



K.RAMAKRISHNAN
COLLEGE OF ENGINEERING

An Autonomous Institution

Permanently Affiliated to Anna University Chennai, Approved by AICTE New Delhi,
ISO 9001:2015, 14001:2015 certified institution, Accredited by NBA and with A grade by NAAC

Samayapuram, Tiruchirappalli – 621 112, Tamilnadu, India.

Cli



RESUME PARSER SYSTEM

PRESENTED BY

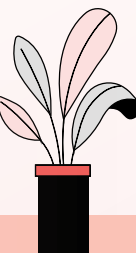
NAME : SWATHI S

REGISTER NO :8115U23AM53

ROLL NO :AM2353

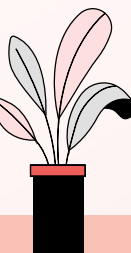
YEAR :II

Guided By:M.KAVITHA



TABLES OF CONTENTS

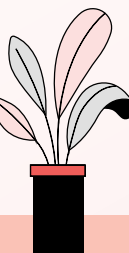
- ABSTRACT
- INTRODUCTION
- OBJECTIVE
- LITERATURE SURVEY
- EXISTING SYSTEM
- PROPOSED SYSTEM
- SYSTEM ARCHITECTURE
- MODULES DESCRIPTION
- IMPLEMENTATION SCREENSHOT
- CONCLUSION
- FUTURE ENHANCEMENT
- REFERENCES





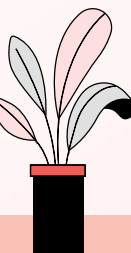
ABSTRACT

- The Resume Parser Project is an advanced system designed to automate the extraction, organization, and analysis of information from resumes. Utilizing Natural Language Processing (NLP) techniques and machine learning algorithms, the project processes resumes in various formats (PDF, Word, etc.) and extracts key details such as personal information, educational qualifications, work experience, skills, and certifications.
- The parser provides a structured output that can be easily integrated into Human Resource Management Systems (HRMS) or Applicant Tracking Systems (ATS), significantly reducing the time and effort required for manual resume screening.



INTRODUCTION

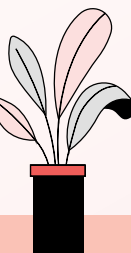
- In the modern recruitment process, organizations receive a large volume of resumes, making manual screening a time-consuming and labor-intensive task. A Resume Parser serves as an intelligent solution to streamline this process by automating the extraction and structuring of data from resumes.
- The Resume Parser Project leverages advanced technologies such as Natural Language Processing (NLP) and machine learning to efficiently analyze resumes in various formats, including PDF, Word, and text files.
- This project not only speeds up the recruitment process but also reduces errors associated with manual data entry. By standardizing resume data, it enables recruiters to focus on candidate evaluation rather than administrative tasks.





OBJECTIVE

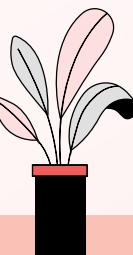
- The primary objective of the Resume Parser Project is to automate and optimize the recruitment process by accurately extracting, organizing, and presenting relevant information from resumes.
- To reduce the time and effort involved in manually reviewing and processing resumes by providing a structured output of candidate details.
- To enable faster shortlisting of candidates by integrating parsed data into Applicant Tracking Systems (ATS) or other HR tools.
- To minimize human errors in data entry and ensure consistent extraction of information regardless of resume format or template.
- To facilitate future enhancements like automated candidate-job matching, ranking candidates, and generating insights for strategic hiring decisions.





