Prediction of Academic Performance Using Data mining Techniques

T. Shilpa[1], Dr.K.Venkataramana[2]

[1] *Student, Department of MCA, K.M.M Institute of P.G Studies, Tirupati*
[2] *Associate Professor, KMM Institute of P.G. Studies, Tirupati*

Abstract- **This paper displays the aftereffects of applying an instructive information mining way to deal with demonstrate scholarly weakening (loss of scholarly status) at the Universidad Nacional de Colombia. Two information mining models were characterized to dissect the scholarly and nonacademic information; the models utilize two order strategies, gullible Bayes and a choice tree classifier, all together to gain a superior comprehension of the weakening amid the to begin with enlistments and to evaluate the nature of the information for the arrangement errand, which can be comprehended as the expectation of the loss of scholastic status because of low scholarly execution. The models mean to anticipate the steady loss in the understudy's initial four enlistments. To start with, thinking about any of these periods, and afterward, at a particular enlistment. Chronicled scholarly records and information from the affirmation procedure were utilized to prepare the models, which were assessed utilizing cross-approval and before hand inconspicuous records from a full scholarly period. Test comes about demonstrate that the forecast of the loss of scholarly status is enhanced when the scholarly information are included.**

Index Terms- **data mining, prediction, attrition**

I. INTRODUCTION

The use of Analytics in the instructive setting has expanded in the most recent decade. Ferguson shows in three drivers for this to happen: to start with, the volumes of information that are gathered in instructive organizations have enormously enlarged, regardless of whether from Course or Learning Management Systems, or Student Information Systems; the second driver is the utilization of e-learning: in spite of the fact that it has helped gather information, it additionally has brought some learning issues, for example, a conceivable absence of inspiration and challenges for the teachers to get immediate criticism in regards to the temperament, level of premium, or even the comprehension of the understudies; at long last, the political concerns: nations are improving comprehension of the significance of advanced education in their improvement and governments have an enthusiasm for enhancing it to offer better learning openings that prompt better scholarly outcomes. Under these conditions, Data Mining systems have been connected in both Learning and Administrative/ approach arranged issues however with a more noteworthy spotlight on computerized revelation. In Learning, the procedure can be viewed as student arranged or instructor situated. In the first, the attention is on supporting the understudy to take in more viably by recommending new substance in the last mentioned, the objective is to furnish teachers with a device to enable them so they can direct the student all the more adequately. Then again, cases of utilizations in the Administrative issues can be Data Warehouses or different Business Intelligence apparatuses to help the basic leadership process.

The utilization of Data Mining in the Educational setting is alluded as Educational Data Mining (EDM) and is characterized by the International Educational Data Mining Society in as "a rising order, worried about creating strategies for investigating the one of a kind sorts of information that originate from instructive settings, and utilizing those techniques to better comprehend understudies, and the settings which they learn in". As to techniques for EDM, Baker proposes in a grouping as takes after: Prediction, Clustering, Relationship mining, Distillation of information for human judgment, and Discovery with models. Romero and Ventura, then again, propose in an alternate scientific categorization in light of the accompanying instructive assignments: Analysis and Visualization, Providing input,

Recommendation, Predicting Performance, Student Modeling, Detecting Behavior, Grouping understudies, Social Network Analysis, creating Concept Map, Planning and Scheduling, and Constructing Courseware. Be that as it may, being EDM a use of Data Mining, its assignments are only the same: Classification, Clustering and Association Rules Analysis and the exploratory errands are additionally incorporated into the Knowledge Discovery Process.

For ensuring the security, the plain text is converted to cipher text and the process is called encryption. Although this conversion idea is old, the way of encryption should not be vulnerable to attacks. Objective of this project is to develop multi-level encrypting software that can be used to encrypt top-secret files including text, images and multimedia files in the secondary storage devices. Encryption is one of the principal means to guarantee security of sensitive information. Encryption algorithm performs various substitutions and transformations on the plain text (original message before encryption) and transforms it into cipher text (scrambled message after encryption). Many encryption algorithms are widely available and used in information security. Encryption algorithms are classified into two groups: Symmetric-key (also called secret-key) and Asymmetric-key (also called public-key) encryption. Symmetric key encryption is a form of cryptosystem in which encryption and decryption are performed using the same key. It is also known as conventional encryption. Asymmetric encryption is a form of cryptosystem in which encryption and decryption are performed using the different keys – one a public key and one a private key. It is also known as public-key encryption.

II. RELATED WORK

Tinto's proposal of a higher education dropout model is an early theoretical framework in the field of dropout research . The model proposed that a student's family background, individual factors and previous education are prerequisite factors for student dropout. Whether or not student will drop out depends on the interaction between the student and the learning environment during the learning process. These interactions primarily include academic integration associated with academic performance and intellectual development as well as social integration associated with peer interaction and student- teacher interaction. Tinto's model, although not intended for E-learning, has a relatively important guiding significance for subsequent studies. Based on Tinto's research results, Kember introduced cost-benefit analysis for explaining the decision-making process of students dropping out, emphasizing that variables would change dynamically over time and that there would be multiple occasions in the students' learning process when they decide whether or not to terminate or continue their studies .

