

## **Resume Analysis and Grading Tool - Project Overview**

### **Project Overview**

#### **Executive Summary:**

This project focuses on building a text-based resume analysis and grading tool using Natural Language Processing (NLP) techniques. The tool evaluates resumes based on their keyword relevance to a job description, readability, and the presence of key skills to generate a "fit score."

#### **Purpose:**

The purpose of this tool is to streamline the resume screening process for HR professionals and job applicants. By automating the analysis of resumes, this tool can help recruiters quickly identify the most suitable candidates for a job opening, saving time and reducing bias.

#### **Methodology:**

1. Text Preprocessing: The resumes and job descriptions are preprocessed to remove special characters and stop words for efficient keyword extraction.
2. Keyword Matching: The tool identifies and counts the most relevant keywords present in the job description and matches them against the resume.
3. Readability Analysis: The readability of the resume is calculated using the Flesch Reading Ease score, which assesses the complexity of the text.
4. Scoring: The tool generates a final score based on a weighted average of keyword matching and readability, providing a quantitative measure of a resume's fit for the job.

#### **Expected Impact:**

- **\*\*For Employers:\*\*** Automates the tedious task of manual resume screening, enabling recruiters to focus on the most qualified candidates.
- **\*\*For Job Seekers:\*\*** Helps applicants understand the importance of tailoring resumes to specific job descriptions, ultimately improving their chances of success.
- **\*\*For Businesses:\*\*** Enhances the efficiency and objectivity of the recruitment process, contributing to better hiring decisions.

#### Future Enhancements:

- Support for PDF and DOCX file formats for a wider range of resume inputs.
- Incorporation of machine learning models to improve scoring accuracy and add context-aware skill extraction.
- A user-friendly web interface for broader accessibility and ease of use

#### Criteria

##### Understanding the Scales and Interpretation of the Scores

##### Keyword Score:

- Scale: The score represents the percentage of relevant keywords found in the resume compared to the keywords extracted from the job description. In your implementation, it is calculated as:

Keyword Score = (Number of matched keywords/Total number of keywords in the job description) × 100

- Interpretation: A keyword score of 145.00 suggests that the resume contains a significant number of relevant keywords—possibly more than the keywords extracted from the job description. This could happen if the resume has repeated instances of the relevant keywords, indicating a strong alignment with the job requirements.
- Note: Since a percentage score generally ranges from 0 to 100, a score above 100% means that the resume may contain multiple instances of each keyword. This could suggest overemphasis on certain keywords, which might not always be beneficial.

##### Readability Score (Flesch Reading Ease):

- Scale: The Flesch Reading Ease score typically ranges from 0 to 100, where:
  - 90–100: Very easy to read (suitable for an 11-year-old).
  - 60–70: Easily understood by 13- to 15-year-old students.
  - 30–50: Difficult to read (suitable for college-level students).
  - 0–30: Very difficult to read (best understood by university graduates).
- Interpretation: A readability score of 12.70 indicates that the resume is quite complex and might be hard to read for the average reader. This might be due to technical jargon, long sentences, or complex phrasing. While some complexity is expected in professional resumes, an overly complicated resume might not effectively communicate key information to recruiters.

##### Final Score:

- Scale: This is a custom score calculated as a weighted average of the keyword score and readability score. In your case, it uses the formula:

Final Score = ( 0.7 × Keyword Score ) + ( 0.3 × Readability Score )

- Interpretation: The final score of 105.31 is a combined measure of how well the resume matches the job description in terms of relevant keywords and how readable it is. A score above 100 in this context suggests that, although the resume is well-aligned with the job requirements (high keyword match), its readability could be a concern, as indicated by the low readability score. This implies that while the content is relevant, the way it is presented might need improvement for better communication.

#### Recommendations

- For Keyword Score: If the keyword score is above 100%, it might indicate over-optimization (keyword stuffing). Review the resume to ensure it contains relevant keywords naturally and in context.
- For Readability Score: Since a readability score of 12.70 is quite low, consider revising the resume to use shorter sentences, simpler words, and clearer formatting.
- Overall: The final score suggests that the resume is strong in terms of relevance but may need stylistic and readability improvements to make it more accessible and engaging for the reader.

Adjust the scoring system if needed to provide a more nuanced evaluation, particularly for keyword density and readability. For example, consider capping the keyword score at 100 to avoid inflated final scores.