PREDICTION OF DROPOUT STUDENTS USING DATAMINING TECHNIQUES FOR IMPROVING THEIR INTELLECTUAL SKILLS BY COLLABORATIVE LEARNING

*Fathimath Shabana

Department of Computer Science, KAHM Unity Women's College, Manjeri, Malappuram District, Kerala, State, India,

*Corresponding Author: fathimathshabanamk@gmail.com

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 pays significant attention towards analysing the higher education and college 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

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 presumed 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.

Introduction

Dropping out students has significant implications for learners and their environment. University dropout is an issue such as, reduced enrolment, decreased college income and The State that finances the inquiries and sets up undergraduate student's financial problems as a social problem. Dropping out is influenced by both the students and institutions. In engineering education, late dropout is becoming more of an issue. 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 wellbeing are impaired by dropping out of institutes that are faced with the fact that they lack skills and expertise to fulfil their wishes. Dropouts can be linked to problems with student success. In certain cases, students are unable to finance these deficits on their own. 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 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, first to identifies an exact factor for the student dropout it can be done through dimension reduction. A dropout early warning system may assist universities or colleges in recognising students who are on the verge of dropping out and encouraging them to act. However, the predicting of children 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.

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.

Methodology

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,

Scanned with OKEN Scanner

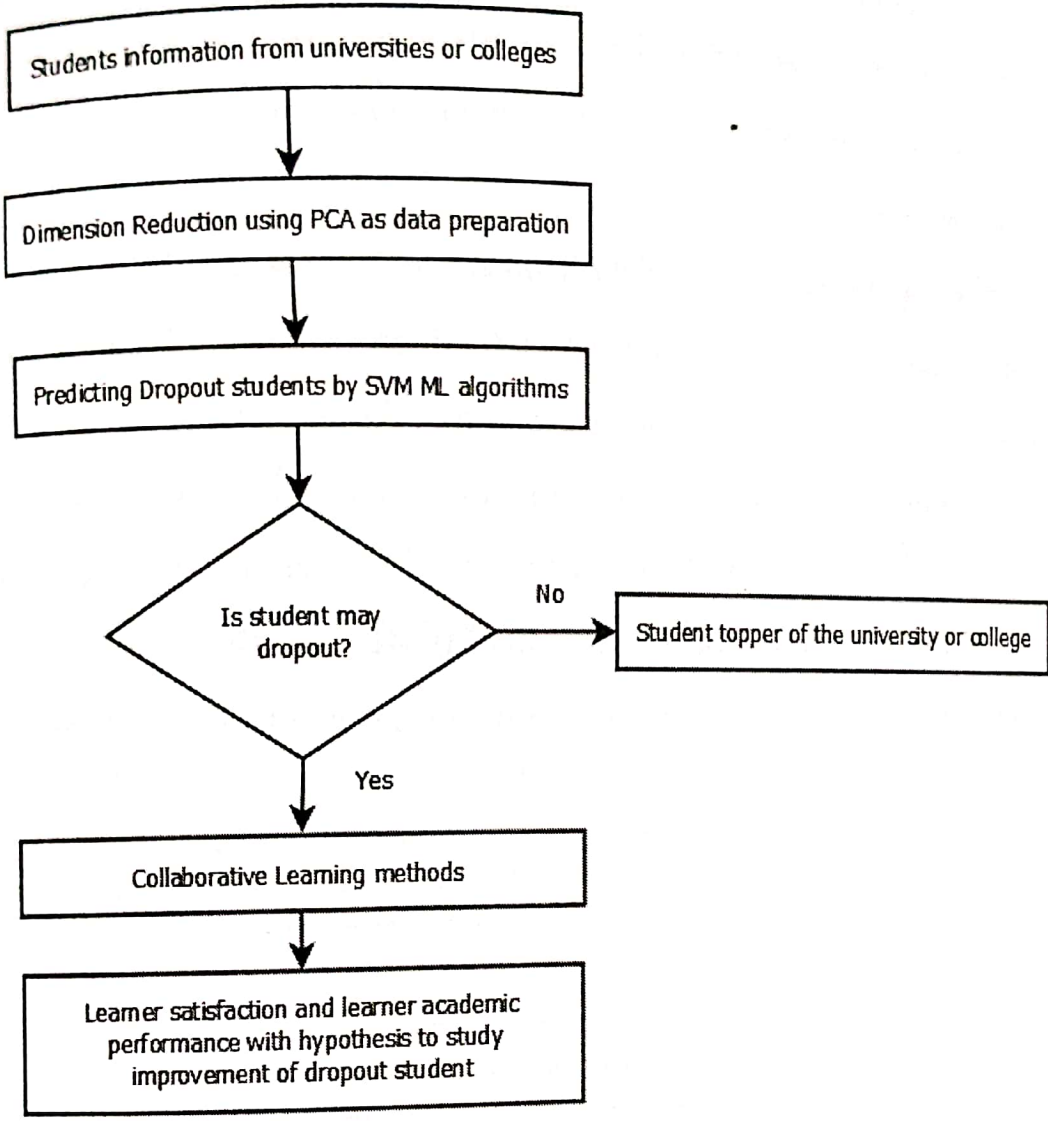respective students to avoid student's dropout rate.

Figure 1 - Overall Proposed Framework For Improving
Dropout Students

In the technique of the extraction process of the functions involved in the conversion 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 data provided by universities or colleges. It helpful to recognize the standard and expectation of the students with several features and it may be 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.

