

# Predicting Dropout on E-learning Using Machine Learning

Akmal\*

Graduated Language Program, Ahmad Dahlan University, Indonesia  
akmal@mpbi.uad.ac.id\*  
\* corresponding author

(Received: July 15, 2020; Revised: July 21 2020; Accepted: August 29, 2020; Available online: September 1, 2020)

## Abstract

After the corona virus outbreak (COVID-19) many various institutions changed how they work. From various sectors, which experienced the biggest change was the education sector. The education sector, which in general requires face-to-face interaction between teachers and students in a place, has now changed to online, which does not require that both parties be in a place. This is certainly a very big change and has an impact. In this paper we will discuss e-learning methods for drop-out prediction, based on three techniques of machine learning.

*Keywords:* Education, Machine learning, E-learning, Dropout, Data Science

## 1. Introduction

E-learning has shown tremendous growth in the last few months and has drawn growing numbers of participants. Contemporary web-based courses use the internet's resources to help and improve more successful conventional education approaches while at the same time providing more creative possibilities. In addition to helping individual students acquire information more flexibly, even without limitations of their presence in the classrooms, this type of education often benefits institutions that may have the ability to offer their educational services to even more students, thereby reducing the expenses they currently incur [1][2]. The e-learning course also draws many adult students with challenging work obligations who are seeking to combine higher education with professional level training.

While e-learning attracts a significant number of students, web-based courses have a much higher dropout rate than conventional education, sometimes exceeding the last with a rate of 10-20% as stated by [3][4]. Additional studies support the veracity of these study and predict an e-learning dropout rate of 25-40% compared to a university average of 10-20% [5]-[9]. Student retention rates are, however, one of the metrics that universities, policymakers, funding agencies for higher education, and educators regard as objective indicators, based on performance outcomes that educational institutions deliver. This paper will recognize deeper about predicting methods & technique to prevent student dropouts.

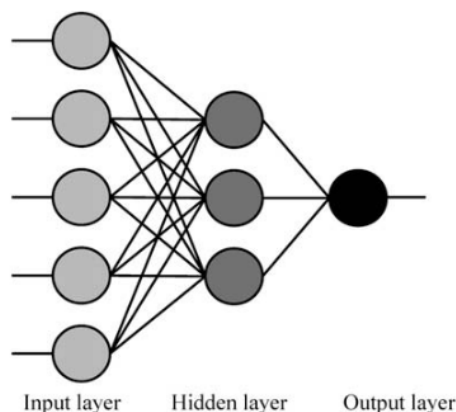
## 2. Literature Review

There have been many studies pertaining to dropouts from academic courses, mostly on online courses. but still, there is no specific definition of dropout from e-learning courses. In this paper we will describe dropout students as students that voluntarily withdraw from courses. Many logical explanations were presented to explain why the dropout rate is higher in e-learning courses. A large number of studies present disappointing results from e-learning courses on dropout rates. Based on [10], research before the internet became a major delivery vehicle for education, rates of e-learning dropouts and correspondence education estimated from 25% to 60%.

Many researchers explored the reasons associated with dropout from courses at higher institutes, while others looked at dropout from correspondence and earlier ways of distance learning. There was a strong agreement in research that dropping out is a difficult and perplexing trend, particularly in distance education. This was critical to be able to understand the factors of dropout for administrators and tutors. There have been many studies referring to dropouts from courses, both university and distance course, but no specific description of dropouts from e-learning courses was received. But when it comes to e-learning courses, this research describes students who dropout (non-completers).

## 2.1. Machine Learning Techniques in Dropout Prediction

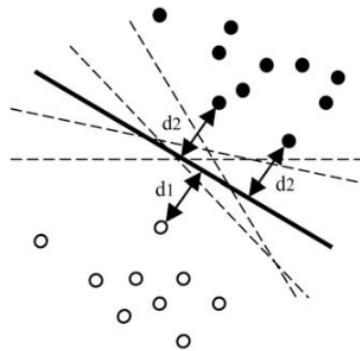
In this study, three popular machine learning techniques, Feed-forward neural networks (FFNN), Support vector machines (SVM), and the probabilistic ensemble simplified fuzzy ARTMAP (PESFAM), were used to predict students dropping out of school. This technique was successfully applied in two phases, the training and testing phase, to solve various classification and function problems. During the training process, a collection of sample data pairs (X,Y) is given to each technique, where X represents the input and Y represents the output of each pair. Y will obtain one of the following values in this sample, 0 if a student completes the class or 1 if a student cancels the course. If the technique is proven to be effective in correctly classifying most of the data in the test set, then the training is considered effective, and the technique of machine learning demonstrates generalizing skill. The following section explains briefly the features of the three methods used in machine learning.



