iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 **Issue:** VII **Month of publication:** July 2020

DOI: https://doi.org/10.22214/ijraset.2020.30378

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue VII July 2020- Available at www.ijraset.com

Resume Analyser: Automated Resume Ranking Software

Asst Prof. Jyothis Joseph[1], Jaimy Sunny[2], Raveena R[3], Blessy Elza Byju[4], Laya K C[5]
[1, 2, 3, 4, 5]*Department. of Computer Science and Engineering College of Engineering Kidangoor*, Kottayam, Kerala

Abstract: ***Recruiting candidates to fit a particular job profile is a task crucial to most of the companies. Due to increasing growth in online recruitment, traditional hiring methods are becoming inefficient. The conventional techniques usually include a labor-intensive process of manually searching through the applied candidates, reviewing their resumes, and then producing a shortlist of suitable candidates to be interviewed. In this era of technology, job searching has become smarter and more accessible at the same time. The companies receive enormous numbers of resumes/CVs, which are not always structured. There have been lots of work done for the job searching process. Whereas, the process of selecting a candidate based on their resume has not been entirely automated.***

Index Terms: resume, job, , recruitment

I. INTRODUCTION

To hire the right person at the right time, recruiters must be able to screen resumes correctly. Resume screening is the process of determining whether a candidate is qualified for a role based his or her education, experience, and other information captured on their resume. The essence of any good recruitment strategy lies in efficient and effective resume screening. The objective of resume screening is to locate the most qualified candidates for a job.In the present system the candidate has to fill each and every information regarding their resume in a manual form which takes large amount of time and then also the candidates, are not satisfied by the job which the present system prefer according to there skills. Our system will act as a handshake between two entities. i.e. the company who prefer the best possible candidate and the candidate who prefers the best possible job according to his or her skills and ability. Our system is an automated resume screening software using NLP and machine learning. This AI powered resume screening software goes beyond keywords and screens resumes contextually. After resume screening, the software ranks candidates based on the recruiters job requirements in real-time. This ranking is relative. The software uses natural language processing and machine learning for matching and ranking candidates in real time.

II. LITERATURE SURVEY

A. Learning to Rank Resumes

Recruiting and assigning right candidate for right job con- sumes significant time and efforts in an organization. Auto- mated resume information retrieval systems are increasingly looked upon as the solution to this problem. In this paper, we focus on problem of learning an end-user specific ranking in such a resume search engine. We provide experimental results of SVM rank algorithm for this task.

B. Resume Ranking using NLP and ML

Using NLP(Natural Language Processing) and ML(Machine Learning) to rank the resumes according to the given con- straint, this intelligent system ranks the resume of any format according to the given constraints or the following requirement provided by the client company. We will basically take the bulk of input resume from the client company and that client company will also provide the requirement and the constraints according to which the resume should be ranked by our system. Beside the information provide by the resume we are going to read the candidates social profiles (like LinkedIn, Github etc) which will give us the more genuine information about that candidate.

C. ResumeNet: A Learning-based Framework for Automatic Resume Quality Assessment

Recruitment of appropriate people for certain positions is critical for any companies or organizations. Manually screen- ing to select appropriate candidates from large amounts of resumes can be exhausted and time-consuming. However, there is no public tool that can be directly used for automatic resume quality assessment (RQA). This motivates us to de- velop a method for automatic RQA. Since there is also no public dataset for model training and evaluation, we build a dataset for RQA by collecting around 10K resumes, which are provided by a private resume management company.

By investigating the dataset, we identify some factors or features that could be useful to discriminate good resumes from bad ones, e.g., the consistency between different parts of a resume. Then a neural-network model is designed to predict the quality of each resume, where some text processing techniques are incorporated. To deal with the label deficiency issue in the dataset, we propose several variants of the model by either utilizing the pair/triplet-based loss, or introducing some semi- supervised learning technique to make use of the abundant unlabeled data. Both the presented baseline model and its variants are general and easy to implement. Various popular criteria including the receiver operating characteristic (ROC) curve, F-measure and ranking-based average precision (AP) are adopted for model evaluation. We compare the different variants with our baseline model. Since there is no public algorithm for RQA, we further compare our results with those obtained from a website that can score a resume. Experimental results in terms of different criteria demonstrate effectiveness of the proposed method. We foresee that our approach would transform the way of future human resources management.

