

Determination of profile via CV using Machine Learning

R.Lohitha, P.Silpa, SK.Ayeesha, V.Harshavardhini

*(Under the guidance of **Mr.M.Chennakesavarao**, Assosiate Professor,
Department of Computer Science and Engineering, Tirumala Engineering College)*

ABSTRACT

We describe a system called CV analysis machine, that processes a set of given free-form textual resumes (in English for IT domain), creates a standardized profile for each candidate, and for a given job description.

***Keywords**--e-recruitment, aptitude, personality prediction, credentials, Big five model of Personality*

I. INTRODUCTION

Identifies a ranked shortlist of k candidates, along with a matching score for each. The resume scoring function is hand-crafted, hierarchical, and uses domain-knowledge from recruitment experts. We describe a simple neural-network system that automatically learns some weights used in the scoring function, based on feedback about whether the candidate with their grades on each portion and also calculate overall grade of each candidate and also find the all grades at a time. Companies often receive thousands of resumes for each job posting and employ dedicated screeners to short list qualified applicants. In this paper, we present PROSPECT, a decision support tool to help these screeners shortlist resumes efficiently. Prospect mines resumes to extract salient aspects of candidate profiles like skills, experience in each skill, education details and past experience. Extracted information is presented in the form of facets to aid recruiters in the task of screening. We also employ Information Retrieval techniques to rank all applicants for a given job opening. In our experiments we show that extracted information improves our ranking by 30% there by making screening task simpler and more efficient.

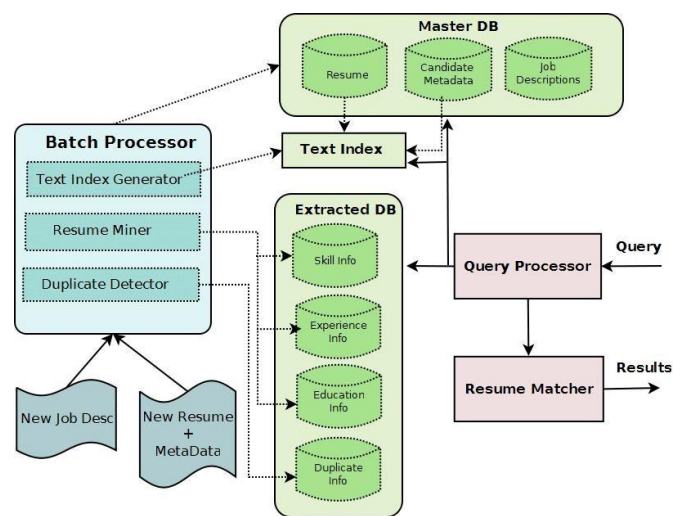
Personality is the most important factor which reflects an individual, which keeps on varying . Tackling them is a tedious task for which we have implemented an approach to identify the personality and also provide with the recommendation. This will enable a more effective way to short list submitted candidate CVs from a large number of applicants providing a consistent and fair CV ranking policy, which can be legally justified. System will rank the experience and key skills required for particular job position. Than system will rank the CV ' s based on the experience and other key skills which are required for particular job profile. This system will help the HR department to easily shortlist the candidate based on the CV ranking policy. This system will focus not only in qualification and in experience but also focuses on other important aspects, which are required for particular job position. This system will help the human resource department to select right candidate for particular job profile, which in turn provide expert workforce for the organization.

II. RELATED WORK

In this section, we present the related work and literature review of various techniques and algorithm used for online selection process.

