

Review Article

Survey on Resume Screening Mechanisms

Tumula Mani Harsha¹, Gangaraju Sai Moukthika², Dudipalli Siva Sai³, Mannuru Naga Rajeswari Pravallika⁴, Satish Anamalamudi⁵

^{1,2,3,4}Bachelor of Technology 4th year, Computer Science Engineering, SRM University, Amaravati, Andhra Pradesh.

⁵Science Engineering, SRM University, Amaravati, Andhra Pradesh ^[5]

Received: 06 March 2022

Revised: 08 April 2022

Accepted: 12 April 2022

Published: 30 April 2022

Abstract - Resume Screening is the primary step in the hiring process. It evaluates the candidates' resumes and determines whether they are qualified for a role based on their education, skill sets, technical stuff, experience, and other information captured in their resume. To make it simple, it's a form of pattern that matches the job requirement and the candidate's qualifications based on their resume. It is a crucial step in the process of hiring. It is the step in which a decision is made to move the candidate to the next level or not. There are multiple processes to perform resume screening. Among all the processes, traditional resume or manual screening is the largest followed, even today. But usually, companies receive thousands of resumes for job applications, which consumes a lot of time and effort. In addition to this, many errors may arise due to human involvement. Multiple ways were introduced to cover all these cons to performing this resume screening process. Various technologies, including Artificial Intelligence and Machine Learning, were involved in searching for solutions. This paper contains a detailed survey report on various methodologies and techniques of resume screening.

Keywords - Resume screening, Artificial Intelligence, Machine learning, Hiring.

1. Introduction

The basic Job Hiring process consists of 5 steps in which resume screening is the first and most effective step to segregate the candidates who are fit for the required job role.

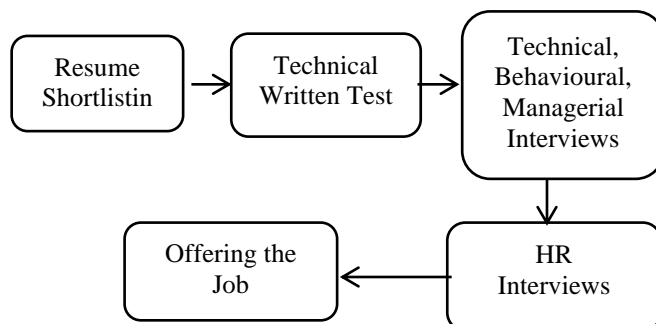


Fig. 1 Basic Job Hiring Process

1.1. What is Resume Screening?

Resume Screening is the process of evaluating the candidates' resumes based on a specific requirement. It is a process of searching for a perfect fit for the required role based on the candidate's qualifications like education, experience, skills, capabilities, and so on mentioned in their resumes.

The qualities of the candidates should fit the requirement for selecting the candidate to enter the next stage of the interview.

1.2. Resume Screening Consists of 4 Basic Steps

- Selecting the resumes with the required credentials, i.e., only the resumes matching the required job description, will be considered.
- Select the resumes with the desired skills, i.e., only the candidates who mentioned the skills that match the job role will be considered.
- Selecting the resumes customized for the job, i.e., considering the resumes that accurately match the job role.
- Checking the applicant's information, i.e., scanning all the candidate's information and considering which matches the job profile properly.

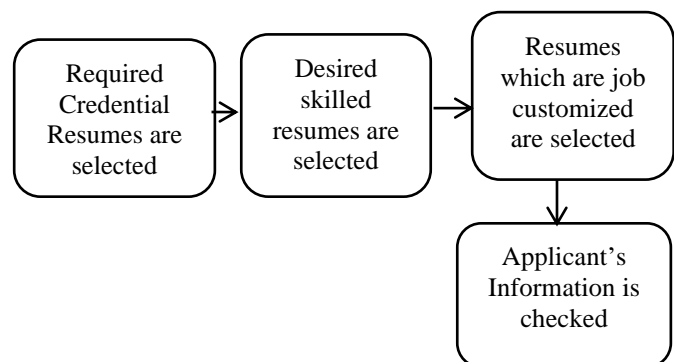


Fig. 2 Steps involved in Resume Screening

Some of the most common difficulties encountered during the resume screening process



- Time Consuming-When there are a lot of resumes, it becomes more difficult.
- Hiring Quality-When there is a bulk of resumes, the hiring quality suffers.
- Hiring Biases-Recruiters may be biased in favor of certain prospects.
- Recruiter Experience-If, the recruiter, is unfamiliar with all of the job-related abilities, the screening procedure may be inefficient.
- Recruiters Search-If, a suitable applicant is located, recruiters may halt the search and skip the rest of the stack.
- Unnecessary recruiter allocation-Separate recruiters should be assigned to the resume screening process, wasting time and resources.

2. Surveyed Mechanisms

1. Manual Screening
2. Resume Screening using Artificial Intelligence
3. Resume Screening using Machine Learning Classifiers
4. Resume Screening using Deep Learning

3. Related Works

3.1. Manual Screening

Screening of resumes is done by some of the company's employees who are going to recruit, i.e., every resume is checked individually, and if the resume is fit for the required job description, then the resume will be selected.

It might be done based on the capabilities they seek, the candidates' work experience, or other factors that are relevant to the job profile.

3.1.1. Problems in this Approach

Time Consuming

All resumes must be referred manually, which takes a long time.

