JobBhartee–Smart Job Search and Recruitment System

[1]Kalpesh Vishwakarma, [2]Vishal Mishra, [3]Nishank Mhatre, [4]Rituja Nanaware, [5]Ranjita Asati
[1][2][3][4] Students at Atharva College of Engineering, Mumbai, [5]Assistant Professor at Atharva College of Engineering, Mumbai

Abstract- Recruitment is a function of human resource management by which an organization can attract potential candidates and select the most appropriate employees for the organization. Now-a-days people are extensively adaptive to the technology and that is why e-recruitment has become a popular practice followed by the organization for hiring employees. Agencies and various firms must deal with a large number of new jobs seeking candidates with various resumes. However, managing large amount of text data and selecting the best-fit candidate is more difficult and time consuming. To keep pace with new technology and tactics, JobBhartee is a web-based program that is designed to give easy and direct interface with the employers and get a chance of being employed for the job seekers. This paper provides an overview of an ongoing Information Extraction System project that helps the recruiters to advertise their vacancies and identify the best candidate by extracting relevant information from their resumes.

1. INTRODUCTION

This application JobBhartee.com is a smart job search and recruitment system which allows the recruiters to conduct the hiring in a smooth way by simply getting the appropriate data of potential candidate and allowing candidates to participate in the process efficiently by applying for the job postings made available by the recruiters. [1] Describe a method for establishing communication between the recruiters and candidates through an easy user friendly interface. [2] Has developed the system which allows the candidate to appear for the test based on his profile and test result is generated by making use of LDA. This report can be compared by the admin and appropriate candidate can be hired. The work presented in [3] this paper is based on an e-recruitment system built for on campus recruitment in an institute with features like recommendations based on various levels of skills, smart multi-criteria search, graduate tracking. [4] This system takes up a resume and perform Named Entity Recognition (NER) on it and summarizes the content. It also ranks the resumes based on company requirements from a group of resumes. [5] makes the hiring process easy and efficient by extracting the required entities automatically by using Spacy NER model from the resumes and then generates a graph displaying the score of each and every resume. [6] This system proposes a two-step resume information extraction approach. In the first step, raw text of resume is identified as different resume blocks. To achieve the goal, a novel feature is designed, Writing Style, to model sentence syntax information. Besides word index and punctuation index, word lexical attribute and prediction results of classifiers are included in Writing Style. In the second step, multiple classifiers are employed to identify different attributes of fact information in resumes. [7] This study proposes an end-to-end pipeline for resume parsing based on neural networks-based classifiers and distributed embeddings. This pipeline leverages the position-wise line information and integrated meanings of each text block. Proposed pipeline joints the text block segmentation with the identification of resume facts in which various sequence labelling classifiers perform named entity recognition within labelled text blocks. [8] This paper presents an effective approach for resume information extraction to support automatic resume management and routing. A cascaded information extraction (IE) framework is designed. In the first pass, a resume is segmented into a consecutive blocks attached with labels indicating the information types. [9] Proposed pipeline joints the text block segmentation with the identification of resume facts in which various sequence labelling classifiers perform named entity recognition within labelled text blocks. Comparative evaluation of four sequence labelling classifiers confirmed BLSTMCNNs-CRF's superiority in named entity recognition task. In [10]this paper, there is study of a novel approach for named entity recognition (NER) and mention detection (MD) in natural language processing. Instead of treating NER as a sequence labeling problem, and propose a new local detection approach, which relies on the recent fixed-size ordinally forgetting encoding (FOFE) method to fully encode each sentence fragment and its left/right contexts into a fixed size representation. In this context, we propose a rule-based matching method from Spacy python library for information extraction. The idea is to extract the information from pdf, doc, docx files using POS tagging from Spacy library and nltk module for lemmatization and tokenization. The choice of the TLC allows to define an image as a set of words based on an alphabet firstly assumed. We developed a web application which allows recruiters to advertise their vacancies which can be approached by job-seeking candidates. This application allows recruiters to select the appropriate candidate to be hired in a smart way. This paper is organized as follows: In section 2, detailed methodology for developing the project is explained. Section 3 gives the experimental results and comparative study. Finally, after discussing the results of proposed system, conclusions and future works are presented.