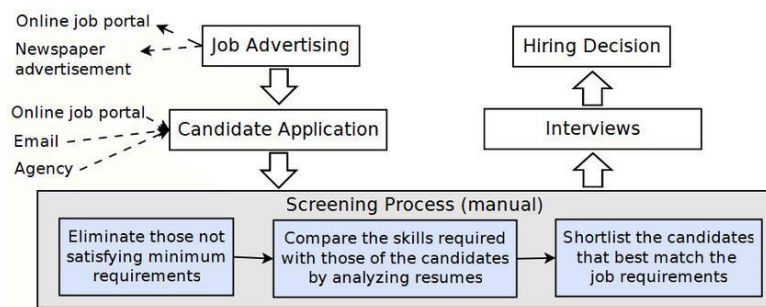
Big Five Personality Model (also known as Five Factor Model) has been used to predict the personality of the candidate which includes Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. For classifying the person Automated Personality Classification is used, which is used to classify the person from a large number of people.

Recommendation using machine learning techniques have been used for the analysis of the CV. In literature, various evaluation tools have been used. One of the approaches has been mentioned in which use a tool called “PROSPECT” for screening candidates for the recruitment.



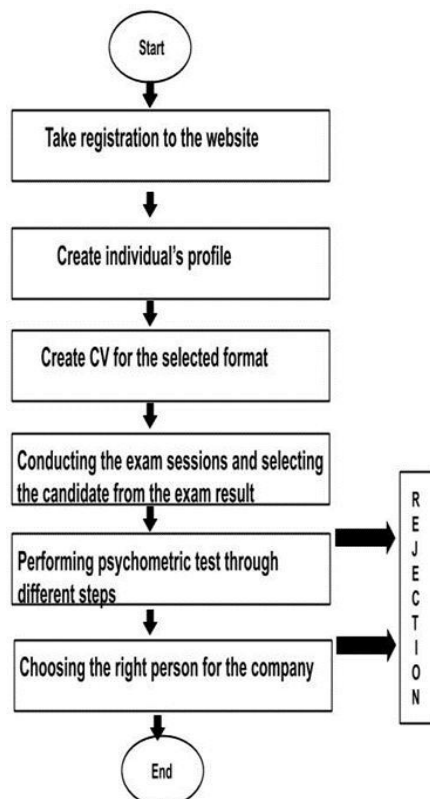
III. EXISTING SYSTEM

The candidates will register themselves with the required details in the CV form and upload the documents for verifying the authenticity of the information provided in the CV registration form. The candidates can then view the requirements and details of the job as specified by the admin or the recruiter. In order to apply for a job the candidates need to be eligible for that job by satisfying all the requirements as stated by the recruiter. If a candidate applies for a specific job, the system checks whether the candidate meets all the requirements/parameters as specified by the recruiter. If the candidate satisfies the requirements/parameters, then the candidate's request for that job would be accepted else the system denies the candidate's request for that job.



Candidate can also give an online test, which will be conducted on personality questions as well as aptitude questions. After completing the online test, candidate can view their own test results in graphical representation with marks. The system would then derive and rank the candidates who were eligible for the job. The rank of each candidate acts as a score of how well the candidate's profile meets the specifications of the recruiters as well as cumulative score of the aptitude test. The recruiter could also analyze the Personality of the candidate based on the result of the personality test. So, based on CV, aptitude test and the personality test the candidate would be selected.

IV. PROPOSED SYSTEM

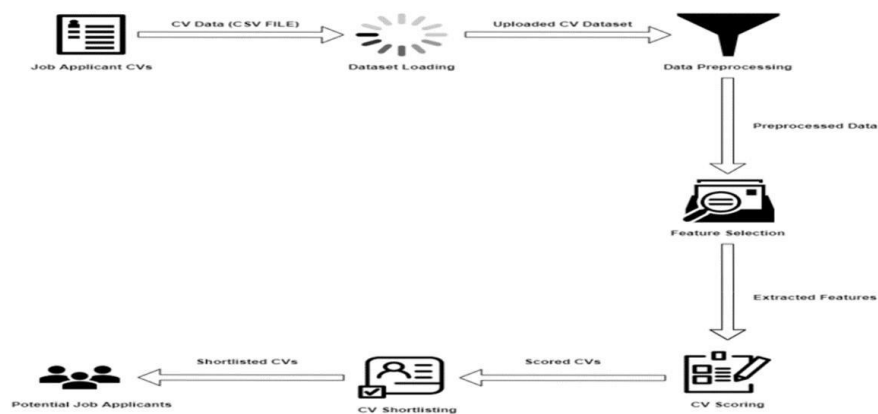


There is a huge workload on the human resource department to select the right candidate for a particular job profile which in turn would provide expert