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

The input section takes data from universities or colleges 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 Support Vector Machine (SVM) 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.
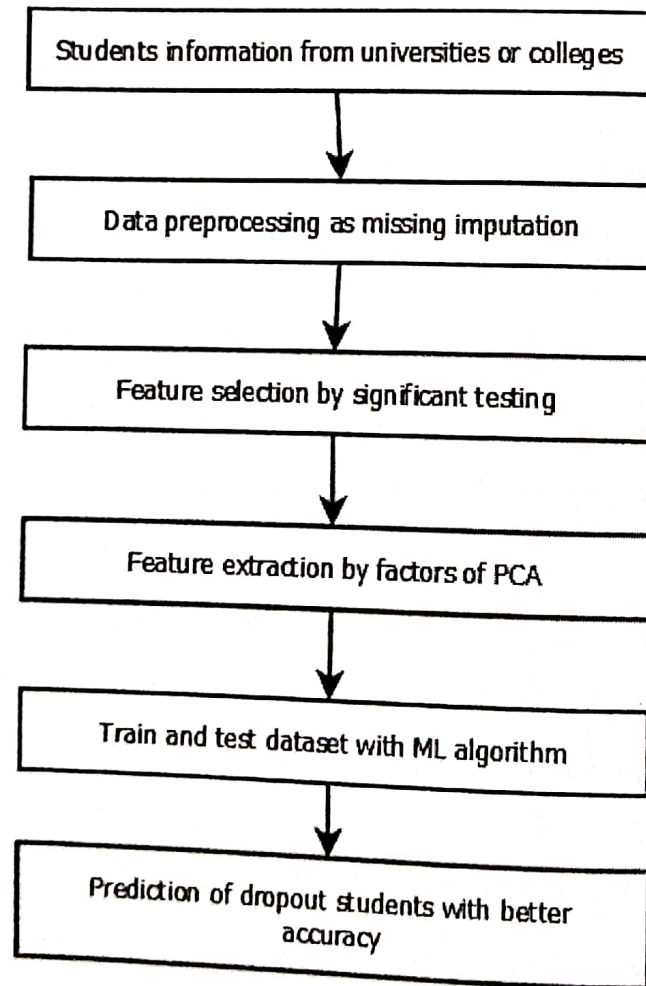
```
┌─────────────────────────────────────────────────┐
│  Students information from universities or colleges │
└─────────────────────────────────────────────────┘
                        │
                        ▼
┌─────────────────────────────────────────────────┐
│      Data preprocessing as missing imputation      │
└─────────────────────────────────────────────────┘
                        │
                        ▼
┌─────────────────────────────────────────────────┐
│       Feature selection by significant testing      │
└─────────────────────────────────────────────────┘
                        │
                        ▼
┌─────────────────────────────────────────────────┐
│         Feature extraction by factors of PCA         │
└─────────────────────────────────────────────────┘
                        │
                        ▼
┌─────────────────────────────────────────────────┐
│      Train and test dataset with ML algorithm       │
└─────────────────────────────────────────────────┘
                        │
                        ▼
┌─────────────────────────────────────────────────┐
│     Prediction of dropout students with better      │
│                     accuracy                        │
└─────────────────────────────────────────────────┘
```

Figure 2 - Improving Dropout Student Skills By Collaborative Learning

This research proposed a collaborative learning 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.

The staffs and mentors are used as Supervisors to monitor for engaging the extra class using concept of Collaborative Learning (CL) to improve Learner Academic Performance (LAP) in which Learner's Satisfaction is essential.

Hardware Environment

The selection of hardware is very important in the existence and proper working of any software. Then selection hardware, the size and capacity requirements are also important.

- Processor: Pentium Dual Core i5 or above
- Primary Memory: 4GB RAM or above
- Storage: 120 GB hard disk or above
- Display: VGA Colour Monitor
- Key Board: Windows compatible
- Mouse: Windows compatible

Software Environment

For the proposed research to work properly, it is necessary that following software are installed and running on the server / client. One of the most difficult tasks is selecting software for the research, once the research requirements is found out then we have to determine whether a particular software package fits for those research requirements.

The application requirement:

- Operating System: Windows 7 or above, Android version 5 or above
- Front end: Python, Android
- Back end: MySQL
- Server: Apache (Wamp)
- IDE: PyCharm (website), Android studio (android app)
- Framework: Flask

App Features

In this publication, we have proposed a dropout warning system, through which we can recognize the students who are on the verge of dropping out in early stage. Hence, we can provide collaborative learning tools for improve their intellectual skills.

The application's main characteristics are:

o It is a web and android combo application.

o The application automates the activities of educational institutions.

o The application detects the students who are going through the dropout situation at earliest.

o It also supports collaborative learning tools.

Conclusion

The framework is based on the constructive concept for adopting CL and engagement type of learning to the dropout students. 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 modes 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 LAP, whereas that exact identification of predicting dropout student can be easily identified by PCA with ML. Hence, through the application, can improve the intellectual skills of such students and by through we can reduce the dropout rate.

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 ... works in Small-Scale Private University". 2017 6th IIAI International Congress on Advanced Applied Informatics ... patterns using Bayesian Networks

[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 RealLife 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 Munoz-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. Asha1 "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]. oderick Lottering, Robert Hans, Manoj Lall "A model for the identification of students at risk of dropout at a university oftechnology" 978-1-7281-6770-1120/$31.00 ©2020 IEEE

[17]. M.Revathy, S. Kamalakkannan, "Collaborative learning for improving intellectual skills of dropout students using datamining techniques" 978-1-7281-9537-7/21/$31.00 ©2021 IEEE

[18]. Axel Bottcher, Veronika Thurner, Tanja H ¨ afner "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