Central Database of Passout and Dropout Students in Higher Universities

Sumaiya Siddique,Payal Chotaliya,Priya Dubey,Devika Rani Roy
Student, Student, Student,Professor
Department of Information Technology
K.C. College of Engineering, Thane, India

Abstract : Student progression database for tracking students of higher education, passouts and dropouts. In this paper, we discuss Data Mining and its application in Higher Secondary Directorate of Assam, as it is problem of Government of Assam.This experiment covers attributes like School Type, Candidate Type, Study Type, Districts, etc. These attribute values, we are using to analyse the result of Higher Secondary Examination. This will improve the performance and data processing speed of Higher Education Directorate. The paper demonstrates the application of Data Mining in Higher Secondary Examination result. This will help further research and will improvise the activity of Higher Secondary Directorate.

IndexTerms - Data Mining, Examination Result Analysis, Data Classification, Higher Secondary Directorate.

I. INTRODUCTION

This is the real time problem faced by Assam Government, that is Lack of central database of Dropout and Passout Students of Higher Education. Assam higher education is formed in 1948.India's Higher Education Policy is mainly governed by the "National Policy on Education" of 1986 (as modified in 1992) .
The 1986 policy and its Programme of Action of 1992 were based on two reports namely, Radhakrishnan Commission Report (1948-49) and Kothari Commission Report (1964-66). The 1986 policy emphasized five main objectives for higher education:

1. Access
2. Equity
3. Quality and Excellence Relevance
4. Value Based Education

The XI and XII plan have laid emphasis on Improving access, equity and fostering excellence in higher education.
In the current paper we have developed an application for face recognition using PCA Algorithm. The overview of the system developed using PCA algorithm is shown in the figure (1).
(A) In particular, XII plan proposed a comprehensive plan for state higher education system for
(B)ensuring the above goals.
This led to the formulation of the Centrally Sponsored Scheme, namely the Rashtriya Uchchatar Shiksha Abhiyan (RUSA). The Government of Assam has consented to participate in the RUSA. Government of Assam has constituted the Assam State Higher Education Council on 7th January 2014 , vide Executive order No: AHE 534/2013/13 dated 7th January 2014. This formulation of the Centrally Sponsored Scheme at our State, namely the **Rashtriya Uchchatar Shiksha Abhiyan (RUSA).**

Data mining has attracted many researchers and analysts in the information industry and in research organizations as a whole in the last years. This is primarily due to the availability of large data and the immediate need for transforming such data into meaningful information and knowledge. The useful knowledge gathered can be applied in many areas such as market survey, customer retention, production control, evolutionary analysis and science and exploration, prediction of students.[1]

The prediction with high accuracy in students performance is beneficial as it helps in identifying the students with low academic achievements at the early stage of acdemics. In universities, student retention is related to academic performance and enrollment system[3].

The steps to assist the low academic performers with better education are:
(a) Generation of data source of students for prediction.
(b) Identification of various features or factors which affects the performance of student's learning during academic career.
(c) Construction of a prediction model with the help of classification data mining techniques on the basis of predictive variables which is readily identified.
(d) Validation of the model which is developed for Universities with students' performance. Image Processing can be put in the educational field to extend for validation.

II. LITERATURE REVIEW

Recent studies and surveys have revealed an emerging need to continuously collect, monitor, analyze, summarize, and visualize relevant information from huge amount of data and user generated content.

2.1 Framework for searching research papers in dataset using soft computing approach.

This paper proposed and tested a framework for searching research papers in dataset using soft computing approach. This framework works in two phases. First, it facilitates classification of given research paper into relevant category in dataset. For categorizing a research paper, paper text is preprocessed first to extract the terms relevant to document nature. To reduce the number of features from the paper, information gain feature selection method is applied on the pre-processed text.

2.2 Application of Data Mining in Higher Secondary Directorate.

In this paper we briefly reviewed the various data mining applications. This review would be helpful to researchers to focus on the various issues of data mining especially in the area of Higher Secondary Directorate activities in Kerala. Extraction of useful information from the datasets is tedious and time consuming. Another challenge is the effectiveness of the existing algorithm.