Recruiters are under a lot of Pressure

Even if all resumes are manually referred to, the procedure will take twice as long if there is no correct fit for the job profile.

Unnecessary Resource Allocation

Recruiters could be working on other projects instead of spending so much time on resume checking.

Inefficient

Once a requirement is identified, recruiters do not review all the resumes.

Errors

Due to many resumes and the little time available for processing, some mistakes may be made.

Biased

Recruiters may favor some candidates over others based on their preferences.

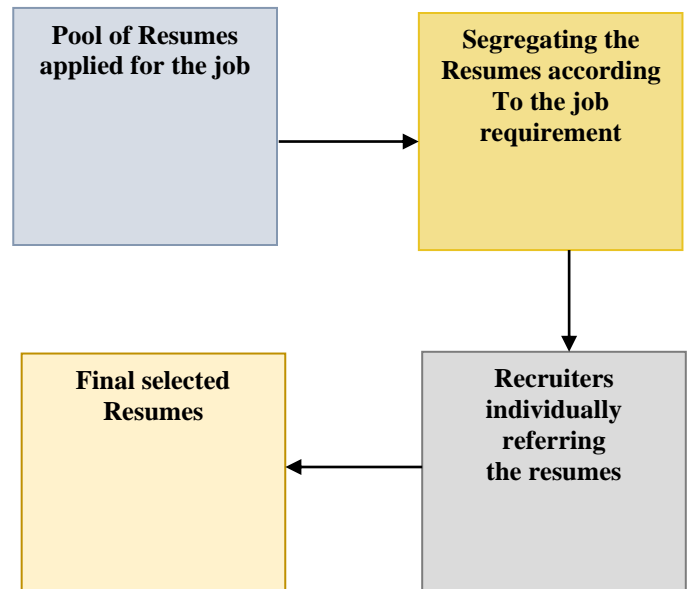


Fig. 3 Manual Screening Process

3.2. Resume Screening using Artificial Intelligence

There are different approaches to Resumes Screening using AI.

3.2.1. Resume Evaluation System based on AI

According to the article Resume Evaluation System based on AI, A Resume Screening Software is created using artificial intelligence, text mining, and processing algorithms. These algorithms filter and rank resumes based on specific keywords to identify which job applications recruiters should consider further.

Their main idea is to build a smart and automated Resume Evaluation System based on AI to overcome manual Resume Evaluation techniques for effectively classifying and shortlisting desired applicants.

The fundamental Working Procedure

The resume should be in PDF format to be viewed and read, and the text extracted. One resume at a time is given to the system. The excess material will be eliminated later, and the keywords, such as prerequisites, will be classified by area. The process is then repeated by calculating the scores for each region and sorting the results before generating the final scores. Finally, a pie chart will be displayed as an output based on the scores, helping recruiters select the required and eligible individuals for the offered job role.

i) Problems in this process

- **Complexity of Time**
It takes a long time to review each resume and calculate the skills.
- **Not user-friendly**
The application can only be used effectively by the programmer.

- **Tightly connected blocks**

Because each code block is interdependent, each change requires disrupting the entire code, which may disrupt the flow.

