YASAR UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

MASTER THESIS

ANALYSIS OF THE OUTCOMES OF YASAR UNIVERSITY E-LEARNING SYSTEM

Erhan KINAY

Thesis Advisor: Assist. Prof. İbrahim ZINCIR

Department of Computer Engineering

Presentation Date:

Bornova - IZMIR

2016

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Dr. Ibrahim Zincir

(Supervisor)

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Dr. Mehmet S. Ünlütürk

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Dr. Samsun M. Başarıcı

Professor Cüneyt GÜZELİŞ

Director of the Graduate School

ABSTRACT

ANALYSIS OF THE OUTCOMES OF YASAR UNIVERSITY E-LEARNING SYSTEM

KINAY, Erhan

MSc in Computer Engineering

Supervisor: Assist. Prof. İbrahim ZİNCİR

Nov 2016,

Data mining is the computational process of analyzing large data sets, discovering the patterns of data groups and reaching the targeted data through processing operation. The process of extraction of the important information through the large amount of data and execution of that data is called "Data Mining". In this research, our aim is to predict a student's grade via implementing machine learning techniques of data mining over Yasar University UFND Dataset. The dataset in question is collected between 2012-2013 Fall Term and 2015-2016 Fall Term and made from the log records of website activity of 10 courses over 5 terms (Yasar University e-learning website https://e.yasar.edu.tr). These logs inherit end user information regarding how and when the website in question is used and studied.

By implementing these data mining techniques, first the proposed framework analyzed the collected data and then tried to successfully guess whether a student will pass or fail from the course in question. In order to achieve this goal, the proposed framework trained the system with 31 different classification algorithms and then a final algorithm was selected for each course, for each term and for the combination of the data of each course over 5 terms. These algorithms are Naive Bayes, Ripper, J48, Bayes Net, Adaboost, AdTree, Attribute Selected Classifier, Bagging, and Classification via Regression, Conjunctives Rules, CV Parameter

Selection, Decision Table, DTNB, END, Filtered Classifier, Grading, IB1, Ibk, K*, Logistic, LogiBoost, LWL, MultiBoostAB, Multiclass Classifier, Multi Scheme, Multilayer Perceptron, SMO, Voted Perceptron, Random Forest, and ZeroR. At the end, all these results are analyzed and then evaluated to achieve the goal of effective prediction of a student's success.

Keywords: data mining, machine learning, e-Learning, classification, C4.5, Naïve Bayes, Ripper, Bayes Net, Adaboost, AD Tree, Attribute Selected Classifier, Bagging, Classification Via Regression, Conjunctive Rule, CV Parameter Selection, Decision Table, DTNB, END, Filtered Classifier, Grading, ZeroR, Random Forest, MultiBoostAB, LogiBoost, Multi Scheme, Multilayer Perceptron, Multiclass Classifier, Voted Perceptron

ÖZET

YASAR ÜNİVERSİTESİ UZAKTAN EĞİTİM SİSTEMİ ANALİZİ

Erhan KINAY

Yüksek Lisans Tezi, Bilgisayar Mühendisliği

Tez Danışmanı: Yrd. Doç. Dr İbrahim ZİNCİR

Kasım 2016

Büyük miktardaki veriler içerisinden önemli olanları bulup çıkarmaya yönelik olarak veriyi işlemeye veri madenciliği denir. Veriler üzerinde çözümleme yapmak amacıyla, veriyi çözümleyip bilgiye ulaşabilmek için veri madenciliği kullanılır. Bu araştırmadaki amacımız, Yaşar Üniversitesi UFND veri kümesi üzerinden Veri Madenciliği ve Makine Öğrenmesi uygulayarak bir öğrencinin notunu tahmin etmek. Bahsedilen veri kümesi 2013-2014 güz dönemi ve 2015-2016 güz dönemi arasında elde edilip web sitesi faaliyetleri olarak log kayıtlarıyla 5 dönem ve 10 ders üzerinden gerçekleştirildi (Yaşar Üniversitesi e-öğrenme web sitesi https://e.yasar.edu.tr). Bahsi geçen bu loglar web sitesinin ne zaman ve nasıl kullanıldığına ve bu web sitesi üzerinde çalışan son kullanıcı bilgilerini içerir. Bu Veri madenciliği teknikleri uygulanırken ilk olarak önerilen system elde edilen bilgiyi inceliyor ve ardından öğrencinin dersten kalıp kalmayacağı ile ilgili doğru tahminde bulunuyor. Bu sonuca ulaşabilmek için önerilen çerçeve uygulama önceden sistemi 31 farklı algoritmik sınıflandırma ile çalıştırmış ve sonar her ders, dönem ve 5 dönem için tüm derslerin dahil edildiği versiyon çalıştırılarak bir algoritma seçilmiş. Analiz esnasında Weka'nın Naïve Bayes, Ripper, J48, Bayes Net, Adaboost, AdTree, Attribute Selected Classifier, Bagging, Classification Via Regression, Conjunctive Rule, CV Parameter Selection, Decision Table, DTNB, END, Filtered Classifier, Grading, IB1, Ibk, K*, Logistic, Logiboost, LWL, MultiBoostAB, Multiclass

Classifier, Multi Scheme, Multilayer Perceptron, SMO, Voted Perceptron, Random Forest and ZeroR sınıflandırma algoritmaları kullanıldı. Sonuçlar üzerinde hangi algoritmanın daha başarılı olacağını birden farklı versiyonlar ile tespit etmek mümkünüdür. Bu tez, bu konuları, uygulamalarını ve sonuçlarını içeren 6 bölümden oluşmaktadır.

Anahtar Sözcükler: Veri Madenciliği, Elektronik Eğitim, Sınıflandırma, J48, Naïve Bayes, JRip, Bayes Net, AdaboostM1, AD Tree, Attribute Selected Classifier, Bagging, Classification via Regression, Conjunctive Rule, CV Parameter Selection, Decision Table, DTNB, END, Filtered Classifier, Grading, ZeroR, Random Forest, MultiBoostAB, LogiBoost, Multi Scheme, Multilayer Perceptron, Multiclass Classifier, Voted Perceptron, K*

ACKNOWLEDGEMENTS

First, I would like to thank to assist. Prof. Ibrahim Zincir, my supervisor. Without his support and patience, it would not have been possible for me to finish this master degree.

Then, I would also like to thank to all the academic staff of Department of Computer Engineering for their assistance on my research.

Finally, I would like to thank my parents for their support and love.

Erhan KINAY

Izmir, 2016

TEXT OF OATH

I declare and honestly confirm that my study, titled "ANALYSIS OF THE OUTCOMES OF YASAR UNIVERSITY E-LEARNING SYSTEM" and presented as a Master's Thesis, has been written without applying to any assistance inconsistent with scientific ethics and traditions, that all sources from which I have benefited are listed in the bibliography, and that I have benefited from these sources by means of making references

Context

1. Introduction

The aim of this thesis is to evaluate and to analyze the classification results of Yasar University E-Learning System (YES) using; Naïve Bayes, Ripper, J48(c4.5), Bayes Net, AdaboostM1, AdTree, Attribute Selected Classifier, Bagging, Classification Via Regression, Conjunctive Rules, CV Parameter Selection, Decision Table, DTNB, END, Filtered Classifier, Grading, IB1, Ibk, Kstar, Logistic, LogiBoost, LWL, MultiBoostAb AB, Multiclass Classifier, Multi Scheme, Naïve Bayes Simple, Multilayer Perceptron, SMO, Voted Perceptron, Random Forest and ZeroR machine learning algorithms. In this research, these algorithms are implemented in order to identify common approaches within the context on data mining via WEKA 3.7 data mining software tool. WEKA software was developed at the University of Waikato in New Zealand. The proposed framework inherits a dataset that consist information regarding ten different courses over five terms that are described in Section 2. In Section 3 previous work is discussed. Following this, in Section 4; how the dataset is preprocessed and converted into arff files are explained (so that it was possible to be implemented via WEKA data mining software tool). WEKA classifiers are implemented separately on full arff files via Net Beans. After completing all the implementations, the results are evaluated according to the error rates. Then, these results are presented and evaluated in Section 5. Finally, conclusion and future work are discussed.

