IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

## IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 5, May 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005239|

# Resume Screening Using KNN

**Eda Grahitha [1], Chilika Alex [2], Annavarapu Venkata Ratna Prasad [3],**

**Alapati Vamsi Krishna Mahesh[4], Pinnamaneni Siva Prasad[5]**

U.G. Student, Department of Computer Science and Engineering, Vasireddy Venkatadri Institute of Technology, Nambur, Andhra Pradesh, India [1,2,3,4]

Assistant Professor, Department of Computer Science and Engineering, Vasireddy Venkatadri Institute of Technology, Nambur, Andhra Pradesh, India [5]

**ABSTRACT:** When applying for a job, many companies require applicants to submit resumes. When the rivalry for a job is severe, a large number of applicants may submit their resumes for a single position. In such cases, the employer is faced with a mountain of resumes. When many job listings are applied for, things grow even more confusing. Because certain jobs may accept applicants with a wide range of educational backgrounds and years of experience, manually processing and shortlisting resumes appears to be a difficult and time-consuming task. To address this issue and make resume processing more efficient, we propose a data mining-based online application that can crawl through all resumes submitted, sort them according to the organization's specifications, and provide them to the administrator. The admin can then send emails to the candidates who have been shortlisted by our web app directly from the admin panel. We chose the KNN Algorithm since it best fits the project's goal and scope. KNN is a machine-based method that consistently outperforms SCORES. Both the Admins and Job Applicants will have access to a web interface built with HTML and CSS. The resumes that are uploaded can be kept in MongoDB, and KNN is written in Python. The FLASK framework is useful for connecting the front end to the database.

## I.INTRODUCTION

Recruitment is a $200 billion industry. It is concerned with selecting the most qualified individuals with the necessary abilities for a specific job profile from a vast pool of applicants. If a company offers a job vacancy for a position, a large number of people apply by sending their resumes to the company[1]. The initial step for any recruiter in the hiring process is to screen all of the job applicants' resumes. Every day, any company with a job opportunity for a certain role will receive thousands of emails from hopeful job applicants. Any recruiter knows how time-consuming it is to choose prospective prospects for a job vacancy from a big pool of candidates.

Approximately 75% of the thousands of resumes sent to the organisation for the given job posting do not highlight the relevant abilities required for the job profile.

[2] As a result, recruiters frequently find it difficult to select the most qualified applicants from a huge applicant pool.

As a result, we created a web application[3] that assists both recruiters and candidates. By logging in with admin credentials, the recruitment team will post a list of job openings. Candidates who wish to apply for a position must first register and then login using their credentials. Then they'll be taken to a page where they may apply for the job of their choice by selecting it from a drop-down menu, entering their email address, and uploading their résumé from their system. The list of resumes submitted for job openings will be visible to the recruitment staff, and they will order the best resumes using the KNN algorithm.

KNN is a machine learning algorithm that screens the best resumes with high precision and accuracy while also taking less time[4], which is a big drawback for many websites.

Finally, the recruitment team will send out selection emails to those candidates who have expressed an interest.

Without having to open the résumé, the recruitment team may see more information about the prospects, such as their talents, experience, and name. This is done with the help of Resume Parser, which saves time while evaluating candidates for this project.

## II.LITERATURE SURVEY

Many people would like to advance their careers. And the majority of these folks have the necessary expertise and abilities to accomplish this goal. Unfortunately, a large number of job seekers have weak communication skills. They are unable to adequately communicate their employment qualifications to potential employers. In a nutshell, they don't know how to look for work. This frequently hinders people from obtaining a high-paying job that they might easily

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005239|

perform. Often, a job will go to someone who is less qualified but has a well-written résumé. Job searchers frequently have a few misconceptions about potential companies. They believe that companies can quickly distinguish between qualified and less qualified job applicants. However, this is most likely not the case. For the same position, there can be anywhere from 30 to 300 resumes. As a result, the interviewer does a quick assessment of all resumes in order to reject as many as possible. The "good" resumes usually make it through the initial screening.

Due to a weak résumé, the best job candidate is frequently screened out. Many skilled people typically apply for the same job in today's corporate sector. What if, out of everyone who applies, one person submits a well-written resume? Which of the candidates do you believe has the best chance of landing the job? Of course, it's the one with the "best" resume. Even though some of the other applicants are more qualified for the position, this is frequently the case.

