



International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206

Volume 9, Issue 4, April 2026



International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Resume Analysis using NLP Models

K. Lakshmi Prasanna , M. Revathi , L. Abhilash Reddy , K. Shasi Vardhan

UG Student, Department of ECE, R.V.R&J.C.C. E, Chowdavaram, India

UG Student, Department of ECE, R.V.R&J.C.C. E, Chowdavaram, India

UG Student, Department of ECE, R.V.R&J.C.C. E, Chowdavaram, India

UG Student, Department of ECE, R.V.R&J.C.C. E, Chowdavaram, India

ABSTRACT: Recruiters have a hard time quickly finding good candidates because manually screening resumes takes a lot of time, is not always accurate, and is subject to bias. Current Applicant Tracking Systems (ATS) depend too much on keyword matching and don't do a good job of capturing semantic relevance and understanding context. This paper proposes an AI-based resume analysis and Recommendation System that combines Natural Language Processing (NLP), Named Entity Recognition (NER), Applicant Tracking System (ATS) scoring, and Retrieval-Augmented Generation (RAG) to overcome these limitations. The suggested system automates the parsing of resumes, pulls out important information like skills, education, and experience, and decides if a candidate is a good fit for the job based on the requirements. It also gives personalized feedback, analyzes skill gaps, and makes up mock interview questions to help candidates get ready. The system makes hiring easier, cuts down on bias, and lets you make decisions based on data. The results of the experiment show that this method is more accurate and better at matching meanings than traditional methods.

KEYWORDS: Applicant Tracking Systems, Natural Language Processing, Named Entity Recognition, Retrieval-Augmented Generation.

I. INTRODUCTION

Hiring is an important part of running a business today, but most of the time, traditional methods for screening resumes are still manual, take a long time, and aren't always reliable. Recruiters often look at a lot of resumes in a short amount of time, which makes it more likely that they will miss qualified candidates. Current Applicant Tracking Systems (ATS) mainly use keyword-based filtering, which doesn't work well for understanding the meaning and context of resumes. Because of this, candidates who might be a good fit could be turned down because their keywords don't match well, which makes the hiring process less efficient and more biased.

With AI and NLP making quick progress, automated resume analysis has become a promising way to deal with these problems. This paper presents an intelligent resume analysis and recommendation system that incorporates NLP techniques, Named Entity Recognition (NER), ATS-based scoring, and Retrieval-Augmented Generation (RAG). The system not only automatically parses resumes and rates candidates, but it also gives personalized feedback, analyses skill gaps, and creates mock interview questions. This method makes hiring more efficient, helps people make better decisions, and helps candidates get ready for job opportunities.

II. LITERATURE SURVEY

A. Traditional Resume Screening Methods

The main ways that traditional resume screening works are by hand and with simple keyword-based filtering. Recruiters look at resumes by matching certain keywords that are related to job descriptions, skills, and qualifications. This method is easy to understand, but it takes a lot of time, isn't always accurate, and is open to human bias. Also, keyword-based filtering doesn't take into account the meaning of the words, which means that it might turn down candidates who could be a good fit but don't use exact keyword matches.

B. AI-Based Resume Screening Systems

Recent improvements in Artificial Intelligence (AI) have made it possible to create systems that automatically screen resumes. The conference paper talks about an intelligent resume screening tool that uses Natural Language Processing



International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

(NLP) and machine learning to find useful information in resumes and sort candidates by job role. These systems make screening faster and easier by cutting down on the amount of work that needs to be done by hand. But most AI-based systems still depend a lot on pattern matching and don't really.

C. NLP and Named Entity Recognition Techniques

Tokenization, stemming, and lemmatization are all examples of Natural Language Processing (NLP) techniques that are commonly used to process resume data. Named Entity Recognition (NER) is very important for getting structured information like a candidate's name, skills, education, and experience from text that isn't structured. These methods make it easier to get accurate data and better candidate profiles. Even with these improvements, many systems still don't do a good job of combining semantic analysis with decision making.