2. PROPOSED SYSTEM

Waterfall model is implemented while developing this project. The requirements of the developing system were collected and the approach for designing the system was prepared.

Waterfall model:

Waterfall approach was first SDLC Model to be used widely in Software Engineering to ensure success of the project. In "The Waterfall" approach, the whole process of software development is divided into separate phases.

1. Requirement Gathering and Analysis:
   There are two user types which demand different requirements.

Candidate requirements:

- Registration: Candidate should be able to register into the system.
- Login: Candidate should be able to login into the system when required and logout when done with the use.
- Make profile: Candidate should be able to create and update his profile time to time.
- Candidate requires to surf through the new job openings and job description published by various recruiters.
- Candidate must be able to apply for interested job advertisements.
- For applying, candidate must be able to upload his/her CV or Resume.
- Candidate must be able to check the status of recruitment process he/she has applied to.

Recruiter requirements:

- Registration: Recruiter should be able to register into the system.
- Login: Recruiter should be able to login into the system when required and logout when done with the use.
- Recruiter should be able to publish job descriptions for the vacancies available.
- Recruiter must be able to see the list of the details of candidates who applied for the job after parsing their resumes.
- Selection of candidates based on their skills is to be done.
- Recruiter must be able to update the candidate’s status for further process.

2. System Design:
   The requirement specifications from first phase are studied in this phase and the system design is prepared. This system design helps in specifying hardware and system requirements and helps in defining the overall system architecture.
   The Hardware requirements for this project are:
   - RAM: 1GB
   - Hard Disk: 80 GB
   - Display: 1024*768, True Type Color-32Bit
   - Mouse: Any working mouse
   - Keyboard: Windows supported keyboard

   The Software requirements are:
   - Database connection: Django
   - Web Server for testing and fixing bugs
   - Technologies: HTML, CSS, JAVASCRIPT
   - Python

3. Implementation:
   The project was divided into various units and designed based on the required functionalities. The resume parser is made with Spacy python library for Natural Language Processing. The website was developed with python, HTML, CSS and JavaScript.
   Both the units were designed to fulfill their functionality requirements. The website is designed user friendly for both the users: candidate and the recruiter. The resume parser is able to parse the details from resume files type .pdf, .doc and .docx.
4. Integration and Testing:
   All the units developed in the implementation phase are integrated into a system after testing of each unit. Post integration the entire system is tested for any faults and failures.
   The parser is checked and tested for its functionality and any bugs found are removed. Similarly, the website is tested and made optimal for use. Both the units are integrated together and then the functionality is tested. The system should work together with Graphical interface easier for use.
5. Deployment of the System:
   Once the functional and non-functional testing is done; the product is deployed in the customer environment or released into the market.
   The system is to be deployed for use by the recruiter and candidates. The system can be used by Training and Placement department of any college or university.
6. Maintenance:
   There are some issues which come up in the client environment. To fix those issues, patches are released. Also to enhance the product some better versions are released. Maintenance is done to deliver these changes in the customer environment.
   Maintenance of the system is to be performed after the deployment for more efficient working.

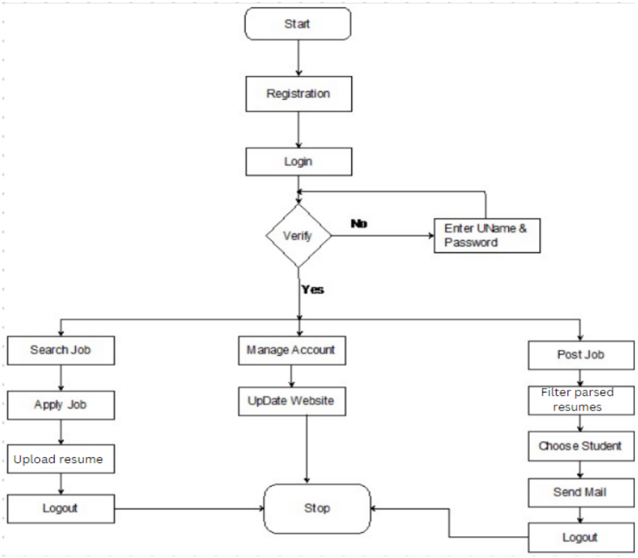


Fig 2.1 CFD (Control Flow Diagram)

