Building Search Engine Using Machine Learning Technique

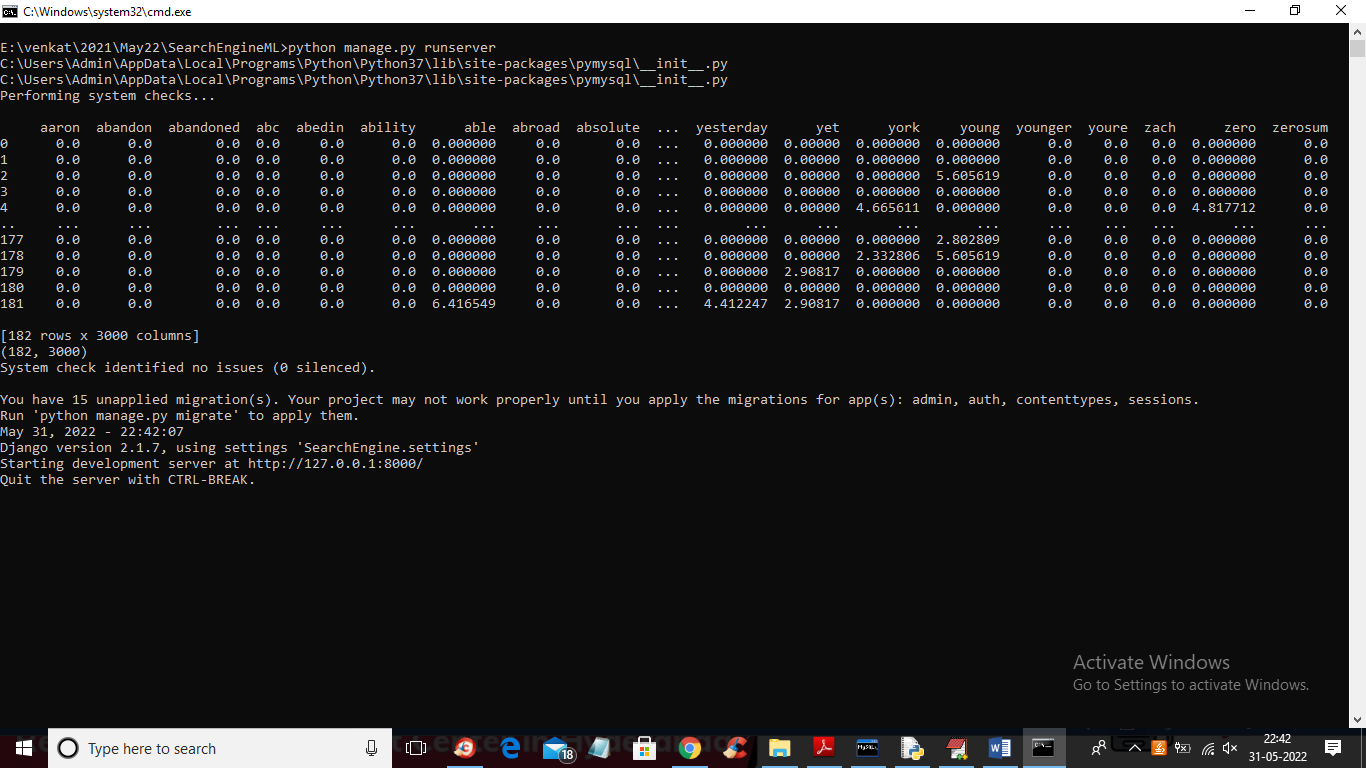
In this paper author is using machine learning algorithms called SVM and XGBOOST to predict search result of given query and building search engine with machine learning algorithms. To train this algorithm author is using website data and then this data will be converted to numeric vector called TFIDF (term frequency inverse document frequency). TFIDF vector contains average frequency of each words.

In this paper author has implemented following modules

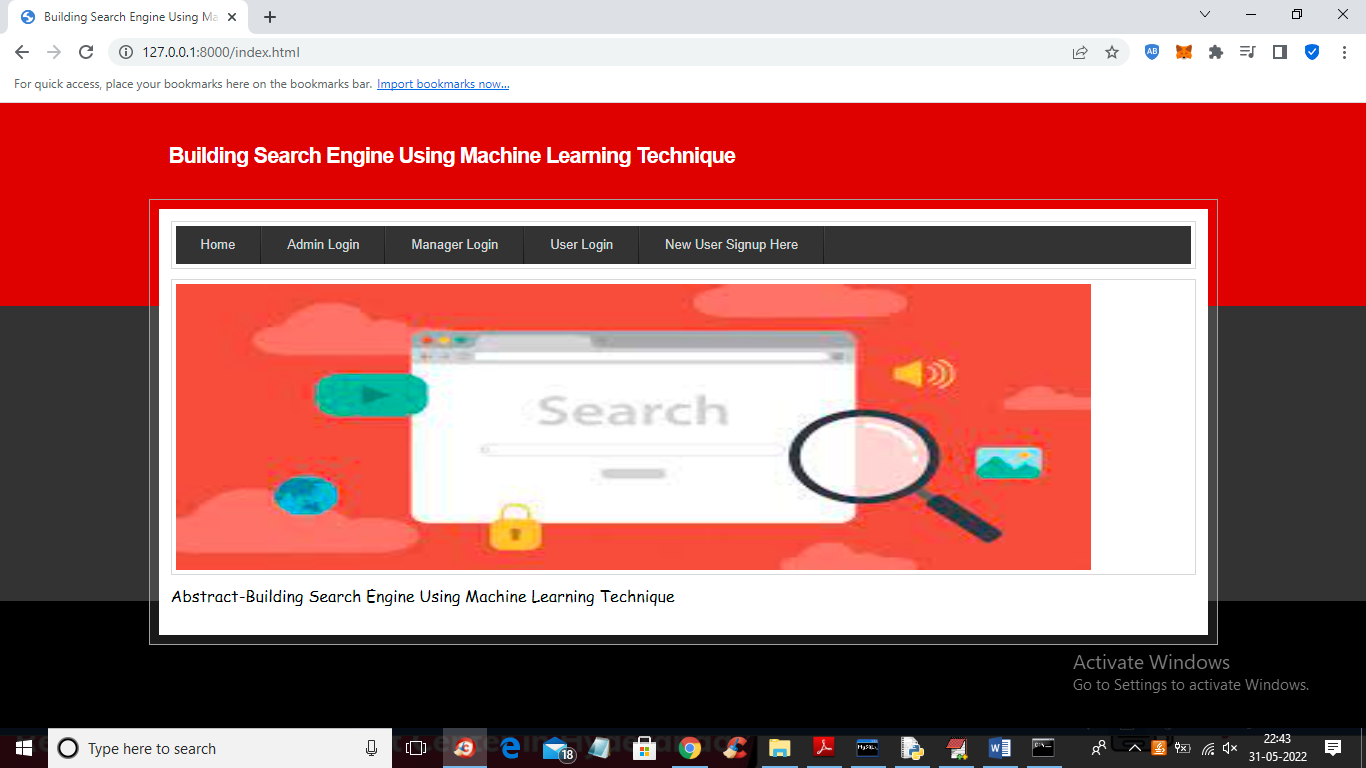
1. Admin module: admin can login to application using username and password as admin and then accept or activate new users registration and then train SVM and XGBOOST algorithm
2. Manager module: manager can login to application by using username and password as Manager and Manager and then upload dataset to application
3. New User Signup: using this module new user can signup with the application
4. User Login: user can login to application and then perform search by giving query.

To run project install MYSQL and python 3.7 and then copy content from DB.txt file and paste in MYSQL to create database.

