

Collaborative learning for improving intellectual skills of dropout students using datamining techniques

M.Revathy Research Scholar,
Department of Computer Science,

Vels Institute of Science, Technology & Advanced Studies
revathysathya85@gmail.com

S. Kamalakkannan Associate Professor,
Department of Information Technology,

Vels Institute of Science, Technology & Advanced Studies (VISTAS)
kannan.scs@velsuniv.ac

Abstract— In each year millions of students drop out without completing their educational course. In such a case, both the individual student and institution will have an effect of dropping out. The proposed research work pays significant research attention towards analysing the higher education and upper primary students to identify their behaviour, which leads them to discontinue in the early stage and stop the dropout by taking necessary action towards the dropout reason. This in turn results in the lack of skilled workspace and weaken the productive system of the country and also student dropouts are more likely to become as the recipients of unemployment subsidies. This research is more focused on the dimension reduction techniques, which involves both the feature selection and feature extraction methods. It also aims to implement better prediction in identifying dropout students and implement collaborative learning with engagement. When PCR measures, the key elements do not look at the reaction but rather at the predictors (by looking for a linear combination of the predictors that has the highest variance). It assumes that, the answer is correlated with the linear combination of the predictors with greatest variance. It is presume that, the regression plane differs when selecting the main variable in the other orthogonal direction, along the line and it does not differ. The second path is disregarded by selecting one component and not the other. Principal Component Analysis (PCA) is a method used for extracting features that use orthogonal linear projections to capture the database. This is illustrated in two phases. First phase is the development of dimension reduction using PCA to identify an accurate prediction variance of dropout students by using various ML algorithms and the second phase involves the developing of collaborative learning with engagement through social media and improves their intellectual skills by performing SVM hypothesis test.

Keywords— SVM, Collaborative learning, dropouts, Hypothesis.

I. INTRODUCTION

Dropping out students has significant implications for learners and their environment. Students who have opted to drop out of schools face stigma, restricted work shortcomings, lower wages and greater participation. within criminal judiciary system. University dropout is an issues such as, reduced enrolment, [14] Decreased college income and The State that finances the inquiries and sets up undergraduate students'

financial problems as a social problem. Dropping out is influenced by both the students and institutions. [9]In engineering education, late dropout is becoming more of an issue. [nine] Despite the fact that, global participation in higher education has increased, investigating the students behaviour, mental health, academic performance and their family background helps to predict the students dropout. Student self-esteem and psychological well-being are impaired by dropping out of institutes that are faced with the fact that they lack skills and expertise to fulfil their wishes.[17] Dropouts can be linked to problems with student success. These typically manifest themselves right at the start of the research process. In certain cases, students are unable to finance these deficits on their own. [12]Over the last fifty years, dropout rates and low graduation rates have become an increasing source of concern for higher education institutions and education authorities, as low graduation rates expand social and economic inequalities and stymie country growth. As a result, students[15] tend to ignore or even negate their issues as long as possible because of a certain helplessness, fearing the pain of exposing themselves to themselves. In this research paper first to identifies an exact factors for the student dropout it can be done through dimension reduction.[11]A dropout early warning system may assist schools in recognising students who are on the verge of dropping out and encouraging them to act. However, the predicting of children dropout can be analyzed by data mining which may assist in identifying the early dropout prediction from institution but need to improve their intellectual skills have become complex and challenging task to the institutions by providing collaborative learning tool for retaining the student in their academic performance. Therefore, the research is focused on those aspects of improving academic performance of those respective students to avoid student's dropout rate. Development of dimension reduction using PCA to identify an accurate prediction of dropout students by various ML algorithms. The early prediction satisfying their requirements. Principal Component Analysis, or PCA, is a technique of dimensionality-reduction that is often used by transforming a large set of variables into a smaller one that still preserves most of the information in the large set to minimise the dimensionality of large data sets.

Naturally, from the mean, or in other words, to see if there is any relationship between them.