D. Limitations of Existing Methods

Even though automated resume analysis has come a long way, there are still some problems. Traditional Applicant Tracking Systems (ATS) don't understand semantics and mostly rely on matching keywords. Many AI-based solutions don't give personalized feedback, analyze skill gaps, or help people get ready for interviews. Also, their decisions are not clear or easy to understand. The proposed system combines NLP, NER, ATS scoring, and Retrieval-Augmented Generation (RAG) to make a more complete and smart solution to these problems.

III. PROPOSED FRAMEWORK

The proposed system is a smart pipeline that will automatically analyze resumes and suggest candidates. It uses Natural Language Processing (NLP), Named Entity Recognition (NER), Applicant Tracking System (ATS) scoring, and Retrieval-Augmented Generation (RAG) to quickly and accurately evaluate resumes and give useful information. The system processes resumes in a set order, turning unstructured text into useful information.

A. System Overview

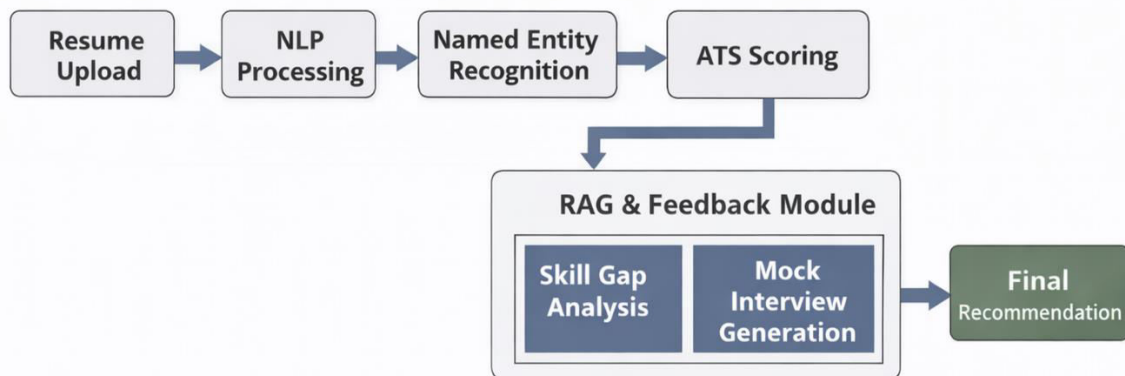


FIG 1: Proposed Framework.

Figure 1 shows the overall design of the proposed system. There are many parts to the system, such as resume input, preprocessing, information extraction, evaluation, and feedback generation. To make sure that resume analysis is accurate and quick, each module does a different job. The pipeline makes sure that data flows smoothly from the raw input to the final output.

B. Resume Input and Preprocessing

{Step 1}: Resume Upload

The system can read resumes in different formats, like PDF and DOCX. These resumes may have unstructured text data that is laid out and styled in different ways.

{Step 2}: Text Extraction



International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Document parsing techniques turn the uploaded resume into plain text. This step makes sure that the content can be processed further.

{Step 3}: Data Preprocessing

The text that was taken out goes through preprocessing steps like tokenization, removing stop words, stemming, and lemmatization. These steps help clean and normalize the data, which makes downstream NLP tasks work better.



FIG 2 : NLP Preprocessing and Feature Extraction Process.

Figure 2 shows the steps that were taken to turn raw resume text into structured features that can be used for more analysis.

C. NLP Processing and Feature Extraction

{Step 4}: Linguistic Processing

We use Natural Language Processing to figure out what the resume text means and how it is put together. This includes tagging parts of speech and breaking sentences down.

{Step 5}: Named Entity Recognition (NER)

NER is used to get important information about candidates, like their name, skills, education, work history, and certifications. These entities are organized and kept for later review. The extracted features serve as the foundation for candidate evaluation

D. ATS Scoring Mechanism.

The ATS scoring mechanism and model structure are shown in Fig. 3.