You must express to the employer that you are ready, willing, and able to complete the job in order to secure a good job. So, if you can write an excellent work resume, you'll have a far better chance of finding a better job. Almost every potential employer will request a copy of your resume. Who gets a job interview is determined on their résumé. Your CV serves as a mini-biography of yourself. The employer should have a better "feel" for you as a person and as a possible employer after reading your CV. Its purpose is to get to know the employer so that they may decide whether they want to learn more about you.

The resume is the initial stage in the process; it serves as your introduction to a potential employer. The importance of first impressions cannot be overstated. You won't get to step two, the job interview, if you make a bad first impression. In effect, your resume should communicate to the company that you have strong skills and are eager to work. This report will assist you in making a positive first impression. All good resumes adhere to the same key principles. A résumé that is bold, engaging, and enticing is what you want. But not too so. You should also have a résumé that is conservative in nature. To put it another way, it must be audacious. It's not flashy. You must project confidence in your abilities without coming across as arrogant. You should appear enthusiastic about the job, but not desperate. You wish to speak in a sophisticated manner. When possible, use everyday language. However, as a general guideline, keep it simple and to the point. As a result, you must avoid becoming overly verbose. It is critical that your CV be as long as possible. Resumes should be between one and three pages long. Even if you have a lot to say, don't make your resume longer than three pages. Remember that a resume is designed to be an overview of your qualifications.

## III.PROPOSED SYSTEM

The proposed resume ranking system is a web application in which the hiring team announces job openings and candidates apply by submitting their resumes and email addresses. Resumes will be shown in the best order possible, according to KNN, a machine learning algorithm, with talents, experience, and other details listed. The hiring team then selects the finest applicant for their needs and contacts them via email. When comparing the proposed system to previously existing systems, it is clear that the suggested system is superior in every regard, including greater accuracy, reduced time, and reduced staffing requirements.

### 3.1 KNN ALGORITHM

The K-Nearest Neighbor algorithm, which is based on the Supervised Learning technique, is one of the most basic Machine Learning algorithms. The K-NN algorithm assumes that the new case/data and existing cases are comparable, and it places the new case in the category that is most similar to the existing categories.

The K-NN algorithm saves all existing data and categorises additional data points based on their similarity. This means that as fresh data is generated, it may be quickly sorted into a suitable category using the K-NN method.

The K-NN approach can be used for both regression and classification, however it is most commonly employed for classification. K-NN is a non-parametric algorithm, which means it makes no assumptions about the underlying data.

It is also known as a lazy learner algorithm because it does not learn from the training set right away; instead, it stores the dataset and performs an action on it when it comes to classification. The KNN algorithm simply stores the dataset during the training phase, and when it receives new data, it classifies it into a category that is very similar to the new data.

White spaces, numerals, and stop words like and, or, etc. are removed from resumes from the Data set. The words in the resumes are then converted to vectors using TF-IDF vectorization. TF-IDF vectorizer is also used to transform the text in the job description to vectors. The Euclidean distance is calculated to determine the degree of similarity between the resume and the job description provided, and the KNN algorithm is then used to find resumes that closely match the JD provided by the recruiters.

KNN is a non-parametric, lazy learning technique. It uses a database of data points divided into discrete clusters to provide conclusions for new samples. Instead of making assumptions about the underlying data distribution,

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005239|

KNN focuses on item feature similarity. KNN evaluates the "distance" between the target and every other resume in its database, ranks the distances, and returns the top K most comparable resume options when inferring about a resume.

## IV.RESULTS AND DISCUSSION

The resumes are pulled from a database of resumes, matched to the job description file, then sorted according to how well they fit the job description. A rank is assigned to the resumes that have been sorted. The best score is used to determine the rank. When the recruitment team picks the job they are hiring for, the sorted resumes are displayed to them on the website's results page. The resumes are displayed in order of the rank assigned with the job matching, starting with the best ones. The first set of facts and resumes displayed are the most suitable for the job that the recruitment team has chosen, followed by the second and so on. When compared to earlier methods for sorting resumes, our results are more accurate and efficient.