III. PROPOSED SYSTEM

The student classification models to predict the loss of academic status due to low academic performance, which uses the student admission data: socioeconomic, demographic and the initial academic information (all gathered from the admission process); and the academic records of previous academic periods. Different settings of the general model were trained and tested to predict the loss of academic status: first, the prediction is done regardless of the enrollment at which it occurs; second, the prediction at a given enrollment is performed based on the initial information (data gathered during the admission process); and then, using the information known before the academic period starts, which includes the grades of the previous academic period when available. The different settings are explained below.

A. Predicting Loss of Academic Status:
The most general case is that in which the interest is to predict the occurrence of the loss of academic status at any time in the first four academic periods based on the initial information.
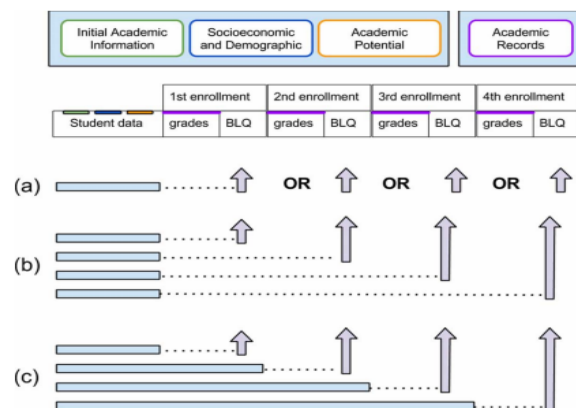
Predicting loss of academic status. (a) At any academic period based on initial information. (b) At a specific academic period based on initial information. (c) At a specific academic period based on initial information and previous academic records

B.Predicting Loss of Academic Status at a Given Academic Period :

First, initial data are used to train a model to predict the loss of academic status at a particular academic period. The model is then complemented by adding academic information to the admissions data. The event of loss of academic status in a given period uses the academic information, grades and previous loss of academic status, all available before the current period. For instance, to make a prediction in the third semester, data from the first two are used, together with the admission data.

The implementation of the models was done using Rapid Miner and the machine learning library for the Bayesian classifier and the decision tree respectively. These methods were selected based on the results of previous work and the need for a predictive model that is descriptive at the same time in order to acquire a better understanding of the event of loss of academic status. In other words, there is an interest beyond the identification of the students: why a student is classified in such a way and what variables are involved.

VI. CONCLUSION

The models were assessed utilizing distinctive situations relying upon the accessible data. These situations incorporated the forecast of the loss of scholarly status because of low scholarly execution whenever in the initial two years, at a particular enlistment utilizing just affirmation information and after that including the scholarly data, i.e. grades furthermore, credits selected. The models were tried with concealed records relating to a full scholarly period Bayes classifier execution enhanced when scholastic information from the primary enlistment were included; be that as it may, the execution diminished after the expansion of the scholastic information of the second enlistment.

REFERENCES

[1] Li Ying, Niu Jian, Ding Xia. A Follow-up Study of the Dropouts from the English Program of Open and Distance Learning (part 2) . Open Education Research, 2012, 18 (6): 80-86.

[2] Ran Jingjing, Guo Xuerong. Analysis and countermeasures of dropouts from distance education. Journal of BUPT (Social Science Edition), 2008, 10 (6): 102-106.

[3] Jiang Yulan, Zhou Lei. A research on the dropout of distance learners: Based on Grade 2000 finance majors of NBTVU. Journal of Ningbo Radio & TV University, 2006, 4 (1): 50-56.

[4] Yang Yongjian, Han Xue, Niu Jian, Li Ying. The factors and costbenefit analysis of student dropout from distance education. China Audiovisual Education, 2011, 7: 64-71.

[5] Chen Lin. Modern distance education high school dropout hazards. Journal of Neimenggu Radio & TV University, 2006,3: 66- 67.

[6] Liu Yongquan, Li Ying. Supporting students from success in open and distance learning: An interview with Prof. Ormond Simpson. Open Education Reserach, 2012,18 (5): 4-10.

[7] Blom, K. and Meyers, D. Quality indicators in vocational education and training: International perspectives. Adelaide, Australia: National Centre for Vocational Education Research(NCVER), 2003.

[8] Ji Zhe, Luo Fafen. Demonstration research on factors affecting learning achievement of distance learners. Open Education Research, 2008, 14 (1): 86-91.

[9] Li Ying, Niu Jian. A Follow-up Study of the Dropouts from the English Program of Open and Distance Learning (part 1). Open Education Research, 2011, 17 (6): 89-97.

[10] Research Group of Jiangxi RTVU. Student attrition in modern open and distance education. Distance Education in China, 2004, 9: 31-36.