2. Background

A) Yasar University e-Learning System

This study is performed on Yasar University e-learning dataset. It is a student database collected at Yasar University E-Learning System; that derives from UFND courses (University Foundation Courses) that have been created with the goal of helping students to build a strong base for their education and life. At Yasar University, there is a research center for open and distance learning. This center offers a range of distance learning courses and to develop university owned learning materials. There are four departments for e-learning development; providing instruction in design, scripting, production and delivery of e-learning content. The course materials in question are delivered via the web system in question [Yasar University].

1) UFND10- Human Sciences

Human Sciences has been devised as a product of Yasar University's institutional scientific identity assertions. It envisages in this assertion that academic and administrative staff, student graduates, all adopts these 'identity values and awareness of human behavior'. The Human Sciences Couse provides the student with the necessary skills to appreciate behavioral sciences (psychology, sociology, social psychology, anthropology, logic, philosophy, semiology, semantics, epistemology, history and philosophy of science etc.) with their basic assertions, principles and philosophies [Yasar University].

2) UFND20A- Research Culture

Research and Methodology has been identified as a component of Yasar University's institutional scientific identity assertions. It envisages in this assertion that academic and administrative staff, students and graduates, all adopts these 'researcher identity values'. The Research and methodology Course provides students with the necessary skills to undertake research in line with scientific principles. This relates to the accurate production of scientific data, verifying and testing the credibility of information, developing and implementing impeccable working practices, testing hypothesis and assumptions, observing and questioning results objectively, following the guidelines to evaluate a write up the findings, as well as fostering a desire to succeed [Yasar University].

3) UFND30A-Design Culture

Design Culture has been identified as a component of Yasar University's institutional scientific identity assertions. It envisages in this assertion that academic and administrative staff, students and graduates, all adopts these 'designer identity values'. The Design Culture Course provides students with techniques for reasoning, quality and orderly presenting thinking skills, creativity, the development of renewable and sustainable projects, the implementation of entrepreneurial techniques and Technologies, as well as the skills to use this knowledge to prepare recommendation reports, which can form the basis of further project applications [Yasar University].

4) UFND40A-Aesthetic Culture

Aesthetics has been devised as a product of Yasar University's institutional scientific identity assertions. It envisages in this assertion that academic and administrative staff, students and graduates, all adopts these values for being 'artistically aware'. The Aesthetics Course equips students with the skills to differentiate between high and low quality work in the art sector; it gives a feel for artistic values; it raises awareness about creating art and the appreciation of an art piece; it allows students to understand the difference between the emotional and the reasoned connection with an art piece and how they are affected by it [Yasar University].

5) UFND50A-Ethics Culture

Ethics has been devised as a product of Yasar University's institutional scientific identity assertions. It envisages in this assertion that academic and administrative staff, students and graduates, all adopts these identity values which reflect 'an ethical consciousness of the environment'. The Ethics Course provides students with the awareness, as well as psychological and behavioral skills, to proactively strive for a quality, secure and structured lifestyle, in which students can take responsibility for protecting social, cultural and environmental values [Yasar University].

6) UFND60A-Project Culture

Project Management has also been identified as a component of Yasar University's institutional scientific identity assertions. It envisages in this assertion that academic and administrative staff, students and graduates, all adopts these 'project identity values'. The Project Management Course provides students with the necessary skills to personally take responsibility and deliver on determining the solution to a problem, based on the developed Project proposal. This course looks to implement the prepared solution in the most efficient, effective and productive way possible, creating new resource rather than repeating existing sources in the transformation process [Yasar University].

B) Data Mining

Data mining explores large, high-dimensional, multi-type data sets that have meaningful structure or patterns with the help of statistical and computational methodologies. Data mining is a practical topic and involves learning in a practical, not a theoretical, sense. Data mining is interested in techniques for finding and describing structural patterns in data as a tool for helping to explain that data and make predictions from it. Data mining is primarily used today by companies with a strong consumer focus - retail, financial, communication, and marketing organizations. With data mining, a retailer could use point-of-sale records of customer purchases to send targeted promotions based on an individual's purchase history. By mining demographic data from comment or warranty cards, the retailer could develop products and promotions to appeal to specific customer segments.

With data mining, a retailer could use point-of-sale records of customer purchases to send targeted promotions based on an individual's purchase history. By mining demographic data from comment or warranty cards, the retailer could develop products and promotions to appeal to specific customer segments.

As it is explained by Alpaydin et al (2004), machine learning is programming computers to optimize a performance criterion using example data or past experience. The model that is generated by machine learning algorithms may then be used to predict the future, or to gain descriptive knowledge from data, or both. Machine learning uses the theory of statistics in building mathematical models, since the core task is to make interpretation from a sample. The role of computer science is twofold: First, in training, efficient algorithms are needed to solve the optimization problem, as well as to store and process the massive amount of data we generally have. Second, once a model is learned, its representation and algorithmic solution for inference needs to be efficient as well. In certain applications, the efficiency of the learning or inference algorithm, namely, its space and time complexity may be as important as its predictive accuracy.

While preparing this research, machine learning is used with data mining. All machine learning algorithms described in Section 4.

3. Previous Work

The prediction of personalized grade prediction was studied by Yannick Meier, Jie Xu, Onur Atan and Mihaela van der Schaar in University of California. They were working on Graduate digital signing processing course. According to research, approximately there are 700 students which are used for prediction. They have 7 homework assignments, one in-class midterm exam taking place after the third homework 7 and the final exam. . They improved their algorithm for predicting data and generated students GPA or Demographic data.

Explaining student grade predicted by a neural network was studied by T. D. Gedeon and H. S. Turner in University of New South Wales. In their study, Undergraduate Computer Science subject is using for prediction. The experiment was performed 153 patterns. The goal of that research is predict the final mark. According to neural network algorithm they had 94% correct output of cases.

The Predicting GPA using Educational Data Mining was studied by Mahdi Nasir, Behrouz Minaei and Fereydoon Vafai in Iran University. In their study, main goal predict GPA of graduated students which use e-learning system and online course.

4. Algorithms and Classification

4.1. Naive Bayes Classifier

Naïve Bayes is a classification algorithm. Naïve Bayes algorithm aims to locate category of data presented to the system with a series of calculations according to defined probability principles. Proportionately processed data is submitted to the system in Naïve Bayes theorem that inherits basic probability principles. Figure 1 describes Naïve Bayes algorithm.