II. REVIEWS OF RELATED WORK

Vinayak Hegde 1 ,Prageeth P On this paper, we present a technique that expects the pupil dropout to use rules in R language of the Naive-Bayes style set of rules. 1st And also observe the cause of falling out of a pupil in an early country and expect the student to drop or not now. As we mentioned above, there are various elements that influence a pupil to devote dropout. Prediction of early dropout assists the business enterprise to keep the students from the respective academic software, which incorporates feature choice and characteristic extraction. Function selection is a grade by grade technique that is used to choose the required characteristic from a collection of available characteristics. The feature extraction method involves the transformation of higher-dimensional records into lower-dimensional records. Feature collection consists of items such as lecturers, demographic variables, psychological variables, health problems, opinion of teacher, and behaviour of scholars. **Naruhiko Shiratori** Based on student learning results, we develop a method to classify learners as either "preliminary dropouts" or "standard" (student history, GPA, etc.). [4] [4] A tentative dropout condition is one in which a student is expected to drop out but has yet to do so; a typical state is one in which the student is expected to drop out but has not yet done so. A state in which the student is not liked. While dropout always occurs while a student is in the preliminary dropout state, there are instances where a return to the normal state can be created by making a successful qualitative shift. We describe this as a successful change of state and examine its distinctive characteristics

Boris perz et al This case study focused on predict dropout student applying data mining techniques. The main objective of this research paper to [11] decrease the dropout rate by identifies the potential causes. They use logistic regression, decision tree and Naive bayes algorithm and compare in order to get the best option. In this experiment they show predicting student from a large, heterogeneous dataset and students demographics and transcript records.

Meenakshi Upreti, Vishal Kumar this paper examines an impact of addiction towards online social network on adolescents and college students to perceive and apprehend [7] the reasons. They analysed the gathered statistics and they perform an experimental paintings on the age between 12-18 years. Because of the development in cellular, Technology in computing, the Internet and electronics at the hand of smart cities and towns, younger technology has become clever. The globalised shifts are adapted by humans and these changes are not part of the life of everyone. [16] [16] In each three months, the large variety of fb customers drops or drive upwards. This paper focusing on software and disadvantages of on-line social media.

Ravindar Ahuja and yash kankane they may be predicting the possibility of students degree crowning glory.this paper makes a speciality of diploma final

touch opportunity of college students at JIIT University. [4]Their preceding educational performance outcomes are in comparison with non-academic performances. They the usage of diverse instructional data mining techniques consisting of Naives bayes, decision tree KNN, and logistic regression. Those algorithms are applied in R language. They use 35 attributes of students encompass educational and non-academic elements.

Kittinan Limsathitwong , Thai-Nichi and Kanda Tiwatthanont this studies is a case study of first 12 months analyzing college students and second 12 months college students at Thau nichu institute of generation. they are analysed to identify the dropout fee and are expecting the scholars with their grades.[3]The prediction fashions have been advanced to use decision Tree algorithms and Random Wooded Area Algorithms were used to boost the performance. The choice was made. The accuracy, don't forget, and F1-degree of the tree classifier were 0.eighty, zero.92, and zero.eighty five, respectively. The results indicate that the predictor's usual effects are good. Dropout college students can understand the application and become conscious of those students who want special attention this is very useful to assist the scholars improving their studying procedure and to screen the scholar overall performance in a systematic manner.

Carlos marquez-vera, cristobal romeo morals they applied a data mining techniques to predicting the faculty screw ups and dropout.on this research paper they use actual statistics on 670 center [5]school college students from mexico. The targets of this look at to enhance their accuracy for predicting the scholars. they used white field classificatoin methods together with induction rules and selection timber for predicting the students as early as feasible.

Alvaro Ortigosa Rosa M. et al this paper focused on From Lab to Production: lessons learnt and real life challenges of an early student-dropout prevention system. The key purpose is to discourage students from leaving the university by seeking to optimize the efficacy of institutional initiatives in this direction by the atmost risk students were the subject of retention acts. [8]For this purpose, we developed predictive models based on the C5.0 algorithm using data from more than 11,000 students gathered over five years. [14]The SPA (Sistema de P P) was then built. [16]Since 2017, the SPA has been in progress and is currently in continuous use in its fourth semester. In order to estimate the dropout probability of more more than 5700 students, it has estimated more than 117,000 risk ratings. There have been around 13,000 retention actions reported. The white-box predictive models used in manufacturing, very similar to those obtained in the laboratory, produced fairly good performance

Naruhiko Shiratori We examine the issue of dropout rates by developing a method for identifying students either as being in a "preliminary dropout" state or a "natural" state based on student learning data (student background, GPA, etc). (student biography, GPA, and so on.) [2]A tentative dropout is a condition in which a student is considering dropping out but has not yet done so;A normal state is a state where the student is unlikely

to drop out, but has not yet; a typical state is a state where the student is not expected to drop out. Although dropout is more likely once a student is in the preliminary dropout stage, successful [6] qualitative adjustment can result in a return to normal status in some cases. It's what we call a good one.