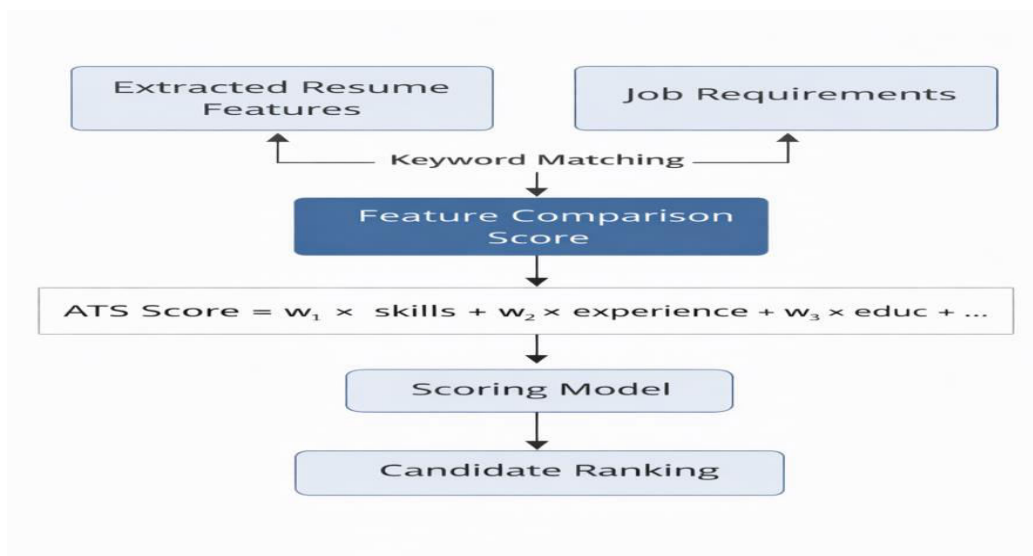


FIG 3. ATS Scoring and Ranking flowchart



International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

{Step 6}: Job Description Matching

The system compares the extracted resume features to the job requirements using keyword matching and semantic similarity techniques.

{Step 7}: Scoring and Ranking

An ATS-based scoring system gives different weights to things like skills, experience, and education. Candidates are ranked based on how well they fit the job description. This helps recruiters find the best candidates quickly.

E. Retrieval-Augmented Generation (RAG) Module

{Step 8}: Contextual Information Retrieval

The RAG module gets useful information from outside sources of knowledge to help people make better decisions.

{Step 9}: Feedback and Recommendation Generation

The system creates personalized feedback based on the analysis, which includes an analysis of skill gaps and suggestions for how to improve. This gives candidates more than just a simple screening.

F. Mock Interview Generation

{Step 10}: Question Generation

Using Natural Language Generation techniques, the system makes interview questions that are specific to each role.

{Step 11}: Candidate Preparation Support

The questions that are made help candidates get ready for interviews. You can also use Text-to-Speech (TTS) modules to give interactive help.

G. System Workflow

Figure 4 shows how the proposed system works from start to finish. The process starts with uploading a resume and then goes through preprocessing, entity extraction, scoring, and generating feedback. The final product includes a list of candidates ranked, a detailed analysis, and help getting ready for interviews.



FIG 4: Work Flow

H. Performance Monitoring

The system's performance is judged by how well it extracts entities, how relevant job matches are, and how useful recommendations are. The model makes sure that NLP processing and scoring work well by optimizing them.

IV. RESULTS AND ANALYSIS

The suggested AI-based resume analysis system is tested by putting it into action and showing the results. The system shows that it works by processing resumes, getting useful information from them, and making smart suggestions.

A. Input Resume

The system can read resumes that are in PDF or DOCX format. Figure 5 shows an example of a resume that was sent to the system. The input has unstructured text that includes personal information, skills, education, and work history



International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Chandra Sekhar Chennupati

✉ cs.chennupati@gmail.com ☎ +91 9949385751 📍 Hyderabad, India 🌐 linkedin.com/in/chandra-sekhar-chennupati

SUMMARY

I am an experienced Data Engineer skilled in PySpark, Google Cloud Platform(GCP), Python, SQL, and Airflow. I specialize in building and optimizing data pipelines. I constantly commit to continuous learning and leveraging new technologies to enhance data solutions.

