International Journal of Research in Advent Technology, Special Issue, March 2019
E-ISSN: 2321-9637
Available online at www.ijrat.org

Prediction & Analysis of CV (PAC)

Mayank Srivastava[1], Akshat Srivastava[2], Krishan Kumar Sharma[3]
[*]*Department of Computer Engineering & Applications,*
GLA University, Mathura,
UP, India
mayank.srivastava@gla.ac.in[1], akshat.srivastava_cs15@gla.ac.in[2], krishan.sharma_cs15@gla.ac.in[3]

Abstract- This paper is aimed at making the process of hiring a semi–automatic one by providing a standard set of questions to the interviewer so that the play field remains even for the appearing candidate. The project aims at taking in PDF format as input then after going into the parser the extraction of the information is done and the data extracted is analyzed for any errors which may be suggested to the applicant as an improvement to the resume. The use of specific words which tells the proficiency of the applicant in the particular subject are taken into special consideration and the question set is suggested in that fashion only. The description of knowhow of the subject by the applicant reveals us the level of hardness an applicant can face an interviewer. Furthermore, the professionalism of the resume such as not missing the indent and various mistakes in the resume tells us the professionalism of the applicant. The more professional the applicant the more sophisticated the interview. Also, the job applicant can be suggested the various jobs just on the choice of subjects he has mentioned in the resume. The profile of the job and the subjects mentioned by the applicant must intersect for the applicant to be eligible for a particular profile.

Keywords: Prediction, document extraction, Spell checker, resume analysis.

1. INTRODUCTION

PAC aims at extracting the information out of a PDF. We have chosen the portable digital format because it does not change across various platforms [1,2]. As designing the parser for various formats was a bit more tedious so we firstly convert the format to PDF then proceed with extraction [3]. The extraction is done with the help of python library named PymuPDF. The result achieved with our extraction were quite satisfactory and we were able to arrange the information in a well-designed manner. After the data is extracted we move forward for checking of various words with the help of spell checker program module of our project to show us both indentation errors and the words which might have been misspelled [4]. The spell checker and indentation error checker is also written in python. After these extraction and arrangement and the suggestion processes we move on to picking out the subjects and the level of expertise if mentioned by the applicant in particular and proceed further on the suggestion module of our project which suggests the questions to the interviewer to be asked to the applicant and it further tells to the interviewer on the basis of resume analysis whether the candidate is eligible for the job profile.

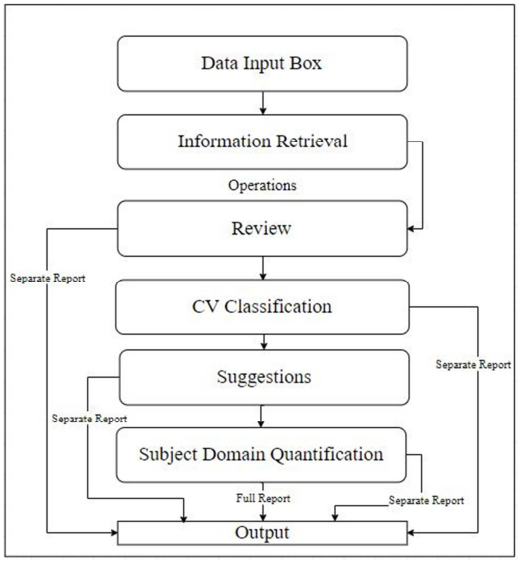

Fig 1. Explaining working mechanism of PAC

2. SCENARIO AND EXISTING SYSTEMS

In the present scenario big companies Like Microsoft Azure's tool, Wipro,Google, Mckinsey are working on this single stop solution but has not achieved it yet with a 100% accuracy. The biggest challenge is posed when we have to get the unstructured data into a structured format and do something useful with it. The problem is an evident one and even big companies with such huge funding's backing them have not been able to achieve

desired results. All of these companies have been able to achieve some significant results but they have yet not found a single stop solution. For example, google has mastered sentiment analysis with staggering level of accuracy but analysis of a resume has yet not been achieved with the accuracy desired at the onset. Same is the case with other companies mentioned above.