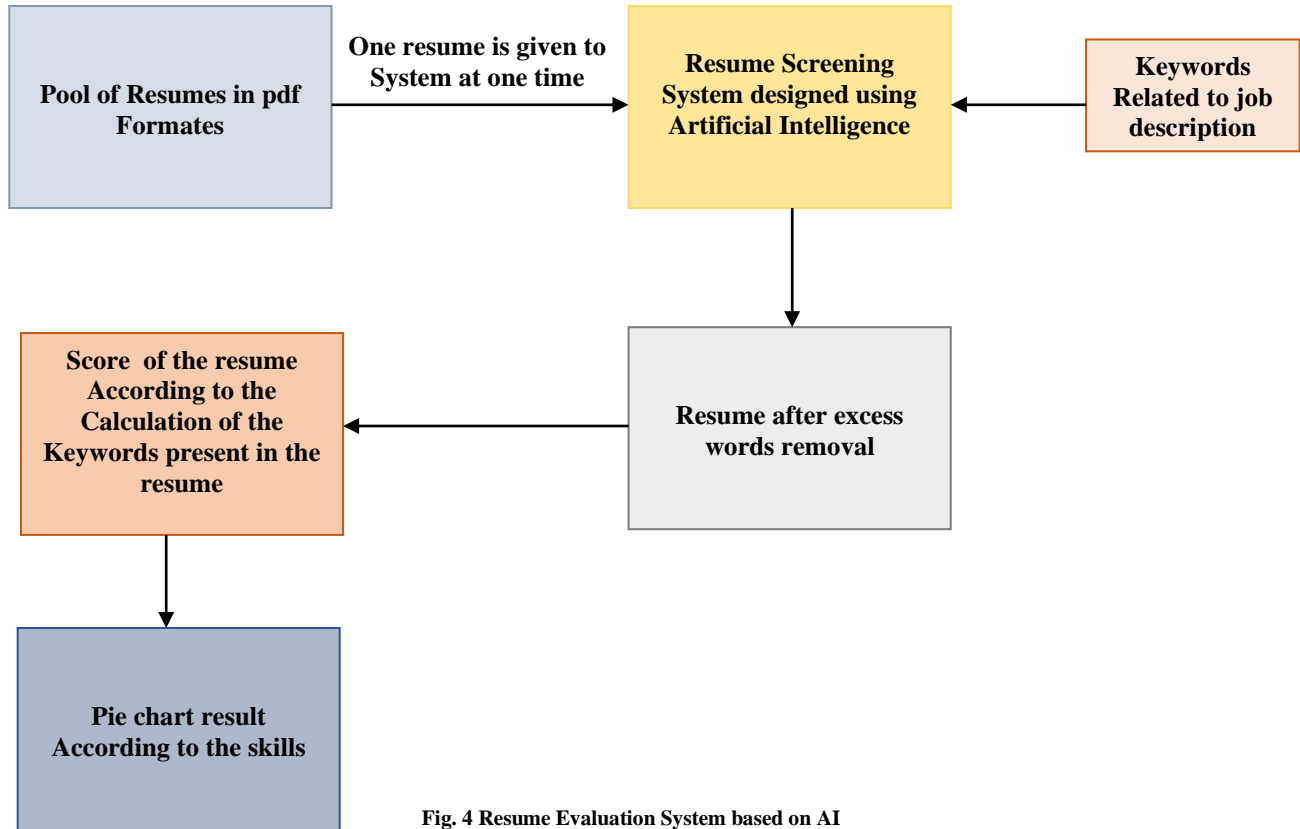


Fig. 4 Resume Evaluation System based on AI

3.3. Resume Sorting using Artificial Intelligence

According to the article Resume Sorting using Artificial Intelligence, A Database is created to store job applicants' resumes. This system is trained using Artificial Intelligence to recognize the words separately from a resume. Some important keywords that fit the job description, like skills, education qualification, and so on, are given to the system separately through some other files. The system is trained to scan the resumes and search for the separately given keywords. By matching the keywords in the resume, the system will shortlist the candidates according to the requirement. The system will select all the resumes that maximum reaches the job description requirement, and the remaining resumes will be rejected.

The entire project is designed to streamline or automate some parts of the recruiting workflow, especially repetitive high-volume tasks.

3.3.1. Advantages

- No need to specially allocate the employees for the screening process
- A bulk number of resumes can be scanned at a time, and fast results can be obtained
- Biased free recruitments
- As the correct fit for the job description is approved, the authorities can easily select candidates to attend the next round of interviews.

3.3.2. Problems in this Process

- **Not user-friendly**

The application can only be used effectively by the programmer.

- **Tightly connected blocks**

Because each code block is interdependent, each change requires disrupting the entire code, which may disrupt the flow.

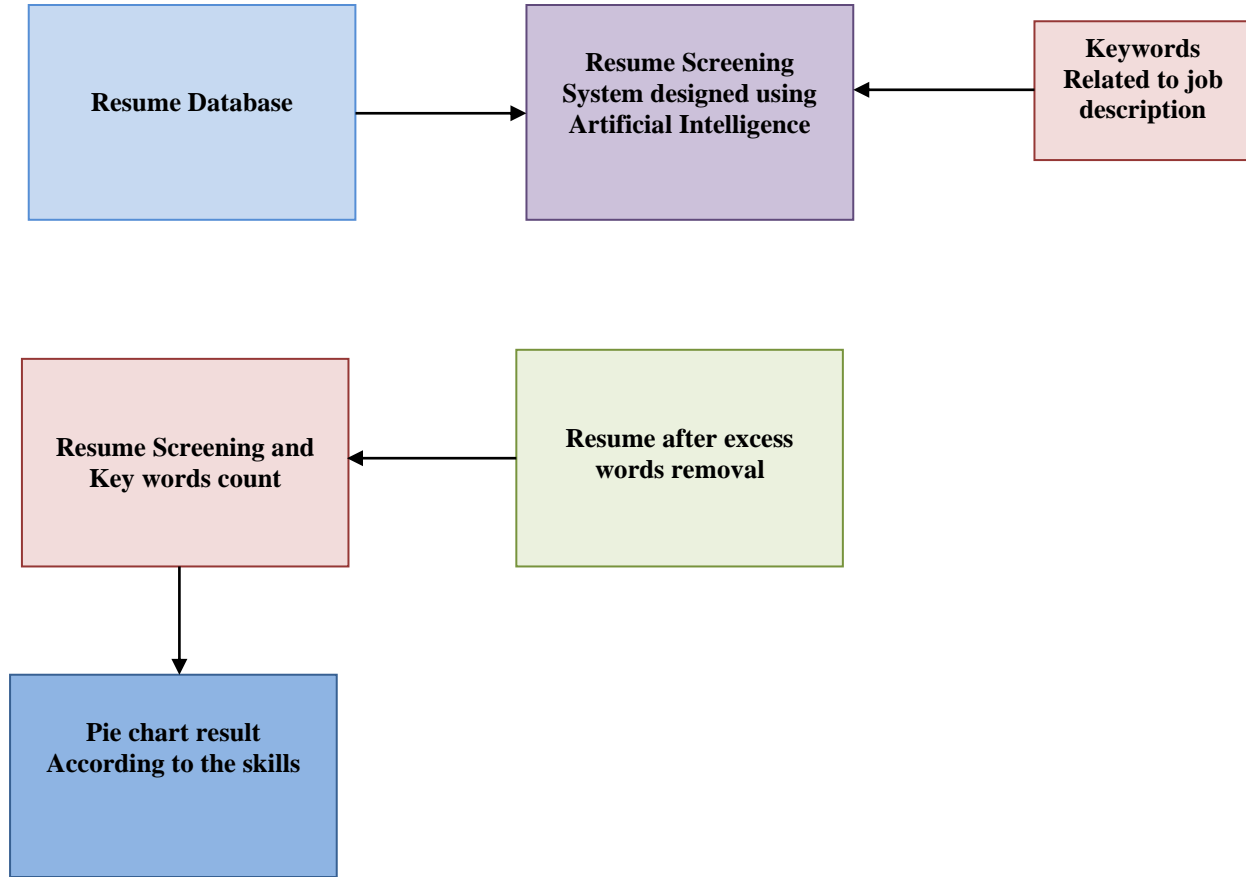


Fig. 5 Resume sorting using artificial intelligence

3.4. Resume Screening using Machine Learning

Resume screening using machine learning has different kinds of approaches. They are as follows:

