Machine Intelligence to Predict Student Performance in Higher Education Institutions

S. Kumaravel
Associate Professor, Department of Computer Science,
A.V.V.M. Sri Pushpam College (Autonomous), Poondi, Thanjavur, Tamil Nadu, India.

Abstract: Predicting students' educational performance in teaching-learning institutions becomes tougher as a result of the large volume of knowledge in educational databases. Advancement in machine intelligence create it attainable to mine these educational databases and to seek out most relevant data that yield innovative ways that of supporting students, educators and better education institutions. The purpose of this paper is to use the options of machine intelligence in teaching establishments and to predict students' educational performance. This paper conjointly aims to specialise in however the prediction rule are often wont to determine the factors that influence the performance of scholars. As a case study, this paper employs machine intelligence ideas within the subject of packages to work out the tutorial performance level of the students so as to supply them a correct problem level of in operation System issues to unravel. To perform the case study, dataset from the educational management system WWW.avvmspclms.com of A.V.V.M. Sri Pushpam College is taken associated analysis is finished using WEKA an Open supply Tool.

Keywords: Machine Intelligence, Leaning Management System, WEKA

I. INTRODUCTION

In India, it's thought of that educational action of students in their education is life deciding issue so as to support student action, educators in educational activity institutions should have the flexibility to predict student performance to a given scenario and measures to enhance his performance. Educators' ability in predicting performance of scholars accurately can support them to classify students into completely different levels and this classification can change them to increase their support supported student level. Since the prediction relies on educators' ability, it's not possible for the educators to predict all the scholars in a very schoolroom in a very short span of your time and lack of man power in educational activity system is additionally another criterion to be addressed in predicting students. Thus, a tool that might mechanically predict students' performance is admittedly vital for educators. The advantage of victimisation the planned system is victimisation the prediction technique for the assessment of student success or failure rate and identification of best students or poor students in a very schoolroom to support teaching colleges. Prediction of best students supports higher officers for tutoring, scholarships or grants. Weak students typically want support from their peers or from concern college and heads of the department. It is additionally helpful to help academics to be able to organize students into seminar teams. On the opposite hand it is difficult once academics build conclusions primarily based solely on this data.

The reminder of this paper is organized as follows: literature survey regarding the coed performance prediction is bestowed in section II, style of the planned methodology is bestowed in section three, Implementation of the planned work is made public in section four, result and discussion of the dataset is highlighted in section five. Finally, section six presents the conclusion.

II. LITERATURE SURVEY

Different studies in terms of educational behaviour [1], student attrition [2] [3], atmosphere issue [4] and student retention [5] are administered by researchers within the domain of student learning and performance prediction [6]. Aljohani, O [7], summarises the prevailing prophetic frameworks, ways and attributes utilized in the prediction of student performance in educational activity establishments. In recent times, stress on the predicting student performance has been on the utilization of their psychological feature ability, additionally because the student demographic attributes. numerous studies [8][9][10] & [11] additionally

emphasizes the role of demographic information in student performance prediction with the support of machine learning techniques. From the literature survey it is noted that the various studies [12], [13] and [14] used survey form techniques to gather the learner intrinsic and temperament information that don't seem to be promptly offered within the information for predicting student performance This man of science measured the consequences of student temperament traits, student learning designs, their learning ways, individual motivation factors and role of psychological well-being on the educational performance of scholars. this sort of form study is currently assisted with log activities in Learning Management System (LMS). Researchers advocates and utilises net primarily based systems additionally as LMS in predicting student performance. Researchers in [15], [16], [17] and [18] foreseen student performance with the support of LMS. In net primarily based system or LMS, researchers build use of variables like the frequency of login , the amount of on-line sessions, the amount of content page viewed , the amount of the initial posts read/ created the amount of follow-up post created, and also the range of posts browse. However, despite the higher than things as an element for the measure of student performance, solely only a few studies agues and went deeper to appear at the standard of participation within the LMS instead of the amount of log details. during this regard, man of science in [19] highlights that a lot of researchers have planned divergent reasonably methodologies and algorithms to enhance the coed educational performance prediction additionally as characteristic the simplest thanks to increase the accuracy and also the potency of the prophetic model. Finally, viewing the literature review of student performance as a full, it's discovered that the varied researchers have shown that the most reasons for low performance of student from educational activity establishments area unit usually not recorded formally. There area unit external factors that area unit out of management of the upper Education establishments. Most of those factors area unit student dependent and intrinsically involve participating the coed in providing answers to them through the utilization of survey or interview is complicated one and obtaining true worth from them is additionally not possible. Moreover, it ought to be noted that these factors that have an effect on and verify student performance don't seem to be solitary in nature however area unit interconnected, reticulated and mutuality. From the higher than survey it's argued that there's a requirement for analysis to develop a brand new framework that's comprehensive and holistic in predicting student performance.

