INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

Automated Resume Analysis with NLP Techniques

Lubdha Borole, Samadhan Zagare, Divya Patil, Saurabh Berad, Priti Sharma

Department of Computer Engineering, SSBT COET, Bambhori, Jalgaon, Maharashtra, India

ABSTRACT: In the era of digital transformation, efficient recruitment processes are pivotal for both applicants and recruiters. This research introduces an "Resume Analyzer System" powered by Natural Language Processing (NLP) to revolutionize candidate screening. By automatically parsing resumes and identifying key keywords using Support Vector Machine (SVM) and Decision Tree algorithms, the system categorizes candidates based on job suitability. Through meticulous data preprocessing and model training, the analyzer significantly reduces manual effort while ensuring high accuracy. This student-led initiative highlights the transformative potential of NLP and machine learning in enhancing recruitment efficiency and transparency, benefiting both employers and candidates alike.

KEYWORDS: Resume Analyzer, Natural Language Processing (NLP), Support Vector Machine (SVM), Decision Tree algorithms.

I. INTRODUCTION

In today's competitive job market, the need for efficient resume screening and evaluation has grown significantly. Both students seeking opportunities and employers aiming to identify the right candidates face challenges in this process. Traditional methods often consume excessive time and are susceptible to oversights, potentially leading to mismatches between candidates and job roles.

This study presents a novel Resume Parsing and Analysis System tailored for students. Leveraging advanced Natural Language Processing (NLP), Support Vector Machines (SVM), and Decision Trees, the system automates the parsing, analysis, and scoring of resumes in comparison to job descriptions. A standout feature of this system is its provision of a resume score for students, coupled with actionable suggestions to enhance their resumes based on the specific job description.

The main goal of this research is to equip students with a tool that simplifies resume screening and provides valuable feedback for improvement. By harnessing the capabilities of NLP, SVM, and Decision Trees, the system aims to make resume evaluation more efficient, accurate, and equitable, thus helping students better align their resumes with job requirements. As the job market continues to evolve towards a more data-driven and automated approach, the integration of advanced technologies in resume parsing and analysis is set to redefine the recruitment landscape. This research aims to make a significant contribution to this field by introducing a user-friendly and effective tool tailored for students, helping them to navigate their career paths with confidence and success.

II. LITERATURE SURVEY

In paper[1] "Resumate: A Prototype to Enhance Recruitment Process with NLP based Resume Parsing" 2023 by S. Mohanty, A. Behera, S. Mishra, A. Alkhayyat, D. Gupta and V. Sharma. highlights the use of natural language processing and machine learning to automate resume review and categorization. The study shows that SVM and XG Boost are the most effective models, with XG Boost outperforming others, offering a novel and unique approach to this task.

Inpaper[2] Resume Parser Using NLP Techniques By Shubham Bhor , Vivek Gupta, Vishak Nair, Harish Shinde, Prof. Manasi S.Kulkarni proposed the approach is to make the work of companies and candidates easier and effective. Basically our aim is to ease the recruitment process. The process will provide the quality of applicants for the companies. The unfair and discriminatory practice in the process will be dampened. Based on the information in the form oftechnical skills the resumes will be ranked in order.

In paper[3] Combination of Neural Networks and Conditional Random Fields for Efficient Resume Parsing By Ayisha Thahira. Research explores successfully segmented and extracted various types of information from personal, educational, and occupational sections of resumes. The output JSON file is generated and provides valuable insights for the recruitment process.

III. SYSTEM ARCHITECTURE

The system architecture for the resume parsing and analysis system is designed to offer a comprehensive solution for students to evaluate and optimize their resumes in alignment with job requirements.

User Interface: The user interface facilitates seamless interaction, allowing users to easily upload their resumes and job descriptions.

Extraction: The resume and job description are parsed to extract skills, experience, education, and other relevant details utilizing techniques such as text extraction, keyword matching to accurately capture skills, work experience, education, and other pertinent information from the documents. The parsed information from the job description acts as a benchmark against which the resume's extracted data is compared, setting the foundation for subsequent analysis and evaluation steps in our system.

Normalization Data: Standardize the extracted information to ensure consistency and facilitate accurate comparison. This involves converting data into a uniform format. and resolving variations in naming conventions, abbreviations, and formats. By normalizing the data, we ensure that the comparison is fair and accurate, providing reliable results for the subsequent analysis and scoring steps in our system.

Clustering: The extracted skills from the resume and job description are grouped into clusters based on similarity to identify patterns and relationships within the extracted information. This helps in identifying related skills and their importance in the job role.