3.4.1. A Machine Learning Approach for automation of the Resume Recommendation system

According to the article A Machine Learning Approach for automation of Resume Recommendation system, a Resume recommendation system is developed. The system starts by cleansing all the unimportant information in the resumes, such as numbers, special characters, single-letter words, etc. They are then placed in one data set where the data is split into tokens using NLTK tokenizes. Resumes in raw format are imported. The resume field data are cleansed using further steps such as stop word removal, stemming, and lemmatization are applied as pre-processors for this filtration process.

The next step, which extracts the features using Tf-Idf, carries out the feature extraction process with the cleansed data. There is a need for a fixed size of the numerical vector as input to the process in machine learning algorithms. So

Tf-Idf (term frequency and inverse document frequency) helps calculate each term present in the data set.

Cleansed data is used in the model to build two models. They are:

Classification

The model has been built to categorize the resume into the appropriate category based on the resume and category.

The classification was done using four different models: Random Forest, Naive Bayes, Logistic Regression, and Linear Support Vector Classifier (Linear SVM). These are used to get an accurate score, and they are recorded.

Recommendation

The model generates a summary of the resume and job description presented by the recruiter and a list of the most eligible resumes based on resume and job description similarity. This model can be approached in two ways:

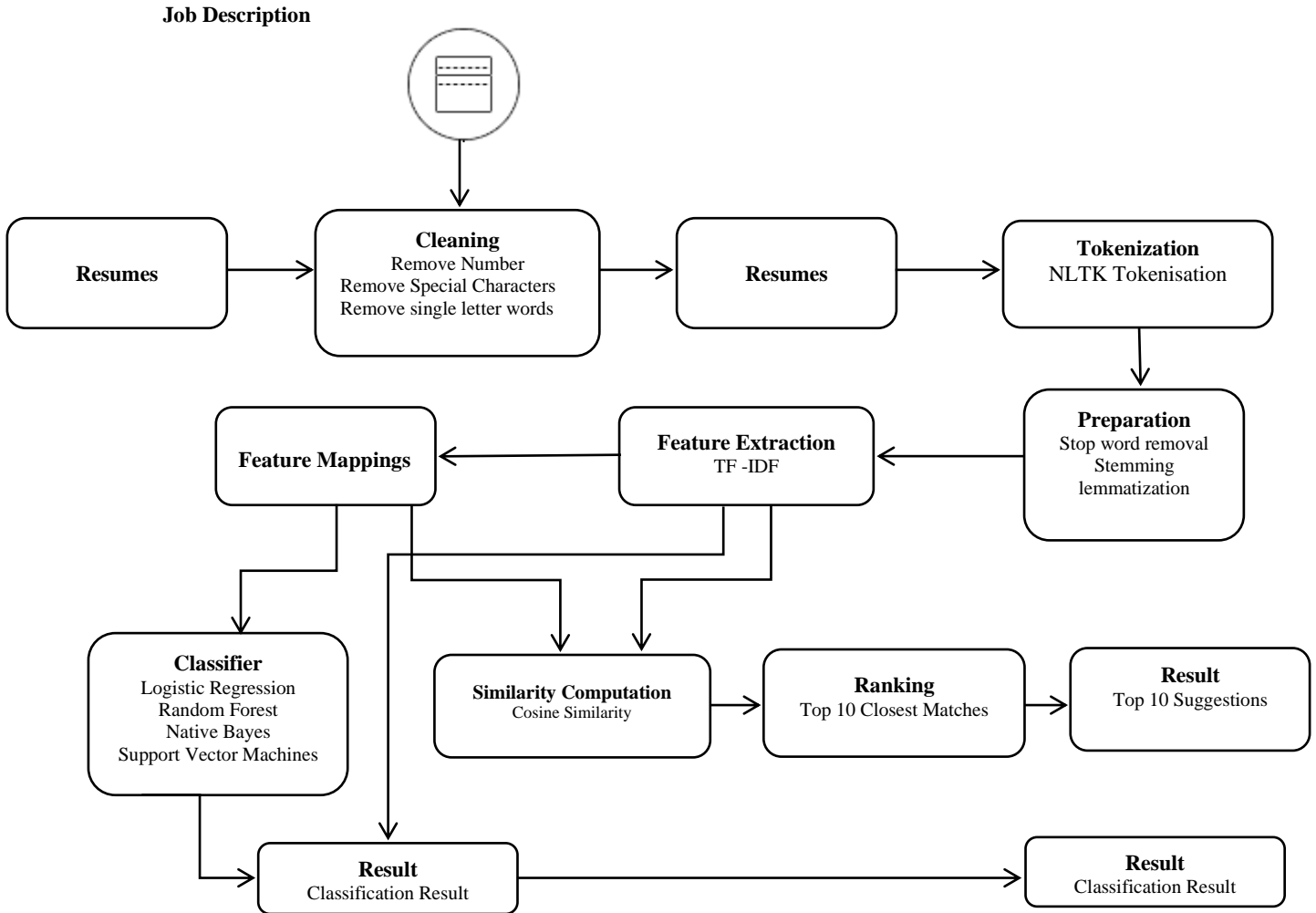


Fig. 6 A Machine Learning approach for automation of the Resume Recommendation system