The KNN algorithm assigns a score based on how well the candidate matches the job description. When compared to other scoring methods, KNN produces the most accurate and reliable scores. In the final result page from resumes, the recruitment team can examine the applicant's major details such as email id, personal information, and experience. This technique is based on a previously developed python library for resume parsing that produces the best results. After checking the candidate's details, the recruitment team can send an email offering the job to the candidate with a single click. The send mail column can be found on the final results page. With respect to the eligible candidates, the recruitment team can send emails by clicking on the button. All of the results are accurate and one step ahead of all other algorithms when compared to typical resume parsing and resume screening algorithms.

We looked at some resumes and manually checked them against the job description before ranking them from best to worst.

After that, we uploaded those resumes to our knn resume screening algorithm, which yielded the same results. as a result, we are getting the most precise findings.

This knn is more suitable for our purpose than other machine learning models because it is basic and provides high accuracy.

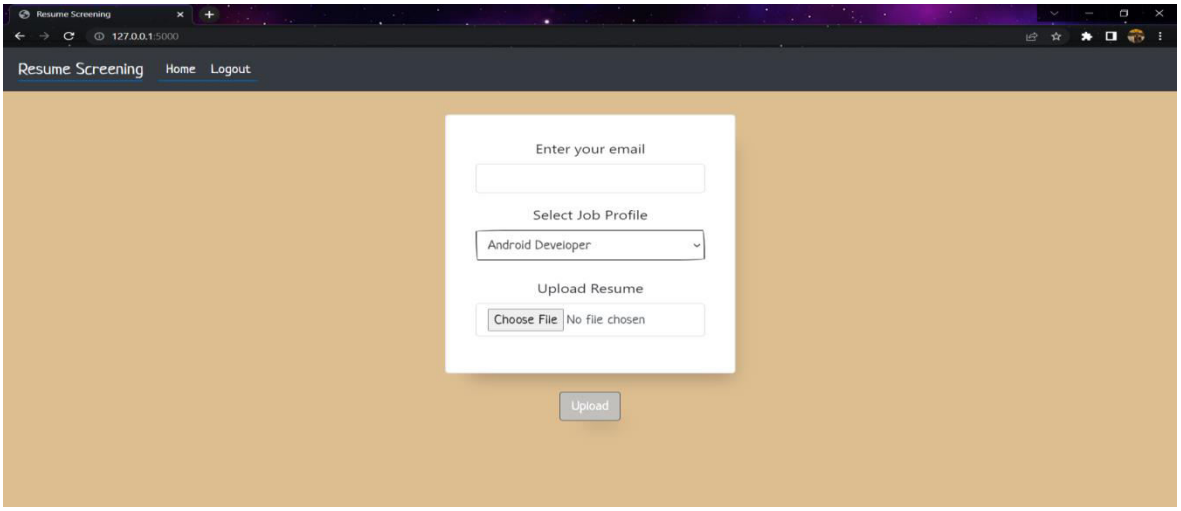


**Fig: 1:Uploading resume**

1. User have to enter their valid email id , select the interested job profile from the drop down menu and have to upload their resume from their local system storage.If an details are not given by the user then the page appears as below with unfilled uploading