Likelihood

Class Prior Probability

$$P(c \mid x) = \frac{P(x \mid c)P(c)}{P(x)}$$

Posterior Probability

Predictor Prior Probability

$$P(c \mid \mathrm{X}) = P(x_1 \mid c) \times P(x_2 \mid c) \times \cdots \times P(x_n \mid c) \times P(c)$$

Figure 1 Naïve Bayes Algorithm

4.2. Bayes Net Algorithm

This method is used to state for data modeling and data migration. It is mentioned as Bayes Network or Belief Network. The characteristics of these networks are to be statistical networks, and they are selected according to statistical decisions of edges that facilitate migration between nodes. Bayes Network is a directed acyclic network.

Bayes Rule:

For any two events, A and B,

P (B|A) = p (A|B) x p (B) / p (a)

4.3. JRIP (Ripper) Algorithm

Ripper is a basic and popular algorithm. Classes are examined in increasing size and an initial set of rules for the class is generated using incremental reduced error. JRip (RIPPER) proceeds by treating all the examples of a particular judgment in the training data as a class, and finding a set of rules that cover all the members of that class. Thereafter it proceeds to the next class and does the same, repeating this until all classes have been covered [Aharwal et al].

Advantages of JRip:

- Easy to interpret
- Easy to generate
- Can Classify new instances rapidly

4.4. ZeroR Algorithm

ZeroR algorithm can be qualified as the simplest algorithm method. This method chooses the dominant classification in the dataset accordingly. The following example is the best definition for ZeroR algorithm. Figure 2 shows ZeroR algorithm.

Construct a frequency table for the target and select it most frequent value.

Number of Yes =9

Number of No=5

Dataset with an accuracy of 0.64

Predictors | Target

Outlook	Temp	Humidity	Windy	Play Golf
Rainy	Hot	High	False	No
Rainy	Hot	High	True	No
Overcast	Hot	High	False	Yes
Sunny	Mild	High	False	Yes
Sunny	Cool	Normal	False	Yes
Sunny	Cool	Normal	True	No
Overcast	Cool	Normal	True	Yes
Rainy	Mild	High	False	No
Rainy	Cool	Normal	False	Yes
Sunny	Mild	Normal	False	Yes
Rainy	Mild	Normal	True	Yes
Overcast	Mild	High	True	Yes
Overcast	Hot	Normal	False	Yes
Sunny	Mild	High	True	No

Figure 2 ZeroR Algorithm

4.5. Decision Tree

Decision Tree is a very popular machine learning algorithm. The decision tree algorithm based on the division of input data to groups repeatedly with assistance of clustering algorithm. Clustering transaction continues deeply until all group members have same category label. Leaned info is modeled on a tree and interior nodes of this tree state one input each. Figure 3 explains Decision Tree algorithm.

Decision tree has three types node:

- Decision nodes
- Leaf nodes
- Root node

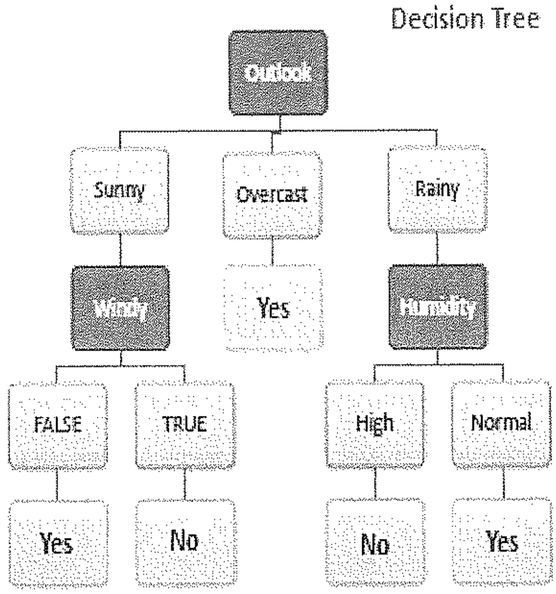

Figure 3 Decision Tree Algorithm

4.6. J48 Algorithm

J48 algorithm generally refers as C4.5 decision tree algorithm. This is the applicable form of ID3 algorithm over a dataset that contains numerical features. It contains thresholding methods which ensure to convert numerical features to categorical status. Main logic is that all values at numerical feature vector are handled as dual as averages of them are tried as threshold. Threshold value is selected if info recovery is the best with which threshold value and feature is categorized ensured according to selection threshold. Figure 4 describes J48 algorithm.

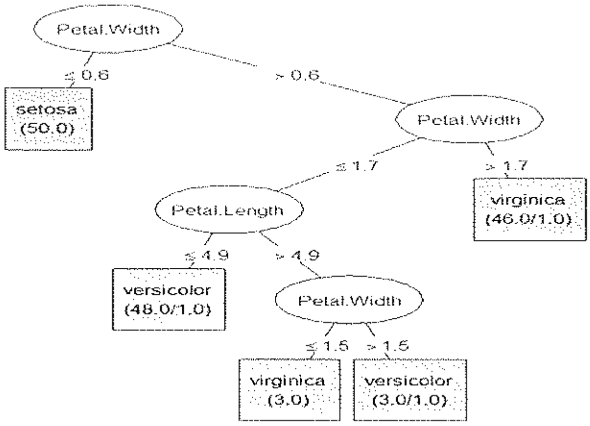

Figure 4 J48 Algorithm

4.7. Bagging Algorithm

The Bagging Algorithm suggested by Breiman focuses on training to generate new methods. The Bagging algorithm divides datasets randomly into N number of train samples and N number of test samples. In this situation, while some of the train samples are to be in only one train cluster, some of them are to be in more than one. This problem is overcome by processing majority vote.

4.8. Boosting

The Boosting algorithm trains the samples with base learners. The decisions are made by implementing clusters generated via success of their weights. Basic learner decisions are associated with success on own education clusters. Classification algorithms are used on using a data cluster and each classification algorithm has weighted according to resulting success of value.

4.9. Random Forest

Random forest is a collective classification algorithm which is developed by Breiman and Cutler. Basic learners are train samples which generated with Bagging algorithm but subset of features are selected randomly instead of using all of features. In this way the production time of decision trees is reduced and differences of trees' decisions are increased with new randomness. The results of basic learners are decided with majority voting again.

4.10. Voted Perceptron

Voted perceptron algorithm is developed by Freund and Schapire. According to Freund and Schapire, voted perceptron algorithm is best for indissoluble data. They suggest this algorithm for inseparable data. While training how method is applied to test examples. Figure 5 present voted perceptron algorithm.

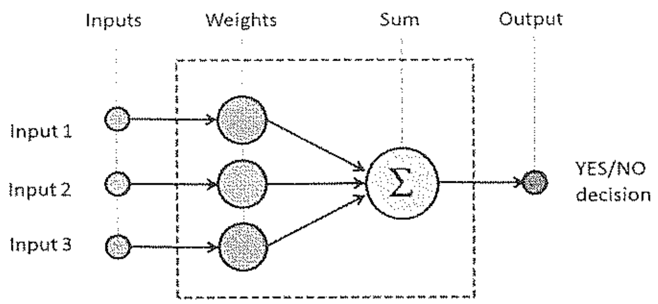

Figure 5 Voted Perceptron Algorithm

4.11. SMO

Sequential Minimal Optimization solves the problem of Support Vector Machine (SVM) expeditiously without being any extra matrix archiving and using numerical optimization steps. This application changes all missing values with new ones globally and transforms nominal attributes to binary ones. Also all features normalize to predefined values.