Roderick Lottering et al. in this paper is focused on the advanced education segment in South Africa, growth has been seen. With this growth, an expansion in the dropout rate is seen. This research discusses the adequacy of dimensional decrease [16] and reflects on the substantial data found in student information for the identification of dropout-threatening students. This research relies on the prediction of dropout target of undergraduates from courses for educational data mining. In the exam, with knowledge from the recognition courses of a University of Technology, the researchers demonstrate promising results.

III METHODOLOGY TO BE ADOPTED

Phase I	Development of dimension reduction using PCA to identify an accurate prediction of dropout students by various ML algorithms. The early prediction satisfying their requirements.
Phase II	Deploying collaborative learning with engagement through social media and improve their intellectual skills to analyze them by SVM hypothesis test.

IV. BRIEF DESCRIPTION ABOUT THE RESEARCH WORK

This research focused on dimension reduction which involves both feature selection and feature extraction techniques while the approach of characteristic choice is a step by step procedure that is used to choose the right attribute from a given attribute units. The scholars at danger are possibly without carefully thought about the awful implications of their decisions, or without getting the ability to speak to experts, they drop out. The early intervention told by the early warning system of dropouts will redirect. However, the predicting of students dropout can be analysed by data mining which may assist in identifying the early dropout prediction from institution but need to improve their intellectual skills have become complex and challenging task to the institutions by providing collaborative learning tool for retaining the student in their academic performance. Therefore, the research is focused on those aspects of improving academic performance of those respective students to avoid student's dropout rate.

1. To identify the exact factors for student's dropout can be done through dimensionality reduction which deals with feature selection and feature extraction from the student's survey.
2. To make it easier to define effective policy recommendations for minimizing dropout rates in

potential cohorts.
 3. To create a technique that provides each student with adaptive input according to their form of collaborative activity and desired gamification elements

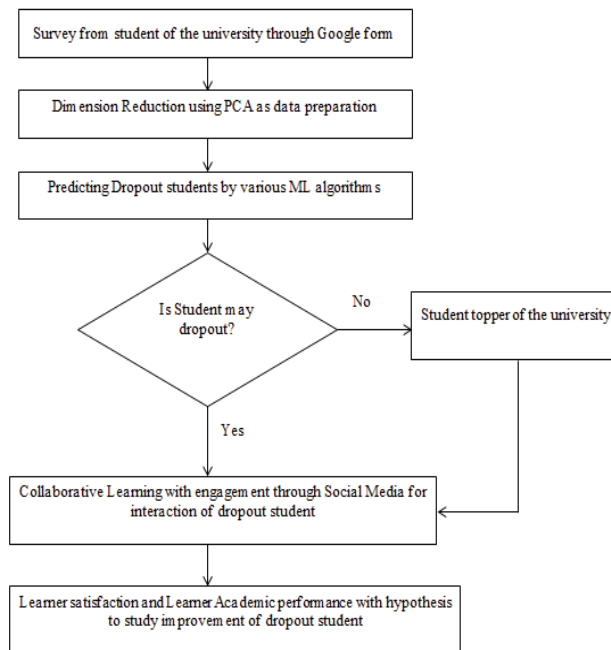


FIGURE 1 OVERALL PROPOSED FRAMEWORK FOR IMPROVING DROPOUT STUDENTS

In the technique of the extraction process of the functions involved in the conversion process of higher dimensional data with regard to lower dimension. In the Context of Science, the implementation of collaborative learning with engagement to the dropout students may assist in sharing of their knowledge and interaction with other students. It has insisted the dropout students to gain knowledge and communicate each other and even present their activity will make them to improve the confidence level of the student who may plan for dropout. This research initially identifies the dropout student from the university by the survey from google form. Once the survey get filled may helpful to recognize the standard and expectation of the students with several features and it may progressed by PCA technique to dimension reduction. The overall flow of the research is shown in figure 1 that illustrate both better prediction in identifying dropout students and implementing collaborative learning with engagement. This is illustrated in two phases namely

- Phase 1 – Identification of Dropout student by ML
- Phase 2 – Improving dropout student skills by Collaborative Learning and Engagement

Phase 1 - Identification of Dropout Student by ML

The input section takes data from the student of a university through Google Form with several variables which may assist to cover all kind of aspect to identify more precisely about dropout planning of future in early stage. Among these variables as features can be selected and extracted by dimensional reduction concept of principle compound analysis (PCA).

Once the data get pre - processed, the dimensional reduction is executed and progressed to various ML algorithms for accomplishing better accuracy in predicting the exact dropout students present in this university. The workflow of phase 1 is shown in the figure 2.