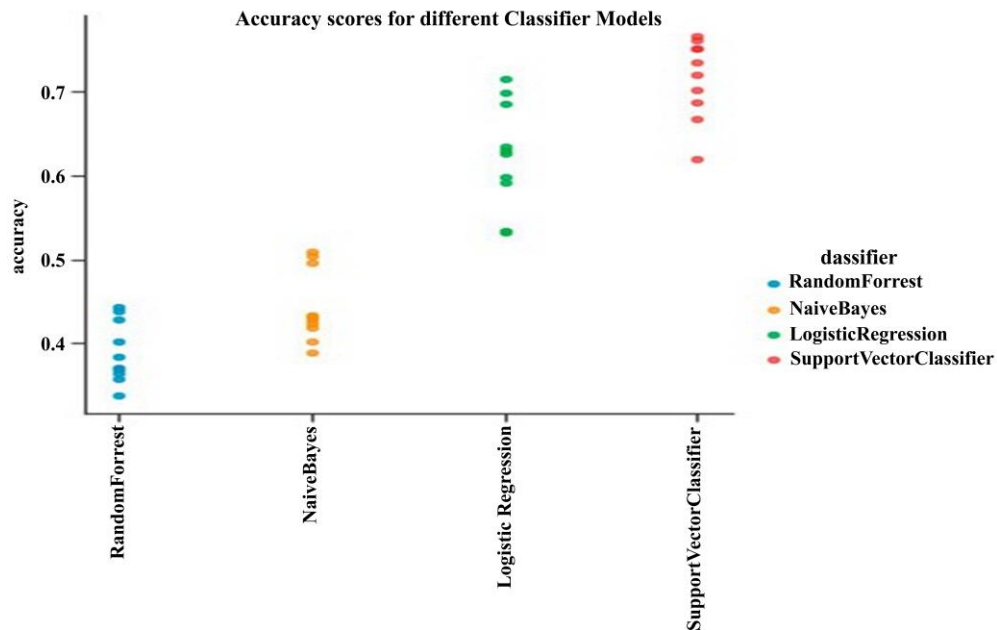


Fig. 7 Based on the above data, Linear SVM performs well in this model

K-Nearest Neighbors

K-NN is used in this model to find resumes that closely match the given job description. So to get similar and accurate resumes matching the job description, an open-source library package called gensim is used to generate the summary of the provided text within the word limit specified. With the help of a summary, the K-NN algorithm is used to locate close matches.

Machine learning classifiers techniques help first classify the right categories using a different classifier; once classification has been done, then as per the job description, top candidates could be ranked using Content-based recommendation, using cosine similarity, and by using k-NN to identify the resumes that are nearest to the provided job description. The system is an automated resume

recommendation system. It takes the features extracted from the applicant's CV as input and categorizes them, then maps the categorized resume to the required job description and recommends the best-suited candidate's profile to HR.

3.5. Resume Screening using Machine Learning

As per the article named Resume Screening using Machine Learning, the fundamental working procedure is resumes should be in CSV format. The screening process should begin with removing garbage terms (unwanted/repeated words). The remaining words are then screened, and skill points are granted. And the skill points will be assigned in the appropriate order. Finally, a graph will be displayed due to the skill points, allowing eligible candidates for the job role to be selected.

