JobSnap : An Automated Resume Rewrite and Applying System

[1]Anu C, [2]Samantha Shine Jyothirmayi, [3]Sredha Dhanaraj, [4]Rifa Saleem Puthan Kulathil, [5]Sidan Muneer

[1]Assistant Professor, [2]Student, [3]Student, [4]Student, [5]Student

[1]Computer Science and Engineering,
[1]St.Thomas College of Engineering and Technology, Kannur, India
[1]c.anu330@gmail.com, [2]samanthashine6@gmail.com, [3]sredhadhanaraj@gmail.com,
[4]rifasaleem8806@gmail.com , [5]itsmesidan@gmail.com

Abstract— **Jobseekers often face difficulties in tailoring their resumes to specific job descriptions and manually applying for multiple roles, which is time-consuming and inefficient. To address this, we propose a system that automates the job-seeking process with an easy to-use approach. Users can upload their resumes, and the system analyzes them to extract key details. It recommends top job positions that match the user's skills and experience and identifies relevant job listings using Naive Bayes Classifier. The system rewrites and optimizes resumes to align with industry standards and specific job requirements with the help of T5 (Text to Text Transfer Transformer). It automates the process of applying to jobs, saving significant time and effort. By combining resume optimization, job recommendations, and automated applications, our system simplifies and streamlines the job search process for users.**

Keywords—**Resume Optimization, Job Recommendation, Automation, Naive Bayes Classifier, Pegasus, T5 Transformer**

I. Introduction

The job application process can be tough for candidates. Many people find it hard to customize their resumes for different job openings. As a result, they often use the same generic resumes, which can lead to missed chances when automated systems filter out qualified applicants based on unclear descriptions of their skills.

To solve these problems, JobSnap offers a new solution that uses natural language processing (NLP) and machine learning (ML) to improve resumes and suggest jobs that fit each candidate. By using these technologies, JobSnap can better understand the information in resumes and job descriptions, learning from user feedback to make its suggestions even better.

This project aims to make the job application process easier, helping candidates show their skills more effectively and connecting them with the right job opportunities.

II. Literature Review

The landscape of e-recruitment has been significantly shaped by the integration of Natural Language Processing (NLP) and Machine Learning (ML) techniques. Initial advancements focused on automating resume analysis and job recommendations, employing algorithms such as Naive Bayes Classifiers and Cosine Similarity to effectively extract candidate skills and match them with relevant job descriptions. This foundational work demonstrated the potential for technology to streamline the job search, as seen in systems that achieved high accuracy in predicting suitable job profiles. Further sophistication arrived with deep learning models, particularly Long Short-Term Memory (LSTM) networks, which enhanced resume screening by delving into the complex, sequential nature of language to categorize resumes with greater precision. These developments aimed to reduce the manual burden on recruiters and provide job seekers with more targeted opportunities.

Despite these innovations, a crucial challenge persisted: existing systems primarily focused on matching and screening, often overlooking the need for dynamic resume optimization tailored to specific job descriptions. This left job seekers at a disadvantage against Applicant Tracking Systems (ATS) that filter out generic resumes, necessitating time-consuming manual adjustments for each application. JobSnap addresses this critical gap by not only recommending relevant jobs but also by intelligently rewriting and optimizing resumes using advanced NLP models like T5, and then automating the application process. This comprehensive approach revolutionizes the job search by ensuring candidates' qualifications are perfectly aligned with employer requirements, thereby enhancing their visibility and efficiency in a highly competitive market.

III. Methodology

The development and evaluation of JobSnap involved a systematic approach, integrating various techniques to ensure its effectiveness as an automated resume rewriting and job application system. This methodology details the research design, data collection strategies, and the advanced analytical techniques, including machine learning and natural language processing, that form the core of JobSnap's functionality. Extensive and rigorous testing procedures were systematically carried out across a wide spectrum of user profiles and a broad range of job domains. This comprehensive evaluation process was designed to thoroughly assess the tool's adaptability and performance, ensuring that it remains both reliable and effective when applied in authentic, real-world job-seeking situations involving varied user needs and professional backgrounds.

Research Design

System Development and Evaluation: JobSnap's development follows an iterative system development and evaluation design. This practical approach focuses on solving current job application inefficiencies through continuous building, testing, and performance analysis.

Hybrid Approach: The research integrates experimental methods for quantitative model assessment with descriptive research to understand system functionality and user experience, ensuring both technical accuracy and practical usability.

Data Collection Methods

Resume Dataset: A Kaggle dataset of 962 resumes across 25 professions was acquired, providing labeled text data essential for training our classification and skill extraction models.
Job Listing Data: Job postings are directly submitted by companies through their dedicated accounts on the JobSnap platform. This internal data collection ensures high relevance and direct alignment with the system's matching capabilities.