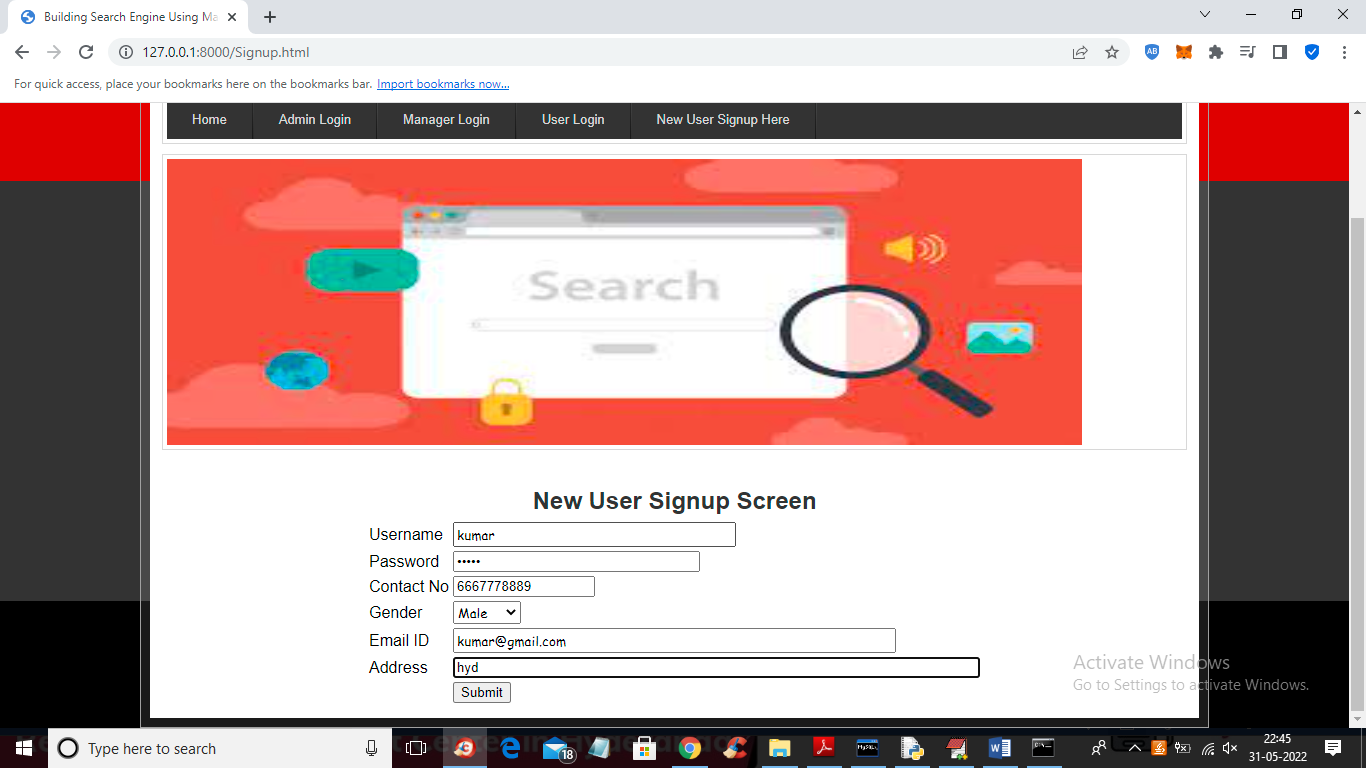
Now double click on ‘run.bat’ file to start python DJANGO server and get below screen



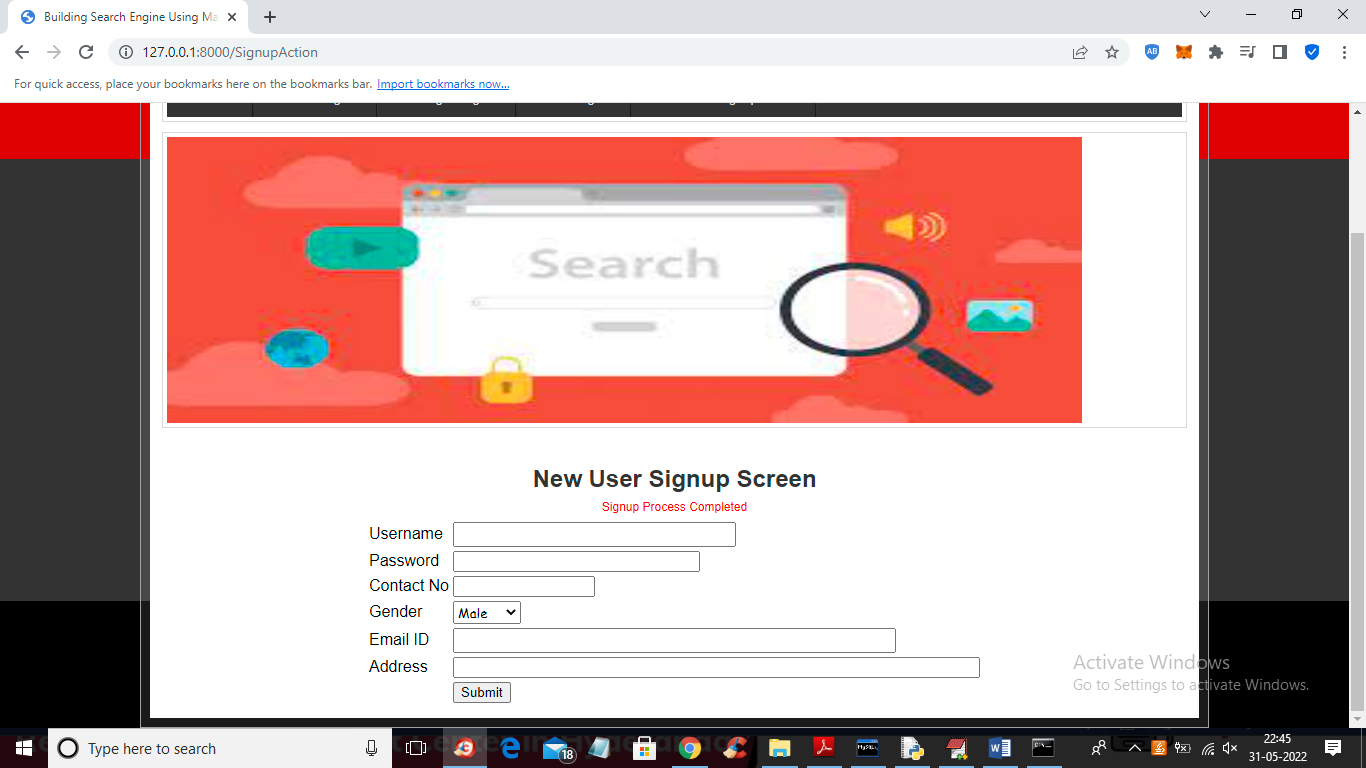
In above screen server started and build a vector from dataset where first row showing word and remaining rows contains TFIDF word frequency. Now open browser and enter URL as <http://127.0.0.1:8000/index.html> and press enter key to get below page



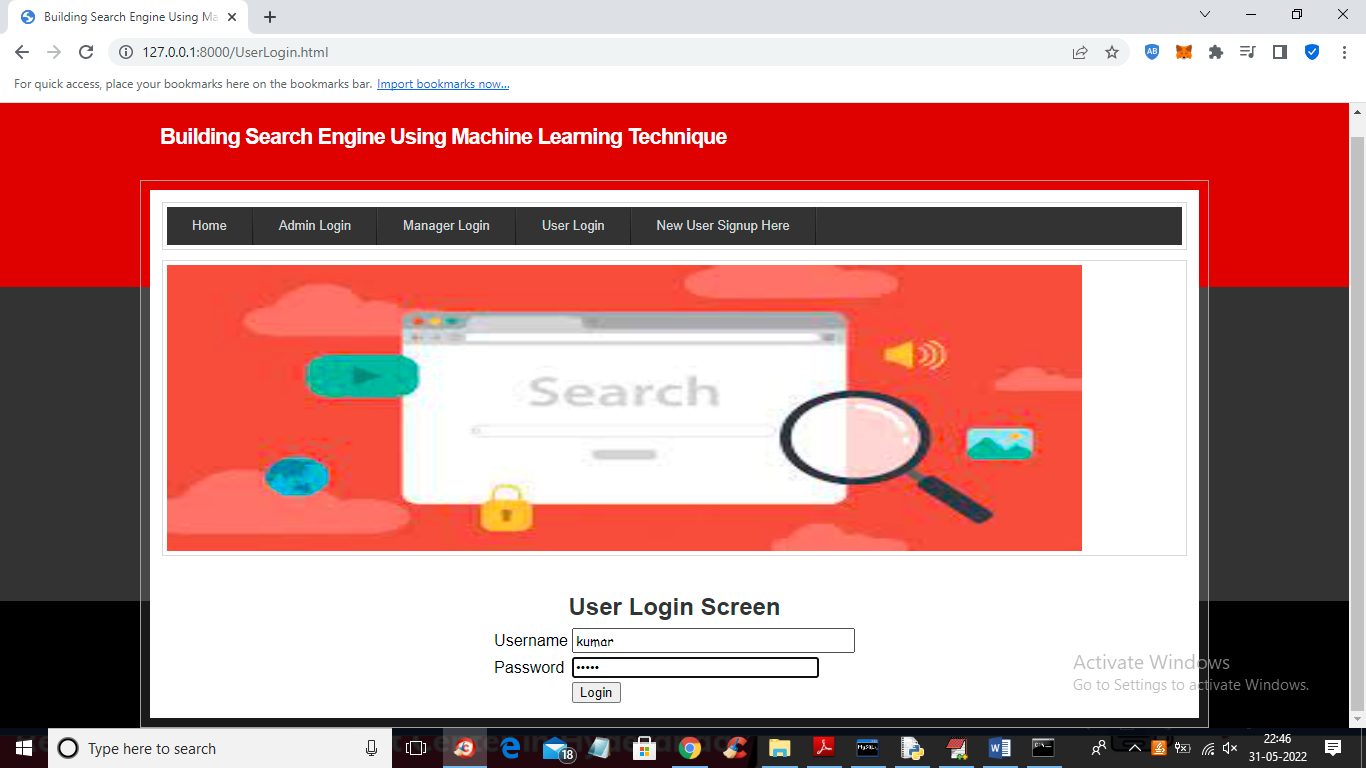
In above screen click on ‘New User Signup Here’ link to get below screen



