**Resume Screener System**



**By**

**AWAIS ZARIF**

**GHULAM MOHI UD DIN**

**SUPERVISED BY:**

**Saifullah Zadran**

**Ameer Hamza**

****

##### **Final Approval**

##### This is to clarify that we have read this report submitted by Awais Zarif, and Ghulam Mohi Ud Din. It is our judgment that this report is of sufficient standard to warrant its acceptance by National Center of Artificial Intelligence for the Artificial Intelligence Development Professional(AIDP) .

**Supervisor**

##### **Acknowledgement**

##### We would like to take this opportunity to first and foremost thank Almighty Allah, for being our strength, ultimate guide and enabling us to complete our project report.

##### We would like to express our deepest gratitude to all those who have contributed to the development and design of this Project.

##### Firstly, we would like to thank the **National Center of Artificial Intelligence** who provided valuable insights and feedback during the development process. Your expertise and guidance were instrumental in ensuring that the project meets the needs of our skincare community.

##### We extend our gratitude to Project Supervisor **Saifullah Zadran** who worked tirelessly to bring this Project to life, using their expertise in Machine learning to create a project that is both visually appealing and highly functional.

##### Awais Zarif

##### Ghulam Mohi Ud Din

**Abstract**

##### In today's competitive job market, manually screening resumes is a time-consuming and inefficient process for recruiters. This project, **"Resume Screener using NLP,"** aims to automate the resume screening process using **Natural Language Processing (NLP)** techniques. The system extracts and analyzes key information from resumes, such as **skills, experience, and education,** to determine their relevance to a given job description. By leveraging **text preprocessing, keyword extraction, and similarity matching techniques,** this model can rank resumes based on their suitability for a role. This beginner-level project demonstrates the practical application of **NLP in recruitment automation,** reducing human effort and improving efficiency in the hiring process.

##### **Keywords:** Resume Screening, **Natural Language Processing (NLP),** Text Processing, Keyword Extraction, Machine Learning (ML), Recruitment Automation, Job Matching, Similarity Matching, AI in Hiring**.**

##### **Background**

Recruitment is a crucial process for any organization, as hiring the right candidate significantly impacts business success. Traditionally, recruiters manually review hundreds of resumes to shortlist suitable candidates, which is time-consuming, prone to human bias, and inefficient. With the increasing number of job applicants, there is a need for an automated system that can **efficiently analyze and filter resumes** based on job requirements.

**Natural Language Processing (NLP),** a subfield of **Artificial Intelligence (AI)**, has shown great potential in automating text-based tasks such as resume screening. By leveraging **NLP techniques like text preprocessing, named entity recognition (NER), keyword extraction, and similarity scoring**, organizations can develop **intelligent resume Screener Systems** that quickly analyze resumes and rank them based on relevance to a given job description.

This project aims to **apply NLP techniques in the recruitment process**, reducing the manual effort required for resume screening while improving accuracy and efficiency. By implementing **Machine Learning (ML) models and NLP algorithms,** this system can **automatically extract key features** such as skills, work experience, education, and relevant keywords to match resumes with job postings. Such an approach enhances the hiring process, allowing recruiters to focus on interviewing top candidates rather than manually reviewing every resume.

**Tabel of content:**

[1. INTRODUCTION: 7](#_Toc9617)

[1.1 Overview: 7](#_Toc20236)

[1.2 Purpose of the System: 7](#_Toc13848)

[1.3 Scope of problems: 7](#_Toc29421)

[1.3.1 Inefficient Screening: 7](#_Toc20781)

[1.3.2 Time-Consuming Manual Review: 8](#_Toc30753)

[1.3.3 Resume Format Variability: 8](#_Toc32349)

[1.3.4 Bias in Candidate Selection: 8](#_Toc31182)

[1.3.5 Lack of Real-Time Analysis: 8](#_Toc7870)

[2. Objectives of the Project: 8](#_Toc25879)

[3. Machine Learning Model Training: 9](#_Toc16830)

[3.1 Dataset Used: 9](#_Toc28453)

[3.2 Model Architecture: 9](#_Toc26000)

[3.3 Model Training & Performance: 10](#_Toc27741)

[ Challenges Faced: 10](#_Toc10235)

[4. Technologies Used: 11](#_Toc6783)

[5. Expected Outcomes: 11](#_Toc21135)