PROFESSIONAL EXPERIENCE

Data Engineer Tata Consultancy Services 07/2021 – present
Hyderabad, India

- Build ETL/ELT/EL batch data pipelines to move data from on-premise to Data Warehouse (**BigQuery**) and vice-versa using **PySpark**, **SparkSQL** and **Apache Airflow**
- Iteratively design and develop a modular, reusable PySpark and Airflow based **framework** to create and orchestrate distributed data processing pipelines
- Fine-tuned and migrated oracle **stored procedures** and **packages** to Bigquery
- Optimize and performance tune big data pipelines to handle TBs of data using Apache Spark and SQL
- Perform proof of concepts (**POC**) using interactive development environments such as Jupyter Notebooks
- Debug erroneous data pipelines including undocumented legacy pipelines and troubleshoot complex data issues and provide solutions
- Create **Bash** or **Python** scripts to **automate** routine tasks

EDUCATION

V R SIDDHARTHA ENGINEERING COLLEGE 2017 – 2021
Bachelor of Technology in Electronics and Communication Engineering Vijayawada, India
CGPA: 8.61

PROJECTS

- Smart AI Doc Miner**
Generative AI, Flask, Python, Google Cloud Platform, pgvector, Google Gemini 1.5
- Implemented **RAG (Retrieval-Augmented Generation)** methodology to enable efficient document retrieval and analysis.
 - Integrated **Google Gemini 1.5 model** for state-of-the-art natural language processing and understanding.
 - Utilized **pgvector** for vector storage and similarity search, enhancing data retrieval capabilities.
 - **Deployment:** Hosted the application on **Google Cloud Run**, ensuring scalability and seamless deployment.

SKILLS

Programming Languages — Python, SQL, Shell Scripting

Frameworks — Apache Spark (PySpark), Apache Airflow

Cloud Services — **Google cloud platform (GCP)** - BigQuery, Google Cloud Storage(GCS), Dataproc, Pub/Sub, Cloud Composer, Google Compute Engine, Google Kubernetes Engine, IAM, Cloud SQL, Cloud Functions, Vertex AI, Cloud Run, Cloud Shell, Cloud Scheduler, Artifact Registry | **Amazon Web Services (AWS)** - S3, Lambda, DynamoDB | **Microsoft Azure** - Azure Data Factory, Azure Databricks, Azure Synapse Analytics

ETL Tools — IBM InfoSphere DataStage

Orchestration Tools — Cloud Composer, Control-M

DBMS — Teradata, postgres, oracle

CERTIFICATES

- Google Cloud Certified Professional Data Engineer ☞
- Google Cloud Certified Associate Data Practitioner ☞
- Google Cloud GenAI L400 Certified ☞

AWARDS

- Employee of the Month x 3**
- On three different occasions, I have been awarded Employee of the Month for my individual contributions
- Project Performance Improvement Award x2**
- As part of a 5 member team, helped the customer save more than \$150k
 - As part of a 2 member team, worked to meet Government's Data Regulatory policies within a strict deadline of 4 months

INTERESTS

Cricket | Volleyball | Badminton | Reading | Photography

FIG 5: Resume input

B. User Interface for Resume Upload



FIG 6: Resume Upload Interface

A user-friendly front-end interface lets users upload resumes and start the analysis process. The system can read files in PDF and DOCX formats, so you can use a variety of input types. Once the resume is uploaded, it is processed automatically without anyone having to do anything. Figure 6 shows the screen where users choose and upload resumes for analysis. The design makes it easy to get around and find results quickly, which improves the overall user experience.



International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

C. Resume Analysis Result

After processing, the system makes a summary of the candidate's profile. The output in Fig. 7 shows the predicted job category, level of experience, and overall resume score. The scoring system based on ATS looks at the resume and gives a numerical score.



FIG 7: Resume Analysis Result

D. Feature-Based Performance Analysis

The system looks at different factors, like skills, education, work experience, and projects. Figure 8 shows these features in a graph, showing how they affect the overall score. This helps you figure out what parts of the resume are good and what parts need work.