4.12. Multilayer Perceptron

Multilayer Perceptron modeling consists of one input layer, one or more hidden layers and one output layer. All of the instruction nodes at one layer depend on all of the transaction nodes a higher layer. The flow of information is forward, and there is no feedback. Because of that it is referred as feed-forward neural networks.

The main aim of multilayer perceptron is to minimize the error between pending output of network and generated output. During training, an assignment of network generates an output for each input that corresponds to this input. Samples are applied to the input layer, and then are manipulated in the hidden layer and finally the output is acquired from the output layer. According to the training algorithm, network weights are modified until the error between the expected output and the actual output is calculated as minimum.

4.13. Multiclass Classification

WEKA supplies parent level classifier in order to approach multiclass data clusters with binary-class classifiers. Additionally this classifier allows being applied error correction exit codes.

4.14. MultiBoostAB

MultiBoostAB algorithm is proposed by Webb in order to support classifiers via using multiboosting algorithm. Multiboosting can be seen as combination of Adaboost and Wagging algorithms. It utilizes standard variance reduction features of Adaboost with superior variance reduction features of Wagging. Parallel working provides more advantage according to Adaboost.

4.15. LWL Algorithm

Atkenson indicate the necessity of Locally Weighted Learning (LWL) algorithm. LWL alleviates independence hypothesis and learn on estimation time. The main advantage of this method compared to other techniques, it is conceptual and it can be calculated very simply as it prefers locally weighted distances [Atkenson et al].

4.16. LogiBoost Algorithm

LogiBoost algorithm inherits Boosting approach and works with applying classification technique in sequence to weighted version of train data and majority voting. Boosting can be seen as an approach to additive logistic model via getting likelihood criteria for binary class classification. This technique implements regression schedule like base learner to approach multiclass problems.

4.17. Logistic Regression

Logistic Regression algorithm was developed by statistician David Cox. The binary logistic model is used to estimate the probability of a binary response based on one or more predictor variables. Logistic regression measures the relationship between the categorical dependent variable and one or more independent variables by estimating probabilities using a logistic function, which is the cumulative logistic distribution. Logistic regression can be binomial, ordinal or multinomial. Binomial or binary logistic regression deals with situations in which the observed outcome for a dependent variable can have only two possible types (for example, "dead" vs. "alive" or "win" vs. "loss").

4.18. K* Algorithm

K* is a sample base classification algorithm. It uses similarity function between train and test sample but differs from other sample base learners with using entropy base distance function.

4.19. IBK Algorithm

IBK algorithm is a derivative of k-Nearest Neighbors algorithm. This algorithm is used for classification and selects a suitable value of K of base neighbors via weighted distance verification.

At the first time, each instance of observation is evaluated as an assemblage. After that, each assemblage joins another to generate new assemblages which in return enables to calculate and then to minimize the distances between observations.

4.20. AdaboostM1

Adaboost algorithm was developed by Freud and Schapire. Boosting and Bagging methods were implemented together with different classifiers and performances of these two methods were evaluated with character recognition and closest neighbor techniques.

4.21. IB1 Algorithm

IB1 algorithm implements closest neighbor classifier technique that uses standardized Euclidean distance in order to find the nearest sample via analyzing the past experience. If more than one sample has the same minimum distance with the previous sample than the first found is used. Euclidean distance is calculated with adjusting the weights according to the ranges defined through the samples of the learning cluster. To calculate the nearest distance at first k will be taken as equal to 1.

4.22. Grading Algorithm

As a Seewald and Fürnkranz, In Grading, the inputs to the Meta learner are base-level predictions that have been marked (i.e. "graded") as correct or incorrect. For each base classifier, Meta learner is learned that predicts when the base classifier will err. Just as stacking may be viewed a generalization of voting, grading generalizes selection by cross-validation. [Seewald et al]

4.23. Alternating Decision Tree (AdTree)

Alternative decision tree consists of decision nodes and estimation nodes. While decision nodes indicate a verb result, estimation nodes contain only one number. Alternative decision trees require estimation nodes as both root and leaf. Classification for one record is predicted when both estimation nodes and decision nodes are calculated true.

4.24. Conjunctive Rule

This class implements a single conjunctive rule learner that can predict for numeric and nominal class labels.

A rule consists of antecedents "AND" together and the consequent (class value) for the classification/regression. In this case, the consequent is the distribution of the *available classes (or mean for a numeric value) in the dataset. If the test instance is not covered by this rule, then it's predicted using the default class distributions/value of the data not covered by the rule in the training data [Xu et al].*

4.25. DTNB

The algorithm for learning the combined model (DTNB) proceeds in much the same way as the one for stand-alone DTs. At each point in the search it evaluates the merit associated with splitting the attributes into two disjoint subsets: one for the DT, the other for NB. We use a forward selection, where, at each step, selected attributes are modeled by NB and the remainders by the DT, all attributes are modeled by the DT initially [Hall et al].

4.26. END Algorithm

Dong, Frank and Kramer supplied superior classifier in order to take configuring Nested dichotomies communities with two class classifiers and multiclass data cluster [Dong et al].

4.27. Attribute Selected Classifier

At selected attribute classifier, size of train and test data is reduced before they transfer to classifier for selection of attribute. The dimensions of the datasets are reduced by attribute selections before they are assigned to a classifier [Witten et al].

4.28. Classification via Regression

In Classification via Regression model tree gets a form of decision tree by using regression function instead of class values as leafs. Attributes with numerical values plays an important role at the regression function. According to Witten et al, their model trees are a kind of their decision trees with linear regression function on leafs. Model trees format principles of last successful technique in order to estimate permanent numerical values.

4.29. CVParameterSelection

In Classical Cross-validation parameter selection is performed by discretizing the parameter space into a grid and searching for the combination of parameters that minimizes the validation error (which corresponds to the upper level objective in the bi-level problem). This is typically followed by a local search for fine tuning the parameters. Typical discretization is logarithmic grids of base 2 or 10 on the parameters [Zurada et al].

4.30. Filtered Classifier

Filtered Classifier consist classes of arbitrary filtered data. Configuration of filter based on only train data as classifier and test samples are then manipulated not to be modified by the filter.

4.31. MultiScheme Algorithm

In MultiScheme algorithm, in order to assign a classifier via 10-fold cross validation performance is measured based on regression or total accuracy.

5. Implementation and Evaluation

5.1. Dataset

As explained in the previous chapters while making this research, Yasar University E-Learning System dataset is implemented. The dataset in question consist information regarding to online activity of Yasar University students and composed from 2 different tables; first one is called "Event" and second one is called "Event Count". Event table contains event logs regarding to when a student accessed to specific files/documents/videos and for how many times. Event count table on the other hand, contains logs regarding to the total amount of times the student accessed to the course website.

All the records kept inside Yasar University E-Learning System (YES) on MySQL database software tool. YES is a heavily modified version of Sakai E-Learning System. The SQL database is approximately 1 GB and inherits 865661 instances/logs. As discussed before, the dataset derives from logs of 10 different courses (UFND10A, UFND10B, UFND10C, UFND10D, UFND10E, UFND20A, UFND30A, UFND40A, UFND50A, UFND60A) over 5 different academic terms (2013-2014 Fall, 2013-2014 Spring, 2014-2015 Fall, 2014-2015 Spring, 2015-2016 Fall). The courses in question are presented below.

