Project Synopsis on

Resume Ranking System

Submitted to

Department of Computer Science and Engineering

Nepal Engineering College

in Partial Fulfillment of the

Requirements for the Degree of B.E. in Computer

By

Roshan Shriwastav (018-326)

Umesh Rana Magar (018-342)

Date: 27/12/2022

Abstract

In current technological world, recruitment process of corporate has evolved to the greater extent. Both the candidates and the recruiters prefer resumes to be submitted as an e-document. Resumes are often the first source of information about candidates and also the first item of evaluation in candidate selection. Thus, it is imperative that resumes are complete, free of errors and well-organized. Validating those resumes manually is not much flexible and effective and time saving. Finding suitable candidates for an open role could be a daunting task, especially when there are many applicants. It can impede team progress for getting the right person on the right time. The team requires more man power to scrutinize the resumes of the candidates. With the assistance of machine learning, a correct and quicker system are often created which might save days for recruiters to scan every resume manually. The aim of our work is to help the recruiters to find the most appropriate resume that match all their requirements. The system allows the recruiter to post his/her requirement as query, and the system will recommend the relevant resume by calculating the similarity between the query and the resume using Random forest, Support Vector Machines etc.

Keyword: Resume Classification, Natural Language Processing, Machine Learning, Text Classification, Recommender System

Table of Contents

List of Figures

Chapter 1: Introduction

1.1 Introduction

In current world, technology has a great impact on all existing as well as newly emerging fields by developing new products to solve problems and fulfill needs of the consumers by its simplicity of usage and efficiency. Also the current recruitment process of corporate world differs from the last generation where the companies receive large set of resumes as e-document. Most of the organizations today are opting for online stages and services for accepting applicant resumes.

Determining a suitable candidate for the job is not a simple task. The conventional recruitment process typically follows manual procedures. The manual recruitment process requires substantial sources such as trained recruiters in the human resource (HR) department, training expenses, etc. Moreover, these recruitment processes also require significant efforts and time to find relevant candidates for the required job positions. Therefore, filtering the most relevant candidates manually from a giant list of prospective candidates is troublesome. Many researchers have developed systems to overcome this difficulty of manual recruitment process. Almost all companies today are opting for online platforms and services for harvesting and managing candidate information. A number of online resume distribution and search sites have widely been in use worldwide today, such as LinkedIn, Indeed, CareerBuilder, Monster, Glassdoor, etc. LinkedIn is undoubtedly, one of the most popular social networking platforms in the world for business networking with over 300 million members.

An effective e-recruiting model frees companies from data overburden and advertisement cost, since it filter out incompetent candidates. The e-recruiting model can also help job seekers effectively access recruitment opportunities and reduce recruitment work. The key module for a unique e-recruiting model is the job matching framework that makes an effort to draw in the jobless who are appropriate to the opportunities to be filled, where appropriate means that a considered employer would be keen on perusing the retrieved resumes, while job

seekers would have a fair chance to be hired. Finally, an automatic resume matching system can be significant in filtering relevant candidates during the recruitment process. Moreover, resume screening is a sensitive subject in biased decision making i.e., ethnic minority application. Since machine learning models are trained using data, and if the data focuses on specific features, then machine learning models will make biased predictions that can have detrimental effects. Therefore, it is vital to ensure that the data is not biased and contains multifaceted classes. For example, training a model on people's resumes in a specific age range will create a biased model that may eliminate a qualified person.

Candidate ranking is based on their resume information which include: candidates' educational qualifications, their previous work experiences and their skills. It will rank a candidate profile against all other suggested candidates in order to find the best candidate that matches their requirement. It will fill the gap in existing candidate recommending systems and it will have a major influence on the career path of employees while condensing the workload of employers when it comes to recruiting.

1.2 Problem Statement

The problem is that the recruitment of employees through the manual procedure is a time and money consuming process. The manual recruitment process could also possibly be erroneous in hiring incompetent individuals. As a result, processing a significant number of applications through conventional methods can lead to the recruitment of clumsy individuals. The manual recruitment process requires substantial sources such as trained recruiters in the human resource (HR) department, training expenses, etc. Moreover, these recruitment processes also require significant efforts and time to find relevant candidates for the required job positions. Therefore, filtering the most relevant candidates manually from a giant list of prospective candidates is troublesome. In the recent time, the major problem being faced in the recruitment process is how to acquire the right talent and right candidate using minimal resources over the internet and in minimal time. As described above there are three major challenges that are required to overcome to increase the efficiency of the complete process.

1. Separating the right candidates from the pack.
2. Knowing the candidates can do the job before you hire them.
3. Making sense of candidate resume.

This system aims to provide a solution to above mentioned challenges by automating the recruitment process. The solution would help to find the right CV from a large group of resumes. It would also provide the list of CVs which are the best match to the job description provided by the recruiter. The proposed solution involves supervised learning to classify the resumes into various categories corresponding to the various domains of expertise of the candidates.