The above figure shows the overall control flow of the system. The diagram describes basic functionality of the web application. The admin can manage all the user accounts and can modify the system for updates. Candidates can have basic features like registration, login, applying for job vacancies, etc. Similarly, recruiters have features like registering, posting job vacancies, filtering the resumes of applied candidates and shortlisting the suitable candidate for further stages.

The below diagram shows the data flow for the project

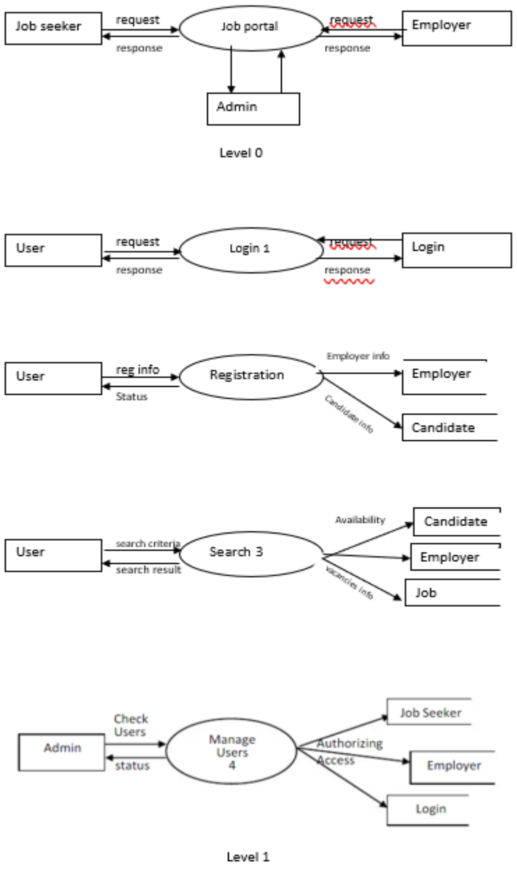


Fig. 2.1 Data Flow Diagram (DFD)

The DFD is developed in two levels. The DFD helps in clarifying system requirements and identify the major transformations that will become programs in system designs. The system provides service for two types of users. The user functionality is explained as below:

Functions of Applicant:

a. Register on the system and setup profile with designation.

b. Login.

c. Know about various companies and the available vacancy for job designation.

d. Select any company to apply.

e. For selected company uploads the Resume/CV for that job profile. (Use NLP to extract
necessary and relevant keywords by converting the pdf/doc files into text files for further
processing and store these extracted details in database)

f. Track status of the recruitment process for that company.

Functions of Recruiter/Company:

a. Register on the system.

b. Login.

c. Fill details for vacancies available according to job profile.

d. See list with details of all candidates who applied for that job.

e. Select candidates based on specific details like skills and update their status of candidates
application process.

f. Track candidate status selected for further rounds.

3. IMPLEMENTATION AND RESULTS

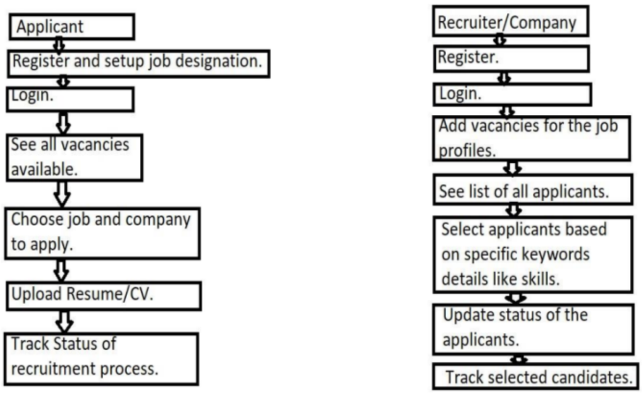

Applicant
Register and setup job designation.
Login.
See all vacancies available.
Choose job and company to apply.
Upload Resume/CV.
Track Status of recruitment process.
Recruiter/Company
Register.
Login.
Add vacancies for the job profiles.
See list of all applicants.
Select applicants based on specific keywords details like skills.
Update status of the applicants.
Track selected candidates.

