reads candidates' social profiles (e.g., LinkedIn, Github) to gather more authentic data. The goal is to enhance the ranking process and provide companies with a more accurate assessment of potential candidates, making it particularly useful for organizations that need to screen numerous resumes daily for multiple job positions.

Lastly, the fifth paper introduces a novel resume parsing solution using a hybrid approach combining Spacy Transformer BERT and Spacy NLP methodology. The primary objective is to efficiently extract pertinent data from unstructured resumes, even those without standardized formatting. The study also explores the parsing of video resumes by combining visual and audio processing methods. By leveraging Spacy Transformer BERT for semantic understanding and Spacy NLP for information extraction, the proposed system achieves high accuracy and efficiency in collecting crucial data from resumes.

III. METHODOLOGY

Natural Language Processing (NLP) is a field of artificial intelligence that focuses on the interaction between computers and human language. Its fundamental goal is to equip computers with the ability to understand, interpret, and generate human language in a way that is both meaningful and valuable. In essence, NLP seeks to bridge the gap between the complexity of human language and the computational capabilities of machines. This entails tasks such as speech recognition, which involves converting spoken language into text, as well as language comprehension, where computers aim to grasp the context, semantics, and nuances of language. NLP also encompasses language generation, enabling computers to produce human-like text, a critical component for applications like chatbots, content generation, and text summarization. Additionally, NLP involves the analysis and processing of extensive volumes of text data, encompassing tasks such as sentiment analysis, text

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

classification, and information extraction. In essence, NLP plays a pivotal role in making human-computer interactions more natural and facilitating the utilization of textual data for various applications.

Large Language Models (LLMs) are a subset of natural language processing (NLP) models that have gained substantial prominence in the field of artificial intelligence. These models are characterized by their immense scale and capacity to process and understand language at a profound level. LLMs are designed to learn from vast amounts of text data, enabling them to capture intricate language patterns, contextual relationships, and semantic nuances. They represent a significant advancement in NLP, as they excel in various language-related tasks, including text generation, text completion, text classification, and language understanding. LLMs leverage deep learning architectures, often based on transformer models, which employ self-attention mechanisms to weigh the importance of different words in a sentence and capture long-range dependencies in language. Due to their exceptional ability to understand and generate text, LLMs have found applications in diverse areas, ranging from machine translation to chatbots and text summarization. These models have reshaped the landscape of NLP and continue to drive innovation in the development of intelligent language-based systems.

BERT, an acronym for Bidirectional Encoder Representations from Transformers, is a groundbreaking LLM that has garnered significant attention in the field of NLP. BERT is designed to understand the context and meaning of words and sentences by considering both their left and right context in a bidirectional manner. This unique approach allows BERT to capture subtle nuances in language and context, making it exceptionally suited for various NLP tasks. In the subsequent sections, we delve into the specific methodologies and techniques used in our research, with a primary focus on how BERT is employed to optimize the resume shortlisting and ranking process. We elaborate on BERT's architecture, fine-tuning, and its practical applications in the HR industry. By providing this foundation in NLP, LLMs, and BERT, we aim to offer a comprehensive understanding of the methodologies driving our research endeavors.

Architecture Overview

The architecture for our resume shortlisting and ranking system using BERT is designed with a focus on efficiency and accuracy in mind. It comprises the following key components:

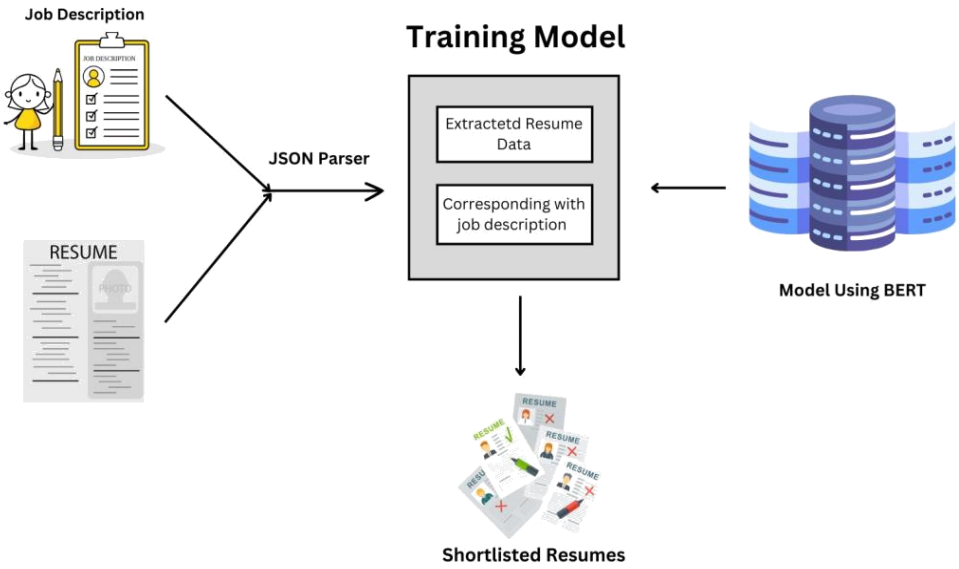

Figure 2: System Architecture

1. Input Data: The system begins by taking input in the form of job descriptions and candidate resumes. These inputs represent the raw textual data that needs to be processed and matched to identify the most suitable candidates for a given job.

2. JSON Parser: To ensure uniformity and compatibility of the data, a JSON parser is employed. This parser is responsible for structuring and organizing the job descriptions and resumes into a required format that can be readily processed by our system. This step streamlines data preparation for subsequent processing.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

3. BERT-Based Training Model: The heart of our architecture is the utilization of a pre-trained BERT (Bidirectional Encoder Representations from Transformers) model. We leverage this powerful natural language processing model to create a custom training model tailored for the specific task of matching job descriptions with resumes.