-UFND10A-Human Resources

-UFND10B-Semiology and Semantics

-UFND10C-History and Philosophy of Science

-UFND10D-Technology and Society

-UFND10E-Philosophy and logic

-UFND20A-Research Culture

-UFND30A-Design Culture

-UFND40A-Ethic Culture

-UFND50A-Aesthetics

-UFND60A-Project Culture

The purpose of this research, as explained before, is to analyze the collected logs so that whether an autonomous system can predict if a student will be able to pass the course or not. In order to achieve this goal, first the dataset in question is modified and multiple datasets are used with different versions. First version is event tables (E); event tables presents the state of being successful or unsuccessful for students according to, which events are done and for how many times, in which day of week those events are done, in which day of month those events are done, and finally in which month of the year those events are done. The second version is event count tables (EC); event count tables present the state of being successful or unsuccessful for students according to, how many times those events are done in total. For event and event count tables 3 different versions are created; course (C), term (TT) and total (T). Then, for the event table from these versions 8 more different versions are generated. These versions are called as ALL, M, W, MD, MW, WMD, MMD and NONE. For the event count tables no other versions are created. These tables' columns are explained below.

Event_ID: This column is used for making analysis depends on what is the event or what events are done.

Event_Count: This column is used for making analysis for how many times events have done.

Grade: This column indicates that students are successful or unsuccessful on the class.

Months: This column is used for making analysis for events being in which month.

Weekday: This column is used for making analysis for activity being in which day of the week.

Monthday: This column is used for making analysis for activity being in which day of month.

Table 1 present those columns and their versions for event tables.

Event Table
Event_ID
Event_Count
Grade{Yes, No}
Months{1-12}
Weekday{1-7}
Monthday{1-31}

Table 1 Columns of Event Table

Total_Event_Count: How many times those events are done in total in one term for each course.

Grade: This column indicates that students are successful or unsuccessful on the class.

Table 2 on the other hand presents which column is used for event count tables.

Event Count Table
Total_Event_Count
Grade

Table 2 Columns of Event Count Table

Table 3 presents the number of students that are registered to a course per each term inside our database.

Courses	2013-2014-I	2013-2014-II	2014-2015-I	2014-2015-II	2015-2016-II
UFND010A	424	462	530	442	156
UFND010B	92	41	94	34	229
UFND010C	93	99	106	28	125
UFND010D	199	99	95	98	53
UFND010E	186	63	90	48	13
UFND020A	771	848	721	623	202
UFND030A	513	663	605	689	381
UFND040A	594	734	505	640	274
UFND050A	643	698	529	662	114
UFND060A	757	710	525	503	105

Table 3 # of Students

As explained before, the first version for event and event count tables is *course* (C). C is used for analysis of each course of each term discretely. Table 4 presents the total number of records for event table version C. Table 5 on the other hand presents total number of records for event count table version C.

Courses	2013-2014-I	2013-2014-II	2014-2015-I	2014-2015-II	2015-2016-I
UFND010A	26095	20344	14570	22160	2369
UFND010B	6117	2070	2495	1683	3973
UFND010C	6658	4331	2701	1217	2196
UFND010D	13284	4903	2655	4509	919
UFND010E	10780	3366	2286	2649	272
UFND020A	41679	35332	16400	30110	2891
UFND030A	27857	28395	13044	31592	5811
UFND040A	32859	30039	11033	8266	3710
UFND050A	33732	28785	11487	27987	1880
UFND060A	39901	28099	9969	20784	1565

Table 4 Number of Records for Event Table Version C

Courses	2013-2014-I	2013-2014-II	2014-2015-I	2014-2015-II	2015-2016-I
UFND010A	424	462	530	442	156
UFND010B	92	41	94	34	229
UFND010C	93	99	106	28	125
UFND010D	198	99	95	98	53
UFND010E	186	63	90	48	13
UFND020A	771	848	721	623	202
UFND030A	513	663	605	689	381
UFND040A	594	734	505	640	274
UFND050A	643	698	529	662	114
UFND060A	757	710	525	503	105

Table 5 Number of Records for Event Count Table Version C

The second version for event and event count tables is term (TT). In this version, all terms (2013-2014 Fall, 2013-2014 Spring, 2014-2015 Fall, 2014-2015 Spring and 2015-2016 Fall) are combined together for each course to see whether any changes will occur through the whole system. Table 6 presents the total number of records for event table version TT. Table 7 on the other hand presents total number of records for event count table version TT.

Courses	#of Records
UFND010A	85538
UFND010B	16338
UFND010C	17103
UFND010D	26270
UFND010E	19353
UFND020A	126412
UFND030A	106166
UFND040A	105907
UFND050A	103871
UFND060A	100318

Table 6 Number of Records for Event Table Version TT

Courses	#of Records
UFND010A	2014
UFND010B	490
UFND010C	451
UFND010D	543
UFND010E	400
UFND020A	3367
UFND030A	2851
UFND040A	2747
UFND050A	2646
UFND060A	2600

Table 7 Number of Records for Event Count Table Version TT

The third version for event and event count tables is total (T). While total combination is formed, All courses (UFND010A, UFND010B, UFND010C, UFND010D, UFND010E, UFND020A, UFND030A, UFND040A, UFND050A, UFND060A) are combined for one term for analysis the effect of the being of successful or unsuccessful. Table 8 presents the total number of records for event table version T. Table 9 on the other hand presents total number of records for event count table version T.

Courses	#of Records
UFNDALL2013-I	238962
UFNDALL2013-II	185664
UFNDALL2014-I	86640
UFNDALL2014-II	170424
UFNDALL2015-I	25586

Table 8 Number of Records for Event Table Version T

Courses	#of Records
UFNDALL2013-I	4271
UFNDALL2013-II	4417
UFNDALL2014-I	3800
UFNDALL2014-II	3767
UFNDALL2015-I	1854

Table 9 Number of Records for Event Count Table Version T

As explained before, event table contains information regarding which type of events are executed by a student, count of these events, when these events are done by the student in question and his/her grade (successful or unsuccessful). For the event table, 8 different versions are created which are described on below at Table 10.

ALL	M	W	MD	MW	WMD	MMD	None
Event_ID	Event_ID	Event_ID	Event_ID	Event_ID	Event_ID	Event_ID	Event_ID
Event_Count	Event_Count	Event_Count	Event_Count	Event_Count	Event_Count	Event_Count	Event_Count
Grade{Yes, No}	Grade{Yes, No}	Grade{Yes, No}	Grade{Yes, No}	Grade{Yes, No}	Grade{Yes, No}	Grade{Yes, No}	Grade{Yes, No}
Months{1-12}	Months{1-12}	Months{1-12}	Months{1-12}	Months{1-12}	Months{1-12}	Months{1-12}	Months{1-12}
Weekday{1-7}	Weekday{1-7}	Weekday{1-7}	Weekday{1-7}	Weekday{1-7}	Weekday{1-7}	Weekday{1-7}	Weekday{1-7}
Monthday{1-31}	Monthday{1-31}	Monthday{1-31}	Monthday{1-31}	Monthday{1-31}	Monthday{1-31}	Monthday{1-31}	Monthday{1-31}

Table 10 Version of Event Table

For version ALL, Table 11 represents how the data is logged. From here once a machine learning algorithm is implemented over the dataset, it is possible to see the effects of activities over a student's grade such as what type of event is executed, for how many times, at which month, at which day of the month, and which day of week.