Analysis: Natural Language Processing (NLP) techniques is used to compare the extracted data. Analyses collectively help gauge how well the resume matches the job description, providing insights into potential areas of alignment or improvement.

Ranking and Score module: Score is calculated based on skill, experience, and education matches, with weights reflecting their importance to the job. Personalized suggestions are generated based on the analysis results to enhance the resume's alignment with the job requirements, providing insights for improvement.
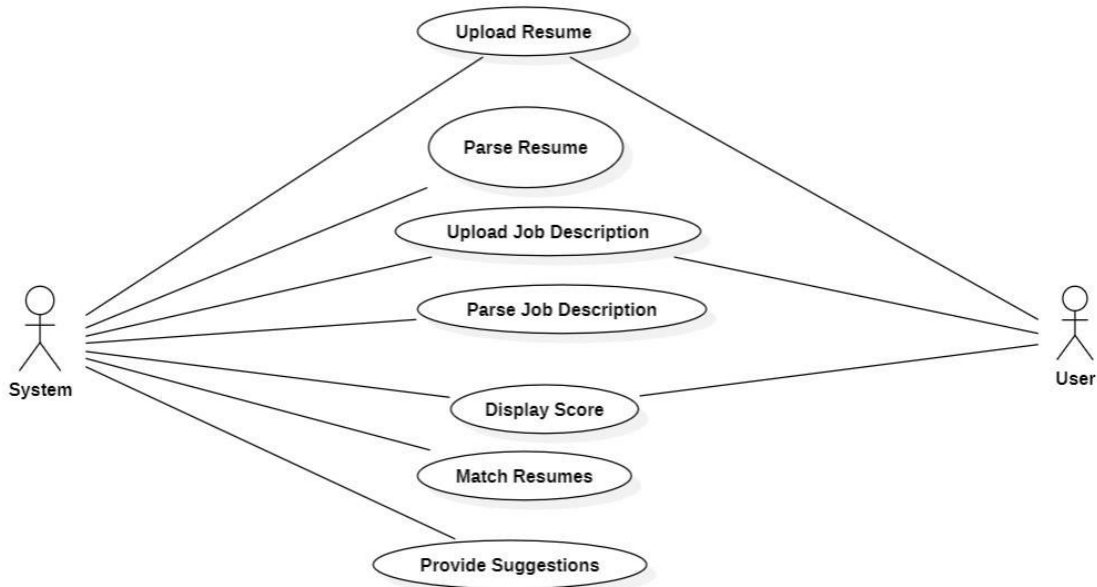
IV. PROPOSED SYSTEM
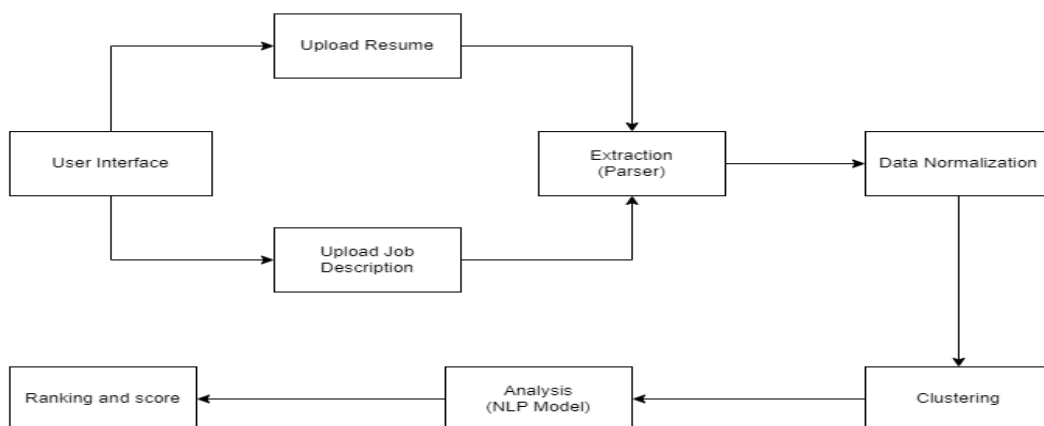
Our proposed system integrates advanced NLP techniques, clustering, and scoring mechanisms to offer a comprehensive solution for resume parsing and analysis. It includes text extraction, data normalization, and skill and experience clustering for accurate information extraction and comparison. NLP-driven analysis matches resume details with job requirements, while a composite scoring system ranks resumes and generates personalized improvement suggestions. This system aims to enhance efficiency and effectiveness in resume optimization and recruitment processes.