Fig. 3.1. Block Diagram of the system

The fig. 3.1 shows the function flow of user client. There are two types of users. One of which is the recruiter who registers into the system and then can publish the vacancies in their company after logging in. The other are job-seeking candidates who can one time register to the system and can search and approach the vacancies based on their skills. The candidate profile gets created based on their education, experience and skills. The recruiter can parse the resumes and hire the appropriate candidate. SpaCy is an open-source package library for Natural language processing in python. SpaCy has pre-trained pipelines with tokenization support. Rule-based matching feature is used for Resume parsing. The system is capable of classifying the candidates based on their skills.

The working of the system can be measured by following feasibility report:

Operational feasibility- Operational feasibility is a measure of how well a proposed system solves the problems, and takes advantage of the opportunities identified during scope definition and how it satisfies the requirements identified in the requirements analysis phase of system development. In this project, we have developed the smart job search and recruitment system which is able to overcome traditional mode of hiring by providing features like two user profiles, classification of the candidates, Information Extraction from resumes, recruitment status update and user-friendly interface.

Technical feasibility- This project requires a system with minimum 8 GB RAM, and intel processor i3 or more. Google chrome is used for testing and bug identification. For development, we require Python 3 and Visual Studio or any other code editor.

Economic feasibility- We are using system with minimum configurations and free software to develop this project. So, we do not require any money for developing this project.

4. Results

The developed project gives user friendly interfaces which makes the application easy for use. The system is perfectly capable of basic user register and login features maintaining the privacy and confidentiality. The system can parse the candidate information and display it in understandable tabular format which will help the employer reduce the time utilized for shortlisting the appropriate candidates.

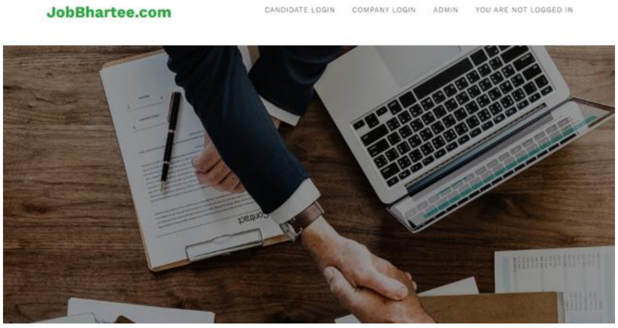


Fig 4.1 Homepage of the webpage

The above figure shows the homepage of the website which provides features for candidate and company login. After logging in, candidates can surf for the job opportunities and company recruiter can publish the job vacancies for their company and shortlist the appropriate candidate for further selection processes.

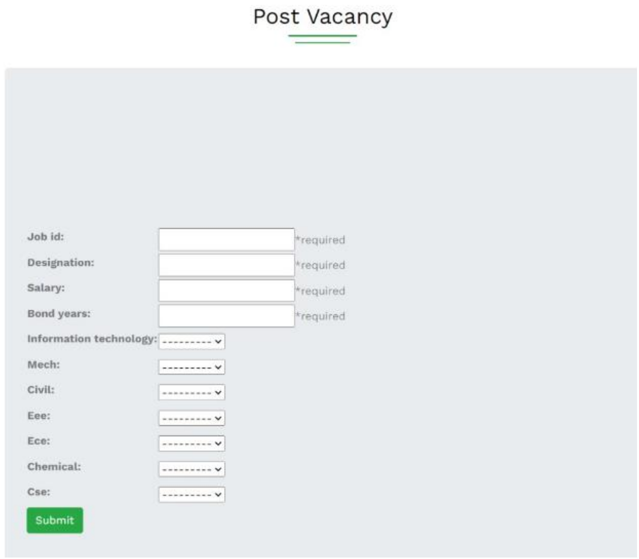


Fig 4.2 Vacancy posting page

The above page allows the recruiter to post about the vacancy in their company or organization. There are text boxes provided which allows the recruiter to provide the job description of vacancy provided.