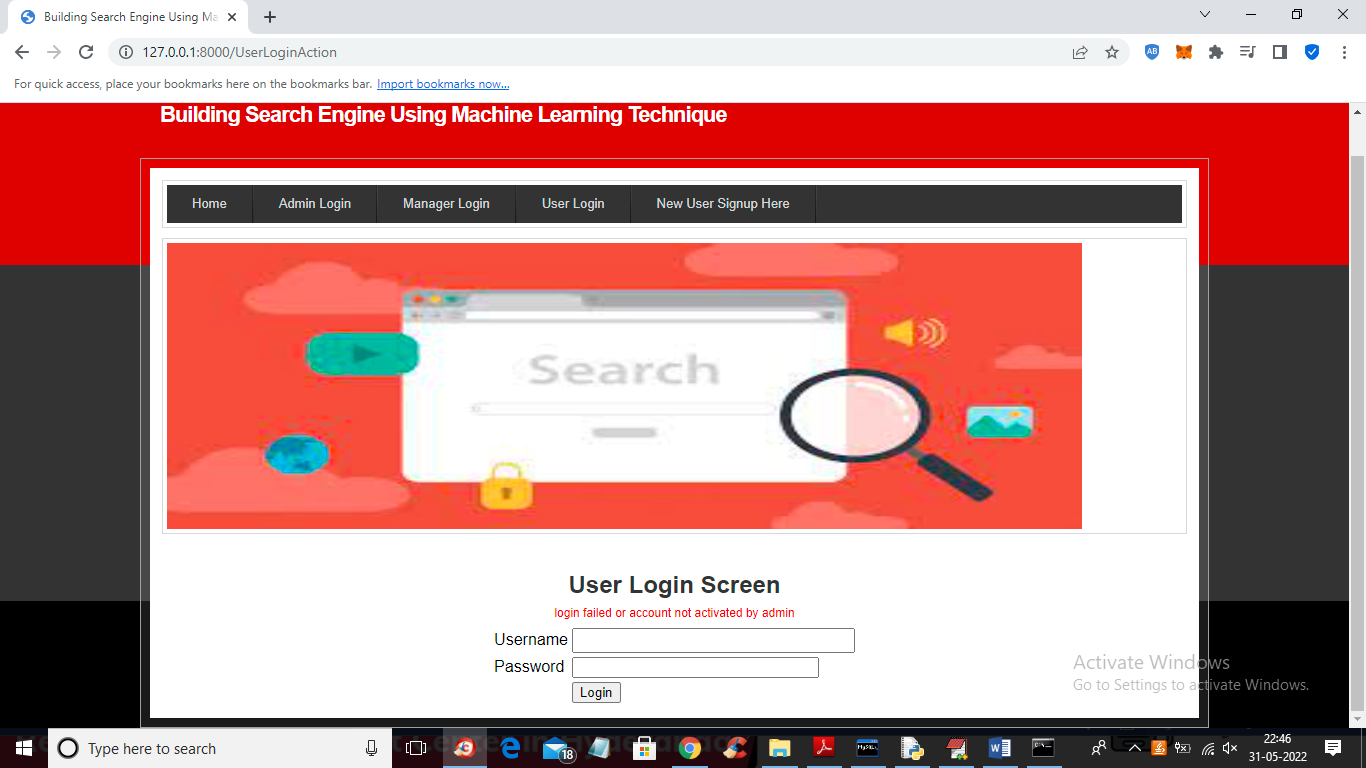
In above screen user is signing up and then press button to get below output



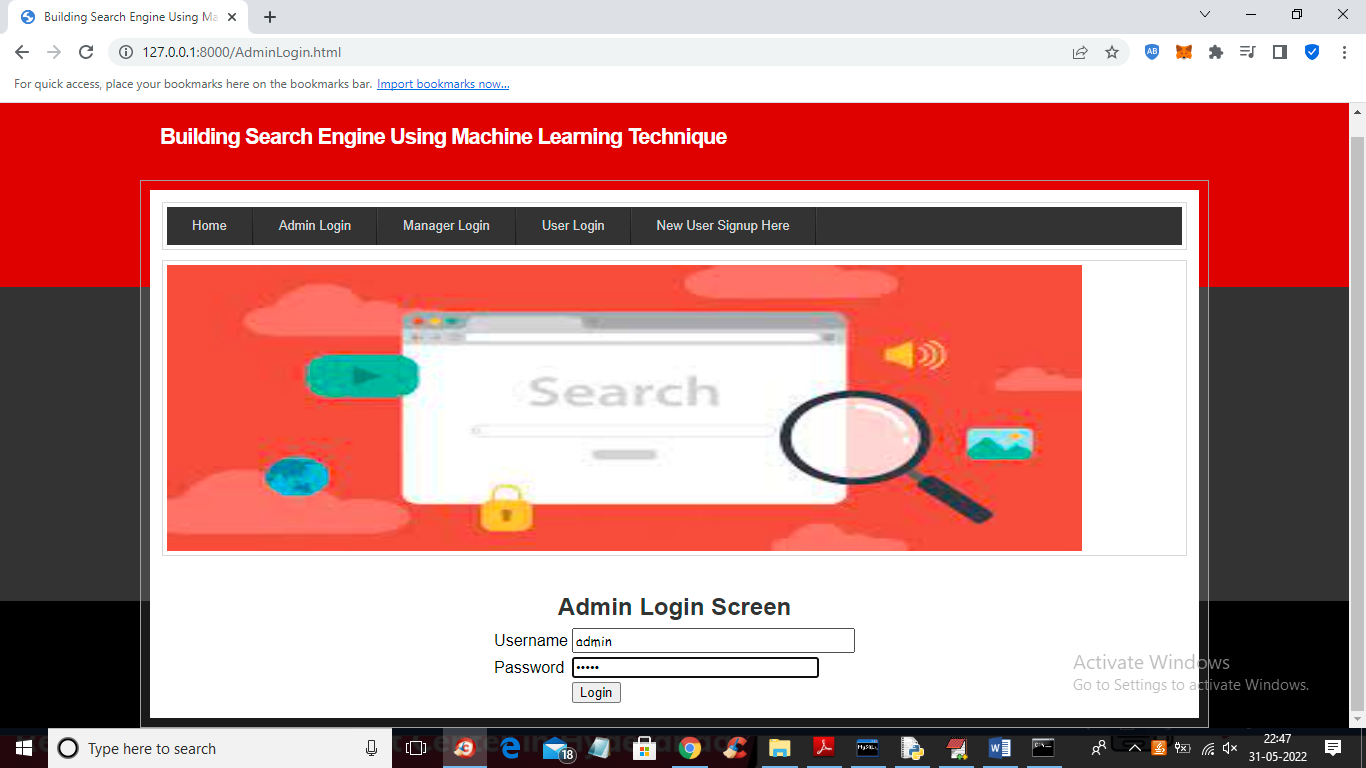
In above screen user signup process completed and now click on ‘User Login’ to get below screen



