ResumeTune - AI to Perfect Your Resume

Garv Makkar, Shubham Pal, Sachin Sharma, Parth Kaushal, Samyak Jain, Rishav Raj
Delhi, India

***Abstract*—In today's competitive job market, many job seekers face the challenge of aligning their resumes with job descriptions and ensuring they pass through Applicant Tracking Systems (ATS). Even qualified candidates often miss out on opportunities because their resumes don't effectively showcase the right skills and experience. This project aims to create an application that helps users improve their resumes using advanced Large Language Models (LLMs) and Natural Language Processing (NLP) techniques. By uploading their resume and a job description, users will receive personalized suggestions for improvement and an ATS-friendly version of their resume. The application will use LLMs, specifically trained on relevant data, to analyze resumes and job descriptions, identifying any gaps in skills or qualifications. The system will provide tailored feedback and utilize Reinforcement Learning with Human Feedback (RLHF) to continuously improve its recommendations. This ensures resumes not only align with job requirements but are also optimized for ATS, increasing the chances of landing a job. The effectiveness of the system will be measured through relevance, job-fit scores, and improved readability.**

***Index Terms*—resume optimization, large language models, natural language processing, applicant tracking systems, job market**

I. Objective

Our goal is to make an interactive application that would use the power of AI for assistance in building and improving Resumes and CVs for the end users, both for job-specific roles and for the overall profile of the User. This would help users not only in making job-specific resumes, but also to guide them for how they should improve themselves for achieving a specific goal in their career. Our application would offer this guidance in the form of recommending the skills which the user should work on, or the projects/works that the user could work on, and hence, show, to build upon a strong profile of the user for the career goals they pursue.

To achieve this, we want to utilize and integrate the powers of AI, specifically the state-of-the-art Language models to build our application, which can deeply analyze and concisely understand both the user's current profile and their prepared resume, and the job description, identifying the merits that user currently has, what they have displayed in their resume, and what might be missing or could be further improved, to fit the Job Description(JD) the user provides. We want our application to provide detailed recommendations and guidance based on contextual analysis and industry-specific language, making a final, optimized resume that could best fit the JD's requirement based on the user's current profile, thereby increasing their chances of success in job applications.

If we achieve this, we may extend our application to further guide the user about their career path by understanding the user's specific aspirations and their current qualifications' portfolio. This would come with our system analyzing and solving the gaps between the user's current standing and their goals, and then suggesting to them on what skills they can work on and what projects they can work on to gain their aspirations and achieve their goals.

II. Literature Review

The literature review highlights recent advancements in resume optimization through the use of Large Language Models (LLMs) and Natural Language Processing (NLP).

The paper "CareerBoost: Revolutionizing the Job Search with Resume Enhancement and Tailored Recommendations" by Asoke Nath et. al proposes a method that enhances the job search process using advanced NLP and machine learning techniques like Random Forest Classifier and K-means Clustering. It features a Resume Enhancer that evaluates resumes and job descriptions to recommend skills, and a Job Recommendation tool that matches job listings to users' qualifications [1].

The paper "Resume Building Application based on LLM" by R. J. Sunico et. al (2023) introduces a resume-building application that helps students, especially from underprivileged backgrounds, create effective resumes using a Large Language Model (LLM). The system converts natural language input into professional resume bullet points, assesses content for improvements, and ensures resumes follow the STAR (Situation, Task, Action, Result) format [2].

The paper "ResumeFlow: An LLM-facilitated Pipeline for Personalized Resume Generation and Refinement" by Zinjad et al. (2024) presents a system that automates resume creation by extracting information from user-provided resumes and job descriptions [3].

In "Comparative Analysis of ML and LLM in Resume Parsing: A Paradigm Shift in Talent Acquisition," the authors discuss the transition from traditional machine learning techniques to LLMs in the context of resume parsing [4].

The study "Generating Synthetic Resume Data with Large Language Models for Enhanced Job Description Classification" by Skondras, Zervas, and Tzimas (2023) explores the use of LLMs to create synthetic resume data aimed at improving job description classification [5].

The article "AI-based Automatic Resume Analysis" gives insights on how HRs of different companies use LLM-based AI applications for helping in their recruitment process [6].

These papers collectively contribute to the understanding of how LLMs can enhance resume optimization, but they also reveal gaps in practical application and user-centric design that the proposed project aims to address.

III. Approach

Our goal is to make an automated robust system that enhances resumes using advanced LLMs and NLP techniques. To create and refine our system we will be gathering all the relevant data (such as CV, JD etc.), using LLMs for parsing our data to analyze and understand the data for the solution, and then generating ATS-optimized resumes as the end result. Our approach is divided into three major phases:

A. Data Collection

The success of this project relies on data from both job descriptions and resumes to train, validate, and fine-tune the LLM-based system. Data sources will include:

- Job Postings: Large datasets of job postings across various industries and roles will be collected from publicly available job boards, APIs, or web scraping.
- Resumes: Diverse resumes covering various industries, experience levels, and job roles will be gathered from publicly available sources.
- ATS Guidelines: Data on Applicant Tracking Systems (ATS), including resume formats, keyword prioritization, and section ordering, will be used to ensure the final output meets industry standards for ATS compatibility.