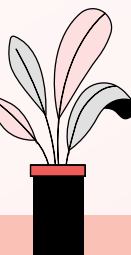
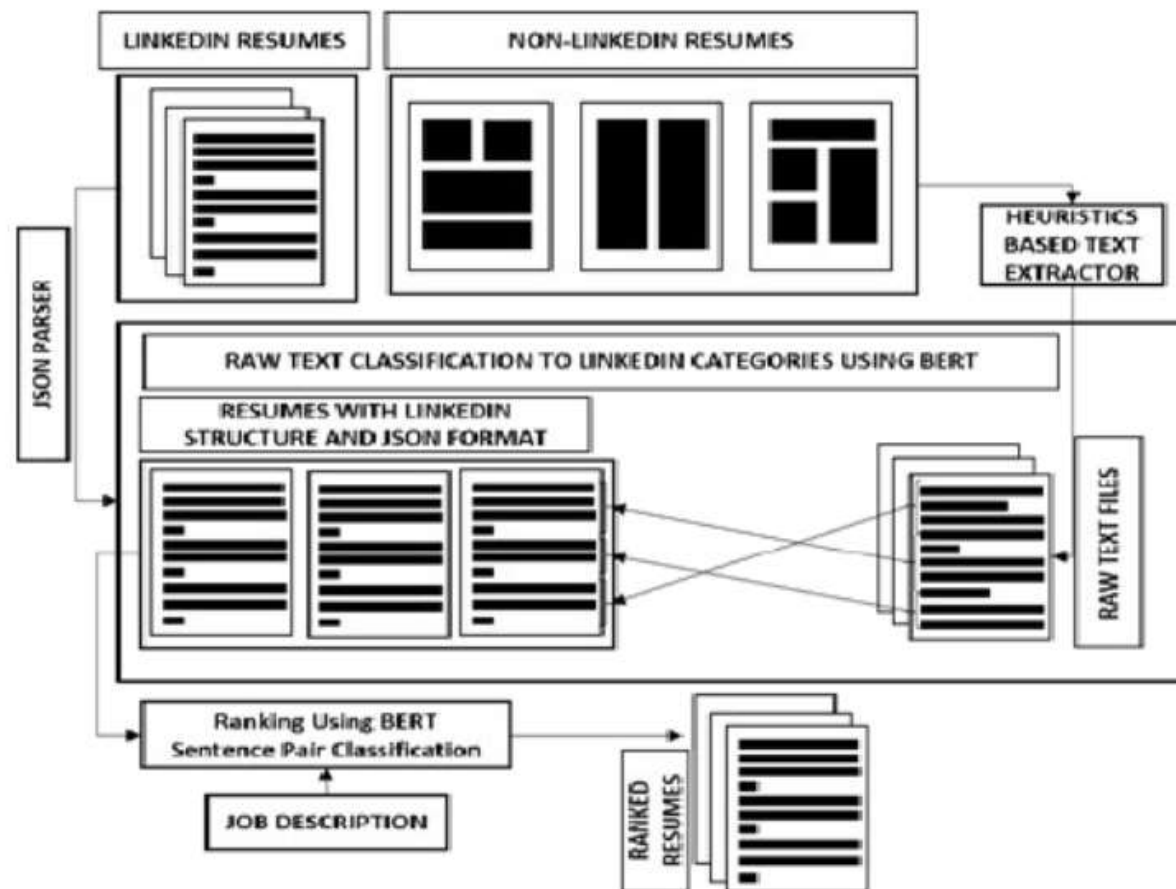
LITERATURE SURVEY