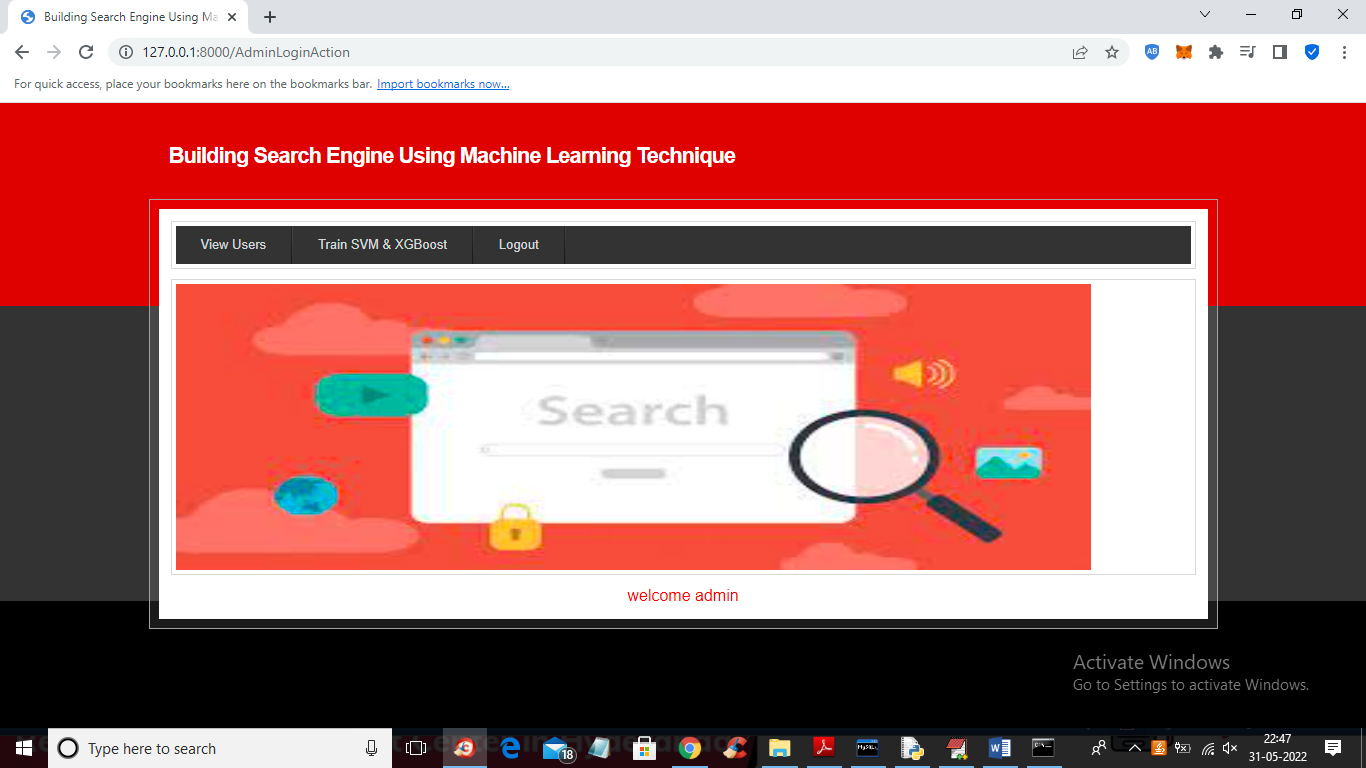
In above screen user is login and will get below output



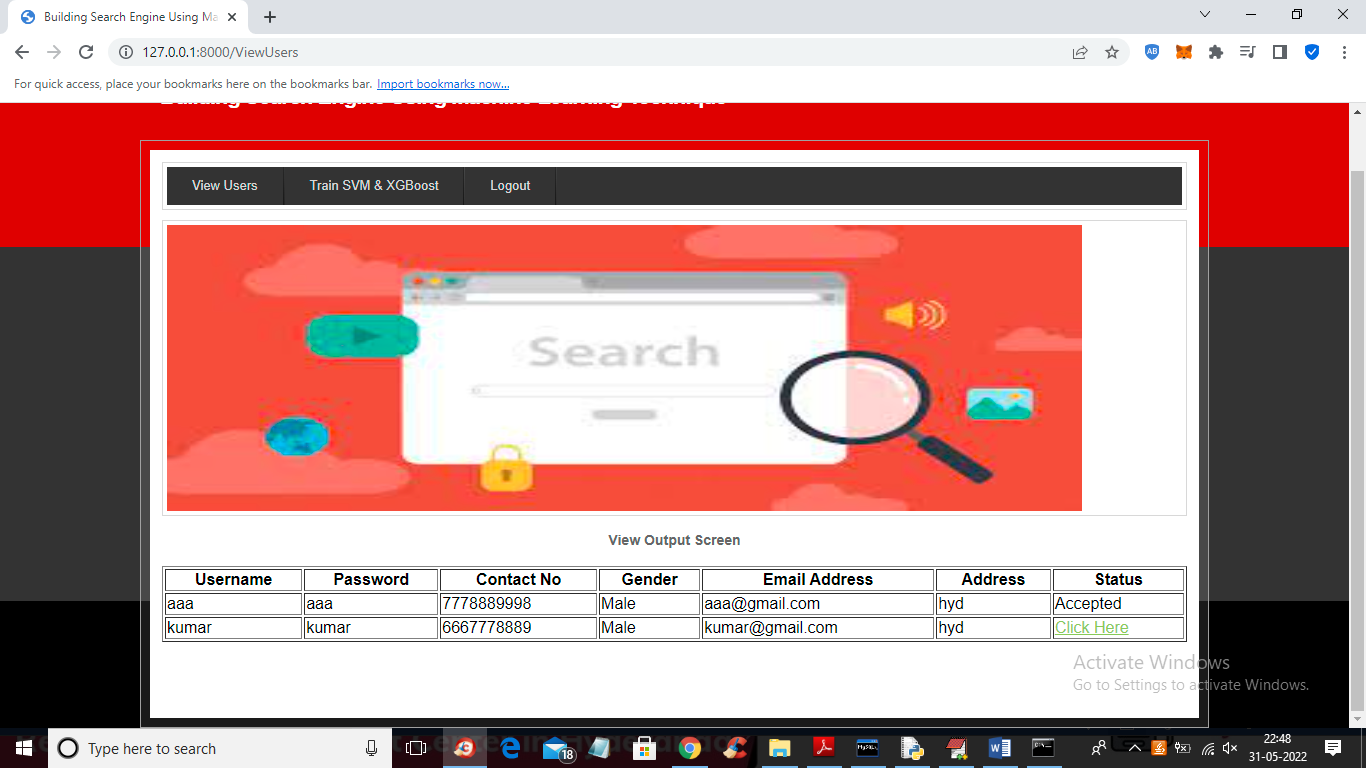
In above screen we gave correct login but account not activated by admin so click on ‘Admin Login’ link to login as admin and then activate user



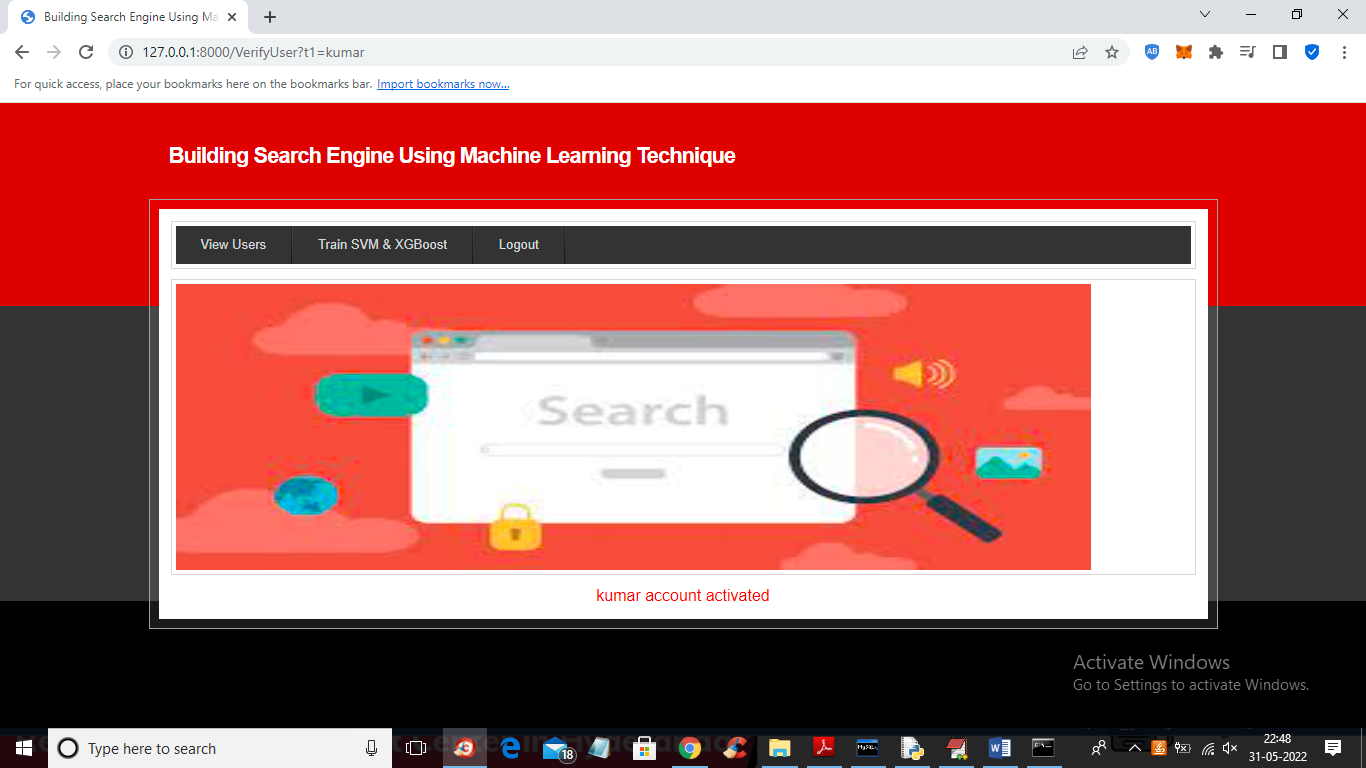
In above screen admin is login and after login will get below screen