B. Resume Parsing, Job Description Analysis, Gap Identification, and Enhancement Suggestions

In this phase, the resumes and job descriptions will be processed and analyzed using advanced LLMs, enhanced by various NLP techniques such as semantic role labeling, dependency parsing, and contextual understanding.

- Model Selection and Fine-Tuning: We will experiment with various open-source LLMs and select the best model for resume parsing and job description comprehension. The chosen LLM will undergo transfer learning, fine-tuned on domain-specific datasets such as job descriptions, resumes, and industry-related documents.
- Task-Specific Fine-Tuning: The LLM will be further fine-tuned for specific tasks such as extracting and parsing information from resumes, understanding job description requirements, and providing relevant suggestions for improvement.
- Gap Identification and Suggestions: The model will identify gaps in the resume compared to the job description, suggesting improvements such as adding missing skills or enhancing descriptions of existing experiences to match job requirements more closely.

To ensure the model evolves and stays current, we will incorporate Reinforcement Learning with Human Feedback (RLHF). By gathering user feedback through the application, we can refine the model iteratively. This feedback loop will allow for continuous updates and improvements, ensuring the system adapts to changing industry trends and user needs.

C. ATS Optimization and Resume Generation

Once the analysis is complete, the system will generate a resume optimized for ATS systems, ensuring both content and formatting are aligned with best practices.

- ATS Optimization: The model will follow the ATS guidelines by ensuring proper keyword usage, section order, and formatting. Key features such as optimized job titles, consistent formatting, and strategic use of keywords will help improve the resume's compatibility with ATS software.
- Final Resume Generation: After users review and accept the suggested enhancements, a final resume will be generated. This resume will be structured to highlight key achievements and orient closely with the job description.

IV. Timeline

A week-wise timeline:

- Week 0: Submit Project Proposal.
- Week 1: Collect the dataset and complete the literature survey.
- Week 2: Implement fine-tuning techniques to build the architecture and obtain baseline results.
- Weeks 3-4: Develop the application.
- Week 5: Refine the strategy and explore additional techniques to improve the model.
- Weeks 6-7: Compare various models and techniques to select the optimal approach.
- Weeks 8-9: Reserve for final improvements and address any outstanding tasks.
- Week 10: Final project submission.

V. Evaluation Criteria

Since this is an application-based project, the evaluation will be done along several dimensions:

- Relevance Precision/Recall: We shall evaluate how many relevant skills, qualifications, and experiences from the job description are included in the resume (precision) and how much of the resume is relevant to the job description (recall).
- Job Relevance Score: Evaluate the quality of resumes using reputed Automatic Resume Screening tools and Application Tracking Systems (ATS) and models that rank resumes based on job fit to compare the original and enhanced versions.
- Readability Score (Flesch-Kincaid): The Flesch-Kincaid Readability Score is used to evaluate how easy a document is to read. It gives a score based on sentence length and the number of syllables per word.
- Pre-trained LLM Baselines: We shall compare our model's performance against baseline pretrained models like Llama3, T5, or Gemma.
- Pre-existing Resume-Job Match Models: We also aim to compare performance with resume parsing and job-matching systems like Jobscan or HireVue.

- Semantic Similarity: Semantic similarity measures how closely the enhanced resume aligns with the job description based on meaning. Common metrics include cosine similarity between text embeddings or BLEU score for comparing n-gram overlap.

By applying these methods, we intend to thoroughly evaluate and refine the performance of our resume enhancement LLM and compare its performance with existing solutions and open-source models.

REFERENCES

[1] A. Nath, S. Saha, S. D. Sarkar, and A. Bose, "CareerBoost: Revolutionizing the Job Search with Resume Enhancement and Tailored Recommendations," International Journal of Scientific Research in Computer Science, Engineering and Information Technology, vol. 10, no. 3, pp. 281-290, 2024.

[2] R. J. Sunico et al., "Resume Building Application based on LLM (Large Language Model)," 2023 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), Greater Noida, India, 2023, pp. 486-492.

[3] S. B. Zinjad, A. Bhattacharjee, A. Bhilegaonkar, and H. Liu, "Resume-Flow: An LLM-facilitated pipeline for personalized resume generation and refinement," Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2024.

[4] E. Kaygin, "Comparative Analysis of ML (Machine Learning) and LLM (Large Language Models) in Resume Parsing: A Paradigm Shift in Talent Acquisition," 2023.

[5] P. Skondras, P. Zervas, and G. Tzimas, "Generating synthetic resume data with large language models for enhanced job description classification," Future Internet, vol. 15, no. 11, p. 363, 2023.

[6] L. J. R&D, "AI-based Automatic Resume Analysis," Medium, Apr. 30, 2022.