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005239|

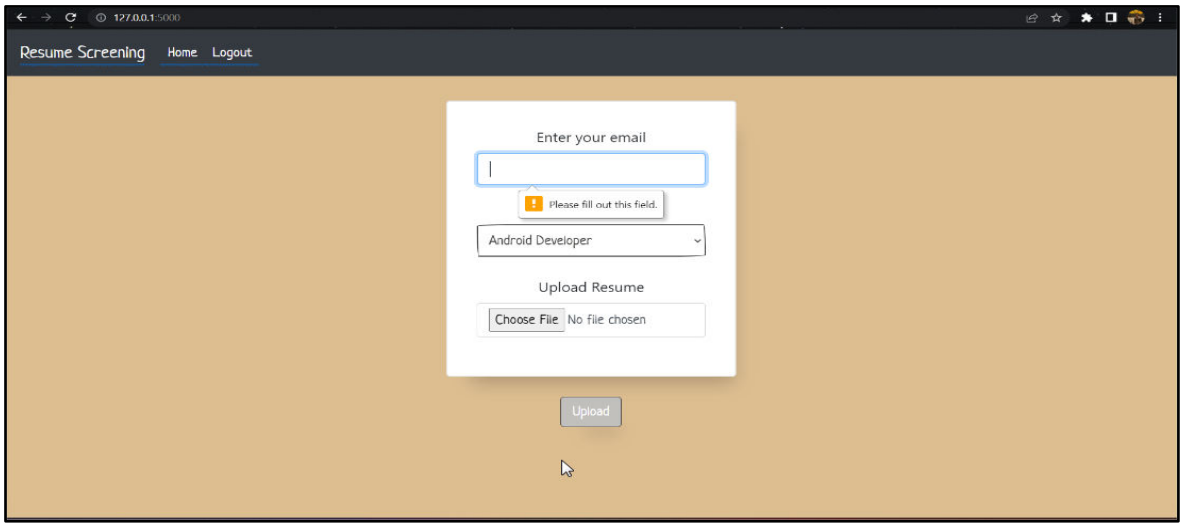


**Fig 2: No email entered**

2. By clicking on upload button , the user applies for the selected role and then he gets a page with message as shown in the following figure