**Figure. 1.** Feed-forward neural networks (FFNN) Structure

## 2.2. Support Vector Machines

The support vector machine technique was pioneered by Vapnik [11] and has been successfully applied to many classification problems in various literature. This technique tries to separate the two data classes using a hyperplane defined by a support vector, which is part of a data set. Through his training, the support vector machines technique looks for Optimal Separating Hyperplane (OSH), which is an optimal hyperplane that maximizes margins between two classes of training datasets.

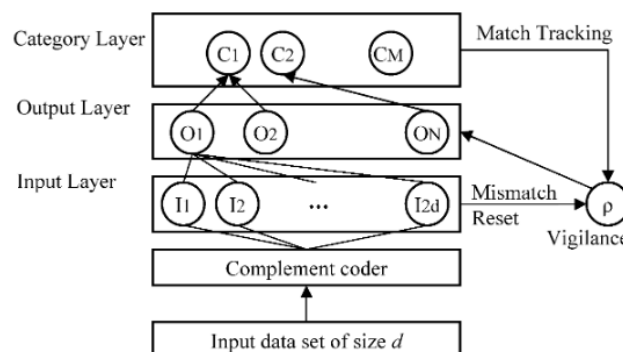


**Figure. 2.** Support vector machines (SVM) concept

Illustration above Provides 2-Dimensional instances. Several rows have been found in this example for successfully separating data into two classes. The best line, OSH, is the middle-of-the-road track. The term "margin" refers to the sum of data distance from the nearest data to the line specified as the Support Vector. Typically, since data on these real-world problems can not be linearly separated, the above techniques are extended explicitly to address inputs that can not be linearly separated at the expense of having a number of misclassifications. The SVM technique further improves its effectiveness by converting data into feature spaces with higher dimensions than inputs before attempting to separate them using a linear discriminator. Use the kernel function, lastly. There are various popular kernel functions, such as linear kernel functions, polynomial functions, and radial bases. A radial basis kernel function is used in this paper.

### 2.3. Probabilistic Ensemble Simplified Fuzzy ARTMAP

Probability ensemble simplified fuzzy ARTMAP (PESFAM) is the last machine learning technique utilized by this paper. PESFAM [13], combines a number of simplified fuzzy ARTMAP modules (SFAM) with a plurality voting strategy based on Lin, Yacoub, Burns and Simske [14] findings. ARTMAP technique is the innovation of the Adaptive Resonance Theory (ART), introduced into self-regulating neural networks by Carpenter and Grossberg [15]. Kasuba. [16] proposed the SFAM technique to reduce the computational overhead and architectural redundancy of the original Fuzzy ARTMAP network. The SFAM network itself is composed of three layers, as shown in the fig. 3, This is the layer of input , output and categories. Each input node in the output layer will be connected to each node and each output layer node will point to a specific category node that correlates to one class of existing classification problems. Firstly, all data entered in the SFAM network is transformed into a range of values (0.1).