Event_ID	Event_Count	Month	Weekday	Monthday	Grade
Content_Read	2	9	3	24	Yes
Pres.begin	1	9	3	24	Yes
Messages.newfolder	4	9	3	24	Yes
Coursebuilder.read	18	9	3	24	Yes
Syllabus.read	1	9	3	24	Yes
Sam.assessment.taken	4	9	3	24	Yes
Sam.assessments.submt	2	9	3	24	Yes
Content.read	3	9	3	24	Yes
Pres.begin	1	9	3	24	Yes
Messages.newfolder	4	9	3	24	Yes

Table 11 Sample of All Dataset for Version C

M: Table 12 presents the columns for M type. This will enable us to analyze the relationship between the grade and the month at which the events are executed.

Event_Id	Event_Count	Month	Grade
Content_Read	2	9	Yes
Pres.begin	1	9	Yes
Messages.newfolder	4	9	Yes
Coursebuilder.read	18	9	Yes
Syllabus.read	1	9	Yes
Sam.assessment.taken	4	9	Yes
Sam.assessments.submt	2	9	Yes
Messages.newfolder	4	9	Yes

Table 12 Sample of M Dataset for version C

W: Table 13 presents the columns for W type. This will enable us to analyze the relationship between the grade and the day of the week at which the events are executed.

Event_Id	Event_Count	Weekday	Grade
Content_Read	2	3	Yes
Pres.begin	1	3	Yes
Messages.newfolder	4	3	Yes
Coursebuilder.read	18	3	Yes
Syllabus.read	1	3	Yes
Sam.assessment.taken	4	3	Yes
Sam.assessments.submt	2	3	Yes
Content.read	3	3	Yes
Pres.begin	1	3	Yes

Table 13 Sample of W Dataset for Version C

MD: Table 14 presents the columns for MD type. This will enable us to analyze the relationship between the grade and the day of the month at which the events are executed.

Event_Id	Event_Count	Monthday	Grade
Content_Read	2	24	Yes
Pres.begin	1	24	Yes
Messages.newfolder	4	24	Yes
Coursebuilder.read	18	24	Yes
Syllabus.read	1	24	Yes
Sam.assessment.taken	4	24	Yes
Sam.assessments.submt	2	24	Yes
Content.read	3	24	Yes
Pres.begin	1	24	Yes
Messages.newfolder	4	24	Yes

Table 14 Sample of MD Dataset for Version C

MW: Table 15 presents the columns for MW type. This will enable us to analyze the relationship between the grade and the month as well as the day of the week at which the events are executed.

Event_Id	Event_Count	Month	Weekday	Grade
Content_Read	2	9	3	Yes
Pres.begin	1	9	3	Yes
Messages.newfolder	4	9	3	Yes
Coursebuilder.read	18	9	3	Yes
Syllabus.read	1	9	3	Yes
Sam.assessment.taken	4	9	3	Yes
Sam.assessments.submt	2	9	3	Yes
Content.read	3	9	3	Yes

Table 15 Sample of MW Dataset for Version C

WMD: Table 16 presents the columns for WMD type. This will enable us to analyze the relationship between the grade and the day of the month as well as the day of the week at which the events are executed.

Event_Id	Event_Count	Weekday	Monthday	Grade
Content_Read	2	3	24	Yes
Pres.begin	1	3	24	Yes
Messages.newfolder	4	3	24	Yes
Coursebuilder.read	18	3	24	Yes
Syllabus.read	1	3	24	Yes
Sam.assessment.taken	4	3	24	Yes
Sam.assessments.submt	2	3	24	Yes
Content.read	3	3	24	Yes
Pres.begin	1	3	24	Yes

Table 16 Sample of WMD Dataset for Version C

MMD: Table 17 presents the columns for MMD type. This will enable us to analyze the relationship between the grade and the month as well as the day of the month at which the events are executed.

Event_Id	Event_Count	Month	Monthday	Grade
Content_Read	2	9	24	Yes
Pres.begin	1	9	24	Yes
Messages.newfolder	4	9	24	Yes
Coursebuilder.read	18	9	24	Yes
Syllabus.read	1	9	24	Yes
Sam.assessment.taken	4	9	24	Yes
Sam.assessments.submt	2	9	24	Yes
Content.read	3	9	24	Yes

Table 17 Sample of MMD Dataset for Version C

NONE: Table 18 presents the columns for NONE type. This will enable us to analyze the relationship between the grade and the total event count.

Event_Id	Event_Count	Grade
Content_Read	2	Yes
Pres.begin	1	Yes
Messages.newfolder	4	Yes
Coursebuilder.read	18	Yes
Syllabus.read	1	Yes
Sam.assessment.taken	4	Yes
Sam.assessments.submt	2	Yes
Content.read	3	Yes
Pres.begin	1	Yes
Messages.newfolder	4	Yes

Table 18 Sample of NONE Dataset for Version C

As explained before event count tables (EC) present a student's total event count with respect to his/her grade. Table 19 shows a sample dataset regarding to 2013-2014 Fall Term.

168	YES
130	YES
7	NO
143	YES
171	YES
118	NO
157	NO
121	NO
90	YES
257	YES

Table 19 A sample of Event Count Table Version C

Found in WEKA data mining tool described Section 4; J48, Naïve Bayes, JRip, Bayes Net, AdaboostM1, AD Tree, Attribute Selected Classifier, Bagging, Classification via Regression, Conjunctive Rule, CV Parameter Selection, Decision Table, DTNB, END, Filtered Classifier, Grading, ZeroR, Random Forest, MultiBoostAb, LogiBoost, Multi Scheme, Multilayer Perceptron, Multiclass Classifier and Voted Perceptron machine learning algorithms are implemented over the datasets by using 10-fold cross validation as well 20-fold cross validation techniques to predict grade of a student. To simply put; in X fold cross validation, the dataset splits to X groups randomly. Then first group is reserved for testing, and a model is built for training with the remainder groups. The same process is repeated for X times and average accuracy rate is calculated at the end. For 10-fold cross validation first the data is divided into 10 groups randomly and while 9 of them are used for training purposes 1 is left for testing. This process repeats 10 times. To apply all of these 31 algorithms over YES dataset, a program in Java language is coded. Once the program is executed the results regarding to total accuracy (TA), true positive (TP) and true negative (TN) rates are calculated and then recorded in excel files. These files then are used for the evaluation of the proposed framework. TA means the accuracy of the proposed framework. TP shows test results decided true when the condition is present. TN shows test results decided true when the condition is absent. FP shows test results decided false when the condition is absent. FN shows test results decided false when the condition is present.

After the completion of the implementation, an algorithm has to be selected as a winner to predict a student's pass of fail. In order to do that, all versions and tables have to be analyzed thoroughly and individually. Once these assessments are completed, then a decision can be made for one algorithm to implement over one version of the dataset. To draw a conclusion, total accuracy, true positive and true negative rates are necessary (confusion matrix is used). Approximately 1000000 experiments were completed and 10-fold cross validation technique is selected over 20-fold cross validation technique since it presents better results.

Confusion matrix contains information about actual and predicted classes once a machine learning algorithm is run. Performance of such system is evaluated using the results in the matrix.