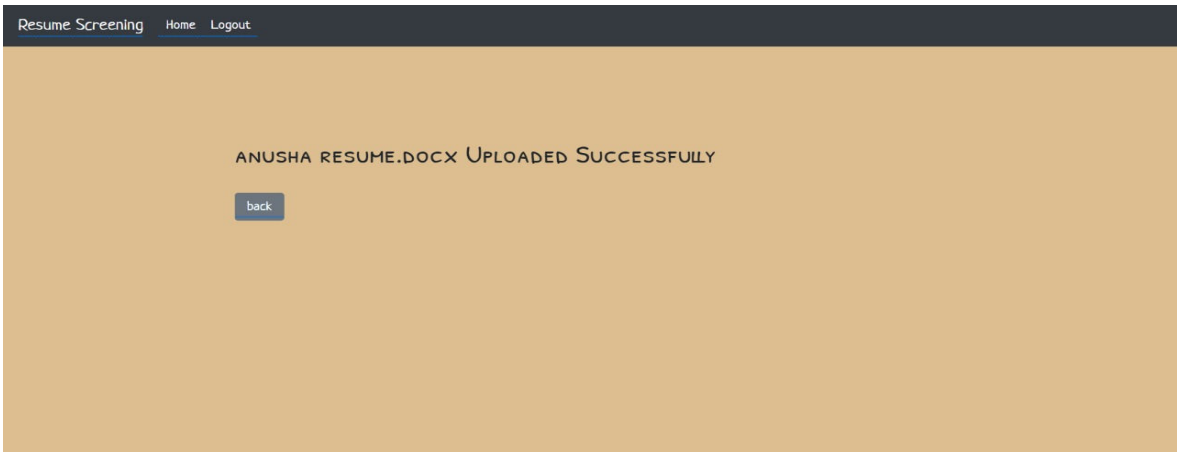


**Fig 3: Resume uploaded successfully**

By clicking back button the previous page appears for another upload

3. Email Sent by the admin will appear as follows

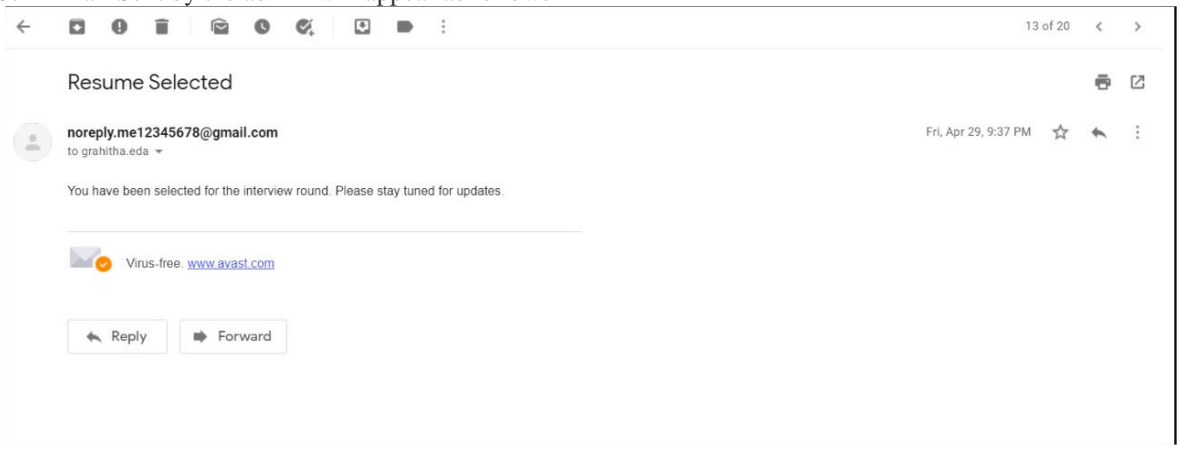


**Fig:4: Selection mail received by user**

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005239|

4. Email received by the user will appear as follows

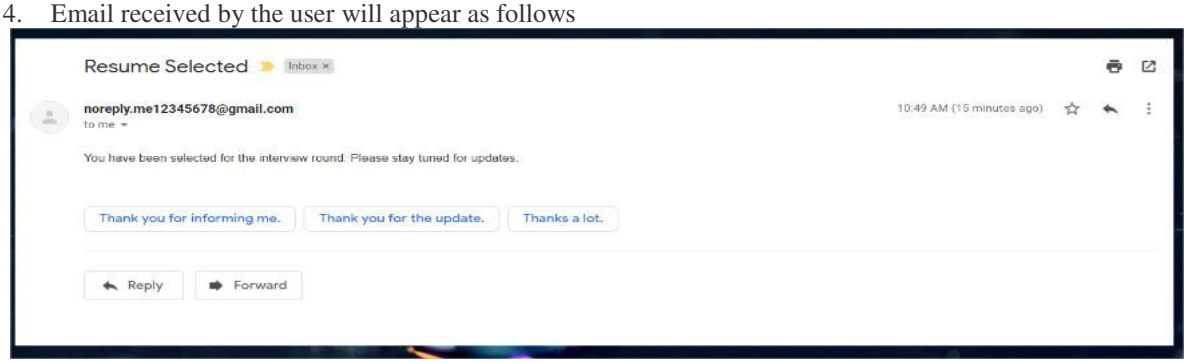


**Fig 5: Selection mail sent by admin**

5. MongoDb Database

We are storing tables in the database by creating the "resume screening" database and under that creating the tables "users" and "resumes". It looks like as follows.Instead of rows and columns they store in the form of json objects which helps in storing unstructured data

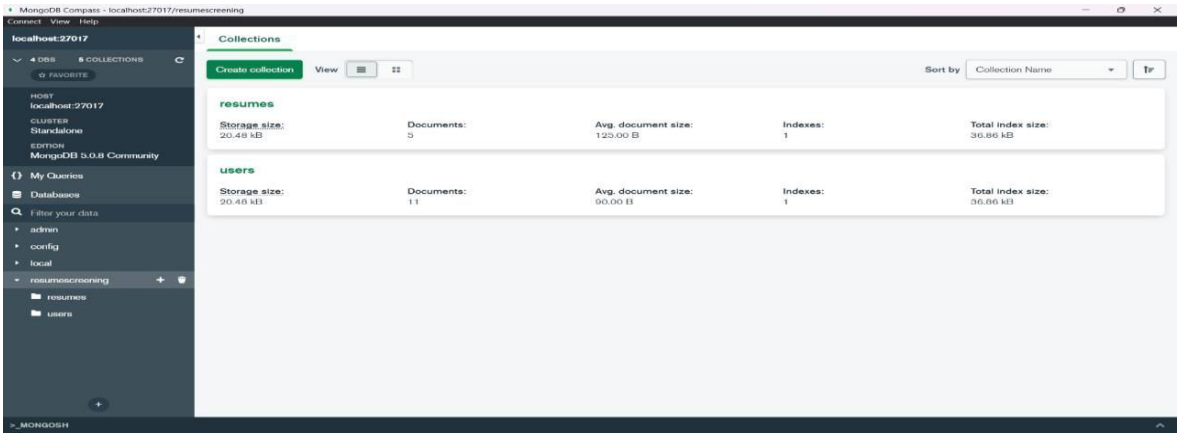


**Fig6: Database showing resumes and user's credentials**

6."resumes" table

This table will store the details of the candidates like generated id for retrieval , email, job profile selection and their resume .