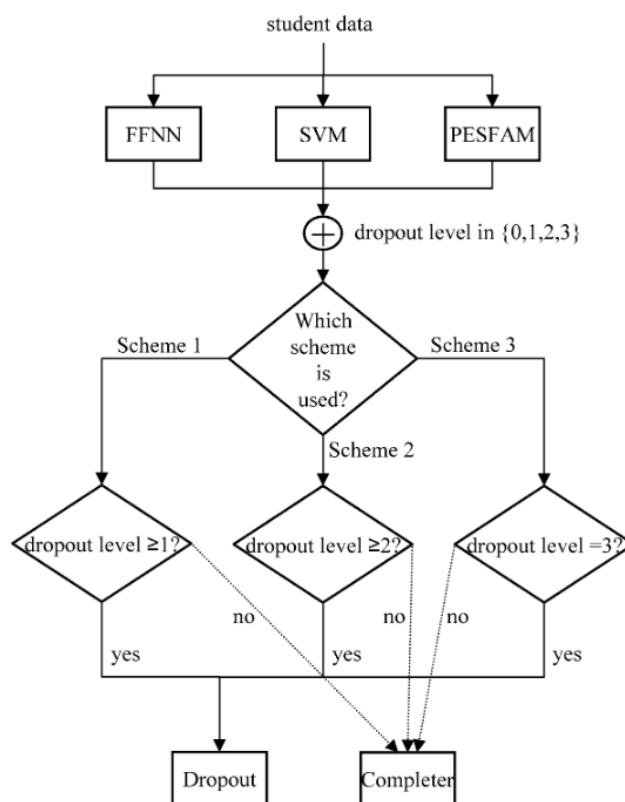


**Figure. 3.** Simplified Fuzzy ARTMAP Structure

### 3. Research Method

Because there is a possibility that one machine learning technique might fail to detect some students who dropped out, while others might succeed, the results of combining 3 networks will also be examined.

First, the results from the three networks are added up to calculate the dropout rate of each student in the dataset. The results of each network are binary, where '1' represents the case a student has been categorized as a dropout and '0' indicates that this student has been identified as a complement (or not a dropout). Therefore, a student's dropout rate will receive four different integer values ranging from 0 (complement) to 3 (all networks have categorized this student as dropout). After a student's dropout rate has been calculated, these students can be categorized as dropouts or who finish using three different decision schemes defined as follows:



**Figure. 4.** Machine learning techniques using decision schemes.

Decision Scheme 1: A student is considered to have dropped out of school if at least one technique had classified this student as such. In other words, a student is categorized as a dropout if the dropout rate is greater than or equal to '1', otherwise, it is classified into a supplementary category.

Decision Scheme 2: If this scheme is selected, a student can be considered a dropout if at least two techniques indicate this student is a dropout, when the student dropout rate is calculated to be greater than or equal to '2'.

Decision Scheme 3: In this scheme, a student is considered a dropout if all the techniques agree that this student is a dropout. In this case, students are identified as dropouts only if their prediction level is equal to '3'.

### 4. Discussion

This research takes advantage of the entirely computerized aspects of e-learning, using data gathered dynamically from the LMS database. Each type of data includes very detailed student records to instructors, which include scores

of students in multiple choice assessments and projects, the willingness of students to reach assignment deadlines, and their performance during the course. Each of these aspects have been shown to be important to the process of dropout prediction, since they represent specific situations for students. For instance, if a student can sign it into the platform and frequently participate in the forums, this student could get low grades and possibly drop the course. The reverse scenario can also be seen in the data collection, which illustrates the situation of a student who does not engage actively in the course, indicates low levels of participation, and eventually ended up completes the course with good grades. As illustrated in the data set, the student may indicate progress by performing various choice tests and applying projects, but the student having problems in fulfilling the requirements of the course, thereby slowing the application of his projects and eventually dropping the course. The use of these specific student data also leads to more precise analyses of dropout.

But as far as machine learning methods are worried, they have completed well on all assessment parameters for all e-learning courses. Machine learning methods have all the strength of being data-driven rather than model-driven. They do not follow an specific data relationship model as do model-based linear or nonlinear methods. Alternatively, the structure of the model and the process parameters they are using are extracted from the problem's real data collection. Additionally, in this study we used three decision schemes that combine the results of the machine learning techniques to overcome individual physical problems. The results indicated that Decision Scheme 1 appears to become the most effective method for improving and sustaining high precision, flexibility and quality outcomes in predicting at-risk students, since this scheme yielded the best outcomes for both courses in nearly all courses.

## 5. Conclusion

This paper presents methods for the prediction of early and accurate dropout of students in e-learning courses. This refers to a detailed student education log (history), obtained from the learning system that organizes elearning, to make estimates dynamically and adjust them to students progress in the course.

This method uses three techniques for machine learning, namely FFNN or Feed-forward Neural Networks, SVM or Support Vector Machines, and PESFAM or probabilistic ensemble simplified Fuzzy ARTMAP. In order to resolve potential inaccuracies in a methodology in the detection of dropouts, we used three separate judgment schemes to aggregate their figures.

This proposed method is intended to encourage the easy recognition of students at risk by educators or teachers and to concentrate on their needs, thereby increasing the degree of engagement in e-learning. An example of another problem that still needs to be explored is the ability to produce improved outcomes using different characteristics of students. Furthermore, more methods can also be explored for predicting dropouts in terms of individual and joint performance. Finally, what needs to be investigated in the future is the potential for incorporating proposed methods into retention strategies for educational institutions to help increase the retention of higher students.