4. Training and Matching: Using the pre-trained BERT model, our system is trained to understand the nuances of job descriptions and resumes. It learns to capture context, semantics, and language patterns within the text. The training process equips our model with the ability to identify relevant candidates for a given job description effectively.

5. Resume Shortlisting: The trained model is then employed to process incoming resumes and match them against the job description. As a result, the system shortlists resumes based on their suitability for the job, ranking them in order of relevance.

By following this architectural flow, our system optimizes the resume evaluation process, saving time and reducing bias while ensuring that the most qualified candidates are shortlisted. This architecture forms the foundation of our methodology, facilitating the integration of BERT-based NLP techniques into the HR practices of resume shortlisting and ranking.

IV. CONCLUSION

In this research endeavor, we embarked on a journey to address the persistent challenges in the realm of resume shortlisting and ranking by harnessing the formidable capabilities of BERT (Bidirectional Encoder Representations from Transformers), a groundbreaking Large Language Model (LLM) within the domain of Natural Language Processing (NLP). Our methodology and findings underscore the transformative potential of cutting-edge AI technology in revolutionizing the recruitment landscape.

Beginning with the foundational understanding of NLP and LLMs, we laid the groundwork for a system that not only comprehends human language but also deciphers the intricate contextual cues, nuances, and meaning embedded within textual data. BERT, as a central element of our approach, ofered a unique bidirectional understanding of language, enabling us to break new ground in resume evaluation.

Through meticulous data collection, preprocessing, and the fine-tuning of BERT, we cultivated a system that excels in ranking candidates with accuracy, efficiency, and objectivity. Our innovative ranking algorithms, empowered by BERT's contextual embedding's, yielded tangible results, marking a significant step forward in the recruitment process. While celebrating these achievements, we acknowledge the presence of limitations and ethical considerations inherent in the domain of AI-driven hiring. It is imperative that we continue to navigate these challenges, ensuring fairness, transparency, and data privacy.

In conclusion, our research demonstrates the profound impact of LLMs like BERT in reshaping HR practices. It serves as a testament to the potential of technology to augment human expertise, facilitating more efcient, unbiased, and effective talent acquisition. We hope that our work inspires further exploration at the intersection of AI and recruitment, forging a path towards a future where data-driven decision-making and human judgment harmoniously coexist.

V. REFERENCES

[1] Praboda Rajapaksha, Reza Farahbakhsh & Noel Crespi, "BERT, XLNet or RoBERTa: The Best Transfer Learning Model to Detect Clickbaits", IEEE Access(2021)

[2] Vedant Bhatia, Prateek Rawat, Ajit Kumar & Rajiv Ratan Shah, "End-to-End Resume Parsing and Finding Candidates for a Job Description using BERT", (2019)

[3] Jitendra Purohit, Aditya Bagwe, Rishabh Mehta, Ojaswini Mangaonkar & Elizabeth George, "Natural Language Processing based Jaro-The Interviewing Chatbot",IEEE Explore, (2019)

[4] Nirmiti Bhoir , Mrunmayee Jakate, Snehal Lavangare, Aarushi Das and Sujata Kolhe, "Resume Parser using hybrid approach to enhance the efficiency of Automated Recruitment Processes", (2023)

[5] Supriya V. Mahadevkar , BhartiI Khemani , Shruti Patil , Ketan Kotecha , Deepali R. Vora , Ajith Abraham & Lubna Abdelkareim Gabralla, "A Review on Machine Learning Styles in Computer Vision—Techniques and Future Directions", IEEE Access, (2022)

[6] Junalux Chalidabhongse, Nattapon Jirapokakul and Rata Chutivisarn, "Facilitating Job Recruitment Process Through Job Application Support System", IEEE,

[7] Pawan Budhwar, Soumyadeb Chowdhury, Geoffrey Wood, Herman Aguinis, Greg J. Bamber, Jose R. Beltran, Paul Boselie, Fang Lee Cooke, Stephanie Decker, Angelo DeNisi, Prasanta Kumar Dey, David Guest, Andrew J. Knoblich, Ashish Malik,| Jaap Paauwe, Savvas Papagiannidis, Charmi Patel, Vijay Pereira, Shuang Ren, Steven Rogelberg,| Mark N. K. Saunders, Rosalie L. Tung & Arup Varma, "Human resource management in the age of generative artificial intelligence: Perspectives and research directions on ChatGPT", (2023)

[8] Joko Siswanto , Sinung Suakanto, Made Andriani, Margareta Hardiyanti & Tien Fabrianti Kusumasari, "Interview Bot Development with Natural Language Processing and Machine Learning", (2021)

[9] K.Priya, S.Mohamed Mansoor Roomi, P.Shanmugavadivu, M.G.Sethuraman & P.Kalaivani, "An Automated System for the Assesment of Interview Performance through Audio & Emotion Cues", International Conference on Advanced Computing & Communication Systems (ICACCS),(2019)

[10] Pasindu Senarathne, Malaka Silva, Ama Methmini, Dulaj Kavinda & Prof.Samantha Thelijjagoda, "Automate Traditional Interviewing Process Using Natural Language Processing and Machine Learning", International Conference for Convergence in Technology (I2CT), (2021)

[11] WEI Wei, Bin Wang, Beibei Zhang, Rafal Scherer & Robertas Damasevicius, "Online Job Search and Recruitment Platform for College Students Based on SSH", International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI),(2020)

[12] Muhammad Yusuf & Kemas M Lhaksmana, "An Automated Interview Grading System in Talent Recruitment using SVM", International Conference on Information and Communication Technology (ICOIACT), (2020)