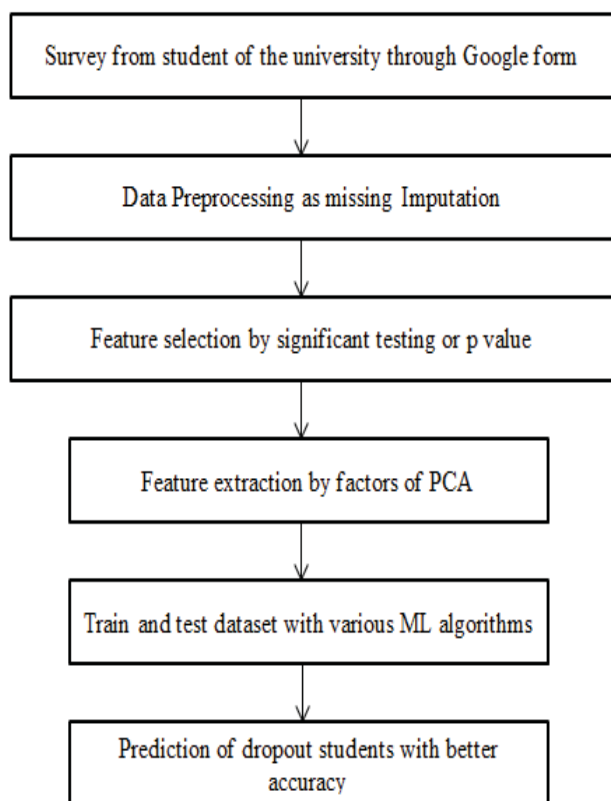


FIGURE 2 - IMPROVING DROPOUT STUDENT SKILLS BY COLLABORATIVE LEARNING AND ENGAGEMENT

This research proposes a framework of collaborative learning with engagement through social media which mainly focused for predicted dropout student along with university topper. This analysis is used to identify the student who can able to improve their intellectual skills by collaborating and communicating with engagement activities through SM are selected as a Research Learner (RL). The staffs and mentors are used as Supervisor (S) to monitor for engaging the online learning class through Engagement Events (EEs) in social media using concept of Collaborative Learning (CL) to improve Learner Academic Performance (LAP) in which Learner's

Satisfaction (LS) is essential has shown in Figure 3. This study has ten hypotheses

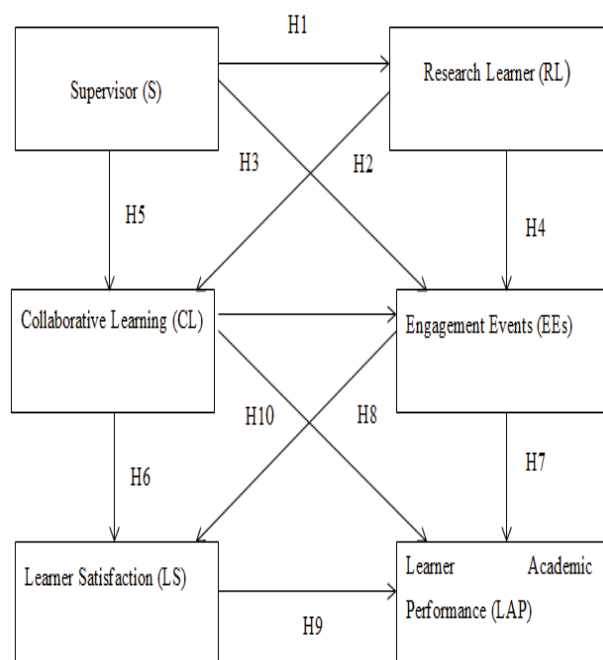


FIGURE 3 PROPOSED FRAMEWORK OF COLLABORATIVE LEARNING FOR IMPROVING DROPOUT STUDENT

V EXPECTED OUTCOME

However, an arrangement of online SM environment may assist in communicating with their peer of dropout students in solving problems or they are permitted to organizer an event in a collaborative manner. Hence, the CL and engagement event placed by instructor plays an essential role improving the intellectual skills of dropout student to obtain better in academic performance. The benefit of SM with unnecessary activity gets avoided by engaging the student with academic track may provide confidence and motivated to focus on their academic. Thus, the linkage of peer has encouraged presence of all students particularly who are dislike for discussing matter face-to-face. In future work we do collaboration work instead of individual learning using linear regression and logistic regression algorithm. Individual learning improve the intellectual skills of dropout students using data mining techniques.

