

# **Smart Resume Parser Project Report**

## **Introduction**

The Resume Parser project is designed to automatically extract key details such as Name, Contact, Information, Education, Skills, and Experience from resumes in various formats (PDF, DOCX). This solution helps streamline recruitment processes by reducing manual effort.

## **Abstract**

The project leverages Natural Language Processing (NLP) techniques using spaCy, regular expressions, and file parsing libraries to process resumes. It extracts structured information into JSON-like formats for easy consumption by recruitment systems or analytics tools.

## **Tools Used**

1. Python
2. spaCy (NLP)
3. pdfplumber (PDF text extraction)
4. python-docx (DOCX parsing)
5. Streamlit (User Interface)

## **Steps Involved in Building the Project**

1. Collect resumes in PDF, DOCX format.
2. Extract raw text using file-specific parsers.
3. Apply regex and spaCy NLP models for entity recognition.
4. Identify key sections such as Education, Skills, and Experience.
5. Format the extracted data into structured JSON or CSV output formats.
6. Display results through a user-friendly Streamlit interface.

## **Conclusion**

The Resume Parser successfully demonstrates how NLP and Python libraries can automate the extraction of structured information from resumes. This system can significantly reduce manual work in recruitment pipelines and can be extended with advanced ML models for improved accuracy.