III. OBJECTIVE AND SCOPE OF WORK

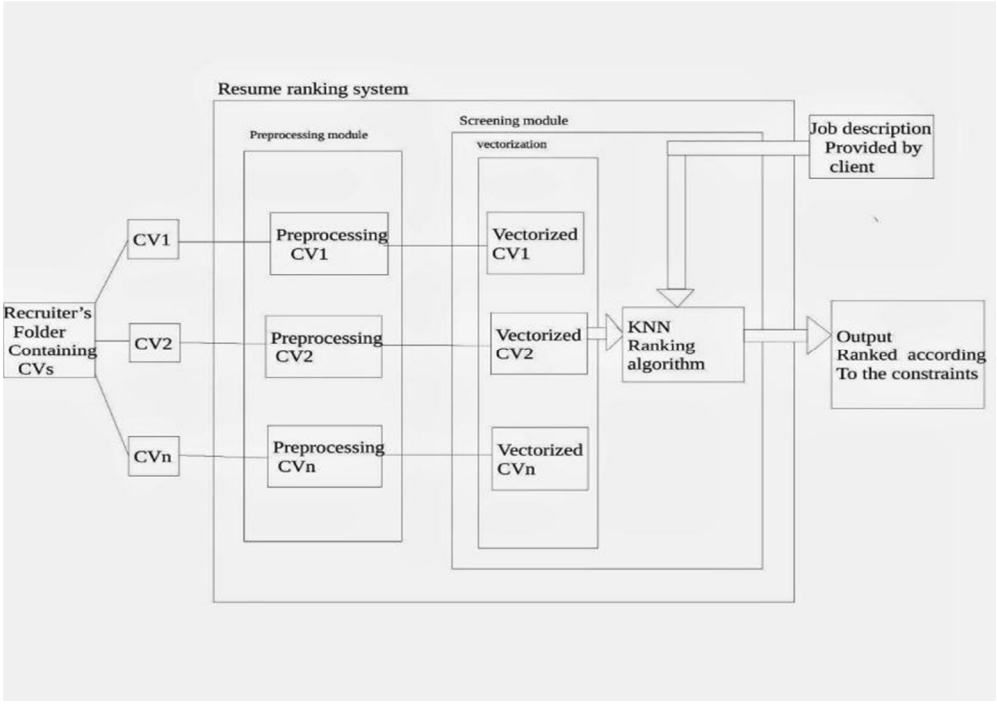
The objective is to identify the most qualified candidates for a job. To hire the right person at the right time, recruiters must be able to screen resumes correctly. The essence of any good recruitment strategy lies in efficient and effective resume screening. The objective of resume screening is to locate the most qualified candidates for a job.

IV. PROPOSED SYSTEM

Our system is an automated resume screening software using NLP and machine learning. This AI powered resume screening software goes beyond keywords and screens resumes contextually. After resume screening, the software ranks can- didates based on the recruiters job requirements in real-time. This ranking is relative. The software uses natural language processing and machine learning for matching and ranking candidates in real time.In our proposed system, resumes are analysed by the software, indexed and stored in a specific format. This makes our search process easy. The analysing system works on the algorithm that uses Natural Language Processing, subdomain of Artificial Intelligence. It reads the resumes and understands the natural language/format created by the candidate and transforms it into a specific format. This acquired knowledge is stored in the knowledge base. And the system ranks the processed resumes relatively based on the job requirements provided by the recruiter by using Natural Language Processing and Machine Learning.

V. SYSTEM DESCRIPTION

A. Architecture



VI. SYSTEM REQUIREMENTS

A. Software Requirements

1) *Python:* Python is an interpreted, high level, general purpose programming language. Its language constructs and object oriented approach aim to help programmers write clear, logical code for small and large scale projects.:
2) *Flask:* Flask is a micro web framework written in Python. It is classified as a micro framework because it does not require particular tools or libraries.[3] It has no database abstraction layer, form validation, or any other components where pre-existing third-party libraries provide common functions. However, Flask supports extensions that can add application features as if they were implemented in Flask itself. Extensions exist for object-relational mappers, form validation, upload handling, various open authentication technologies and several common framework related tools. Extensions are updated far more frequently than the core Flask program.:
3) *Textract:* Amazon Textract is a service that automati- cally extracts text and data from scanned documents. Amazon Textract goes beyond simple optical character recognition (OCR) to also identify the contents of fields in forms and information stored in tables. Textract uses machine learning to instantly "read" virtually any type of document to accurately extract text and data without the need for any manual effort or custom code. With Textract you can quickly automate document workflows, enabling you to process millions of document pages in hours.:
4) *Gensim:* It is a robust semantic modeling library that is useful for many applications.:
5) *NLTK:* NLTK is one of the leading platforms for working with human language data and Python, the module NLTK is used for natural language processing. NLTK is literally an acronym for Natural Language Tool Kit.:
6) *Machine Learning tool:* Scikit-learn (Python Package) It is a Python module integrating classic machine learning algorithms in the tightly-knit scientific Python world (numpy, scipy, matplotlib). It aims to provide simple and efficient solutions to learning problems, accessible to everybody and reusable in various contexts: machine-learning as a versatile tool for science and engineering.:

B. Hardware Requirements

Linux: GNOME or KDE desktop GNU C Library (glibc) 2.15 or later, 2 GB RAM minimum, 4 GB RAM recom- mended, 1280 x 800 minimum screen resolution. Windows: Microsoft R Windows R 8/7/Vista (32 or 64-bit) 2 GB RAM minimum, 4 GB RAM recommended, 1280 x 800 minimum screen resolution, Intel R processor with support for Intel R VT-x, Intel R EM64T (Intel R 64) Execute Disable (XD) Bit functionality.