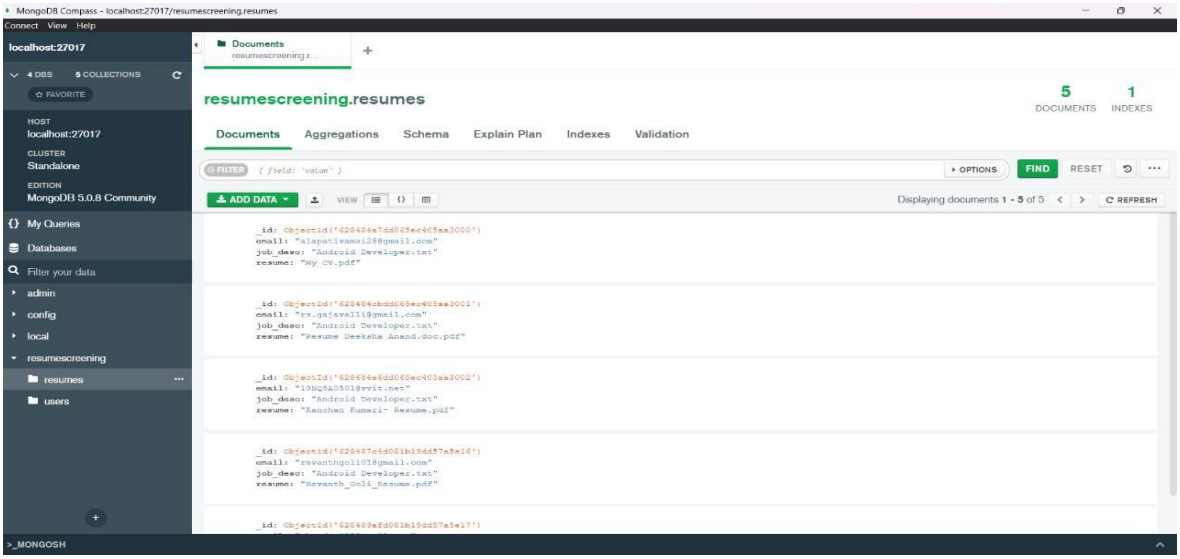


**Fig:7: How resumes data is stored**

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005239|

7.”users” table:

This table contains the usernames and passwords of the registered users and admin for validation of credentials while login .The password is in encrypted format for security . And it also store generated id for easy retrieval of data .Here shows how it looks like