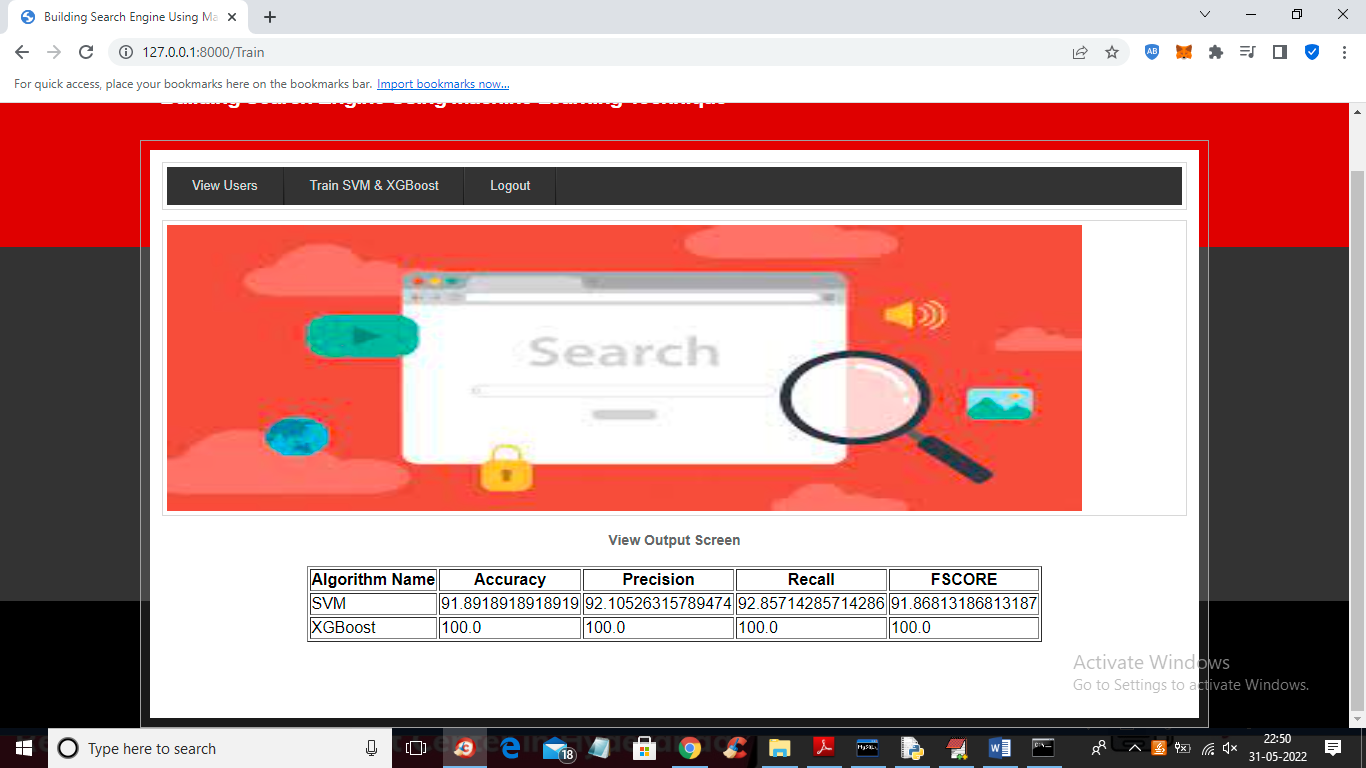
In above screen admin can click on ‘View Users’ link to view all users



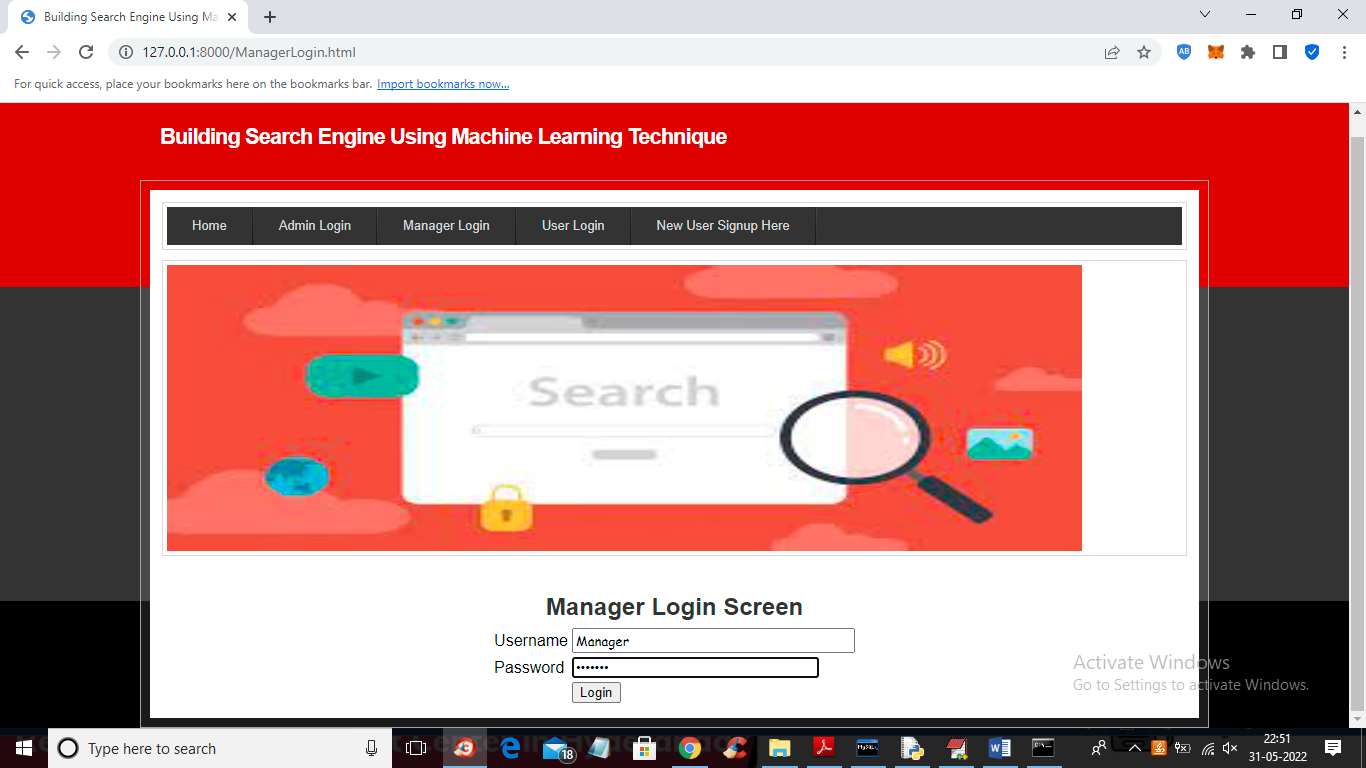
In above screen admin can click on ‘Click Here’ link to activate that user account