3. IMPLEMENTATION ALGORITHM

Algorithm:

- Take in the resume in any format
- Convert it into the PDF
- Send the PDF to the extraction module
- Check the extracted information with The help of the spell checker for indentation and spelling errors
- Suggest any changes to the applicant for improvement in the resume
- Send the data to the suggestion module of the program
- Suggest the interviewer the set of questions and the job eligibility of the candidate
- Reflect on the professionalism of the applicant

Our approach differs from others in the manner it suggests the questions to the interviewer and selects the various candidates for various profiles on the basis of the subject mentioned in the resume. Other software or tools of various companies are busier in finding the sentiment or just analyzing the resumes. None of them were using the information gained to refine the interview process. These organizations were more focused on analyzing the interview process rather than automating or semi-automating it. May be in future they will move towards it and there are signs of it from companies like Wipro and Microsoft but right now no one has been doing it so we dared to go on this unchartered plain and explore for ourselves what challengges might lie ahead of us.


CONTACT
WORKING KNOWLEDGE
EDUCATION
WORKSHOPS
Trainee: Web Developer
Attended Workshop on Machine Learning using python at
GLA University, Mathura.
SUMMER INTERNSHIP PROGRAM
INTUIT INTEGRATED MARKETING| 1 JUNE'17-15 JULY'17
GLA University | 2015 - 2019
BACHELOR OF TECHNOLOGY, MAJOR
IN COMPUTER SCIENCE AND ENGINEERING.
SUMMARY
ACHIEVEMENTS
Nominated among top 10 projects titled as "Automatic
Vehicle Registry System" in Technotsav'18.
Organized 2-day workshop on Data Science at GLA
University.
Subject of Interest: Data Structure, DBMS, Software
Engineering.
Language: C, Python, Java, HTML.
SKILL SET
AKSHAT SRIVASTAVA
COMPUTER SCIENCE ENGINEERING
STUDENT
LinkedIn: linkedin.com/in/akshat-
srivastava-089b2b136
Github : github.com/
Akshatsrivastava700
SOCIAL MEDIA
PROJECTS
Biometrics Analysis Based on Ear Recognition. (Python):
Researched based project to find the difference between two
ears.
Automatic Vehicle Registry System(Python & OpenCV):
A system which collects information of the vehicles passing
through a gate by scanning the QR code.
INTERMEDIATE
Blue Bird Senior Secondary School | 2014
Blue Bird Senior Secondary School | 2012
HIGH SCHOOL

Fig 2. Text retrieved from a PDF

4. CONCLUSION

This project will help the users to develop a better understanding of their resume by showing them the errors and insight information hidden within the data. Since this type of project is also being done by the companies this project helps us build our knowledge. This project will also be able to ease the process of recruitment by giving them the insight information in a short period of time. Thank You.

REFERENCES

[1] Vishnu M Menon and H A Rahulnath , A novel approach to evaluate and rank candidates in recruitment process by estimating emotional intelligence through social media data, IEEE 2016 International Conference on Next Generation Intelligent Systems (ICNGIS), 1-3 Sept. 2016.

[2] Kanupriya Sharma and Amanpreet Kaur, A review of the existing state of Personality prediction of Twitter users with Machine Learning Algorithms, IOSR Journal of Computer Engineering (IOSR-JCE), Volume 17, Issue 4, Ver. III (July –Aug. 2015), PP 35-38

[3] R.C. Warnacula, Y. Munas, C. S. Karunasekara, T. A. Fernando and N. Vithana, Personality Based E-Recruitment System, International Journal of Innovative Research in Computerand Communication Engineering, Vol.5, Issue 12, December2017.

[4] Gayatri Vaidya, Pratima Yadav, Reena Yadav and Chandana Nighut, Personality Prediction By Discrete Methodology, IOSR Journal of Engineering (IOSRJEN), Volume 14, P P 10-13.