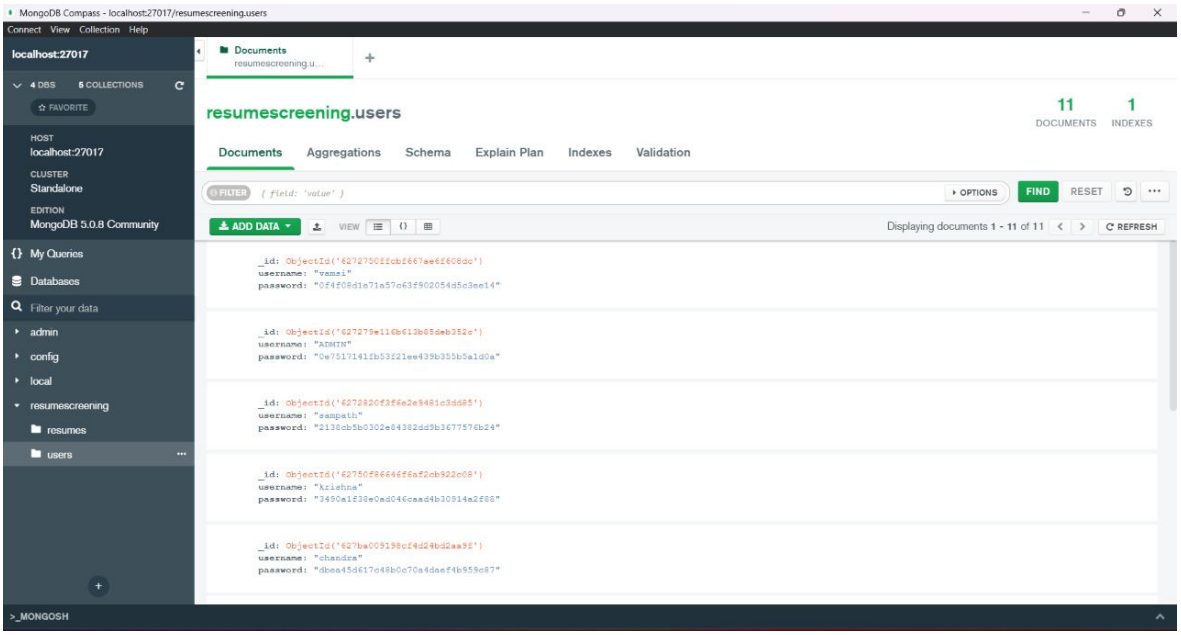


**Fig:8: How user’s credentials are stored**

## V.CONCLUSION

The suggested system screens resumes from best to worst in order using the knn algorithm, which is a supervised machine learning technique. This technique will undoubtedly assist recruiters in weeding out the most qualified individuals for further consideration in the recruiting process based on their resumes. It will relieve recruiters of the strain of manually reviewing each and every resume from a vast pool of candidates. This also makes it easier for people to apply for jobs by simply uploading their resumes and email addresses under the appropriate roles. Recruiters will find it simple because the resumes are shown in chronological order, and they will use smtplib to send emails to those individuals who have been chosen. We also used a resume parser to display information from the candidates' resumes, such as talents, experience, and so on. The candidates who have been chosen will thereafter receive their selection emails.

The logins are made for the recruitment team so that only authorised individuals can access the site; this adds to the site's security. Only approved recruitment teams have access to the candidates' data, which is securely saved in the database. The recruitment staff protects the candidate data, which is non-sharable. Before applying for a position, each applicant must create a new account using his or her email address. They can sign in without having to join up the next time they apply for the same company. Users' data and account information, such as email addresses and passwords, are safely saved in the database. Passwords are encrypted and stored in the database.

## REFERENCES

1. E Faliagka, L Iliadis, I Karydis, M Rigou, S Sioutas, A Tsakalidis, and G Tzimas, “ On-line consistent ranking on e-recruitment: seeking the truth behind a well-formed CV,” The Artificial Intelligence Review, 42(3), 515, 2014.
2. S Al-Otaibi and M Ykhlef, "Job Recommendation Systems for Enhancing E-recruitment Process", in Proceedings of the International Conference on Information and Knowledge Engineering (IKE), Las Vegas Nevada, USA, pp. 433-439, 2012
3. J. M. Vara Mesa, ―ATL/AMW use case modeling Web applications: Detailed description and user guide‖, (2009, access date), http://www.eclipse.org/m2m/atl/usecases/webapp.modeling/resources/User.Guide.pdf
4. Wei Zheng, HaiDong Wang, Lin Ma and RuoYi Wang, "An Improved k-Nearest Neighbor Classification Algorithm Using Shared Nearest Neighbor Similarity", Metallurgical & Mining Industry, no. 10, pp. 133-137, 2015.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005239|

5. Benymol Jose, Sajimon Abraham," Exploring the Merits of NoSQL: A Study Based on MongoDB", 978-1-5090-6590-5/17/$31.00 ©2017 IEEE
6. Minhaz Fahim Zibran," Cryptographic Security for Emails: A Focus on S/MIME" 2011 .
7. Everett et al., "Software testing: testing across the entire software development life cycle". John Wiley & Sons, 2007.
8. J Chen, Z Niu and H Fu, "A Novel Knowledge Extraction Framework for Resumes Based on Text Classifier", Proceedings of the International Conference on Web-Age Information Management, pp. 540-543, 2015.
9. "Introduction to UML diagrams" [online] available http://www.tutorialspoint.com/uml/
10. By Armin Ronacher, "Flask Web Development, One Drop At A Time", http://flask.pocoo.org/,March 20 th 2015

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.165

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

## IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details