[5.1 Expected Outcomes: 11](#_Toc15484)

[5.1.1 Accurate Resume Filtering: 11](#_Toc17803)

[5.1.2 Improved Candidate-Job Matching: 12](#_Toc7754)

[5.1.3 Reduced Manual Effort: 12](#_Toc7934)

[5.1.4 Fair and Unbiased Screening: 12](#_Toc2042)

[5.1.5 Scalability for Large Recruitment Needs: 12](#_Toc10219)

[5.1.6 Integration Possibilities: 12](#_Toc8924)

[6. Target Audience: 13](#_Toc8044)

[6.1 Job Seekers: 13](#_Toc10366)

[6.2 Recruiters & HR Professionals: 13](#_Toc31798)

[6.3 Job Portals & Hiring Platforms: 13](#_Toc12984)

[6.4 Companies & Enterprises: 13](#_Toc16100)

[6.5 AI & NLP Enthusiasts: 13](#_Toc32052)

[6.6 Educational Institutions & Career Counselors: 14](#_Toc1014)

[7. Conclusion: 14](#_Toc11277)

# INTRODUCTION:

## Overview:

Traditional resume screening is time-consuming, inefficient, and prone to human bias. Recruiters manually review numerous resumes, often leading to inconsistencies and overlooked candidates. To address these challenges, the "Resume Screener System Using NLP" automates the screening process using Natural Language Processing (NLP) and Machine Learning (ML). This system analysis resumes, extracts key details like skills, experience, and education, and matches them with job descriptions. By automating resume filtering, the project enhances efficiency, accuracy, and fairness in recruitment.

## Purpose of the System:

This project aims to develop an AI-powered resume screener system that Analyzes resumes using NLP techniques like text preprocessing and keyword extraction.

* Matches resumes with job descriptions based on relevant qualifications.
* Reduces manual effort and accelerates the hiring process.
* Improves fairness by minimizing human bias in candidate selection.

## Scope of problems:

#### Inefficient Screening:

* Traditional keyword-based screener lacks contextual understanding, leading to inaccurate shortlisting.

#### Time-Consuming Manual Review:

* Recruiters spend excessive time reviewing resumes, delaying hiring decisions.

#### Resume Format Variability:

* Different file formats and structures make data extraction challenging.

#### Bias in Candidate Selection:

* Human biases and limited AI training data can affect fair hiring decisions.

#### Lack of Real-Time Analysis:

* Most systems do not analyze career progression or adapt to changing job market trends.
* This project overcomes these challenges with an intelligent, automated, and unbiased resume screener solution, ensuring fast and accurate recruitment.

# Objectives of the Project:

The main objectives of this project are:

* **Train an NLP-based model** to analyze and extract key information from resumes, such as **skills, experience, and education**.
* **Use Machine Learning (ML) techniques** to rank resumes based on relevance to job descriptions.
* **Implement Named Entity Recognition (NER)** **and keyword extraction** to improve resume screening accuracy.
* **Enhance fairness and efficiency** in recruitment by reducing bias and **automating filtering.**
* Lay the foundation for future enhancements, such as real-time resume analysis and integration with applicant tracking systems (ATS).

# Machine Learning Model Training:

## Dataset Used:

For training the NLP-based resume screener model, we used a dataset containing labeled resumes and job descriptions. The dataset includes:

* **Resumes with different formats (PDF)** containing skills, experience, and education.
* **Job descriptions** specifying required qualifications, responsibilities, and preferred skills.
* **Labeled data** indicating whether a resume is a good match for a given job.

Resumes were **preprocessed using NLP** **techniques** such as **tokenization, stopword removal, lemmatization, and vectorization** to improve the model's ability to extract meaningful information.

## 3.2 Model Architecture:

We designed a Natural Language Processing (NLP) model for resume screener with the following components:

* **Text Preprocessing:** Tokenization, stopword removal, and lemmatization to clean resume text.
* **TF-IDF / Word Embeddings:** Converts text data into numerical vectors for better analysis.
* **Named Entity Recognition (NER):** Extracts key details like skills, education, and experience.
* **Similarity Matching:** Compares resumes with job descriptions using cosine similarity or semantic matching.
* **Fully Connected Layers:** Applies classification to rank resumes based on job relevance.