4	<i>Alharthi, Alhothali & Moria (2021)</i>	More than 10,000 Arabic tweets collected with Twitter API	Long Short Term Memory (LSTM) with word embedding feature representation	Time requirement to classify the tweets is very less compared to the state-of-the art methods	System classification accuracy depends on tweet length	Accuracy-0.97 Precision-0.98 Recall-0.95 F1-score-0.97
5	<i>Liu, Pang & Wang (2019)</i>	97,839 Restaurant (RES) and 31,317 Hotel review dataset (HOS)	Machine Learning (ML) techniques and Bi-LSTM	Could capture sophisticated spammer activities using multimodal neural network model	There is a need to analyze the use of other effective features to improve the performance	Recall-0.80 Precision-0.82 F1-score-0.81
6	<i>Fuslier et al. (2015)</i>	Hotel review corpus consisting of 1,600 reviews	SVM, K-Nearest Neighbor and Naïve Bayes (NB)	Lexical content and stylistic information were captured better using character n-grams	Need to build a hybrid feature set combining character and word n-grams	F1-score-0.87
7	<i>Wu et al. (2017)</i>	10 day real-life Twitter dataset of 1,376,206 spam and 6,73,836 non-spam tweets	RF, Multi-Layer Perceptron (MLP) and Naïve Bayes	Variations in spamming activities are captured within a short span of time.	The model needs to be adaptable to new characteristics	Accuracy-99.35 Recall-91.03% Precision-95.84% F-measure-93.37%





EXISTING SYSTEM

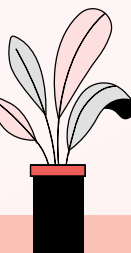


: PROPOSED SYSTEM

The proposed Resume Parser System is an intelligent, automated solution designed to extract, organize, and present resume information in a structured format. The system leverages advanced technologies to address the inefficiencies and challenges of manual resume screening. Below are the key components and functionalities of the proposed system:

System Features

1. Automated Data Extraction
 - Extracts essential details such as personal information, contact details, education, work experience, skills, certifications, and project highlights.
2. Multi-format Support
 - Handles resumes in various formats, including PDF, Word, text, and HTML.
3. Template Independence
 - Processes resumes regardless of layout or formatting style using Natural Language Processing (NLP).



PROPOSED SYSTEM

3. Multilingual Capability

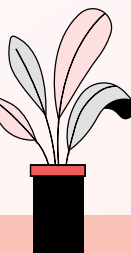
- Supports resumes in multiple languages to accommodate diverse candidate pools.

4. Error Handling

- Identifies and manages ambiguities or incomplete information in resumes.

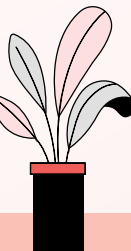
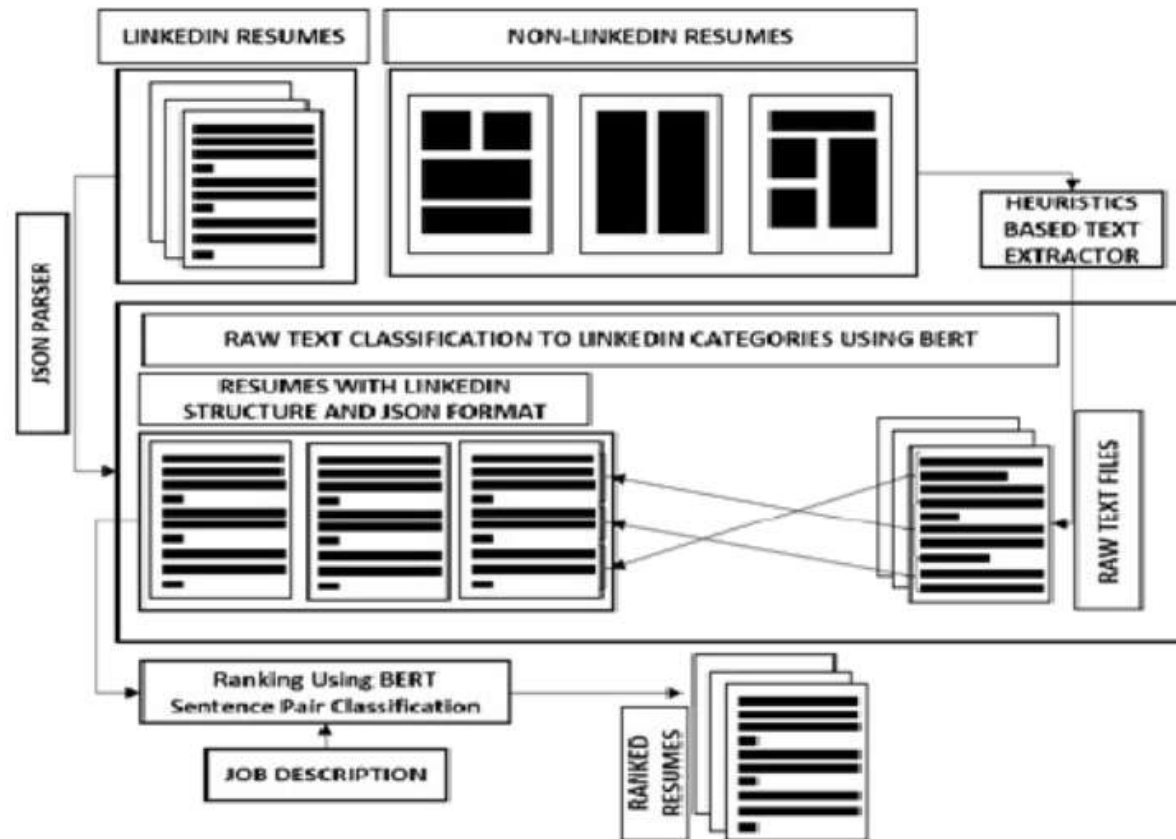
5. Scalability