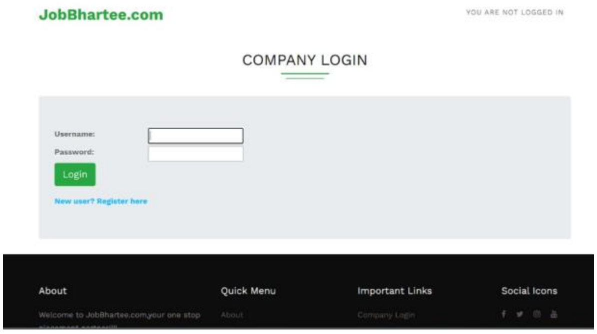


Fig 4.3 Login page for the recruiter

The above figure shows the login page for the company recruiter. In similar way, login page for candidates is available. Without logging in, candidate and recruiter cannot have access of the system and would not be able to use the further functionalities and services.

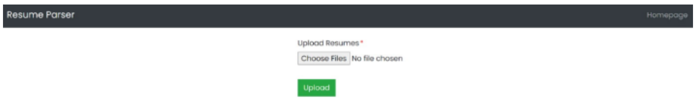
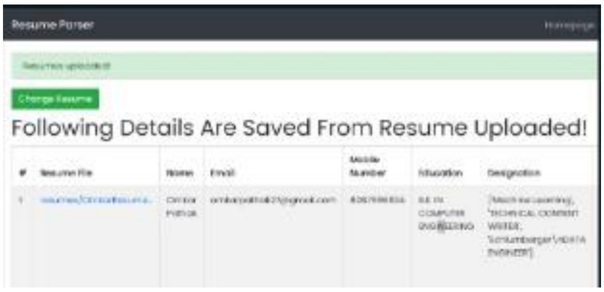


Fig 4.4 Webpage for uploading the CV/resume

The above shown figure shows the webpage through which candidates can upload their CV or Resume in .pdf, .doc or .docx format. These files will get parsed and the extracted information can be seen in tabular form by the recruiter. The workload of the recruiter is reduced by the use of parser.



Resume Parser Homepage

Resumes uploaded!

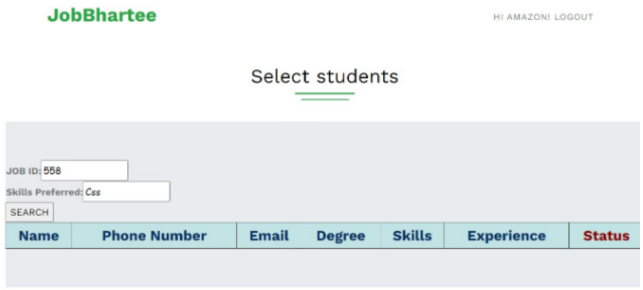
Change Resume

Following Details Are Saved From Resume Uploaded!

#	Resume File	Name	Email	Mobile Number	Education	Designation
1	resumes/[illegible]	Omkar Pathak	omkarpathak27@gmail.com	[illegible]	B.E. IN COMPUTER ENGINEERING	['Machine Learning', 'TECHNICAL CONTENT WRITER', 'Schlumberger\'DATA ENGINEER']

Fig 4.5 Parsed data in tabular format

The fig. 4.5 shows the successful result of parsed data from the resume uploaded. This information is useful for the recruiter to shortlist pertinent candidates in a less time for a single job vacancy.



JobBhartee HI AMAZON! LOGOUT

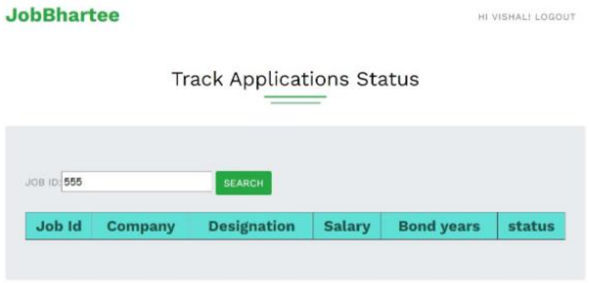
Select students

JOB ID: 558
Skills Preferred: Css
SEARCH

Name	Phone Number	Email	Degree	Skills	Experience	Status

Fig 4.6 Selecting students by keywords

The recruiter can search the candidates with required skill for that job using keywords. The list of candidates with details is given as a successful search result.