For JRIP:

A	B	← Classified As
22910	81	A
250	2854	B

- A is the number of correct predictions when the condition is positive,
- B is the number of incorrect predictions when the condition is positive,

TA Rate = (22910+2854) / (22910+2854+81+250) = 0,987

TP Rate = 22910 / (22910+81) = 0,996

TN Rate = 2854 / (250+2854) = 0,919

FP Rate = 250 / (250+2854) = 0,080

FN Rate = 81 / (22910+81) = 0,004

For Decision Table:

A	B	←Classified As
22891	100	A
173	2931	B

TA Rate = (22891+2931) / (22891+2931+173+100) = 0,989

TP Rate = 22891 / (22891+100) = 0.996

TN Rate = 2931 / (2931+173) = 0,944

FP Rate = 173 / (2931+173) = 0,056

FN Rate = 100 / (22891+100) = 0,004

For Voted Perceptron:

A	B	←Classified As
22796	195	A
198	2906	B

TA Rate = (22796+2906) / (22796+2906+198+195) = 0,984

TP Rate = 22796 / (22796+195) = 0,992

TN Rate = 2906 / (2906+198) = 0,936

FP Rate = 198 / (2906+198) = 0,064

FN Rate = 195 / (22796+195) = 0,008

To choose the right algorithm for the specific dataset, an analysis is made. First the TP value is taken into consideration, and then TN value is evaluated. The final decision is made by deciding on which algorithm will be the most successful in both cases. Hence according to this, Decision Table is the most successful algorithm amongst the three. This logic is then implemented over the whole system. Tables 20, 21, 22, 23, 24 and 25 present the final results of the evaluation of the system. In the end, other than the winner algorithms presented, BAGGING algorithm is chosen as a general solution.

Event Table (E) – Version Course (C)		
Courses	**Version**	**Rules**
UFND010A-2013-2014-I	ALL-MMD	END-J48
UFND010A-2013-2014-II	ALL-MMD	END-J48
UFND010A-2014-2015-I	ALL	BAGGING
UFND010A-2014-2015-II	WMD	BAGGING
UFND010A-2015-2016-I	MD	MLAYERPERCEPTRON
UFND010B-2013-2014-I	ALL-MMD	END-J48
UFND010B-2013-2014-II	ALL	END-J48
UFND010B-2014-2015-I	WMD	JRIP
UFND010B-2014-2015-II	MD	IBK
UFND010B-2015-2016-I	MW	BAGGING
UFND010C-2013-2014-I	WMD	BAGGING
UFND010C-2013-2014-II	MW	BAGGING
UFND010C-2014-2015-I	MMD	CLASSIFICATIONVIAR
UFND010C-2014-2015-II	ALL	DECISION TREE
UFND010C-2015-2016-I	M	IB1
UFND010D-2013-2014-I	ALL-MMD	BAGGING
UFND010D-2013-2014-II	MMD	BAYES NET
UFND010D-2014-2015-I	M-MW	NAÏVE BAYES SIMPLE
UFND010D-2014-2015-II	ALL	RANDOM FOREST
UFND010D-2015-2016-I	MW	IB1
UFND010E-2013-2014-I	ALL-MMD	END-J48
UFND010E-2013-2014-II	ALL	DECISION TABLE
UFND010E-2014-2015-I	ALL	DNTB
UFND010E-2014-2015-II	ALL	KSTAR
UFND010E-2015-2016-I	ALL	JRIP
UFND020A-2013-2014-I	ALL-MMD	BAGGING
UFND020A-2013-2014-II	ALL-MMD	DECISION TABLE

UFND020A-2014-2015-I	M	END-J48
UFND020A-2014-2015-II	MMD	DECISION TABLE
UFND020A-2015-2016-I	MW	MLAYERPERCEPTRON
UFND030A-2013-2014-I	ALL	DECISION TABLE
UFND030A-2013-2014-II	ALL-MMD	BAGGING
UFND030A-2014-2015-I	ALL-MMD	BAGGING
UFND030A-2014-2015-II	ALL	RANDOM FOREST
UFND030A-2015-2016-I	ALL	BAGGING
UFND040A-2013-2014-I	ALL	BAGGING
UFND040A-2013-2014-II	ALL	BAGGING
UFND040A-2014-2015-I	ALL-MMD	DECISION TREE
UFND040A-2014-2015-II	ALL-MMD	BAGGING
UFND040A-2015-2016-I	ALL	DECISION TABLE
UFND050A-2013-2014-I	ALL-MMD	END-J48
UFND050A-2013-2014-II	ALL-MMD	BAGGING
UFND050A-2014-2015-I	ALL-MMD	END-J48
UFND050A-2014-2015-II	ALL-MMD	BAGGING
UFND050A-2015-2016-I	ALL	JRIP-LOGIBOOST
UFND060A-2013-2014-I	ALL-MMD	BAGGING
UFND060A-2013-2014-II	ALL-MMD	BAGGING
UFND060A-2014-2015-I	ALL	MLAYERPERCEPTRON
UFND060A-2014-2015-II	ALL-MMD	BAGGING
UFND060A-2015-2016-I	ALL-MMD-WMD	DECISION TREE

Table 20 Winner Algorithms of Event Table for Version C

Event Table (E) – Version Term (TT)		
COURSES	**VERSION**	**RULES**
UFNDALL2013-I	ALL-MMD	BAGGING
UFNDALL2013-II	ALL-MMD	BAGGING
UFNDALL2014-I	ALL	END-J48
UFNDALL2014-II	ALL	BAGGING
UFNDALL2015-I	ALL-M-MW-MMD	DECISION TREE

Table 21 Winner Algorithms of Event Table for Version TT

Event Table (E) – Version Total (T)		
COURSES	**VERSION**	**RULES**
UFND010AALL	ALL	BAGGING
UFND010BALL	ALL	DECISION TREE
UFND010CALL	ALL	BAGGING
UFND010DALL	ALL	CLASSIFICATIONVIAREGRESSION
UFND010EALL	ALL-MMD	END-J48
UFND020AALL	ALL	CLASSIFICATIONVIAREGRESSION
UFND030AALL	ALL	BAGGING
UFND040AALL	ALL	BAGGING
UFND050AALL	ALL	BAGGING
UFND060AALL	MMD	BAGGING

Table 22 Winner Algorithms of Event Table for Version T

Event Count Table (EC) – Version Course (C)	
Courses	**Rules**
UFND010A-2013-2014-I	MULTICLASS CLASSIFIER
UFND010A-2013-2014-II	MULTILAYER PERCEPTRON
UFND010A-2014-2015-I	BAGGING
UFND010A-2014-2015-II	DECISION TABLE