- Designed to process large volumes of resumes efficiently, suitable for enterprise-level recruitment.





SYSTEM ARCHITECTURE





MODULES DESCRIPTION

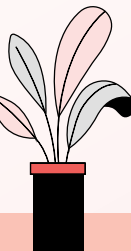
1. Input Module

2. Parsing Engine

3. Data Validation Module

4. Output Module

5. Multilingual Processing Module



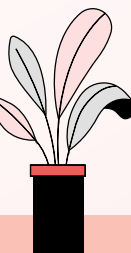


MODULE 1

1. Input Module

Purpose: Handles the uploading and pre-processing of resumes.

- Features:
 - Accepts resumes in various formats such as PDF, Word, and text.
 - Ensures compatibility with different file sizes and structures.
 - Pre-processes resumes (e.g., text extraction from PDF/Word files).



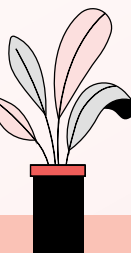


MODULE 2

2. Parsing Engine

Purpose: Extracts structured data from resumes.

- Features:
 - Utilizes Natural Language Processing (NLP) to analyze and segment resume content.
 - Personal Details: Name, contact information, address.
 - Educational Qualifications: Degrees, institutions, graduation dates.
 - Work Experience: Job titles, organizations, duration, responsibilities.
 - Skills: Technical and soft skills.
 - Certifications and Projects: Achievements and portfolio highlights.
 - Handles various templates and unstructured layouts.



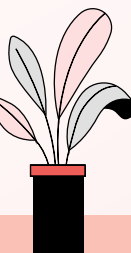


MODULE 3

3. Data Validation Module

Purpose: Validates the extracted data to ensure accuracy and completeness.

- Features:
 - Checks for missing or incomplete fields (e.g., missing email addresses or invalid dates).
 - Uses machine learning models to identify and correct extraction errors.
 - Ensures data is normalized (e.g., consistent formats for dates, phone numbers).



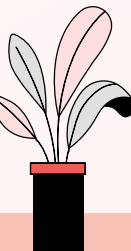


MODULE 4

4. Output Module

Purpose: Formats and delivers the parsed data.

- Features:
 - Converts extracted data into structured formats like JSON, XML, or database entries.
 - Provides options to download parsed data or directly integrate with Applicant Tracking Systems (ATS) or Human Resource Management Systems (HRMS).



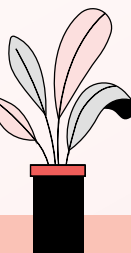


MODULE 5

5. Multilingual Processing Module

Purpose: Supports resumes in different languages.

