

# **[PRD] Resume Parser, Sumariser and Rater**

**Summary:** Extract useful informations from the Resumes

| **Created:** July 20, 2021  **Owner:** rishabh.hit@gmail.com  **Contributors:** rishabh.hit@gmail.com | **Status:** WIP | **In Review** | Approved | Obsolete  **RFC:** Link to RFC when created |
| --- | --- |



The goal of the project is to build a automatic resume summariser and Rater to assist people is easy filtering of relevant resumes.

## Problem: Reading resumes and filtering out the relevant resumes is a tedious process. The aim of project is to come up with an AI based solution to automatically rate and extract relevant informations from the resume.

Requirements:

1. Extract Person Name
2. Extract Persons Phone and email mentioned
3. Extract Skills mentioned in Resume
4. Extract past Organisation and Experiences
5. Extract college studies and current Location
6. Rate the resume based on Experience and study

Approach:

1. Collect sample resumes from open source platforms
2. Tag resumes manually or automatically for the target entities
3. Get list of IT skills and Organisation names, Find there position in the resume and mark the span as the type of entity found
4. Similarly mark Current Location and University manually
5. Prepare the data in word,tag\_type format where each line has 1 word and entity type separated by space. New line is marked by an empty line. Mark the entities in BIE format
6. Analyse the entities extracted to understand their distribution and check if certain entity are represented more.
7. Tag more data to bring a balance between different entities
8. Train different models and report the accuracy metrics
9. Write a heuristic function to rate the resumes based on entities extracted
10. Use the output from the heuristic function as the training data for the rater system
11. Check model accuracy on new samples and analyse points of improvement