workforce for the organization from a large pool of candidates. The proposed system will enable a more effective way to short list submitted candidate CVs from a large number of applicants providing a consistent and fair CV ranking policy. This can be legally justified. System will rank the experience and key skills required for a particular job position than system will rank the CV's based on the experience and other key skills which are required for particular job profile. This system will help the HR department to easily shortlist the candidate based on the CV ranking policy.

This system employs a machine learning algorithm namely "Logistic Regression" which helps to choose fair decisions to recruit a suitable candidate. It has the potential to transform research and assessment in personality psychology. Algorithms can handle vast datasets, including thousands of attributes, without succumbing to collinearity issues. Moreover, ML algorithms are highly efficient in recognizing patterns in datasets that humans cannot even perceive. It does not consume more time to identify the personalities. This predictive personality makes the recruiters understand if a candidate will be a top performer and fit to the culture of the company easily.

V. SYSTEM ARCHITECTURE



VI. DESIGN ALGORITHM

In this section we state the machine learning algorithm (TF-IDF) for CV analysis. The TF-IDF Algorithm is used to find out the important keywords in a document/CV. Below, we give the working of TF-IDF in detail.

Step 1: Calculate TF (Term Frequency) Term Frequency (TF) - Number of times a keyword appeared in a document is calculated by Term Frequency. $TF(\text{keyword}) = \frac{\text{number of times 'keyword' appears in document}}{\text{Total number of keywords in the document}}$. Here, the term 'keyword' signifies any job specific skill which the algorithm is searching for.

Step 2: Calculate IDF (Inverse Document Frequency) value. The problem of rare and frequent words is resolved. This helps our system to give more priority to the required word or skills. IDF sets the log value=1 for the required CV as per skill sets and log value=0 for the unwanted CV.

$IDF('keyword') = \log(\text{total number of CV} / \text{Number of document with term 'keyword'})$

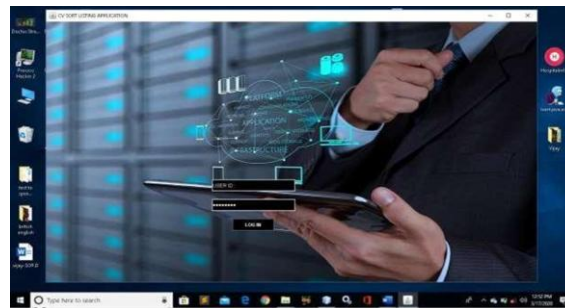
Step 3: Calculate TF-IDF weight

$\text{Weight} = TF('keyword') * IDF('keyword')$

Higher the weight, more relevant is the CV and lower the weight, less or no relevance of the CV for the selection process. This step returns the CV with highest and lowest weight values which is further useful for classification

VII. RESULTS

Admin login:

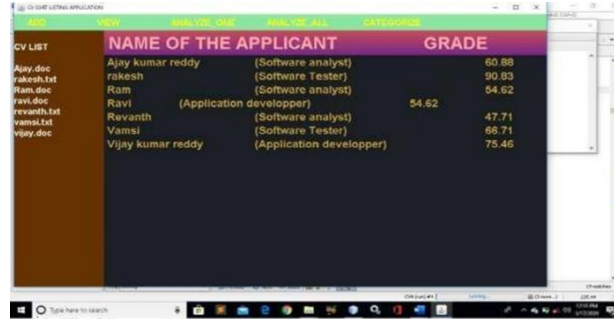


To analyze the cv at first we need to load the file containing the information about the person who wants to analyze. We load any number of cv by just clicking the button add and it shows the popup window and we select the file using the path of the file by file chooser. The first main idea coming from this section in this section the company provides a basic email and they offer to all persons who are eligible for the job they need to send their resume to the particular email after completion of the time period the company head collect the all resumes and store those files anywhere in the system.