The model uses natural language processing to understand the resume and then parse the information from it. Once information is parsed it is stored in the database. Quite productive for applicants because it gives predictions, tips and recommendations based on their resume information. System works proper when the uploaded resume is in traditional chronological.

V. USECASE DIAGRAM

VI. ALGORITHM

Our proposed system incorporates the following algorithms:
Support Vector Machine (SVM) is a supervised machine learning algorithm used for classification and regression tasks. In the context of our system, SVM can be utilized for skill and experience matching by classifying resumes based on their features to determine their relevance to a job description. SVM works by finding the hyperplane that best separates the data into different classes, maximizing the margin between them. It is effective for handling high-dimensional data and can be trained to categorize resumes accurately by learning from labeled training examples. The decision function for an SVM classifier can be represented as:

$$f(x)=\text{sign}(\sum_{i=1}^{n}\alpha_i y_i K(x,x_i)+b)$$

Where:
- α_i are the dual coefficients
- y_i are the labels
- $K(x,x_i)$ is the kernel function
- b is the bias term

Decision Tree is a supervised learning algorithm used for both classification and regression tasks. A decision tree consists of nodes that represent feature tests, branches representing decision rules, and leaf nodes representing the

outcome or target variable. In our system, decision trees can be used for ranking resumes based on their feature values, such as skills, experience, and education. Decision trees are interpretable and easy to understand, making them suitable for generating transparent and explainable ranking criteria. They recursively partition the data into subsets based on feature values to make decisions, making them effective for resume ranking and classification tasks in our system

VII. IMPLEMENTATION

The implementation phase involved translating the system design and technology stack into a functional application. Below is a detailed explanation of the implementation process and the technologies used.

Frontend Development: HTML and CSS are utilized to create the user interface, providing a user-friendly and responsive design for seamless navigation.

Backend Development: Python served as the backend language, handling the core logic, while Streamlit facilitated the integration with the frontend to create interactive web applications.

Database Management: MySQL is used for database management, providing a robust and scalable solution for storing and managing resume data, user profiles, and system configurations.

Development Environment: Visual Studio Code is the selected IDE for development, offering a suite of tools including a debugger, compiler, and text editor for efficient coding and testing.

Core Libraries Utilized: The project utilized Streamlit for web application development, Pandas for data manipulation, PyMySQL for database interactions, and Plotly for data visualization. ResumeParser and PDFMiner3 handled resume PDF parsing, Geopy's Nominatim managed geocoding, PIL was employed for image processing, and NLTK for natural language processing tasks.