2.3 Prediction of Students Performance using Educational Data Mining.

In this paper, the classification is employed in student information to predict the students division on the premise of previous information. As there are several approaches that area unit used for knowledge classification, Naive theorem is employed here. Information like group action, class test, seminar and assignment marks were collected from the students' previous information, to predict the performance at the top of the semester.

2.4 Importance of data mining in higher education system.

In this paper, Data mining process can be used in a mining process to develop the student database using k-means clustering algorithm to improve the student knowledge in the result of data set table. Implementation of data mining educational system can use to help the data instruction set in a student database system.

III .IMPLEMENTATION

3.1PROJECT IEDA:-

Creation of web portal which will maintain the records/data regarding all pass out and drop students of an Institution/University. Also the data will be sorted category wise for each and every student their by enabling a efficient data search and modification in database (if required).

Portal will also feature a section for fetching teachers vacancy in every institution as per specified method. User can search his institution in the portal as per location filtering and can fill form for his academic credentials with individual log in and also upload the specified documents as a scan copy, at the very moment he will be prompted (if he enters invalid or duplicate data) by Admin .

User will be provided/asked a section to upload/enter his current work state or CV (only if graduated) and notification form as per his field of interest or updates regarding his previous academics (for drop outs or mid-way stoppers only).User can also see different types of courses offered by near by institutions as per his filtered location.

3.2 USE CASE DESCRIPTION :-

- Users can search their institution directly also, on the portals homepage wherein they will be directed to their login/register page of institution to update their details.
- Followed by a specified result as per their search.
- Through OTP Verification a user can verify himself to update his/her data/credentials in database server.
- Recently searched tab will be accessing the most recently updated/searched universities or institution through the dynamically created database for every new entry, hence making the search more reliable.
- All information regarding current technology trends, government policies and jobs, carrier scopes and e-scholarships for deserved candidates will be modified regularly for keeping the portal updated.
- If a user tries to make a entry for his log in twice, he will be prohibited to do so, their by making portal more consistent, reliable and avoid data duplication, though an user has an access to update his data (if required).
- During all this, data will be minimized so as to efficiently load webpage in less network zones, their by making website robust as per user location. Offline Data Cache Manifest will be used to load a page without effecting its content.

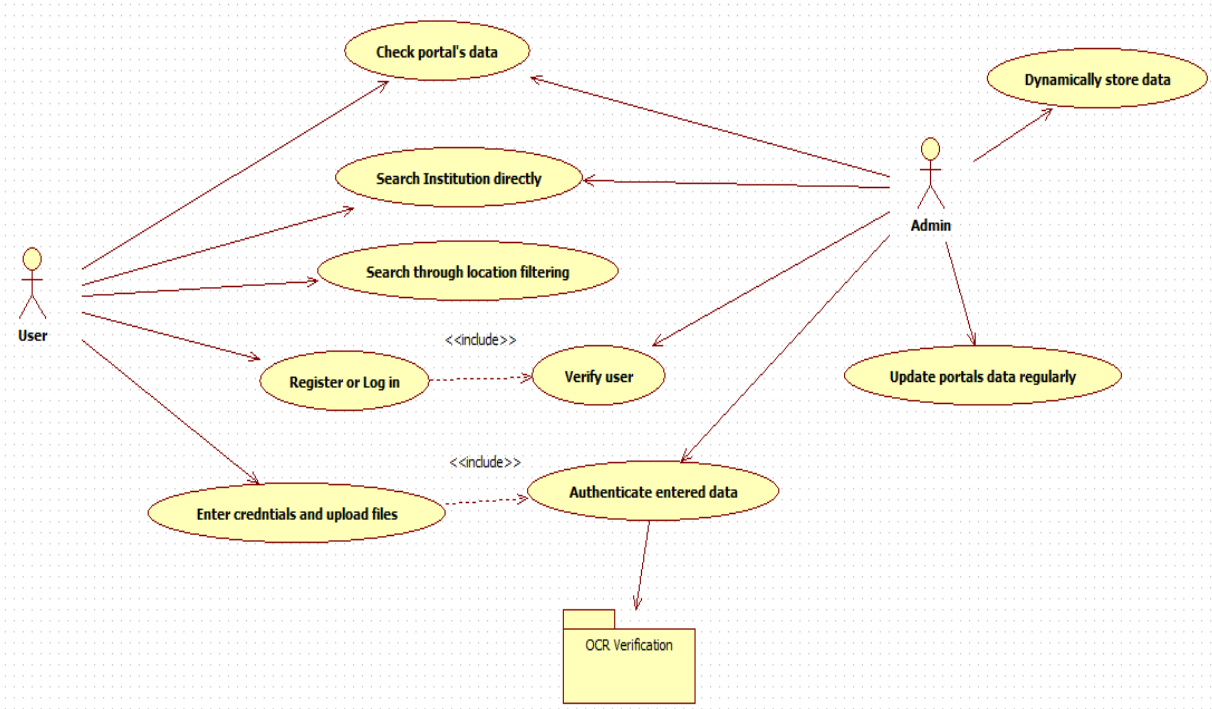

Fig. 1-Usecase Digram.

3.3 OCR TECHNIQUE:-

Forms submitted by users will be verified with scan copy uploads for the same using OCR (Optical Character Recognition) technique and that institutions existing records thereby providing security assurance before final submission.

For authentication process, we will be using **OCR technique**. **Tesseract** is an optical character recognition engine for various operating systems. Tesseract is used for text detection on mobile devices, in video, and in Gmail image spam detection.

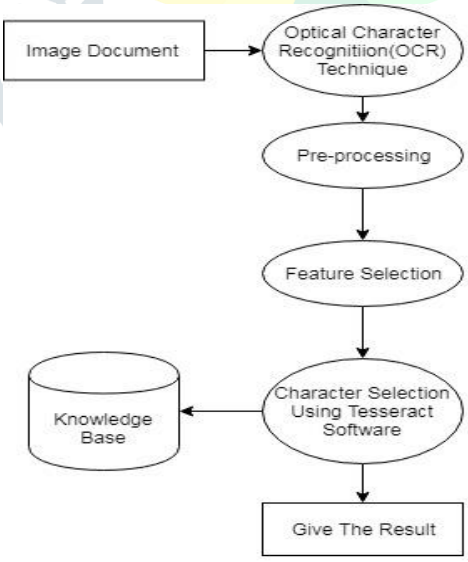

Fig 2. FlowChart of Image processing.

3.4 SOFTWARE USED:

The software used for developing isTesseract. In Our Case We have used it for converting upoaded scan document into text format.We used it in such way that from text format for feature selection.

Tesseract is an optical character recognition engine for various operating systems. It is free software, released under the Apache License, Version 2.0, and development has been sponsored by Google since 2006. In 2006, Tesseract was considered one of the most accurate open-source OCR engines then available.

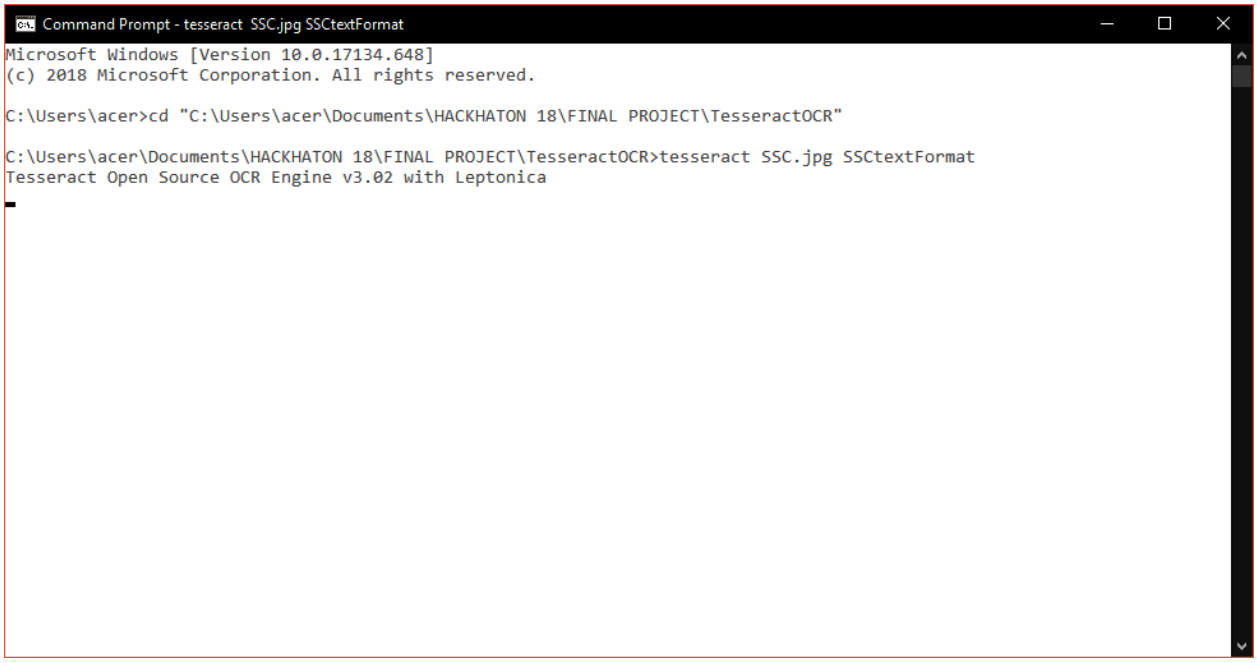

Fig3. Converting Scan image into Text File Using Tesserct OCR.

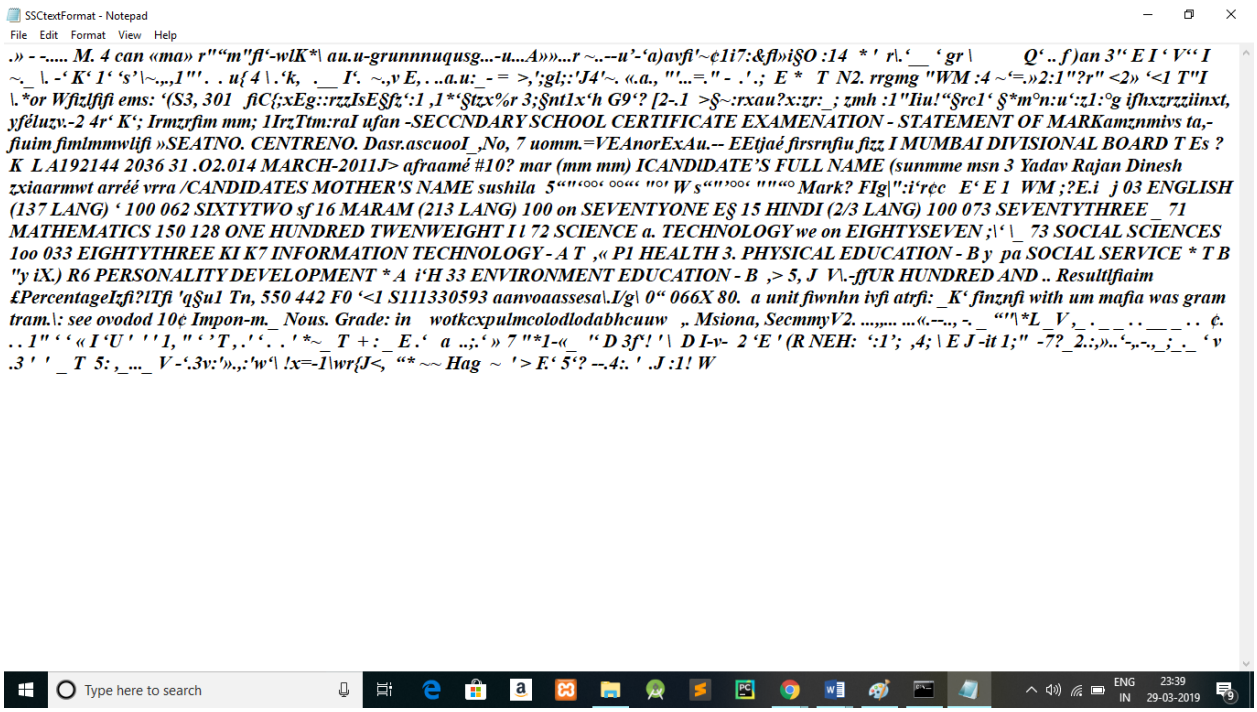

Fig4 .Converted Scan image into Text file.

III. RESULTS AND DISCUSSION

.The problem faced by Government of Assam(i.e Lack of Central Database of Dropout and Passout students of Higher Education.) will be focused by reviewing this report and would be helpful to determine the solution based on the problem of data mining.

By using data analysis approach,this paper proposed and tested a framework for searching
passout and dropout student in dataset. And last but not the least by using Image Processing Technique we'll be verifying the students accordingly.

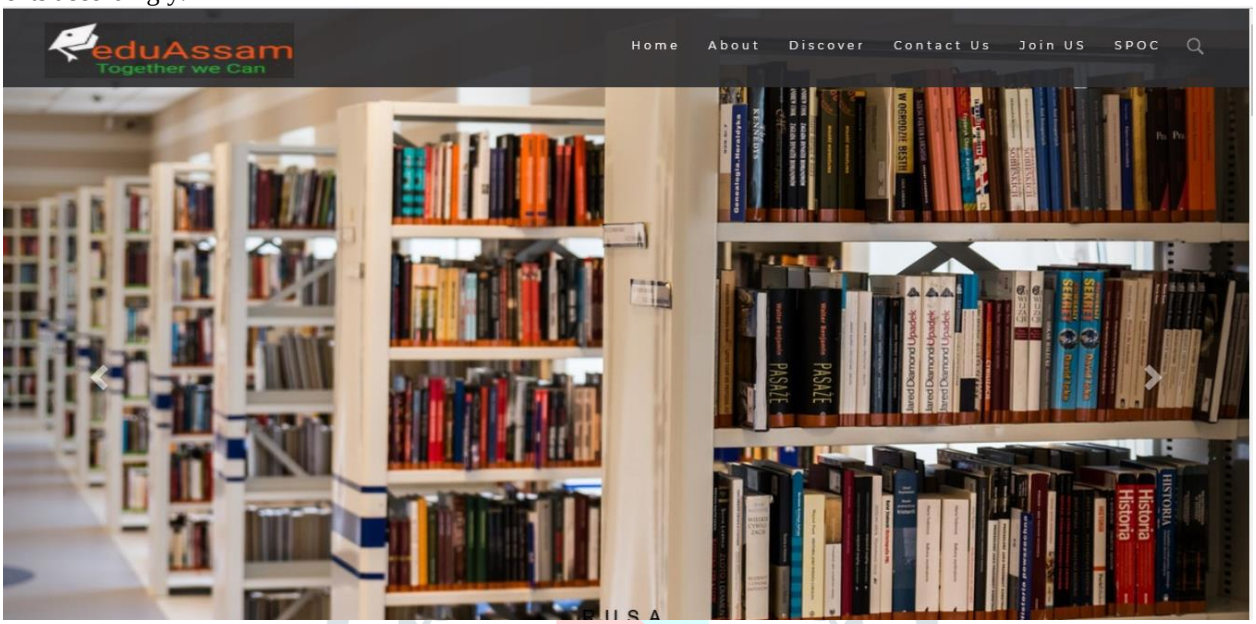

Fig.5- GUI

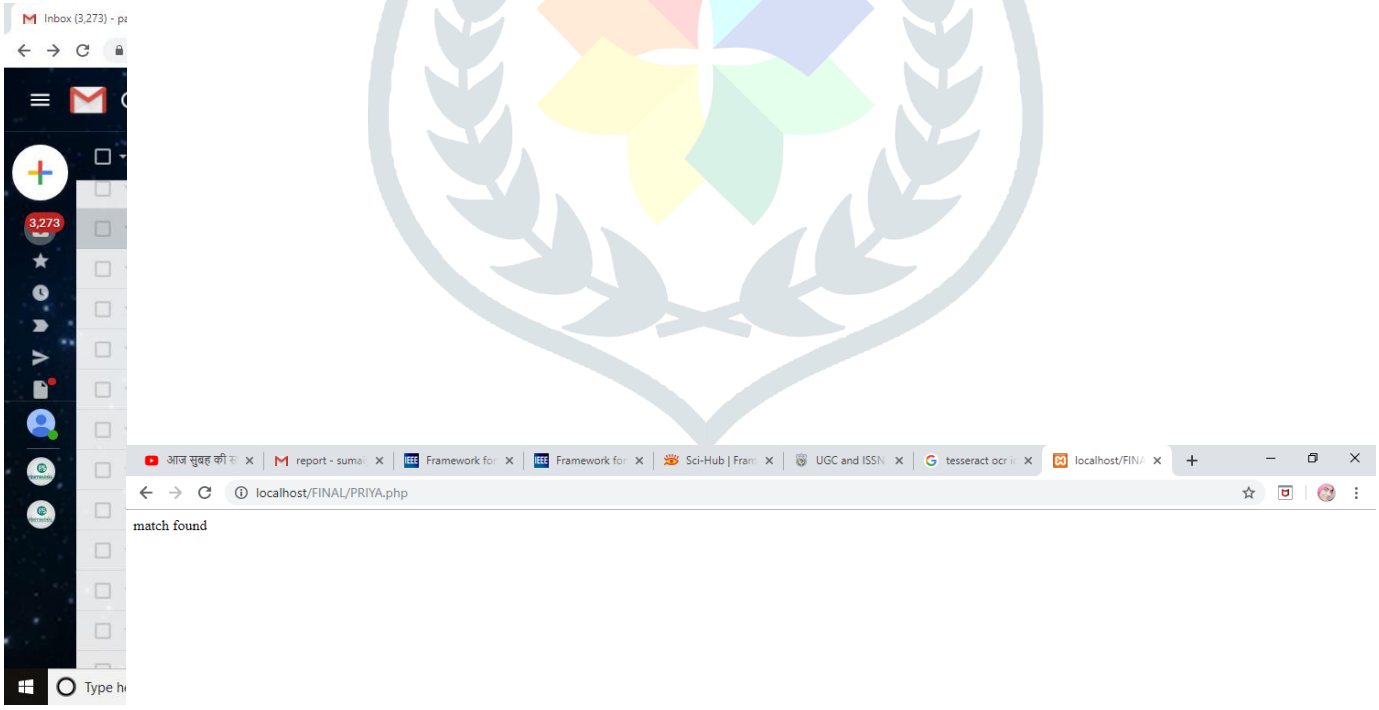

Fig.6- Verification link send to SPOC.

Fig 7- Authentication BySPOC(Admin)

IV. ACKNOWLEDGMENT

We would like to thank our project guide Mrs.Devika Rani Roy for encouraging us to work on this project and a special thanks to our parents for the constant support through this endeavor.

REFERENCES

[1]Application of Data Mining in Higher Secondary Directorate of KeralaSajan Mathew ; John T. Abraham2016 International Conference on Data Mining and Advanced Computing (SAPIENCE)(2016).
[2]Framework for searching research papers in dataset using soft computing approach Juhi Ganwani ; Jayant Gadge
2016 International Conference on Computing Communication Control and automation (ICCUBEA)(2016)
[3] "Prediction of Students Performance using Educational Data Mining" Ms.Tismy Devasia1 ,Ms.Vinushree T P2 , Mr.Vinayak Hegde3 Department of Computer Science Amrita Vishwa Vidyapeetham University,Mysuru Campus Mysuru, Karnataka, India .
[4] "Data Classification Of Student Perception Analysis Based Onnaive Bayes And J48 Algorithm"Dr. Ramchandra G Pawar.

[5] Higher Secondary Examination Manual (2012). Published by HigherSecondary Directorate, Govt. of Kerala.

[6]Mohammed M.Abu Tair and Alaa M.El-Halees, 'Mining Educational Data to Improve Students' Performance: A case Study', International journal of information and Communication Technology Research, ISSN: 2223-4985, vol.2 no.2, February 2012.

[7] Bhise R.B, Thorat S.S and Supekar A.K, "Importance of data mining in higher education system", IOSR Journal of Humanities and Social Science, ISSN: 2279-0837, vol.6, iss 6, January-February 2013.

[8] Kehar Singh , Dimple Malik and Naveen Sharma. "Evolving limitations in K-means algorithm in data mining and their removal", IJCEM International Journal of Computational Engineering & Management, April 2011.

[9] A.F . El Gamal, "An educational data mining model for predicting student performance in programming course", International Journal of Computer Applications(0975-8887), Vol.70, No.17, May 2013.