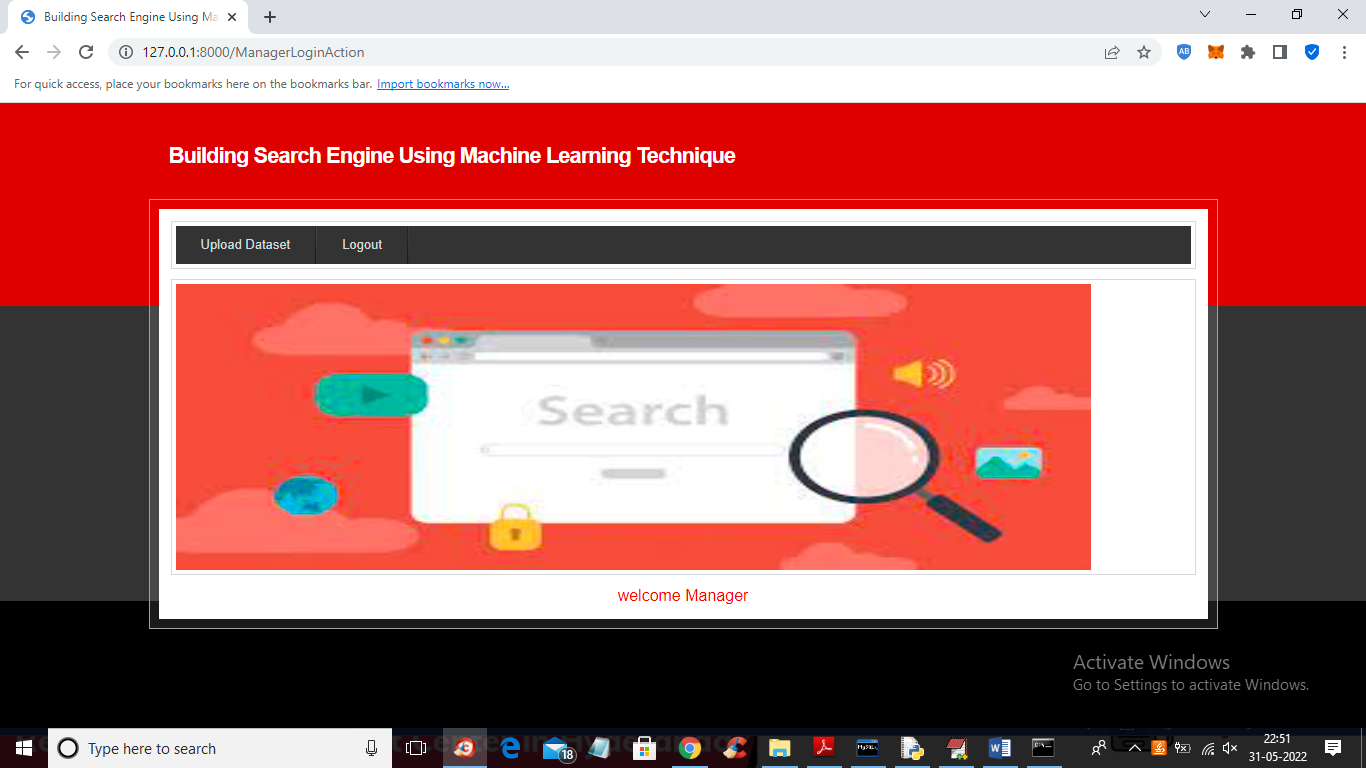
In above screen we can see admin activated kumar user account and now admin can click on ‘Train SVM & XGBOOST’ link to train machine learning SVM and XGBOOST algorithm and get below output



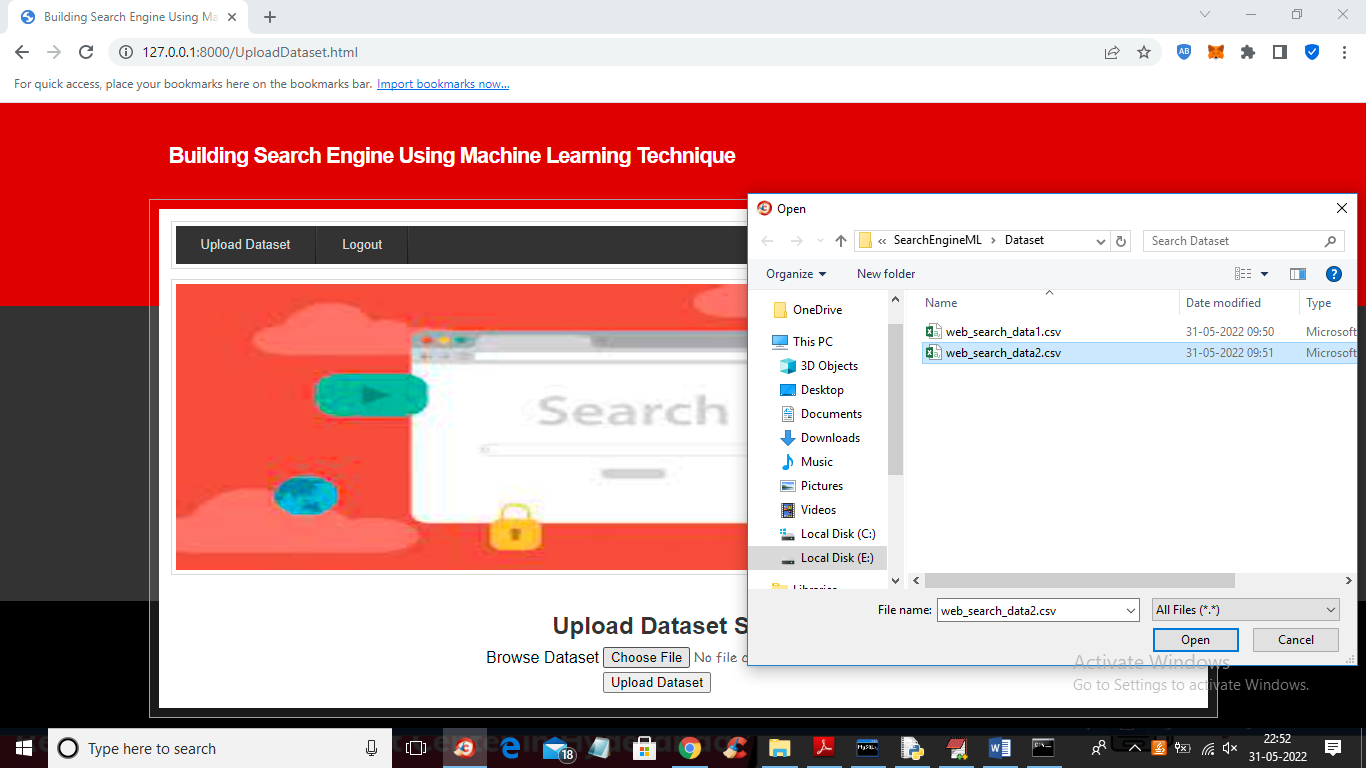
In above screen we can see SVM and XGBOOST accuracy and in both algorithms XGBOOST got high accuracy and now logout and login as Manager



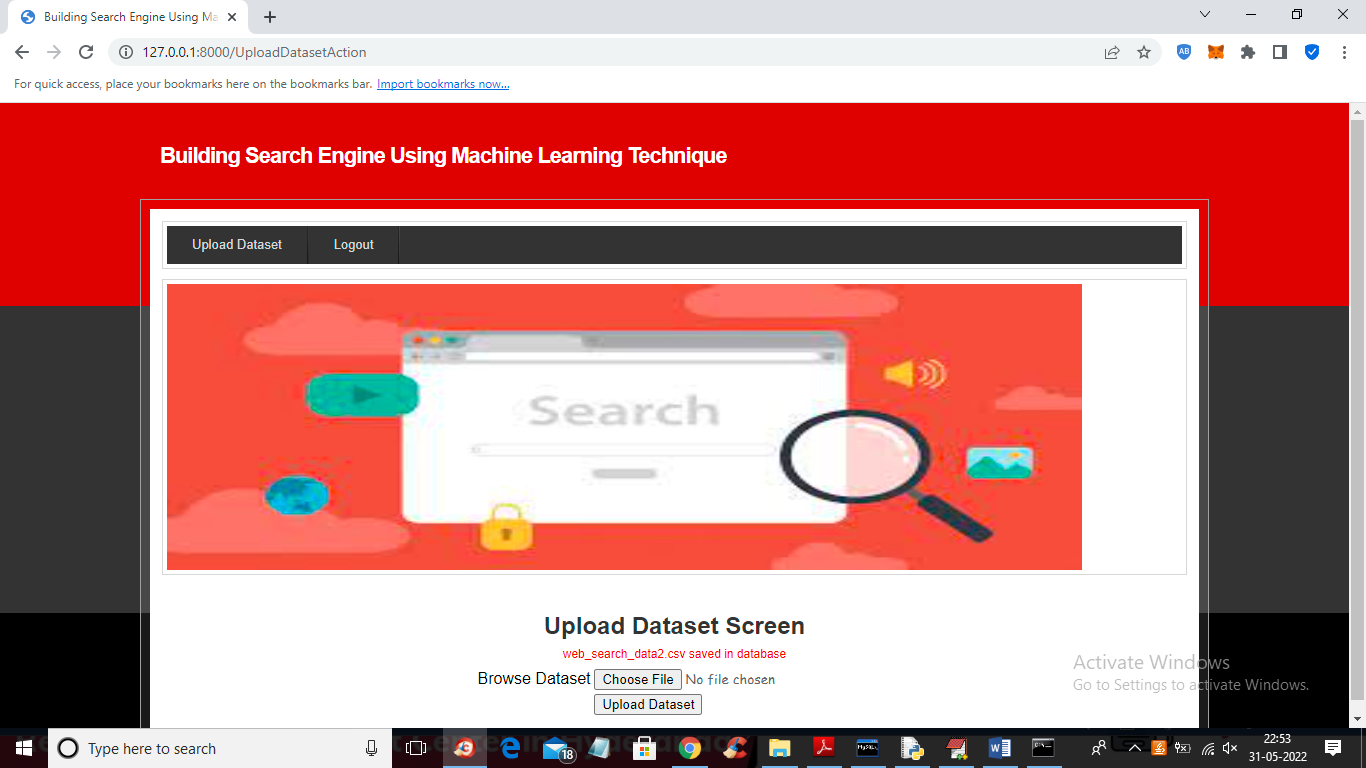
In above screen manager is login and after login will get below screen