Functionalities of Resume Analyser are listed below:

1) Screening
2) Elimination of irrelevant candidates
3) Ranking of resume
4) Scoring of resumes

VII. METHODOLOGY

The method proposed in this work offers an approach for screening the resumes based on job description. This is a web application to help recruiters by analysing resumes and CVs, surfacing candidates that best match the position and filtering out those who don't. Software uses recommendation engine techniques such as KNN, content based filtering for fuzzy matching job description with multiple resumes. Resumes are ranked based on the score on comparison with the job descrip- tion. We retrieve the CVs from the recruiter. We will retrieve it using textract which is a library in python. The CVs will be of different formats like doc,docx,pdf etc. We convert all these CVs into text format which will be easy to analyse. Each word of the CV will be separated and put in a list and also the count of the words are also stored. Stop words are words which have no use in the project. Hence the stop words will be removed. Each word in a preprocessed CV will be mapped into vectors of real numbers using word embeddings. Word embedding is the collective name for a set of language modeling and feature learning techniques in natural language processing where words or phrases from the vocabulary are mapped to vectors of real numbers.

In the scoring ranking module, we use K nearest neighbor algorithm to find the vector distance between the vectorized features and the ideal vector which in this case is the job description provided by the recruiter. We assign a score for every feature according to their vector distance between the ideal vector. Then we rank these resumes based on their total assigned score. The output will be the ranked resumes with their respective scores which will be given to the recruiter.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue VII July 2020- Available at www.ijraset.com

VIII. RESULT AND DISCUSSION

The end product will have the ability to analyze the Re- sumes of every user and give relative ranking and a score. The product will be flexible as per the requirements of the users. On an average the system shows an accuracy of 90.53 percentage for parsing and 77.66 percentage for scoring.

Accuracy Rate

Number of Resumes	Correctly Parsed	Percentage (%)
10	10	100
15	13	86.6
20	17	85

Number of Resumes	Correctly Scored	Percentage (%)
10	9	90
15	11	73
20	14	70

IX. RELEVANCE

Resume screening software eliminates the need for unreli- able manual screening, making sure no candidate is missed. The essence of any good recruitment strategy lies in efficient and effective resume screening.