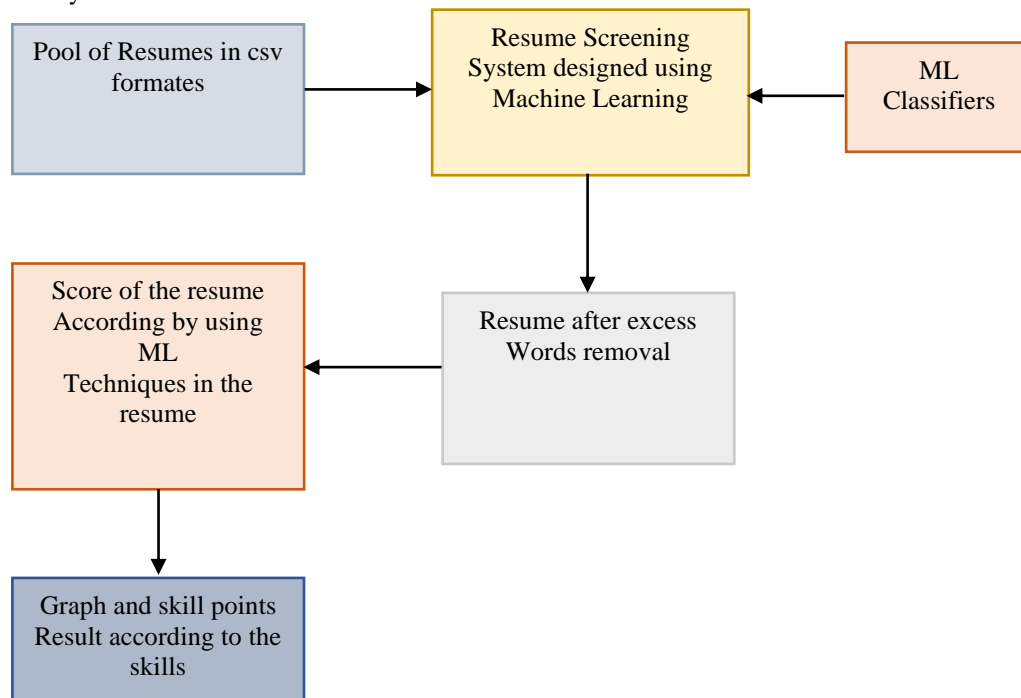


Fig. 8 Resume screening using machine learning

3.5.1. Advantages

- A large number of resumes can be scanned at a time.
- The correct fit for the job role can be identified easily.
- It takes less time to complete the process.

Limitations of the Machine Learning Approach

i) Issues with Formatting

Only CSV format resumes will be evaluated, which is not always practicable because the most common resume formats are Word or PDF.

ii) A Method that is Extremely Difficult To Implement

The concept of dealing with these types of approaches will be complex, and not everyone will be able to grasp them.

iii) Tightly Connected Blocks

Because each code block is interdependent, each change requires disrupting the entire code, which may disrupt the flow.

iv) Data loss

Some approaches, such as gensim (a library for filtering data), may lose crucial data, lowering the engine's performance.

3.6. Resume Screening Using Deep Learning

Resume Screening techniques based on Deep Learning are mentioned as follows:

3.6.1. Resume Screening Using Deep Learning

As per the process mentioned in the article Resume Screening using Deep Learning, the data set primarily has 2 columns in this project. They are Category and Resume. The category represents the fields like Java, Database, Blockchain, and so on. We must allocate the resume into one of the categories by utilizing it as input.

- We could get the frequency-wise distribution of distinct categories in our dataset by analyzing the data using value counts on categories.
- Then pre-processing is required to remove all the irrelevant information in the resume.

- Furthermore, stop words are needed to be removed using nltk, which does not impact the information.
- Following the cleansing and pre-processing, the next step is splitting the data and tokenising features and labels. Due to this, the most recurring words are given less weightage, and less frequent words are given more prominence. It makes the concise words less important, and the unique words are more useful.
- Now training and evaluating the model's test score and accuracy are produced, which helps generate the graphs.

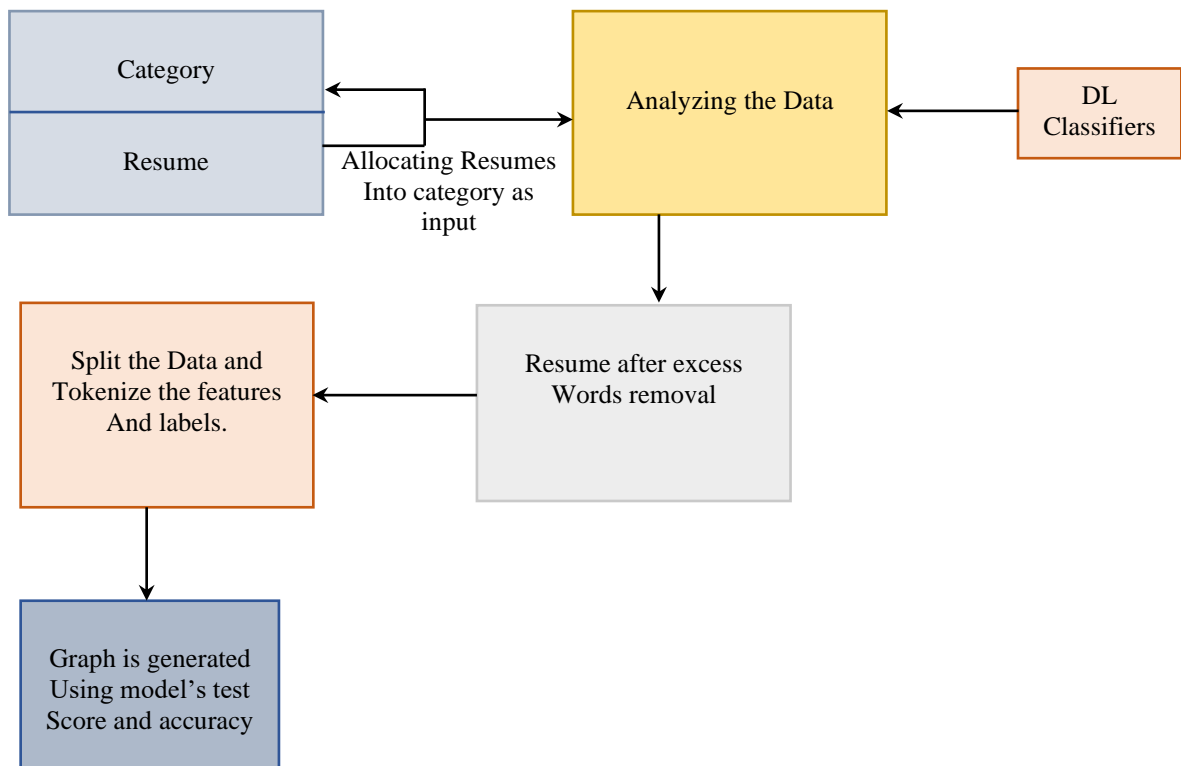


Fig. 9 Resume screening using deep learning

3.6.2. Resume Screening with Python

It was created using the Deep Learning method and learns from its prior data.