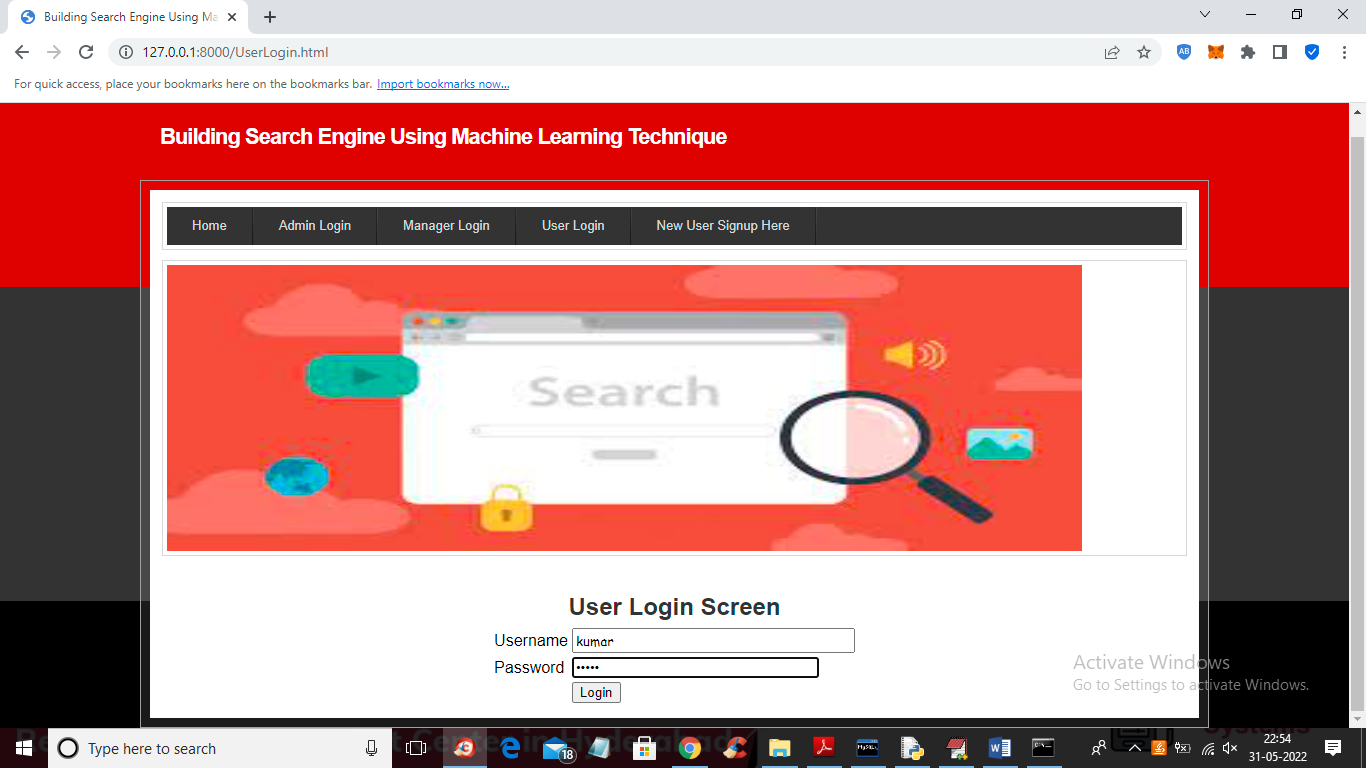
In above screen manager can click on ‘Upload Dataset’ link to upload dataset or documents



In above screen manager is browsing and uploading dataset and this file you can find inside ‘Dataset’ folder and now press button to saved dataset at server database



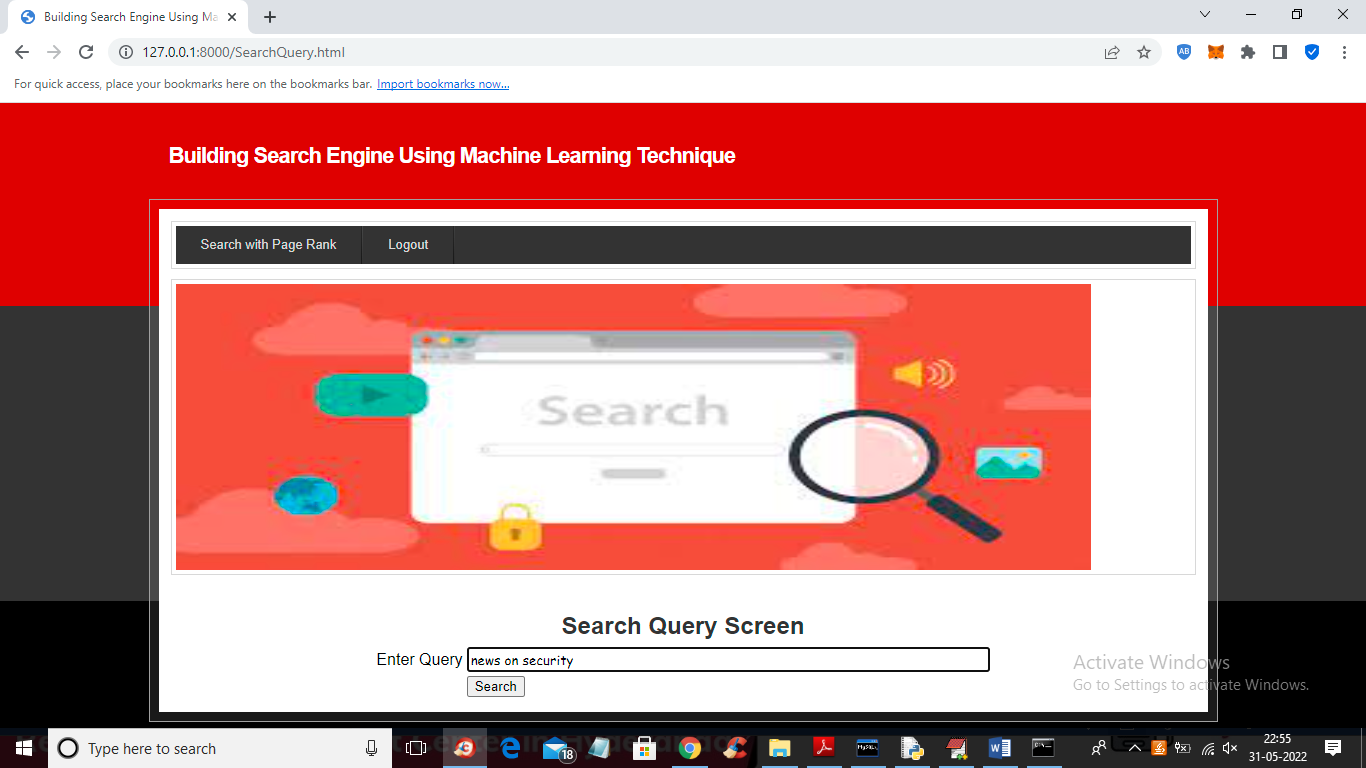
In above screen dataset file saved in database and now logout and login as user to perform search



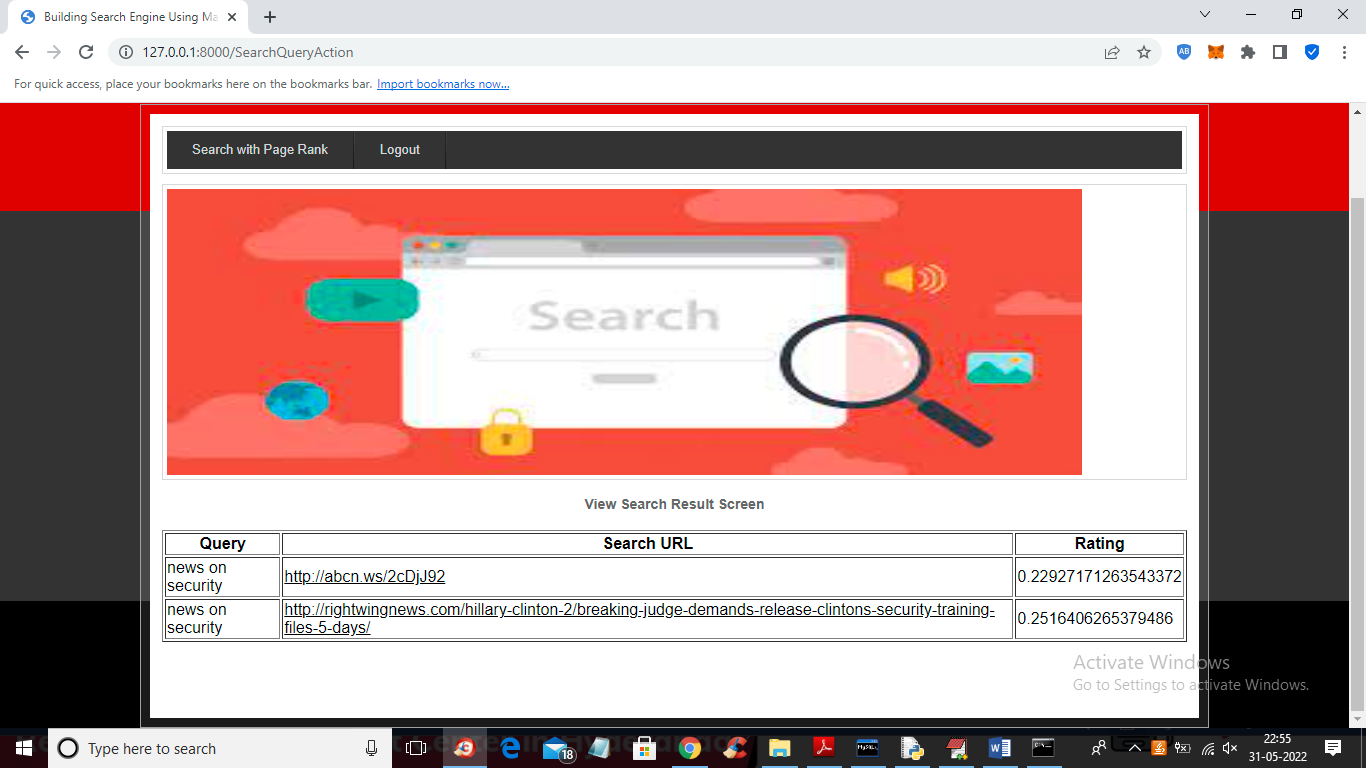
In above screen user is login and after login will get below output