VI CONCLUSION

The present framework is based on the constructive concept for adopting CL and engagement type of learning to the dropout students through SM. Principal component analysis Principal Component Analysis is the Unsupervised Learning Algorithm. PCA implements the Dimensionality Reduction technique. The goal of PCA is the removal of irrelevant features while developing a model. PCA extracts the most dependent features contributing to the output. This research majorly focused on the student who gets stressed and planned for dropping. However, the proposed method

may assist in improving the mental ability, interactive and communicating with each other without hesitation in sharing their thought. One of the interactive mode is collaboration for the students to share their knowledge and constructing their problem solving ability by interactive with instructor and the group members. Thus the overall outcome of dropout predicted student can be evaluated through hypothesis support of LS and LAP, whereas that exact identification of predicting dropout student can be easily identified by PCA with ML.

REFERENCES

- [1] Vinayak Hegde "Higher Education Student Dropout Prediction and Analysis through Educational Data Mining Proceedings of the Second International Conference on Inventive Systems and Control (ICISC 2018) IEEE Xplore
- [2] Naruhiko Shiratori "Modeling dropout behavior patterns using Bayesian Networks in Small-Scale Private University". 2017 6th IIAI International Congress on Advanced Applied Informatics
- [3] Kittinan Limsathitwon Thai-Nichi "Dropout Prediction System to Reduce Discontinue Study Rate of Information Technology Students 2018 5th International Conference on Business and Industrial Research (ICBIR), Bangkok, Thailand
- [4] Ravinder Ahuja Yash Kankane "Predicting the Probability of Student's Degree Completion by Using Different Data Mining Techniques" 2017 Fourth International Conference on Image Information Processing (ICIIP)
- [5] Carlos Márquez-Vera, Cristóbal Romero Morales, and Sebastián Ventura Soto "Predicting School Failure and Dropout by Using Data Mining Techniques
- [6] Naruhiko Shiratori "Derivation of student patterns in a preliminary dropout state and identification of measures for reducing student dropouts", 2018 7th International Congress on Advanced Applied Informatics.
- [7] Meenakshi Upreti, Vishal Kumar "Learning the student's sufferings using Social Networks" International Conference on Computing, Communication and Automation (ICCCA2017)
- [8] Alvaro Ortigosa, Rosa M. Carro, Javier Bravo-Agapito, David Lizcano, Juan Jess Alcolea, From Lab to Production: Lessons Learnt and Real Life Challenges of an Early Student-Dropout Prevention System IEEE Transaction technologies, VOL. 12, NO. 2 April 2019
- [9] Juan Pablo Salazar-Fernandez, Marcos Sepúlveda Pontificia, Jorge Muñoz-Gama Influence of Student Diversity on Educational Trajectories in Engineering High-Failure Rate Courses that Lead to Late Dropout 978-1-5386-9506-7/19/\$31.00 ©2019 IEEE
- [10] P. Perchinunno, M. Bilancia, and D. Vitale, "A Statistical Analysis of Factors Affecting Higher Education Dropouts," Soc. Indic. Res., no. 0123456789, 2019.
- [11] S. Lee and J. Y. Chung, "The machine learning-based dropout early warning system for improving the performance of dropout prediction," Appl. Sci., vol. 9, no. 15, 2019.
- [12] Martín Solís Tania Moreira Roberto González Perspectives to Predict Dropout in University Students with Machine Learning
- [13] S. C. Tsai, C. H. Chen, Y. T. Shiao, J. S. Ciou, and T. N. Wu, "Precision education with statistical learning and deep learning: a case study in Taiwan," Int. J. Educ. Technol. High. Educ., vol. 17, no. 1, 2020.
- [14] Dr.P.Ashal Predicting University Dropout through Data Analysis Proceedings of the Fourth International Conference on Trends in Electronics and Informatics (ICOEI 2020) IEEE Xplore Part Number: CFP20J32-ART; ISBN: 978-1-7281-5518-0
- [15] Marina Andreia Gomes Pereira Collaborative transmedia learning environments for the promotion of students' autonomy and motivation A study developed with students at risk of dropout and school failure 2020 15th Iberian Conference on Information Systems and Technologies (CISTI) 24 - 27 June 2020, Seville, Spain ISBN: 978-989-54659-0-3
- [16] Roderick Lottering, Robert Hans, Manoj Lall A model for the identification of students at risk of dropout at a university of technology 978-1-7281-6770-1/20/\$31.00 ©2020 IEEE
- [17] Axel Bottcher, Veronika Thumer, Tanja Hoffner Applying Data Analysis to Identify Early Indicators for Potential Risk of Dropout in CS Students 978-1-7281-0930-5/20/\$31.00 ©2020 IEEE