1.3 Objectives

This project aims to propose an effective e-recruiting tool to suggest the best possible candidates for the job postings. We propose to investigate the following objectives:

1. To provide a web portal where the HR can post a job and see the list of shortlisted candidates.
2. To propose an e-recruiting architecture that considers Resume matching by combining knowledge bases with a pre-trained transformer-based machine learning model.
3. To provide an explainable report to the stakeholders of the ranking of the matching decision.
4. To select and rank the most able and deserving candidates for the job postings on the basis of his knowledge, skills, qualification and work experiences.
5. To acquire the right talent and right candidate using minimal resources over the internet and in minimal time.
6. To notify the shortlisted candidate by the trained models of our systems.

1.4 Scope and Application

Resume Ranking system is a really versatile system as it can be implemented in various sectors which involves a large number of candidates applying for the job

postings. In the recent days, it is used in every large companies or organizations with the sole purpose of recruiting most able and deserving candidates based on their skills, knowledge, qualification and past work experiences. Some of the major application areas of the resume ranking are listed as below:

a) It can be useful to ease the workload of a job recruitment company or a company's human resource team who are looking for best matching candidates for IT related vacancies.
b) It can be used to provide an explainable report to the stakeholders of the ranking of the matching decision.
c) Recruiters can also utilize this system to choose employees for different kind of jobs, one where only knowledge and skills play an essential part and another where all the factors are critical.
d) Employers can make use of this system to find the relevant resumes whereas job seekers can use to search the most relevant job matches their resumes.
e) It can also be used to save the valuable time, effort and money of the recruiters as these recruitment processes require significant efforts, trained recruiters, training expenses and a lot of time to find relevant candidates for the required job if carried out manually.

1.5 Feasibility Study

From all the study done regarding the feasibility of the proposed system, it can be said that the system is moderately feasible. Feasibility study on the project can be categorized in the following three types which are:

1. Technical Feasibility
2. Operational Feasibility
3. Economic Feasibility

1.5.1 Technology and system feasibility

The proposed system is technically moderately feasible. It requires Dataset and CNN based on the attributes and features of the data extracted during its acquisition. The proposed system will require machine learning (KNN, RF, and NB).

1.5.2 Economic Feasibility

The main expenses for this project are the computer, cloud services, GPU and computation power. As we will be using our own computer for designing, training and operating the model, the cost of the computer is preserved. Also, we will be using Google colab for GPU resources and Google Drive for storage, which may cost some money only if we use pro version. Our main expense is for hosting and deployment of model in cloud services. Considering all these factors, our project is economically feasible.

1.5.3 Operational Feasibility

We intend to develop our system with interactive UI and UX, so our system will be easy to operate with. In this proposed system, user will be able to input their resume then the system will collect all the resumes analyze them and rank them and recommend the best one. So, this system is operationally feasible

Chapter 2: Literature Review

Resume Ranking System is a summary and evaluation of the existing research on the topic of using artificial intelligence and machine learning techniques to rank and recommend resumes for job openings. It typically involves reviewing and synthesizing the findings of multiple studies and papers to provide an overview of the state of the field and identify any gaps or areas for further research[1].

In a resume ranking system, the literature review might cover topics such as the effectiveness of different algorithms and techniques for analyzing resumes, the impact of different data sources and features on the accuracy of the system, and the ethical and legal considerations related to using AI in Resume Ranking. It could also include a discussion of related work in areas such as natural language processing, information retrieval, and machine learning, and how these techniques can be applied to the problem of Resume Ranking. Overall, the purpose of a literature review in the context of a resume ranking system is to provide a comprehensive and up-to-date understanding of the field and inform the design and development of the system[2].

A literature review may be included in a research paper or report on a resume ranking system, or it may be a standalone document that is used to provide an overview of the current state of knowledge on the topic. It is an important step in the research process, as it helps to ensure that new research is grounded in the existing body of knowledge and contributes to the overall understanding of the topic. In the context of a resume ranking system, a literature review is a review of existing research on the topic of resume ranking. It involves identifying and analyzing relevant studies and articles in order to gain a better understanding of the current state of knowledge on the topic and to identify areas for further research[3].

The purpose of a literature review in this context is to provide a summary of the current state of knowledge on resume ranking systems and to identify any gaps or limitations in the existing research. It can help to inform the development of new Resume Ranking Systems by identifying best practices and areas for

improvement. A literature review in the context of a Resume Ranking System would involve researching and examining the existing body of knowledge on the topic of automated resume evaluation systems. This might include reviewing research papers, articles, and other published materials on the subject, as well as considering the various approaches and techniques that have been developed and used in the development of such systems.