## References

- [1] A. Cohen and R. Nachmias, "A quantitative cost effectiveness model for Web-supported academic instruction," *Internet High. Educ.*, vol. 9, no. 2, pp. 81–90, 2006, doi: 10.1016/j.iheduc.2006.03.007.
- [2] J. Mendling, G. Neumann, A. Pinterits, and B. Simon, "Revenue Models for E-Learning at Universities," *Wirtschaftsinformatik 2005*, pp. 827–846, 2005, doi: 10.1007/3-7908-1624-8\_43.
- [3] A. Joshi, P. Desai, and P. Tewari, "Learning Analytics framework for measuring students' performance and teachers' involvement through problem based learning in engineering education.," *Procedia Comput. Sci.*, vol. 172, pp. 954–959, 2020, doi: 10.1016/j.procs.2020.05.138.

- 
- [4] J. Littenberg-Tobias and J. Reich, "Evaluating access, quality, and equity in online learning: A case study of a MOOC-based blended professional degree program," *Internet High. Educ.*, vol. 47, p. 100759, 2020, doi: 10.1016/j.iheduc.2020.100759.
  - [5] V. Carter, "Do media influence learning? Revisiting the debate in the context of distance education," *Open Learn.*, vol. 11, no. 1, pp. 31–40, 1996, doi: 10.1080/0268051960110104.
  - [6] W. Doherty, "An analysis of multiple factors affecting retention in Web-based community college courses," *Internet High. Educ.*, vol. 9, no. 4, pp. 245–255, 2006, doi: 10.1016/j.iheduc.2006.08.004.
  - [7] R. Damadian, "Abnormal phosphorus metabolism in a potassium transport mutant of *Escherichia coli*," *BBA - Biomembr.*, vol. 135, no. 2, pp. 378–380, 1967, doi: 10.1016/0005-2736(67)90137-X.
  - [8] B. Holder, "An investigation of hope, academics, environment, and motivation as predictors of persistence in higher education online programs," *Internet High. Educ.*, vol. 10, no. 4, pp. 245–260, 2007, doi: 10.1016/j.iheduc.2007.08.002.
  - [9] M. Xenos, "Prediction and assessment of student behaviour in open and distance education in computers using Bayesian networks," *Comput. Educ.*, vol. 43, no. 4, pp. 345–359, 2004, doi: 10.1016/j.compedu.2003.09.005.
  - [10] J. Culpeper and M. Gillings, "Pragmatics: Data trends," *J. Pragmat.*, vol. 145, pp. 4–14, 2019, doi: 10.1016/j.pragma.2019.01.004.
  - [11] M. F. Jefferson, N. Pendleton, S. Lucas, M. A. Horan, and L. Tarassenko, "Neural networks," *Lancet*, vol. 346, no. 8991–8892, p. 1712, 1995, doi: 10.1016/S0140-6736(95)92880-4.
  - [12] N. H. Farhat, "Photonit neural networks and learning mathines the role of electron-trapping materials," *IEEE Expert. Syst. their Appl.*, vol. 7, no. 5, pp. 63–72, 1992, doi: 10.1109/64.163674.
  - [13] C. K. Loo and M. V. C. Rao, "Accurate and reliable diagnosis and classification using probabilistic ensemble simplified fuzzy ARTMAP," *IEEE Trans. Knowl. Data Eng.*, vol. 17, no. 11, pp. 1589–1593, 2005, doi: 10.1109/TKDE.2005.173.
  - [14] X. Lin, S. Yacoub, J. Burns, and S. Simske, "Performance analysis of pattern classifier combination by plurality voting," *Pattern Recognit. Lett.*, vol. 24, no. 12, pp. 1959–1969, 2003, doi: 10.1016/S0167-8655(03)00035-7.
  - [15] G. A. Carpenter and S. Grossberg, "A Massively Parallel Architecture for a Self-Organizing Neural Pattern Recognition Machine C . Stability-Plasticity Dilemma : Multiple Interacting Memory Systems The properties of plasticity and stability are intimately related . An adequate," *Pattern Recognit.*, vol. 115, pp. 54–115, 1987.
  - [16] M. A. de Oliveira and D. J. Inman, "Performance analysis of simplified Fuzzy ARTMAP and Probabilistic Neural Networks for identifying structural damage growth," *Appl. Soft Comput. J.*, vol. 52, pp. 53–63, 2017, doi: 10.1016/j.asoc.2016.12.020.