III. METHODOLOGY

Higher education educators and learners area unit a great deal curious about sharing info resources and providing learning tasks that might improve educators and learners information thus the planned system aims to provide and monitor users moreover as offer them an easiest method to enhance the general establishment performance. The planned system is shown in figure 1.

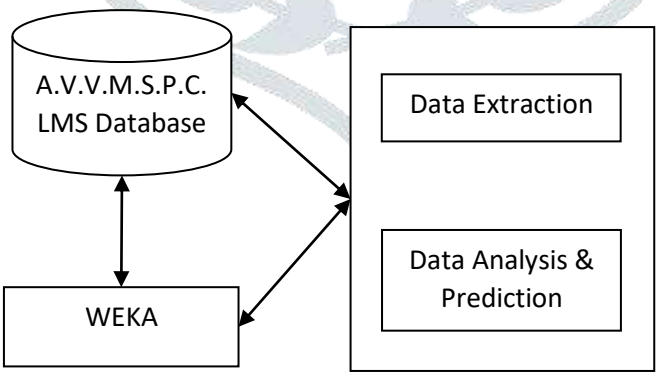

Figure 1: Proposed System

This method consists of four major elements specifically A.V.V.M.S.P.C LMS information, information Extraction, information Analysis & Prediction and wood hen.

This method consists of four major elements specifically A.V.V.M.S.P.C LMS information, information Extraction, information Analysis & Prediction and wood hen.

A.V.V.M.S.P.C. LMS information

The A.V.V.M.S.P.C LMS information may be a varied system that's wont to support not solely education at A.V.V.M. Sri Pushpam school. This information aids several functions within the establishment and its

main practicality involves managing students' learning-related records. The LMS information contains numerous sorts of information, e.g. students' grades, exams, electronic questionnaires, course enrolments, discussion teams, chatting, games, calendar, e-learning materials, journal or conference publications, course management, etc. what is more, the system keeps the complete log details of the user requests.

Data Extraction

This half extracts information from the LMS information. This information extraction generates a batch files to begin the computations mechanically. This batch files area unit handled by the information analysis and prediction module to analyse and predict the aforesaid batches. the opposite necessary role of this information extraction half is to sporadically outline the necessary characteristics that area unit required to analyse and reserve it in information. This information extraction can handle standard information pre-processing mechanism to cut back computation temporal arrangement.

Data Analysis and Prediction

The information generated by the extraction module as a batch files area unit additional processed during this data analysis and prediction module. To extract relevant options from the batch files, this half uses standard feature choice algorithms to extract the foremost relevant options from the batch files. additional this module obtains alternative basic statistics concerning the options, and it's ready to run machine learning algorithms per wood hen. This information analysis and prediction can give a prediction likelihood supported historical log information and alternative machine learning algorithms. This work utilizes 2 completely different machine learning algorithms enforced within the wood hen tool that features naive Thomas Bayes (NB) and support vector machines (SVM).

IV. IMPLEMENTATION

As it has been already mentioned, the information is obtained by the extraction module from A.V.V.M.S.P.C LMS information. every student is represented by a collection of attributes that exactly characterize student's qualities, potentials and interests. This work depends on 2 sorts of information: the learning-related information and data concerning the antecedently passed courses. the foremost helpful information from the learning-related information area unit average of grades, weighted average of grades, variety of credits to be gained, gained credits in an exceedingly course and quantitative relation between the last 2 attributes, programme details, and field of study. The antecedently passed course is chosen for every investigated course with relevance the rule that just about fifty students area unit to be listed therein course. This advocates and helps to cut back the quantity of courses to be foretold and to pick the foremost reliable courses. These attributes area unit added to the datasets to seek out out if there's any correlation between the courses and therefore the course to be investigated. This work predicts student interest in software subject victimisation the higher than criteria and provides connected materials to the scholars.

V. RESULT AND DISCUSSION

This work determines whether or not students have problem in learning and determination the software issues at hand or not. This work uses naive Thomas Bayes (NB) and support vector machines (SVM) rule to predict that a learner can have problem if the worth is quite zero.4, therefore requiring the intervention of the skilled system to work out an appropriate problem level so as to come up with drawback for the learner to unravel. If the output price is a smaller amount than or equal zero.4, then it's assumed that the learner can create four or less errors on a given drawback, won't need the help of the colleges and students will complete this drawback of his own. beginning with the system description as mentioned within the previous sections, we have a tendency to trained it with information collected from the log files from last version of A.V.V.M.LMS for the year of 2016. This work tests the software issues for the learner to unravel and at constant time predict the educational performance of the learner for constant drawback and store each results. meaning the particular learner result and therefore the foretold results area unit recorded in an exceedingly file for crucial the accuracy of the planned system the typical incorrect student tutorial performance predictions was 11 November. the typical prediction accuracy of the system was eighty nine.