Filechooser:



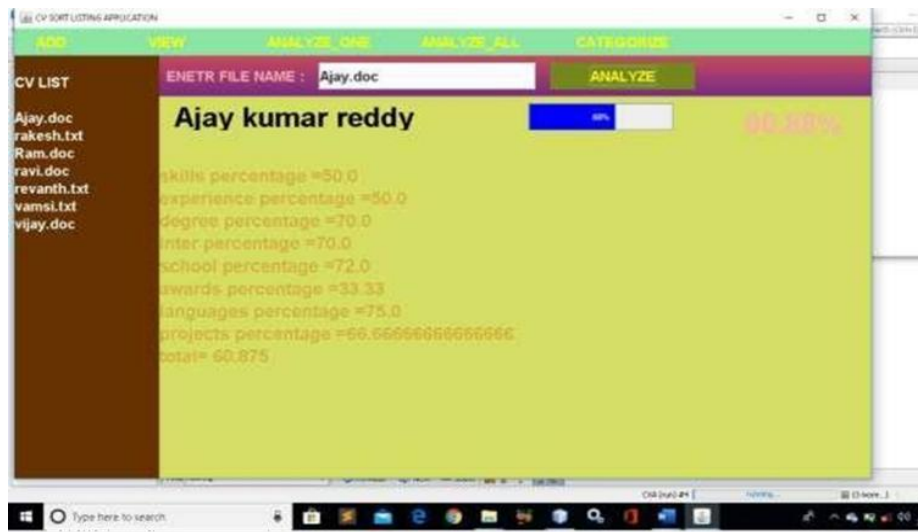
Read the content:



CV LIST	NAME OF THE APPLICANT	GRADE
Ajay.doc	Ajay kumar reddy (Software analyst)	60.88
rakesh.txt	rakesh (Software Tester)	90.83
Ram.doc	Ram (Software analyst)	54.62
ravi.doc	Ravi (Application developer)	54.62
revanth.txt	Revanth (Software analyst)	47.71
vamsi.txt	Vamsi (Software Tester)	66.71
vijay.doc	Vijay kumar reddy (Application developer)	75.46

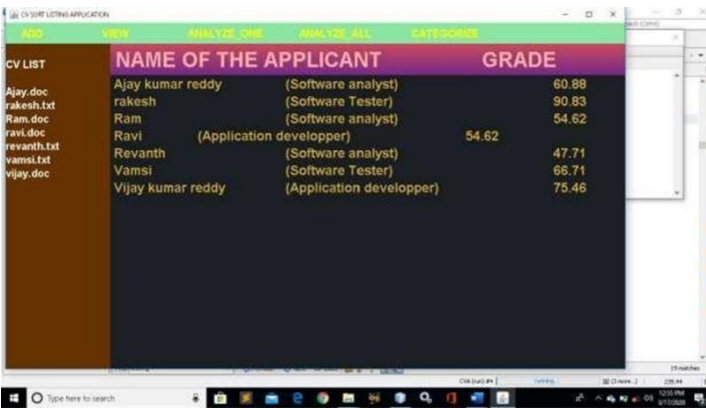
This module enables the user or officer to view the content in the document and observe the format of the cv is suitable for your application and this module offers to head officials to read a particular cv and check for there is any tricks are apply for the grade.

Analyze one cv:



We can also analyze the one cv at a time by using the module analyze one panel , we simply type the file name in the text field provided in the same panel.After loading the file into the panel it analyze the cv and give the information about the candidate and his overall grade and overview in all fields.