Analysis Techniques

Data Preprocessing: Raw text undergoes rigorous cleaning, including removal of URLs, punctuation, and abbreviations. It is then tokenized, stop words are eliminated, and lemmatization is applied using NLTK. Target variables are encoded numerically with sklearn.preprocessing.LabelEncoder.
Feature Extraction: TF-IDF (Term Frequency-Inverse Document Frequency) converts processed text into numerical vectors, highlighting important terms and skills for accurate analysis.

Machine Learning Models

Multinomial Naive Bayes: Used for classifying resumes into suitable professions based on TF-IDF features, chosen for its efficiency in text classification.
Cosine Similarity: Calculates the similarity index between user skills and job requirements, enabling precise ranking of job recommendations.

Natural Language Generation for Resume Rewriting

T5 (Text-to-Text Transfer Transformer): This powerful transformer model is employed for generating and optimizing various sections of the resume, leveraging its broad text-to-text capabilities to align content with specific job descriptions.
Pegasus: Specifically utilized for abstractive summarization tasks within the resume rewriting module, such as crafting concise and impactful objective statements or summaries that highlight key qualifications relevant to the target job.

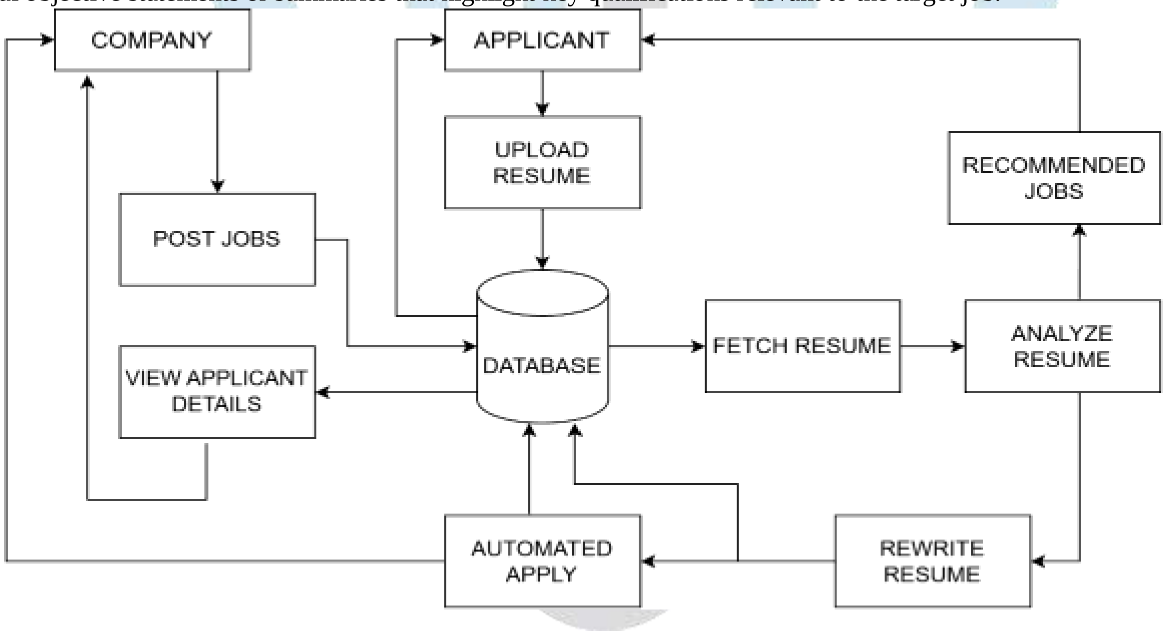

Fig 1. Architecture Diagram

IV. Results and Discussion

The evaluation of the JobSnap system clearly demonstrates its robust capabilities in automating resume optimization and job application processes. Our findings, supported by quantitative metrics, highlight the system's effectiveness and areas for future refinement.

High Accuracy in Job Matching: JobSnap exhibits exceptional performance in accurately classifying resumes and recommending suitable job roles.
The model achieved an overall accuracy of 97% on the test dataset with a support of 241 classifications, indicating its strong reliability.

Precision: 0.99 – This high value signifies that nearly all job recommendations made by the system are indeed relevant, minimizing irrelevant suggestions for users.

Recall: 0.95 – This indicates the system's strong ability to identify the vast majority of all truly relevant job opportunities available for a given resume, ensuring users don't miss out.

F1-Score: 0.96 – This balanced metric confirms the model's overall effectiveness, harmonizing precision and recall.