VIII. RESULT

The outcomes include performance metrics for algorithms such as SVM and Decision Tree in ranking resumes and generating personalized suggestions. We evaluated the system using a dataset of resumes and job descriptions, measuring accuracy, precision, and recall as evaluation metrics. The output showcases ranked resumes with scores and personalized improvement suggestions. Our findings indicate that SVM and Decision Tree algorithms effectively categorize and rank resumes, with Decision Tree demonstrating higher interpretability and SVM showing better accuracy in certain tasks
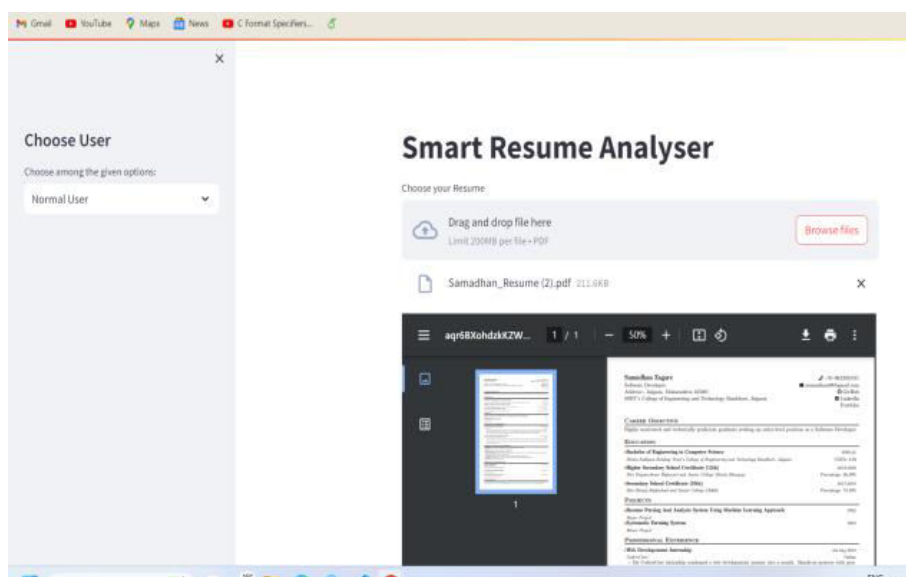
OUTPUT:-
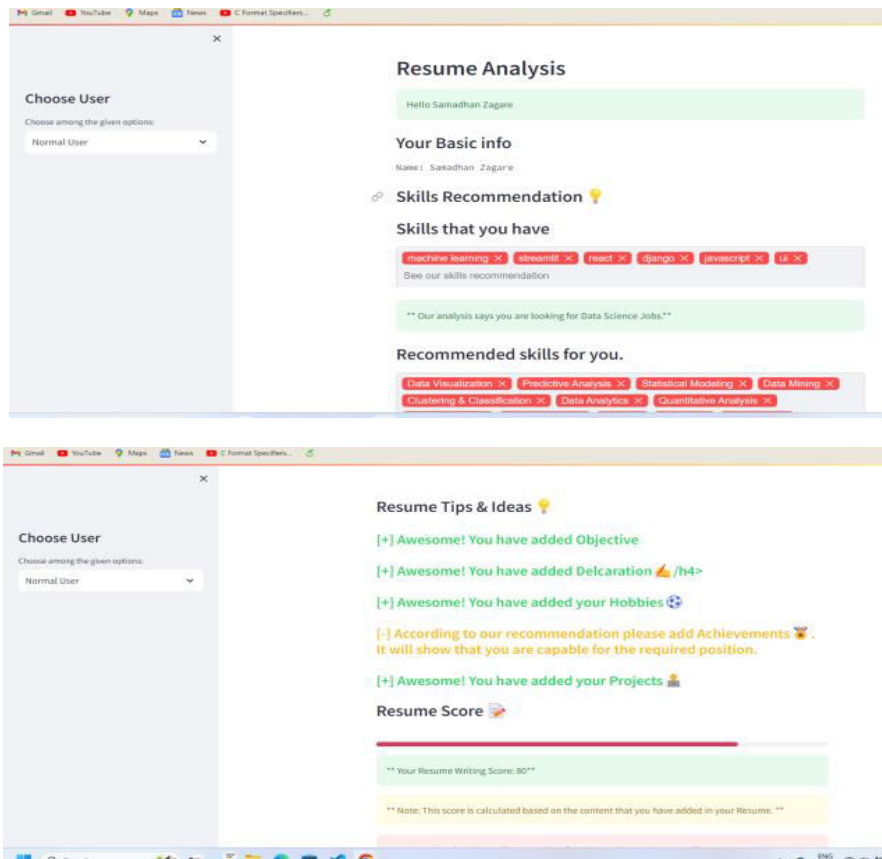
IX. CONCLUSION

In conclusion an applicant cum recruiter-based Quick and easy to use Resume Analyzer. That analyze resume data and extract it into machine-readable output. Helps applicants with recommendations, prediction and analytics. Helps recruiter by automatically store, organize, andanalyze resume data to find the best candidate. Can be widely used by any organization to analyze and get insights of a resume. This project underscores the power of cutting-edge technology in revolutionizing human resources practices, making the hiring process more intelligent, swift, and data-driven.

REFERENCES

1. "Resumate: A Prototype to Enhance Recruitment Process with NLP based Resume Parsing" 2023 4th International Conference on Intelligent Engineering and Management (ICIEM) By S. Mohanty, A. Behera, S. Mishra, A.Alkhayyat, D. Gupta and V. Sharma.
2. Resume Parser Using NLP Techniques By Shubham Bhor, Vivek Gupta, Vishak Nair, Harish Shinde, Prof. Manasi S.Kulkarni
3. Combination of Neural Networks and Conditional Random Fields for Efficient Resume Parsing By Ayisha Thahira
4. A Survey on Resume Parsing and Job Matching Techniques" by A. Choudhury and R. Dhar. (2019).
5. RESUME PARSER AND ANALYZER USING NLP By Nimish Patil, Shubham Yadav, Vikas Biradar (2023).
6. "NLP based Extraction of Relevant Resume using Machine Learning ", IJIITEE By Nirali Bhaliya, Jay Gandhi, Dheeraj Kumar Singh.(2020)
7. "A CV parser Model using Entity Extraction Process and Big Data Tools ",IJIITCS by Papiya Das, Manjusha Pandey, Siddharth Swarup Rautaray.(2018)

ISSN
INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

INNO SPACE
SJIF Scientific Journal Impact Factor

doi
cross ref

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL
OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

📱 9940 572 462 ⬛ 6381 907 438 ✉ ijircce@gmail.com

Scan to save the contact details