JobBhartee HI VISHAL! LOGOUT

Track Applications Status

JOB ID: 555 SEARCH

Job Id	Company	Designation	Salary	Bond years	status

Fig. 4.7 Track application status

Candidates can track the status of their application and are able to see if they are selected for further recruitment process or not.

5. CONCLUSION

The website will be useful to ease the job search and recruitment process and will surely be saving lot of unnecessary physical movements and unnecessary time wastage throughout the process of recruitment unlike to the traditional system where the company visits the campus for face-to-face recruitment or invite the candidates to their offices for the hiring process.

The website will be useful to a job applicant to find out the various job profiles/ roles of the companies available via the job postings made by the recruiting company and apply for the position he/she wish to and further he/she can upload their resume/CV which would be stored in the database in backend using NLP based techniques.

The Recruiter would be able to see, track and update status of the applicants and can query for details of specific applicant based on details like skills. etc. This will be beneficial to find the candidates based on the company's requirements and interests. Also, the recruiters will be having a lot more of control on the schedules and then they can plan the hiring process according to their interests.

The system can have integrated environment for aptitude tests conduction and other tests conduction for the various rounds during the recruitment process which will avoid the need to visit any third-party application for this purpose.

REFERENCES

[1] Keethana Kopuri, Gulam Mujtaba Hussain Aqueel , Azbar Sadiqa Jabeen , Dr.T.K.Shaik Shavali, "A Online Job portal management system" , IJIRT ,Volume 3 Issue 9 ,ISSN: 23496002 February 2017.

[2] Shubham Shendage, Tushar Shinde, Ketki Govilkar Pune. Jitendra Chavan,"SMART RECRUITMENT SYSTEM USING MACHINE LEARNING", JETIR June 2019, Volume 6,Issue 6,2019.

[3] Vijay Yadav, Suman Khatri, Ujjwal gewali, shree Ram Rauniyar, "Smart Job Recruitment Automation: Bridging Industry and University ", November 2019 DOI: 10.1109/AITB48515.2019.894744.

[4] Narendra G O and Hashwanth S, "Named Entity Recognition based Resume Parser and Summarizer", IJARSCT,Volume 2,issue 1,march 2022.

[5] Dr. K .Satheesh, A.Jahnavi, L.Iswarya, K.Ayesha, G.Bhanusekhar, K.Hanisa, "Resume Ranking based on Job Description using SpaCy NER model",International Research Journal of Engineering and Technology (IRJET),Volume:07,issue:05,may 2020.

[6] Agnieszka Wosiak-"Automated extraction of information from Polish resume documents in the IT recruitment process", 25th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems.

[7] Jie Chen, Chunxia Zhang, and Zhendong Niu (2018) "A Two-Step Resume Information Extraction Algorithm" Mathematical Problems in Engineering pp1-8.

[8] Kun Yu, Gang Guan, and Ming Zhou (2005) "Resume Information Extraction with Cascaded Hybrid Model" In Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL'05), Stroudsburg, PA, USA, June 2005, pp 499-506.

[9] Shicheng Zu and Xiulai Wang,"RESUME INFORMATION EXTRACTION WITH A NOVEL TEXT BLOCK SEGMENTATION ALGORITHM",International Journal on Natural Language Computing (IJNLC) Vol.8, No.5, October 2019.

[10] Mingbin Xu, Hui Jiang, Sedtawut Watcharawittayakul, "A Local Detection Approach for Named Entity Recognition and Mention Detection", Proceedings of the 55th Annual meeting of the Association for Computational Linguistics, pages 12371247 Vancouver, Canada, July 30 - August 4, 2017.