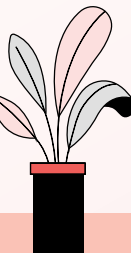
- Features:
 - Leverages NLP techniques for language detection.
 - Processes content in multiple languages and translates if required.





IMPLEMENTATION

Test Data:	Test Data	Expected Result	Actual Result
	Upload the resume	Show name, email, mobile number, technical skills, soft skills, education, Language, experience, address	Shown name, email, mobile number, technical skills, soft skills, education, Language, experience, address
Test Results	The above result indicates that the necessary information from the resume has been parsed.		





RESULT

+ Code + Text

✓ 51s

```
print("Unsupported file format. Please upload a PDF or DOCX.")
return

# Parse resume
parsed_data = parse_resume(text)

# Display in application format
display_application_format(parsed_data)

if __name__ == "__main__":
    main()
```

Enable browser notifications in Settings to get alerts when executions complete

OK No thanks

↩

Requirement already satisfied: PyPDF2 in /usr/local/lib/python3.10/dist-packages (3.0.1)
Requirement already satisfied: python-docx in /usr/local/lib/python3.10/dist-packages (1.1.2)
Requirement already satisfied: lxml>=3.1.0 in /usr/local/lib/python3.10/dist-packages (from python-docx) (5.3.0)
Requirement already satisfied: typing-extensions>=4.9.0 in /usr/local/lib/python3.10/dist-packages (from python-docx) (4.12.2)
Upload your resume file (PDF or DOCX):
Choose Files kishorkumar.pdf

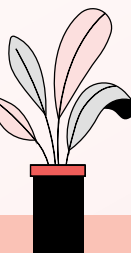
- **kishorkumar.pdf**(application/pdf) - 107936 bytes, last modified: 11/29/2024 - 100% done

Saving kishorkumar.pdf to kishorkumar.pdf

--- Job Application ---

Name: KISHOR KUMAR. M
Email: kishor.m2024@outlook.com
Phone: 90187530591
Skills: ADAS, Embedded C and C++, Microcontroller, Embedded Developer, Verification and
Experience: Not Found

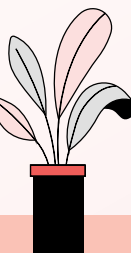
--- End of Application ---





CONCLUSION

The Resume Parser Project provides an innovative solution to streamline and enhance the recruitment process by automating the extraction, organization, and analysis of resume data. Through the integration of advanced technologies such as Natural Language Processing (NLP) and machine learning, the system can process resumes in various formats, extract key information, and deliver structured data that can be seamlessly integrated into Applicant Tracking Systems (ATS) or Human Resource Management Systems (HRMS).



⋮ **FUTURE ENHANCEMENT**

1. Enhanced Machine Learning Models

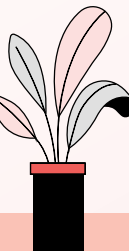
- Implement advanced deep learning models like BERT or GPT for more accurate context-based information extraction.

2. Candidate-Job Matching

- Develop algorithms to automatically match parsed resumes to job descriptions based on skills, experience, and qualifications.

3. Multilingual and Localization Support

- Expand support for additional languages, including regional dialects.





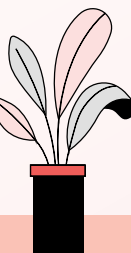
FUTURE ENHANCEMENT

4. Cloud-Based Deployment

- Deploy the system on cloud platforms for improved scalability and accessibility.
- Enable real-time resume parsing for recruiters working remotely or in different locations.

5. Integration with External Platforms

- Develop APIs for seamless integration with popular recruitment platforms, Applicant Tracking Systems (ATS), and Human Resource Management Systems (HRMS).
- Allow direct import of job descriptions from HR systems for candidate-job matching.





20XX/01/01



THANK YOU....

Any queries???