The fundamental working procedure

Category: Specific types of jobs where resume fits.

Resume: Applicant's Resume

Using the specified categories and resumes, the screening process has begun. Each resume is taken in and screened, then used to create output for that specific resume. As a result of natural language processing, a pie chart is produced. For each resume, a new graph will be provided. With each run, it will become more efficient.

3.6.3. Problems

- **There will be no bulk resume screening.**

Only one resume is screened in each run, and a report is generated. It is the result of a single resume.

- **Time Consumption**

The more resumes you have, the longer it will take for screening.

4. Things Observed from the Literature Survey

- Resume Screening is a very important step in the hiring process
- Even today, many companies are following the traditional resume screening process

- Even though many efficient systems were proposed to perform resume screening operations, due to some problems, the systems are not efficient as expected.
- Resume Screening using Machine Learning approaches has some issues like a time consuming, not being user-friendly, etc.
- Resume Screening using Artificial Intelligence approaches has some issues with input/output format, single resume acceptance for a time (which leads to a lot of time to screen all the resumes), etc.
- Resume Screening using Deep Learning has issues like

complex architecture (like coupling between any two blocks of code), inefficient, etc.

5. Proposed Idea After Literature Survey

As mentioned in the literature survey, Resume Screening using different approaches like Machine Learning, Artificial Intelligence, and Deep Learning has some efficiency issues and is not user-friendly. So, the existing ideas can be improved for better performance.

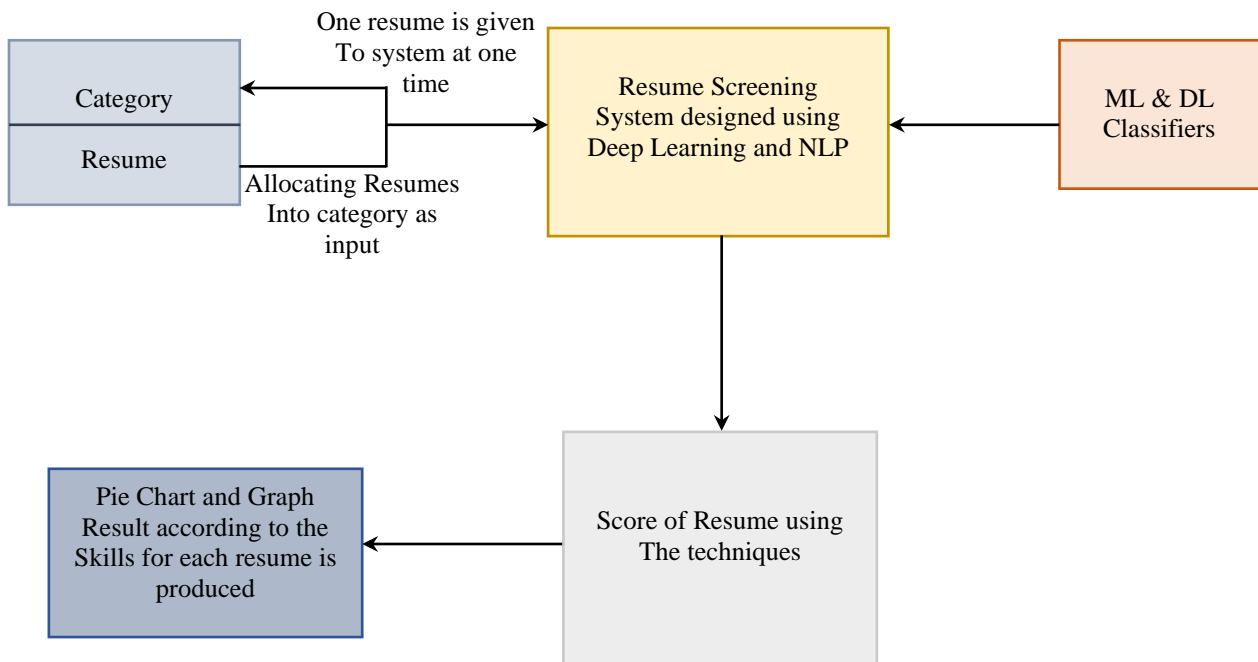


Fig. 10 Resume screening with python

The solution's efficiency can be increased by following the algorithms like NLP (Natural Language Processing) and design patterns like MVC (Model View Controller).

The assumed system can be user-friendly by following the API-led approach.

An external dataset can be used to reduce the dependencies and hard coding. The external dataset can be given with the required job skills by which the system gives the desired output instead of a generalized output. A user-friendly output can be obtained by using python plotting.

The assumed system can follow a folder structure and movement of the files between the folders to make the system clear to the user.

With these improvements, the system for Resume Screening can be made more efficient and user-friendly.

6. Conclusion

Resume Screening is one of the most critical steps in the recruitment cycle. Many approaches have been developed to carry out this method of screening resumes. Technology has made this Resume Screening technique very easy for recruiters. Even a large pool of CVs can be screened easily and efficaciously with the help of new technologies. But due to some problems, these methods cannot be used widely.

Finally, the assumed idea can be a very effective approach developed using efficient technologies like python and NLP, which saves time and resources in the recruiting process for any job role.

Acknowledgment

Here we'll take a moment to thank our professor Mr. Satish Anamalamudi for guiding and helping us with this survey.