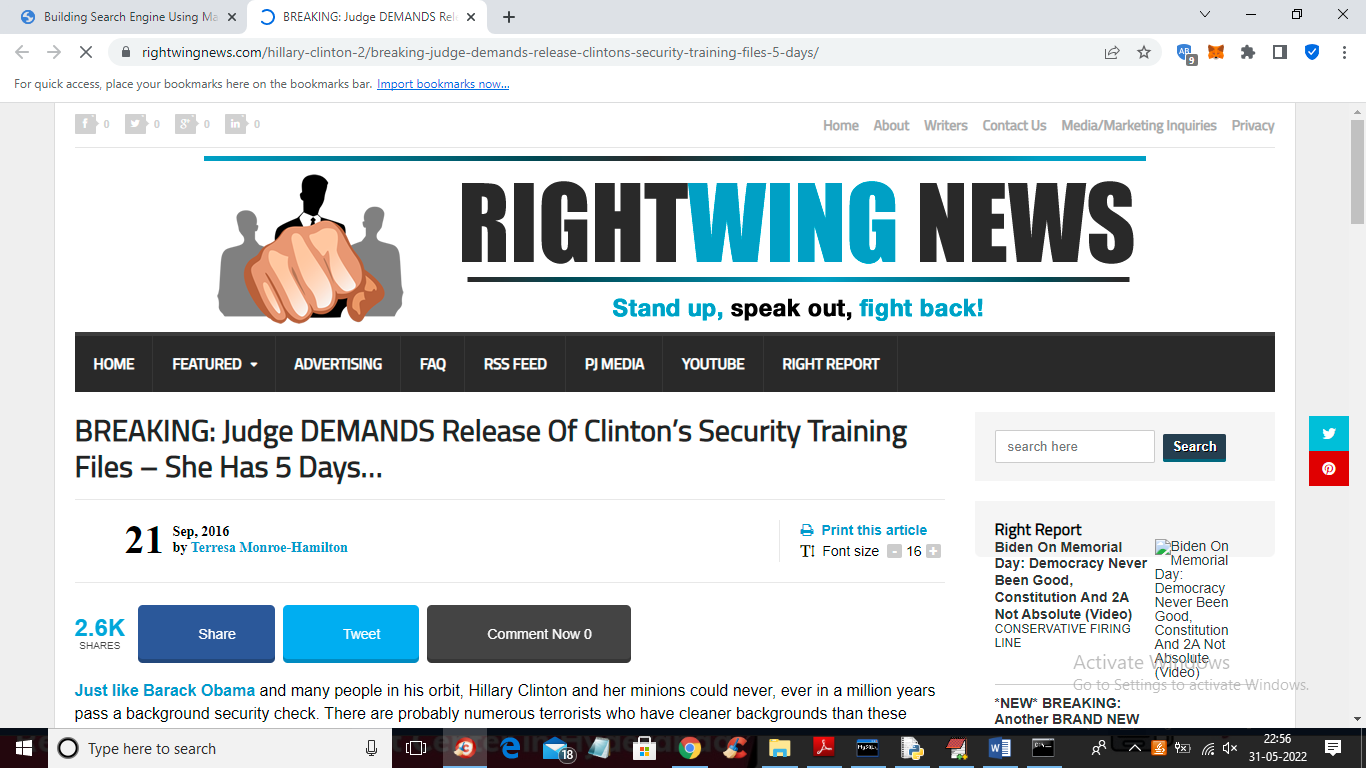
In above screen user can click on ‘Search with Page Rank’ link to search any data



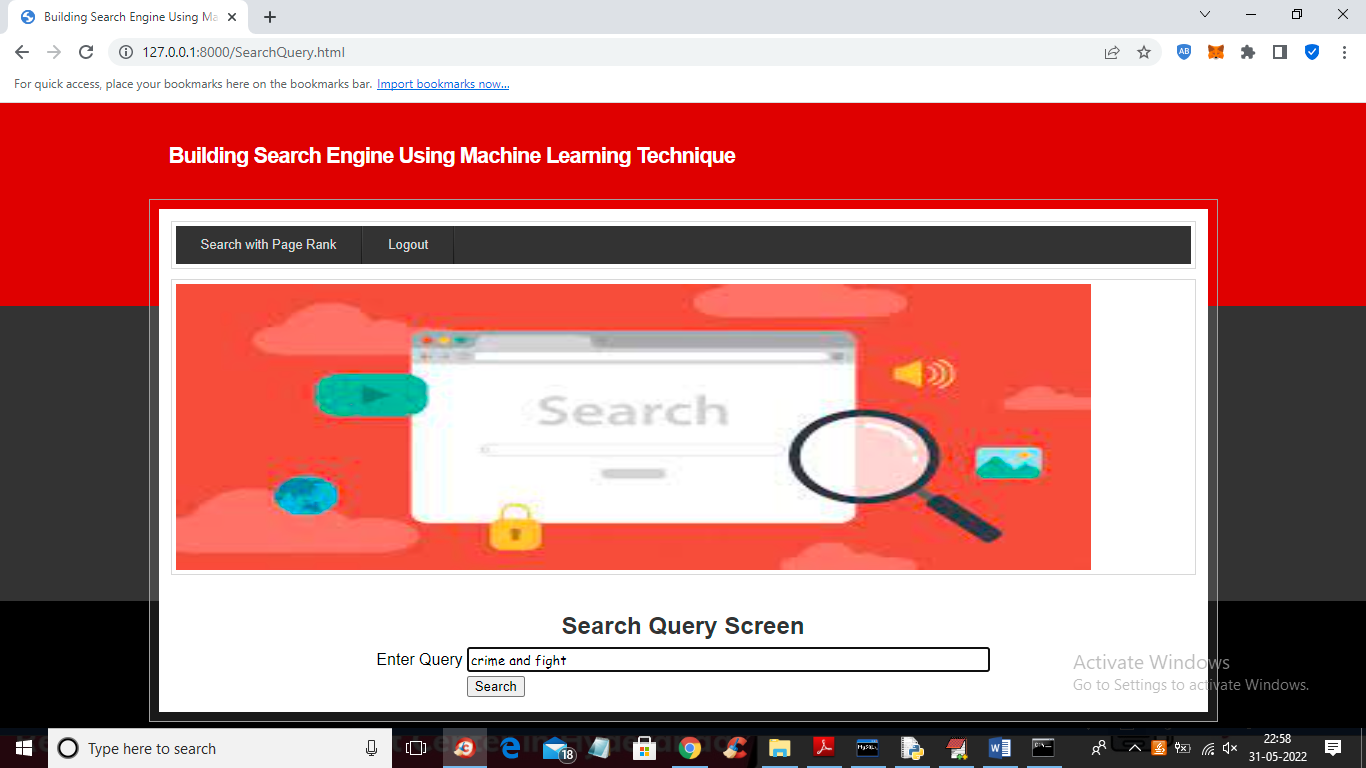
In above screen I entered query as ‘news on security’ and press button to get below search result



In above screen machine learning algorithm predicts two URLS for given query and user can click on those URLS to visit page



In above screen by clicking on URL link user can visit and view page. Similarly user can give any query and if query available in dataset then he will get output



For above query we got below result