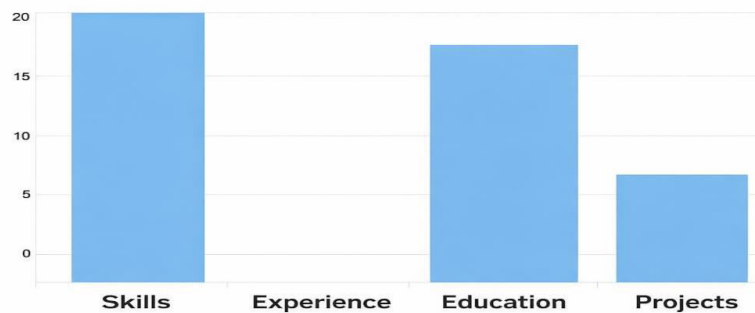


FIG 8: Resume Analysis Result

E. Extracted Skills and Domain-Specific Recommendations

By utilizing NLP and NER modules, the system extracts relevant skills from resumes to identify deficiencies and provide tailored preparation guidelines as illustrated in Figure 9. Beyond basic skill assessment, the system generates domain-specific recommendations, such as the Data Science interview topics and questions shown in Figure 10, to help candidates meet industry standards.

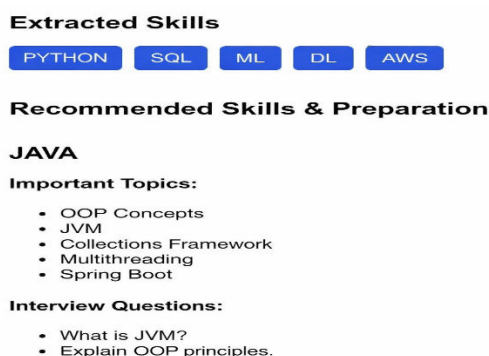


FIG 9: Extracted Skills and Recommendations



FIG 10: Data Science Recommendations



International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

F. DevOps Domain Analysis

In the same way, Fig. 11 shows suggestions and interview questions for the DevOps field. The system changes its suggestions based on the skills it sees and the job role it thinks the person will have.

DEVOPS

Important Topics:

- CI/CD
- Docker
- Kubernetes
- Jenkins
- Linux Commands

Interview Questions:

- What is CI/CD?
- What is Docker?
- Explain Kubernetes.
- What is Jenkins?
- What is Infrastructure as Code?
- Explain Git workflow.
- What is containerization?
- Difference between VM and container?
- Explain monitoring tools.
- What is Blue-Green deployment?

FIG 11: DevOps Recommendations

V. CONCLUSION

This study has created and tested an intelligent AI-based system for automatically analyzing resumes and recommending candidates in order to make hiring more efficient. The system effectively analyzes unstructured resume data and generates meaningful insights by combining NLP-based preprocessing, NER-driven information extraction, ATS-based scoring, and RAG-based recommendation modules.

The analysis of the proposed framework leads to the following key observations:

1. Using NLP and NER techniques, the system successfully pulls out useful information like skills, education, and experience, which makes sure that candidate profiles are correct.
2. The ATS-based scoring system is a reliable and organized way to rank candidates based on how well they fit the job.
3. The combination of RAG makes it possible to give personalized feedback, analyze skill gaps, and make recommendations that are specific to the field, which helps candidates get ready.
4. The system shows that it is more efficient by cutting down on manual work and speeding up decision-making in hiring processes by using data.
5. The modular architecture ensures scalability and adaptability across multiple domains such as Data Science, DevOps, Testing, and Web Development.

The proposed system is a smart and effective way to deal with modern resume screening problems. It makes both the recruiter's job easier and the candidate's experience better.

REFERENCES

- [1] J. Chen, C. Zhang, and Z. Niu, "A Two-Step Resume Information Extraction Algorithm," 2018.
- [2] A. Kumar and R. Singh, "AI Based Skill Extraction from Resumes for Career Path Recommendation," International Journal of Computer Applications, vol. 185, no. 7, pp. 25–31, 2023.
- [3] K. L. Abhishek et al., "Developing an Intelligent Resume Screening Tool With AI-Driven Analysis and Recommendation Features," Applied AI Letters, vol. 6, 2025.
- [4] Y. Chaudhari, P. Jadhav, and Y. Gupta, "An End-to-End Solution for Automated Hiring," in Proceedings of the 4th International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT), 2022.
- [5] A. Géron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly Media, 2019.
- [6] D. E. Goldberg, Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley, 1989.
- [7] S. Bird, E. Klein, and E. Loper, Natural Language Processing with Python. O'Reilly Media, 2009.
- [8] T. Mikolov et al., "Efficient Estimation of Word Representations in Vector Space," arXiv preprint arXiv:1301.3781, 2013.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com