VI. CONCLUSION

Improving student performance in education is one in every of the foremost necessary focuses of any school administration. to enhance student performance, predicting student tutorial performance is most significant which will be helpful to support the academics, learners and alternative stakeholders in instruction establishments. This work points out the importance of student performance predictions to the assorted stakeholders, reasons to spot the scholar performance degradation earlier to produce necessary support and intervention for them with the goals of accelerating student performance.

REFERENCES

[1] Summerskill, J. (1962) Dropout from college. In N.Sanford(ed). The American college, New York, Wiley

[2] Spady W.G., (1971). Dropouts from higher education: Toward an empirical model. Interchange, Vol.2, No.3, pp.38-62

[3] Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. Review of educational research, Vol.45, No.1, pp.89-125

[4] Cabrera, A. F., Castaneda, M. B., Nora, A., & Hengstler, D. (1992). "The convergence between two theories of college persistence". The Journal of Higher Education, Vol.63, No.2, pp.143-164.

[5] Bean, J. (1980). "Dropouts and turnover: The synthesis and test of a causal model of student attrition. Research in Higher Education", Vol.12, No.2, pp.155-187. http://dx.doi.org/10.1007/BF00976194

[6] Bekele, R. & McPherson, M., (2011). "A Bayesian performance prediction model for mathematics education: A prototypical approach for effective group composition". British Journal of Educational Technology, Vol.4, No.3, pp.395-416.

[7] Aljohani, O., (2016). "A comprehensive review of the major studies and theoretical models of student retention in higher education. Higher Education Studies,Vol. 6, No.2, p.1.

[8] Kotsiantis, S.B. & Pintelas, P.E., (2005), July. "Predicting students marks in Hellenic open university". In Advanced learning technologies, 2005. ICALT 2005. fifth IEEE international conference on (pp. 664-668). IEEE.

[9] Oladokun, V. O., Adebanjo, A. T. & Charles-Owaba, O. E. (2008). "Predicting students' academic performance using artificial neural network: A case study of an engineering course". The Pacific Journal of Science and Technology, Vol.9, No.1, pp.72-79.

[10] Hoe, A.C.K., Ahmad, M.S., Hooi, T.C., Shanmugam, M., Gunasekaran, S.S., Cob, Z.C.& Ramasamy, A., (2013) "Analyzing students records to identify patterns of students' performance". In Research and Innovation in Information Systems (ICRIIS), 2013 International Conference on (pp. 544-547). IEEE

[11] Ikbal, S., Tamhane, A., Sengupta, B., Chetlur, M., Ghosh, S. & Appleton, J. (2015). "On early prediction of risks in academic performance for students. IBM Journal of Research and Development, Vol.59, No.6, pp.5-1.

[12] Sembiring, S., Zarlis, M., Hartama, D., Ramliana, S., & Wani, E.(2011). Prediction of student academic performance by an application of data mining techniques. In International Conference on Management and Artificial Intelligence IPEDR, Vol.6, pp.110-114.

[13] Fariba, T.B. (2013). "Academic performance of virtual students based on their personality traits, learning styles and psychological well-being: A prediction". Procedia Social and Behavioral Sciences, Vol.84, pp.112-116.

[14] Gray, Geraldine, Colm Mcguinness, and Philip Owende. (2016) "Non-Cognitive Factors of Learning as Early Indicators of Students at-Risk of Failing in Tertiary Education." In Non cognitive Skills and Factors in Educational Attainment, pp. 199-237. Sense Publishers.

[15] Minaei-Bidgoli, B., Kashy, D.A., Kortemeyer, G. & Punch, W., (2003). "Predicting student performance: an application of data mining methods with an educational web-based system". In Frontiers in education, 2003. FIE 2003 33rd annual (Vol. 1, pp. T2A-13). IEEE

[16] Romero, C., López, M.I., Luna, J.M. & Ventura, S., (2013). "Predicting students' final performance from participation in on-line discussion forums".Computers & Education, Vol.68, pp.458-472

[17] Agudo-Peregrina A.F., Iglesias-Pradas S., Conde-Gonzalez M.A., and Hernandez-Garcia A. (2014). "Can we predict success from log data in VLEs? Classification of interactions for learning analytics and their relation with performance in VLE-supported F2F and online learning" Computers in Human Behavior, Vol.31, No.1, pp. 542-550.

[18] Cerezoa, R., Sánchez-Santillánb, M., Paule-Ruizb,M.P., & Núñeza J. (2016). "Students' LMS interaction patterns and their relationship with achievement: A case study in higher education", Computer and Education Vol. 96, May 2016, pp. 42–54

[19] Bekele, R. & McPherson, M., (2011). "A Bayesian performance prediction model for mathematics education: A prototypical approach for effective group composition". British Journal of Educational Technology, Vol.4, No.3, pp.395-416.