UFND010A-2015-2016-I	RANDOM FOREST
UFND010B-2013-2014-I	MULTICLASS CLASSIFIER
UFND010B-2013-2014-II	MULTILAYER PERCEPTRON
UFND010B-2014-2015-I	MULTILAYER PERCEPTRON
UFND010B-2014-2015-II	LOGISTIC
UFND010B-2015-2016-I	IBK
UFND010C-2013-2014-I	NAIVEBAYES
UFND010C-2013-2014-II	KSTAR
UFND010C-2014-2015-I	MULTILAYER PERCEPTRON
UFND010C-2014-2015-II	JRIP
UFND010C-2015-2016-I	IB1
UFND010D-2013-2014-I	MULTICLASS CLASSIFIER
UFND010D-2013-2014-II	BAGGING
UFND010D-2014-2015-I	BAGGING
UFND010D-2014-2015-II	MULTILAYER PERCEPTRON
UFND010D-2015-2016-I	NAÏVE BAYES
UFND010E-2013-2014-I	CLASSIFICATION VIA REG.
UFND010E-2013-2014-II	BAGGING
UFND010E-2014-2015-I	J48
UFND010E-2014-2015-II	BAGGING
UFND010E-2015-2016-I	RANDOM FOREST
UFND020A-2013-2014-I	CLASSIFICATION VIA REG.
UFND020A-2013-2014-II	MULTILAYER PERCEPTRON
UFND020A-2014-2015-I	MULTICLASS CLASSIFIER
UFND020A-2014-2015-II	CONJUNCTIVE RULES
UFND020A-2015-2016-I	IB1
UFND030A-2013-2014-I	MULTIBOOST AB
UFND030A-2013-2014-II	ADTREE
UFND030A-2014-2015-I	ADTREE

UFND030A-2014-2015-II	J48
UFND030A-2015-2016-I	IB1
UFND040A-2013-2014-I	CONJUNCTIVE RULES
UFND040A-2013-2014-II	LWL
UFND040A-2014-2015-I	BAGGING
UFND040A-2014-2015-II	JRIP
UFND040A-2015-2016-I	IB1
UFND050A-2013-2014-I	MULTILAYER PERCEPTRON
UFND050A-2013-2014-II	ADABOOST
UFND050A-2014-2015-I	JRIP
UFND050A-2014-2015-II	CONJUNCTIVE RULES
UFND050A-2015-2016-I	IB1
UFND060A-2013-2014-I	RANDOM FOREST
UFND060A-2013-2014-II	JRIP
UFND060A-2014-2015-I	DECISION TABLE
UFND060A-2014-2015-II	BAGGING
UFND060A-2015-2016-I	ERROR

Table 23 Winner Algorithms of Event Count Table for Version C

Event Count Table (EC) – Version Term (TT)	
COURSES	**RULES**
UFNDALL2013-I	MULTIBOOST AB
UFNDALL2013-II	MULTIBOOST AB
UFNDALL2014-I	J48
UFNDALL2014-II	JRIP
UFNDALL2015-I	IB1

Table 24 Winner Algorithms of Event Count Table for Version TT

Event Count Table (EC) - Version Total (T)	
COURSES	**RULES**
UFND010AALL	BAYES NET
UFND010BALL	IB1
UFND010CALL	IBK
UFND010DALL	CONJUNCTIVE RULES
UFND010EALL	CLASSIFICATION VIA REGRESSION
UFND020AALL	LOGIBOOST
UFND030AALL	MULTILAYER PERCEPTRON
UFND040AALL	JRIP
UFND050AALL	ADABOOST
UFND060AALL	BAYESNET

Table 25 Winner Algorithms of Event Count Table for Version T

6. CONCLUSION AND FUTURE WORKS

In this thesis, the performance of 31 algorithms (Naïve Bayes, JRip, J48, Bayes Net, Adaboost M1, AdTree, Attribute Selected Classifier, Bagging, Classification Via Regression, Conjunctive Rules, Cv Parameter Selection, Decision Table, DTNB, END, Filtered Classifier, Grading, IB1, IBK, K*, Logistic, LogiBoost, LWL, MultiBoostAb, Multiclass Classifier, Multi Scheme, Naïve Bayes Simple, Multilayer Perceptron, SMO, Voted Perceptron, Random Forest and ZeroR) are applied on Yasar University E-Learning dataset and then the results of the generated models are analyzed. These algorithms are implemented using 10 cross fold validation and 20 cross fold validation technique via WEKA Data Mining Tool. While making this research, approximately 1000000 different training models were generated.

While applying those algorithm methods, 6 different datasets were used. First, for event and event count tables, all courses and all semesters were studied and analyzed separately and these algorithms were applied to all of them. Second, separate courses and total semesters were studied and analyzed and these algorithms were applied too. Third, the tables were studied and analyzed which contain all courses with in each semester and all algorithms were applied for each dataset. For event count tables, all courses and all semesters were studied and analyzed separately and these algorithms were applied all of them. Separate courses and total semesters were studied and analyzed and these algorithms were applied too. Total were studied and analyzed which contain all courses with in each semester and all algorithm were applied for each dataset.

The final evaluation presents that although the accuracy of all the algorithms are very close to each other, it is seen that 10-fold cross validation is slightly better than 20-fold cross validation technique and BAGGING is the most successful algorithm.

For future work, in order to move this research one step further by implementing these machine learning algorithms over a larger dataset, which inherits more information regarding the interaction of the user with the system (i.e. the exact time of activity in a day, 18:21.45 20/11/2016), the authors of this research believe that it is possible to predict all the grades of a student during a whole term of the course (i.e. midterm, quiz, coursework or final grades).

References

Ralput A., Aharwal R., Dubey M, Saxena S., Raghuvanshi M., 2011, "J48 and JRIP Rules for E-Governance Data", International Journal of Computer Science and Security (IJCSS), p. 2-6.

Yannick M, Jie X., Atan O., Van der Schaar M, 2015, "Personalized Grade Precision: A Data Mining Approach", University of California, Los Angeles,p. 1-6.

Yasar University "eYasar Learning Environment" https://e.yasar.edu.tr

Alpaydin E., 2010, "Introduction to Machine Learning", The MIT Press, Cambridge, Massachusetts, London, England, p. 3.

Nasiri M., Minaei B., Vafaei F., 2012, "Predicting GPA and Academic Dismissal in LMS Using Educational Data Mining: A Case Mining", 6th National and 3rd International Conference of E-Learning and E-Teaching (ICELET2012), p. 1-6.

Atkenson C.G., Moore A.W., Schaal S., 1997, "Locally Weighted Learning for Control Artificial Intelligence Review", p. 75-113.

Seewald A., Fürnkranz J., 2001, "An Evaluation of Grading Classifiers", Austrian Research Instute for Artificial Intelligence, p. 1-5.

Xu X., http://weka.sourceforge.net/packageMetaData/conjunctiveRule/index.html, 28/07/2016.

Hall M, Frank E., 2008, "Combining Naïve Bayes and Decision Tables", Proceedings of the Twenty First International FLAIRS Conference, p. 318-319.

Dong L., Frank E., Kramer S., 2005, "Ensembles of Balanced Nested Dichotomies for Multi-class Problems In: PKDD", p. 84-95.

Frank E., Kramer S., 2004, "Ensembles of nested dichotomies for multi-class problems", 21 International Conference on Machine Learning.

Witten I., Frank E., 2005, "Data Mining Practical Machine Learning Tools and Techniques", Second Edition, Elsevier.

Zurada J. M., Yen G. G., Wang J., 2008, "Computational Intelligence: Research Frontiers", First Edition, Springer.

Appendix

Part of Appendix is given with CD.

1) 10 Fold
    - A) Courses
    - B) Term
    - C) Total

2) 20 Fold
    - A) Courses
    - B) Term
    - C) Total

3) Total Count 10 Fold
    - A) Courses
    - B) Term
    - C) Total

4) Code

CURRICULUM VITEA

Erhan KINAY was born in 1990, Konak-Izmir. He graduated from Yasar University, Computer Engineering with a CGPA 2.61 in 2013. His professional career was started in 2013 as an Infoera Yazilim.