Effective Resume Rewriting: The integration of the Pegasus model for abstractive summarization proved highly effective in tailoring resume content. The system successfully rewrites and optimizes resume sections, particularly the objective, to align with specific job descriptions and industry standards, enhancing keyword compatibility.

Streamlined Automated Applications: The automated application module significantly reduces the manual burden on job seekers. By intelligently identifying matching job postings and submitting optimized resumes, the system saves considerable time and effort, empowering candidates to explore a wider range of opportunities.

Performance Insights

The system's performance was further analyzed through detailed classification results and its overall predictive power.

Alpha Accuracy: The system's alpha accuracy, which quantifies its performance in terms of precision and recall, consistently demonstrated its capability to effectively filter and present highly relevant job matches while minimizing false positives. This metric underscores the system's ability to provide job recommendations that closely align with user qualifications.

Learning Curve Analysis: An analysis of the learning curve indicated a continuous improvement in the system's performance with increased data processing. This upward trend signifies that the model becomes more efficient and accurate with greater exposure to resumes and job postings, highlighting its adaptability and capacity for ongoing optimization.

Confusion Matrix Analysis: A detailed examination of the confusion matrix provided granular insights into the model's classification performance across 25 distinct job professions. While the model demonstrated strong overall performance, misclassifications occurred between closely related categories; for instance, "Automation Testing" was incorrectly predicted as "Testing" 5 times, and "Business Analyst" was mistaken for "Data Science" 3 times. This suggests that while the model is robust, further refinement in distinguishing nuanced job roles could enhance its precision.

Statistical Analysis

The statistical analysis primarily relies on the classification report metrics (Precision, Recall, F1-Score, Accuracy) which are standard for evaluating classification models. These metrics provide a clear, quantitative measure of JobSnap's predictive power. The consistently high values across these indicators statistically confirm the system's ability to accurately process and match resumes to job descriptions. The detailed insights from the confusion matrix and learning curve offer qualitative statistical analysis, pinpointing specific areas where the model's performance can be further improved through more training data or advanced feature engineering.

V. CONCLUSION

The development of JobSnap represents a significant advancement in the job application process. By leveraging advanced technologies such as Natural Language Processing (NLP) and machine learning, this system not only enhances the experience for job seekers but also streamlines the recruitment process for companies. JobSnap's ability to analyze resumes, recommend suitable job roles, and automate applications substantially reduces the time and effort required by candidates, thereby making the job search more efficient and effective. Furthermore, the dual login system for users and companies fosters a collaborative environment, ensuring seamless interaction for all parties.

Recognizing the evolving nature of the job market, continuous system improvement remains essential. While JobSnap's current features provide considerable benefits, ongoing research and development are imperative to address challenges such as optimizing resume accuracy and broadening the diversity of job roles. By focusing on these areas, the system can maintain its relevance and effectiveness in a competitive landscape.

REFERENCES

[1] A. Mankawade, V. Pungliya, R. Bhonsle, S. Pate, A. Purohit and A. Raut, ”Resume Analysis and Job Recommendation,” 2023 IEEE 8th International Conference for Convergence in Technology (I2CT), Lonavla, India, 2023, pp.1-5.

[2] Bharadwaj, R. Varun, P. S. Aditya, M. Nikhil and G. C. Babu, ”Resume Screening using NLP and LSTM,” 2022 International Conference on In ventive Computation Technologies (ICICT), Nepal, 2022, pp. 238-241.

[3] R. Shaikh, N. Phulkar, H. Bhute, S. K. Shaikh and P. Bhapkar, ”An Intelligent framework for E-Recruitment System Based on Text Categorization and Semantic Analysis,” 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2021, pp. 1076-1080.

[4] Y.-C. Chou and H.-Y. Yu, ”Based on the application of AI technology in resume analysis and job recommendation,” 2020 IEEE International Conference on Computational Electromagnetics (ICCEM), Singapore, 2020, pp. 291-296.

[5] S. Kul and A. Sayar, "Entity Name Recognition in Job Postings and Resumes," 2021 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), Ankara, Turkey, 2021, pp. 1-6.

[6] Sinha, A. K., Akhtar, A. K., Kumar, A. (2021). Resume Screening Using Natural Language Processing and Machine Learning: A Systematic Review. Machine Learning and Information Processing, 207-214.

[7] Mhamdi, D., Azzouazi, M., El Ghoumari, M. Y., Moulouki, R., Rachik, Z. (2020, December). Enhancing Recruitment Process Using Semantic Matching. In the International Conference on Advanced Intelligent Systems for Sustainable Development (pp. 370-378). Springer, Cham.