## Model Training & Performance:

* The model was trained for 20-30 epochs, using 80% training data and 20% validation data.
* Accuracy achieved: 85-90% on test data (varies based on dataset quality).
* **Challenges Faced:**
  + Dataset imbalance: Some job roles and skill sets had fewer resumes, affecting model performance.
  + Diversity issues: The model performed better on common job titles but struggled with niche roles.
  + Resume formatting variations: Different file formats (PDF, DOCX) required extensive preprocessing..

Future improvements include expanding the dataset for better diversity, fine-tuning hyperparameters, and optimizing NLP techniques to enhance accuracy.

## Technologies Used:

* **Python:** Primary programming language for model development.
* **VS Code:** lightweight and powerful source-code editor used for coding, debugging, and version control across multiple programming languages.
* **Numpy**:Handles arrays and numerical computations.
* **Os:** Used to interact with the operating system, such as handling file paths and directories.
* **PyPDF2:** Used to read and extract text from PDF files.
* **Pandas:** Used for data manipulation and analysis, especially handling structured data.
* **TfidfVectorizer:** An NLP technique that converts text into a numerical format.
* **Jupyter Notebook/Google Colab:** Training and experimentation environment.
* **Cosine\_Similarity:** Measures the similarity between texts.
* **Streamlit:** Used for building interactive web applications with a simple UI.
* **Time:** Used for handling time-related operations, such as delays or timestamps.
* **Logging:** Used for recording logs, debugging, and tracking the execution of code.

Since this is a beginner-level project, deployment on a website or app is not included at this stage.

# Expected Outcomes:

## The **AI-Powered Resume Screener** is designed to **automate** and **enhance** the resume screening process, ensuring accurate and efficient candidate selection. Below are the key expected outcomes and the future scope of this project:

## Expected Outcomes:

#### Accurate Resume Filtering:

The AI model will classify and rank resumes with high accuracy based on job descriptions, extracting key details like skills, experience, and education.

#### Improved Candidate-Job Matching:

Resumes will be matched more effectively with job requirements, ensuring that recruiters receive highly relevant candidates.

#### Reduced Manual Effort:

By automating the initial screening process, recruiters will save time and effort, allowing them to focus on final selection and interviews.

#### Fair and Unbiased Screening:

The AI-powered system will help minimize human bias, ensuring equal opportunities for candidates based on their qualifications rather than subjective judgments.

#### Scalability for Large Recruitment Needs:

The system will be designed to handle large volumes of resumes, making it suitable for corporate hiring, job portals, and HR agencies.

#### Integration Possibilities:

* The AI model can be integrated with Applicant Tracking Systems (ATS), job portals, and HR software to streamline recruitment processes.
* This project sets the foundation for a more efficient, data-driven, and fair recruitment system, with future enhancements focused on real-time resume analysis and advanced AI-driven candidate profiling.

# Target Audience:

#### Job Seekers:

Individuals looking for jobs who want their resumes to be accurately matched with job descriptions, improving their chances of getting shortlisted.

#### Recruiters & HR Professionals:

Hiring managers and recruitment teams who need an efficient, AI-driven resume screeener system to speed up hiring and find the best candidates.

#### Job Portals & Hiring Platforms:

Online job portals and recruitment agencies that want to integrate AI-powered resume screening to enhance candidate-job matching.

#### Companies & Enterprises:

Organizations looking to automate their hiring process and reduce manual effort in resume screening while improving hiring accuracy.

#### AI & NLP Enthusiasts:

Researchers, developers, and AI professionals interested in exploring NLP-based recruitment solutions and enhancing hiring efficiency through technology.

This system aims to benefit a wide range of users in the recruitment ecosystem, ensuring faster, fairer, and more data-driven hiring decisions.

#### Educational Institutions & Career Counselors:

Universities, colleges, and career counseling services can use the system to help students and job seekers improve their resumes by identifying gaps and suggesting enhancements based on industry standards.

# Conclusion:

The **Resume screener System** Using **NLP** enhances recruitment by automating resume screening, reducing manual effort and bias. By extracting key details like skills, experience, and education, it ensures efficient and accurate candidate shortlisting.

With potential integration into job portals and **ATS**, this system can streamline hiring and improve recruitment efficiency. Future enhancements, such as real-time analysis and improved ranking algorithms, will further refine its accuracy and adaptability.