References

- [1] Rutuja Patil et al., "Resume Evaluation System Based on AI," *International Research Journal of Engineering and Technology*, vol. 7, no. 7, pp. 2782-2784, 2020. [[Publisher Link](#)]
- [2] V. V. Dixit et al., "Resume Sorting using Artificial Intelligence," *International Journal of Research in Engineering, Science and Management*, vol. 2, no. 4, pp. 423-425, 2019. [[Google Scholar](#)] [[Publisher Link](#)]
- [3] Pradeep Kumar Roy, Sarabjeet Singh Chowdhary, and Rocky Bhatia, "A Machine Learning Approach for Automation of Resume Recommendation System," *Procedia Computer Science*, vol. 167, pp. 2318-2327, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [4] Gaurav Dutta, Resume Screening Using Machine Learning, 2021. [Online]. Available: <https://www.kaggle.com/code/gauravduttakiit/resume-screening-using-machine-learning>
- [5] Resume Screening Using Deep Learning on Cainvas, AI Technologies and Systems, 2022. [Online]. Available: <https://medium.com/ai-techsystems/resume-screening-using-deep-learning-on-cainvas-d93f84e65e9c>
- [6] Aman Kharwal, "Machine Learning Project on Resume Screening with Python," *Resume Screening with Python*, 2021.
- [7] Fresh Works, Resume Screening - How to Stop Swimming in a Sea of Resumes?, 2016.
- [8] Roberto Salazar, Analyzing Candidates Resumes for Jobs Openings, Resume Screening with Python, 2020. [Online]. Available: <https://towardsdatascience.com/resume-screening-with-python-1dea360be49b>
- [9] Zafar Iqbal et al., "Machine Learning Based Student Grade Prediction: A Case Study," *Computers and Society*, 2017. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [10] G. Vadivu, and K. Sornalakshmi, "Applying Machine Learning Algorithms for Student Employability Prediction Using R," *International Journal of Pharmaceutical Sciences Review and Research*, vol. 43, no. 1, pp. 38-41, 2017. [[Google Scholar](#)] [[Publisher Link](#)]
- [11] Byung-Hak Kim, Ethan Vizitei, and Varun Ganapathi, "Gritnet: Student Performance Prediction with Deep Learning," *Machine Learning*, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [12] Jie Xu, Kyeong Ho Moon, and Mihaela Van Der Schaar, "A Machine Learning Approach for Tracking and Predicting Student Performance in Degree Programs," *IEEE Journal of Selected Topics in Signal Processing*, vol. 11, no. 5, pp. 742-753, 2017. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [13] Sayantani Ghosh, Sudipta Roy, and Samir K. Bandyopadhyay, "A Tutorial Review on Text Mining Algorithms," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 1, no. 4, 2012. [[Google Scholar](#)] [[Publisher Link](#)]
- [14] Vishal Gupta, and Gurpreet Lehal, "A Survey of Text Mining Techniques and Applications," *Journal of Emerging Technologies in Web Intelligence*, vol. 1, no. 1, 2009. [[Google Scholar](#)] [[Publisher Link](#)]
- [15] Falguni N. Patel, and Neha R. Soni, "Text Mining: A Brief Survey," *International Journal of Advanced Computer Research*, vol. 2, no. 4, pp. 243-248, 2012. [[Google Scholar](#)] [[Publisher Link](#)]
- [16] Amrut M. Jadhav, and Devendra P. Gadekar, "A Survey on Text Mining and Its Techniques," *International Journal of Science and Research (IJSR)*, vol. 3, no. 11, 2014.
- [17] Kanya Varathan and S. Geetha, "Information Extraction: A Text Mining Approach," *IET-UK International Conference on Information and Communication Technology in Electrical Sciences*, pp. 1111- 1118, 2007. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [18] Sholom M. Weiss et al., *Text Mining: Predictive Methods for Analyzing Unstructured Information*, Springerscience and Business Media, 2010. [[Google Scholar](#)]
- [19] S. Liritano, and M. Ruffolo, "Managing the Knowledge Contained in Electronic Documents: A Clustering Method for Text Mining," *12th International Workshop on Database and Expert Systems Applications*, 2001. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [20] Deepshikha Patel, and Monika Bhatnagar, "Mobile SMS Classification," *International Journal of Soft Computing and Engineering (IJSCE)*, vol. 1, no. 1, pp. 47-49, 2011. [[Publisher Link](#)]
- [21] Anushree Negi, "A Brief Survey on Text Mining, Its Techniques, and Applications," *SSRG International Journal of Mobile Computing and Application*, vol. 8, no. 1, pp. 1-6, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [22] Chetan Arora et al., "Automated Checking of Conformance to Requirements Templates Using Natural Language Processing," *IEEE Transactions on Software Engineering*, vol. 41, no. 10, pp. 944-968, 2015. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [23] Julia Hirschberg, Bruce W. Ballard, and Donald Hindle, Natural Language Processing,
- [24] Minhwa Chung, and D. Moldevan, "Applying Parallel Processing to Natural-Language Processing," *IEEE Expert*, vol. 9, no. 1, pp. 36-44, 1994. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [25] B.W. Wah et al., "Report on Workshop on High-Performance Computing and Communications for Grand Challenge Applications: Computer Vision, Speech, and Natural Language Processing, and Artificial Intelligence," *IEEE Transactions on Knowledge and Data Engineering*, vol. 5, no. 1, pp. 138-154, 1993. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]