Analyze all:

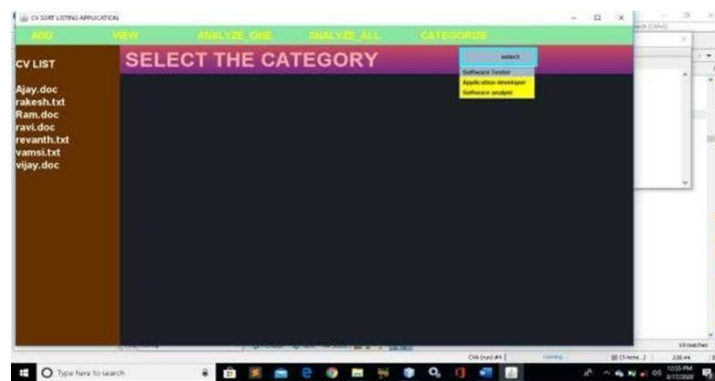


CV LIST	NAME OF THE APPLICANT	GRADE	
Ajay.doc	Ajay kumar reddy (Software analyst)	60.88	
rakesh.txt	rakesh (Software Tester)	90.83	
Ram.doc	Ram (Software analyst)	54.62	
ravi.doc	Ravi (Application developer)	54.62	
revanth.txt	Revanth (Software analyst)	47.71	
vamsi.txt	Vamsi (Software Tester)	65.71	
vijay.doc	Vijay kumar reddy (Application developer)	75.46	

We can also analyze the all cv we loaded into the application in this windows shows the all the grades forthe each applicant.

Categorize:

We can also find out the presongs based on their workcategory



VIII. CONCLUSION AND FUTURE SCOPE

In recent years, all industries and IT companies feel high burden to analyze all resumes provided by the candidates who are willing to participate job requirement processes so we need automated system for analyzing and shortlisting the all resumes at a time and also calculate the grades of the persons and their individual grades in each section like experience ,skills, education background , languages known and number of projects did. So this project will help almost all companies who are willing to analyze all resume at a time without putting any human effect.

In future cv analysis is required for each and every organization and adding Additional features like “ formats of the resumes are independent for analysis and applying for a job “.This will help to reduce the time taken to

analyses all the cv that are received from many candidates who are willing to do the particular job that a company offers . which can predict the traits of an individual using the group of tweets posted by him. It is different from the approaches reviewed in the literature survey, in that it works with the group of tweets and does not take user's profile into account. Further, the system makes use of the Hadoop framework to predict personality traits of multiple individuals at the same time.

IX. REFERENCES

1. Anderson, N. H., & Barrios, A.A. (1961). Primacy effects in personality impression formation. *Journal of Abnormal and Social Psychology*, 63, 346- 350. <https://doi.org/10.1037/h0046719>
2. Arnulf, J., K., Tegner, L., & Larssen, Ø. (2010). Impression making by resume layout: Its impact on the probability of being shortlisted. *European Journal of Work and Organizational Psychology*, 19, 221-230. <https://doi.org/10.1080/13594320902903613>
3. Augustin, H. M. (1991). The written job search: A comparison of the traditional and a nontraditional approach. *Business Communication Quarterly*, 54, 13-14. <https://doi.org/10.1177/108056999105400304>
4. Bendick, M., Jackson, C. W., & Romero, J. H. (1996). Employment discrimination against older workers: An experimental study of hiring practices. *Journal of Aging & Social Policy*, 8, 25-46. https://doi.org/10.1300/J031v08n04_03
5. Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *The American Economic Review*, 94, 991-1013. <https://doi.org/10.1257/0002828042002561>
6. Bird, C. P., & Puglisi, D. D. (1986). Method of resume reproduction and evaluations of employment suitability. *The Journal of Business Communication*, 23, 31-39. <https://doi.org/10.1177/002194368602300303>
7. Blackburn-Brockman, E., & Belanger, K. (2001). One page or two?: A national study of CPA recruiters' preferences for resume length. *The Journal of Business Communication*, 38, 29-57. <https://doi.org/10.1177/002194360103800104>