The purpose in this context would be to gain a better understanding of the state of the art in the field, identify key challenges and issues that have been addressed in the past, and identify areas where further research and development might be needed. By thoroughly reviewing the existing literature on the topic, it is possible to identify gaps in knowledge, areas of disagreement, and trends in the field, which can inform the development and evaluation of a Resume Ranking System.

Chapter 3: System Design

3.1 System block diagram

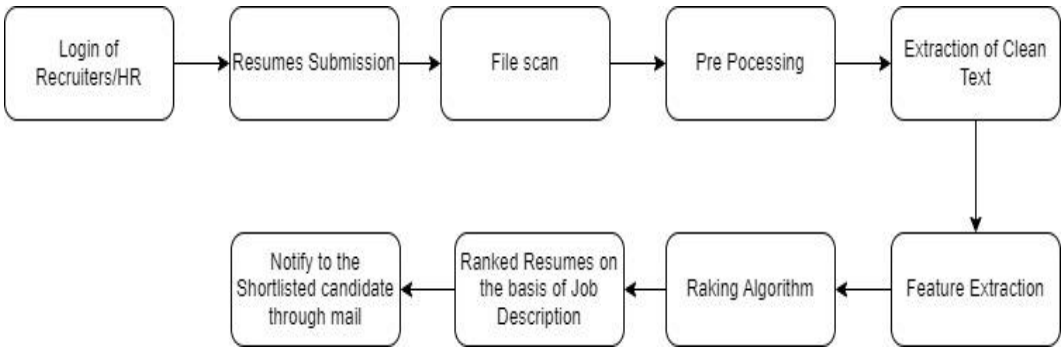

Figure 1: System Block diagram

The System Design Document describes the system requirements, operating environment, system and subsystem architecture, files and database design, input formats, output layouts, human-machine interfaces, detailed design, processing logic, and external interfaces.

For the proposed system, the following steps we are going to follow:

a) Users will submit their CV.

b) After the CV is submitted, the data preprocessing is done in which the different format of the resume are converted to .CSV file.

c) Stemming and lemmatization is done using Normalization or sometimes Word Normalization techniques in Natural Language Processing (NLP) called annotation.

d) Feature extraction is done by using Term Frequency-Inverse-Document-Frequency (TF-IDF).

e) Resume classifier is made to classify the resume files using Random Forest, Support Vector Machines algorithms.

Chapter 4: Expected Output

The proposal project is expected to process hundreds and thousands of resumes of the candidate for a given job description and provide a list or proper ranking of finest and most suited candidate's resume for the given job. The system must be able to classify hundreds of resumes and categorize considering the job description. This system will be able to find the most suited candidate for the job.

The expected output of a resume ranking system would depend on the specific goals and objectives of the system. Some possible outputs could include:

1. A list of resumes ranked in order of relevance or quality, based on the criteria specified by the system.
2. A set of ranking for which resumes should be given priority or further consideration, based on the criteria specified by the system.

Chapter 5: Conclusion

In general, however, the main purpose of a resume ranking system is to help employers and recruiters quickly and efficiently identify the most qualified candidates for a particular position or role. These systems can be particularly useful when there are a large number of resumes to review, as they can help to filter out less relevant or qualified candidates and focus on the ones that are most likely to be a good fit. Overall, the effectiveness of a resume ranking system will depend on the quality of the data it uses, the algorithms and models it employs, and the goals and objectives it is designed to meet.

However, some potential conclusions that could be drawn about these systems include:

1. Resume ranking systems can be useful tools for employers and recruiters, as they can help to streamline the process of reviewing and evaluating large numbers of resumes.
2. These systems can be designed to consider a wide range of factors, including skills, experience, education, and other relevant qualifications, in order to identify the most qualified candidates.
3. However, these systems may also be subject to biases or limitations, depending on the data and algorithms used to design and operate them. It is important to carefully consider these issues and take steps to mitigate any potential biases or limitations.
4. In general, resume ranking systems can be valuable resources for employers and recruiters, but they should be used in conjunction with other forms of assessment and evaluation, such as interviews, reference checks, and skills testing, to ensure that the most qualified and suitable candidates are identified.

Bibliography

[1] I. Ali, N. Mughal, Z. H. Khan, J. Ahmed, and G. Mujtaba, “Resume Classification System using Natural Language Processing and Machine Learning Techniques,” *Mehran Univ. Res. J. Eng. Technol.*, vol. 41, no. 1, pp. 65–79, Jan. 2022, doi: 10.22581/MUET1982.2201.07.

[2] A. Barrak, B. Adams, and A. Zouaq, “Toward a traceable, explainable, and fairJD/Resume recommendation system,” Feb. 2022, doi: 10.48550/arxiv.2202.08960.

[3] P. K. Roy, S. S. Chowdhary, and R. Bhatia, “A Machine Learning approach for automation of Resume Recommendation system,” *Procedia Comput. Sci.*, vol. 167, pp. 2318–2327, Jan. 2020, doi: 10.1016/J.PROCS.2020.03.284.