X. CONCLUSION

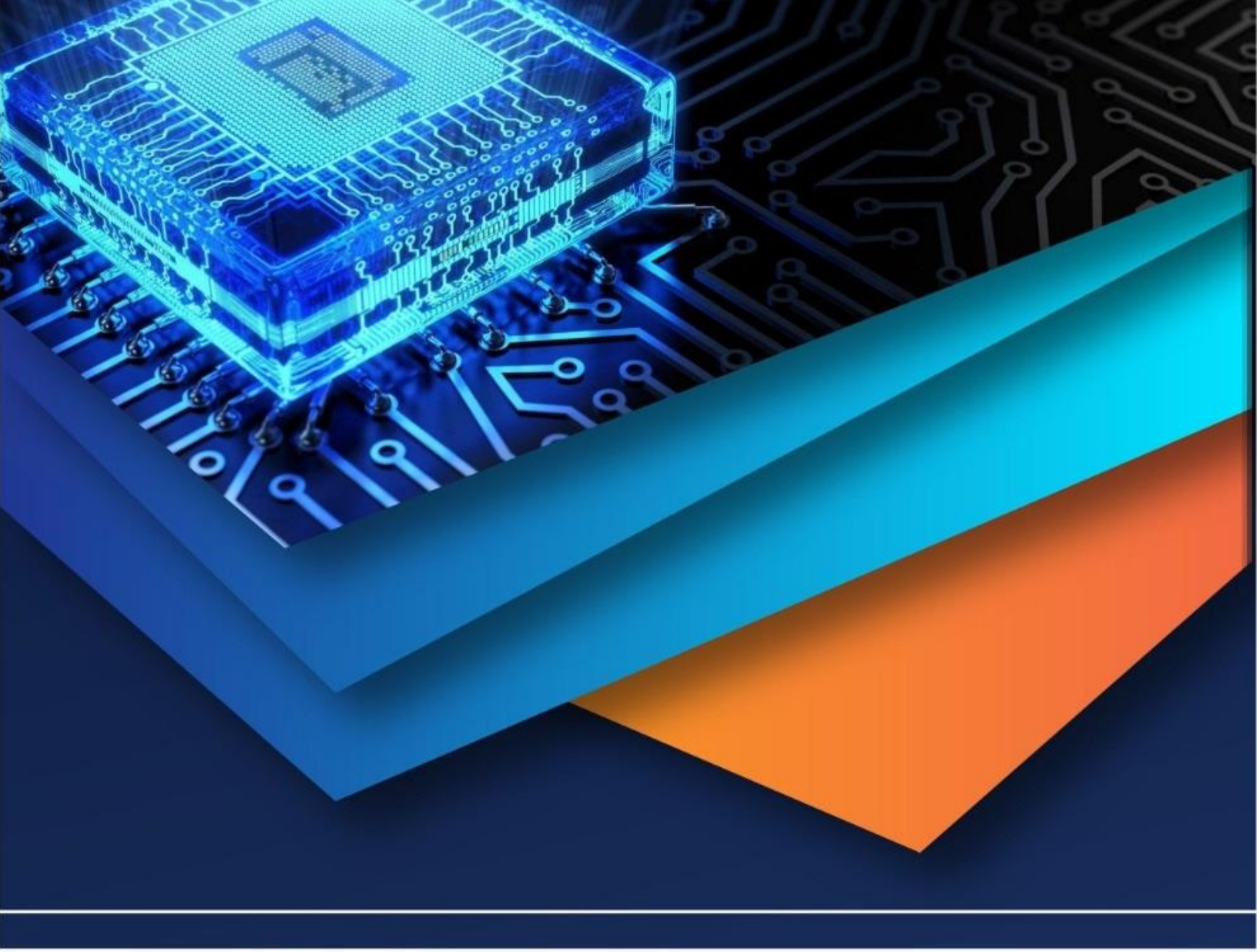
The end product will have the ability to accept or reject a job applicant based on two factors the requirements of the company should match the skills mentioned in the applicants CV and the other factor is the test evaluation which will based on the skills of the applicant which will ensure that the CVs uploaded by the applicant is genuine i.e. the applicant is really well versed in the skills mentioned in their CVs. The project will use machine learning and natural language processing algorithms like KNN classifier which will ensure text categorization and the skills from the CVs of the applicant will be added to the database based on which a test will be given to the applicant. The project also will recommend some courses to non shortlisted applicants based on the test evaluation score. Hence the project will not only benefit the company recruiters but also the applicants as they will be recommend courses if they are lacking any technical skills which will be based on the skills mentioned in the CVs and also the test evaluation score.

XI. ACKNOWLEDGEMENT

The authors would like to thank all who were involved directly or indirectly in the data collection, experiment con- duction and analysis stages of the project.

REFERENCES

[1] Evanthia Faliagka, Giannis Tzimas ," Application of Machine Learning Algorithms to an online Recruitment System" ,ICIW 2012.
[2] Ramjeet Singh Yadav, A.K. Soni, "A study of academic performance evaluation using Fuzzy Logic techniques",March 2014
[3] M.Mochol, H.Wache, and L.Nixon, "Improving the Accuracy of Job Search with Semantic Techniques", Business Information Systems,vol. 4439, 2007,pp.301-313.
[4] E.Faliagka, K.Ramantas, A.Tsakalidis, M.Viennas, E.Kafeza and G.Tzimas,"An Integrated e-Recruitment System for CV Ranking based on AHP" ,"Proc.of WEBIST 2011,May. 2011, pp. 147 -150.
[5] Belong.co, http://www.belong.co [14] ALEX System ,http://www.hireability.com/alex/
[6] S.Amdouni and W. Ben Abdessalem Karaa, "Web - based recruiting", Proc.Of International Conference on Computer Systems and Applica- tions (AICCSA),2010,pp.1-7.
[7] EICE TRANS. INF. SYST., VOL.E94–D, NO.10 OCTOBER 2011 Special Section on Information-Based Induction Sciences and Machine Learning A Short Introduction to Learning to Rank, Hang LI
[8] Referenced Links : Jessica Simko , "How Hiring Managers Make Decisions" http://www.careerealism.com/hiring-managers-decisions/Vinayak Joglekar , "Ranking Resumes using Machine- Learning" https://vinayakjoglekar.wordpress.com/2014/06/24/ranking- resumes-using-machine Peter Gold , "Artificial Intelligence Recruiting" https://www.linkedin.com/pulse/artificial-intelligence-recruiting-peter- gold Turbo Ricruit ,"Automated Application Processing","Better candidate experience","Matching Job Descriptions to Resumes" http://www.turborecruit.com.au/benefits-of-artificial-inteligence-for- recruitment/



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY