

Data Pre-Processing for Web usage Mining

^[1] K.Kiruthika, ^[2] T.Jeevitha, ^[3] S.Kamali Prabha, ^[4] V.S.BabyPriya
^[1] Assistant Professor, ^{[2][3][4]} UG Scholar

Department of Computer Science and Engineering, Panimalar Engineering College, Chennai-600123

Abstract: - Obtaining all the academic input about any individual at one click will be a great comfort. This is the case especially for any college/university or employment agency, getting na authenticated information about the student/candidate is very helpful in making a proper and well informed selection. This application has been developed to support such a requirement. The application will capture all the relevant information about the student on- line and by getting the student to feed the information directly. This also helps in tracking the student performance.

I. INTRODUCTION

Making higher education affordable has a significant impact on ensuring the nation's economic prosperity and represents a central focus of the government when making education policies. There is a good demand for higher education which also gives chances for better employment. Due to this, many students take loans to help in payment of the fees and other expenses connected with education like staying, transportation food expenses. It is said that the loan amount for student loans is more than the credit card or other loans in U.S. Thus there is demand for quality education from a good university. Further, the period by which the student completes his higher education is also a point of consideration, which can also escalate the costs and loan amount. In fact, recent studies show that only 50 of the more than 580 public four-year institutions in the United States have on-time graduation rates at or above 50 percent for their full-time students. To make college more affordable, it is thus crucial to ensure that many more students graduate on time through early interventions on students whose performance will be unlikely to meet the graduation criteria of the degree program on time. Such extended time spent on education increases the cost by way of additional tuition, fees and other essential expenditure. This in turn will increase the debt incurred on education.

The important thing is to ensure that the number of students completing their education within time increases. For this, proper system should be in place for identification of such students whose performance will be unlikely to meet the graduation criteria of the degree program on time and providing support to them. A critical step towards effective intervention is to build a system that can continuously keep track of students' academic performance and accurately predict their future performance, such as when they are likely to graduate and their estimated final GPAs, given the current progress.. Although predicting student performance is being done, it has not been fully tried for a student

obtaining a degree. This is because there a difference in predicting a student at degree level and it has different challenges. Some of them are

- 1) Students differ in terms of background
- 2) There are various courses which can be chosen
- 3) Same course can be chosen by students in different areas
- 4) Students can show different performance levels

Thus, predicting student performance based on a vast range is very difficult

The data received from various sources are also varied. The prediction of the student performance depends of past behaviour is difficult due to data coming from different areas and different student interests. Thus has to be compared with the normal approach of solving problems by following standardised steps, without considering the differences. Further, the assessment are also standard for all the students. The data collection is not to be one time effort, it has to be continuous through the period of study. Thus, the prediction should consider the performance of the student though the entire period when the performance can change.

Faced with the above challenges, a new application for the collecting of the student information, which may be useful for predicting is now proposed.

In the existing system The Society for Human Resource Management recommends that employers to use a trusted background screening company or verify education credentials of applicant's to check their education background and to identify applicants who have inflated their qualifications so that such persons can be removed and to protect against fraud activities like diplomas purchased from and verified by illegitimate universities and also to verify education credentials only through direct contact with the school and colleges.

Job seekers provide the information to the service for verification and pay the required fees. Colleges need to check the entire database to get the details. The process may extend the deadline time. The main advantage of the

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 5, Issue 3, March 2018

proposed system is to create a database which stores student information - academic, personal, certifications, workshops and more for current and passed out student. It relieves the administrative burden which involves hours of staff time and stacks of paper work. Here the colleges can simply get the profile of each student by their University Registration Number so that authorized company can easily recruit the student on the basis of their requirements.

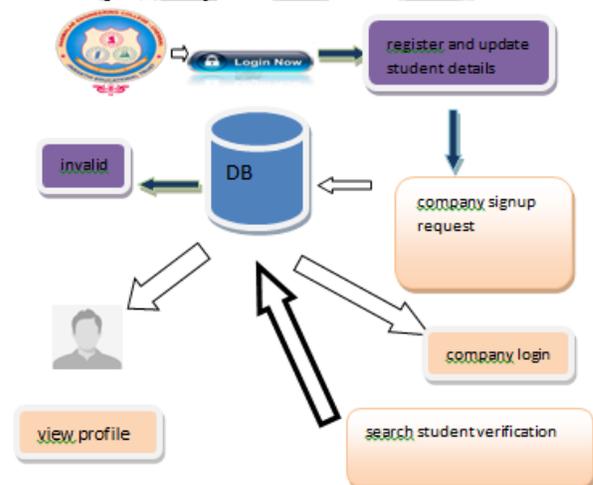
II. RELATED WORK

Artificial learning has given rise to many papers on various methods of predicting including predicting student performance in solving problems or completing courses. Many machine learning techniques, such as decision trees, artificial neural networks collaborative filters and probabilistic graphical models have been applied to develop prediction algorithms. However These techniques do consider the important dynamic of the change the student perception to learning over these periods. There is possibility that the student may improve with time and his approach to learning will be with more interest. To take the temporal/sequential effect into account, a three mode tensor factorization (on student/problem/time) technique was developed for predicting student performance in solving problems. A similarity-based algorithm was proposed to issue predictions of student grades in courses only when a certain confidence level is reached. However, due to the aforementioned substantial differences of predicting student performance in degree programs, these methods are considered successful. The application that is being discussed uses progressive prediction algorithm which is structured from the ensemble learning technique. It uses the Exponentially Weighted Average Forecaster (EWAF) as one of key structure for predicting the capability of the student to complete the course on time. But for the prediction to be reasonably accurate, there is requirement for regular updating mechanism from the student's part on the on-line mode. The EWAF used in this app has the ability to predict based on previous term predictions and multiple predictions on current data and process both together to give us a current prediction, which can be considered as a unique approach. The app is simple, scalable and with comparative prediction capability. However, as the existing predictors consider the current data for its prediction, this app considers all the data current and past with equal values, and thus is able to give a better prediction.

System Model:

The administrator (placement cell) can sign in with allotted user name and password. The administrator can update the student

details like personal, additional information, ratings and reviews. The personal details contain university register number (primary key), student's name, department, Address, Aadhar number, email id, Photo (image) of the student and contact number. Additional information contains academic input like course Type, academic percentage, number of arrears, course status, behaviour, attendance etc. The rating and review field which provides ratings and review of a student on the basis of their academic and behavioural performance. After updating, all the required fields are stored into the database which can be viewed for future verification. The company wishing to recruit will request the placement department .The placement department checks whether the request is coming from an authorized company. If yes, the company will be permitted to view the updated student profile which is a stored in the database. If the company is not authorised, then the request is rejected.



fig(1): System Architecture

System Module Description

Login:

Login purpose is mainly done for authentication of accounts. Primarily, User name and a password is entered to check for valid user if the user name and a password is valid it directs the administrator to the student verification and then the process is done. If it is invalid error occurs and user cannot do the further operation

Register and update student details:

The main purpose of this project is to help the organisation from the tedious paper works. Here we mainly work on updation of student details whereas in realtime it could take hours to do this process. But in this project we can easily update student details not only update we can also view and store them for the future purpose. After entering all the fields click the "submit" button. The updated details gets stored in the database.

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 5, Issue 3, March 2018

Database:

The updated students details gets stored in the database which can be viewed and modified at any number of times, just by entering the students unique registration number we can retrieve the particular students details with minimal amount of time and with less effort. The one advantage is that we can view the profile of a student directly by clicking 'view profile'.

Algorithm:

Linear search or Sequential search algorithm is used in order to find the target value in a list. This algorithm sequentially checks the each and every element until the exact match value is found. The purpose of using this algorithm is in a database it contains multiple student details in the same name in order to collect a particular students details so, linear search algorithm is used.

III. CONCLUSION

Currently though the app has been developed for both student support and administrative comfort by recording student details it can also be used effectively to predict and provide a tool to improve student performance. Further it minimises the time and effort of the administrative staff. This application gives an effective and dependable student verification with less effort and no cost.

REFERENCES

- [1]. Jie Xia, Member, IEEE, Kyeong Ho Moon, student member, IEEE, and Mihaela van der Schaar, Fellow, IEEE, "A Machine Learning Approach for Tracking and Predicting Student Performance in Degree Programs."
- [2]. Gunathilaka T.M.A.U, Fernando M.S.D., Pasqual H. "Identification of the Learning Behaviour of the Students for Education Personalization"
- [3]. Manish Joshi, Ravindra Vaidya, Prawan Lingras, "Automatic determination of learning", 2010.
- [4]. Mpine E. Makoe, "Bridging the distance: The pedagogy of mobile learning in supporting distance learners", 2012.