Computers in Human Behavior 36 (2014) 469–478

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Developing early warning systems to predict students' online learning performance

Ya-Han Hu [a,1], Chia-Lun Lo [a,1], Sheng-Pao Shih [*,b]

[a] Department of Information Management, National Chung Cheng University, 168, University Rd., Min-Hsiung Chia-Yi, Taiwan, ROC
[b] Department of Information Management, Tamkang University, No. 151, Yingzhuan Rd., Tamsui Dist., New Taipei City 25137, Taiwan, ROC

ARTICLE INFO

Article history:

Keywords:
Learning management system
e-Learning
Early warning system
Data-mining
Learning performance prediction

ABSTRACT

An early warning system can help to identify at-risk students, or predict student learning performance by analyzing learning portfolios recorded in a learning management system (LMS). Although previous studies have shown the applicability of determining learner behaviors from an LMS, most investigated datasets are not assembled from online learning courses or from whole learning activities undertaken on courses that can be analyzed to evaluate students' academic achievement. Previous studies generally focus on the construction of predictors for learner performance evaluation after a course has ended, and neglect the practical value of an "early warning" system to predict at-risk students while a course is in progress. We collected the complete learning activities of an online undergraduate course and applied data-mining techniques to develop an early warning system. Our results showed that, time-dependent variables extracted from LMS are critical factors for online learning. After students have used an LMS for a period of time, our early warning system effectively characterizes their current learning performance. Data-mining techniques are useful in the construction of early warning systems; based on our experimental results, classification and regression tree (CART), supplemented by AdaBoost is the best classifier for the evaluation of learning performance investigated by this study.

1. Introduction

Recent technological innovations and the development of e-Learning platforms, such as web-based online-learning and multimedia technologies, not only overcome limitations of time and space, but also reduce learning costs. Educators can monitor students' learning processes by evaluating their learning portfolios using an learning management system (LMS) in an e-Learning environment (Macfadyen & Dawson, 2010). Through the use of an LMS, information about student learning behaviors and activities are retrieved from system logs or databases, and the data is then be analyzed by an early warning system. Educators can assess overall learning performances, and determine how well students are learning and what particular difficulties they might be having, so gaining insight into students that are at-risk of course failure (Campbell & Oblinger, 2007; Lust, Elen, & Clarebout, 2013). Past studies report that, logs of online activity and learner data stored in an LMS can be used to make forecasts of student future performance (Macfadyen & Dawson, 2010; Romero, Ventura, & García, 2008; Valsamidis, Kontogiannis, Kazanidis, Theodosiou, & Karakos, 2012). LMS prompts academically at-risk students to study more effectively, while educators can track the progress of their students and generate timely feedback (Kotsiantis, Pierrakeas, & Pintelas, 2004). However, the development of a precise LMS that can assess student learning performances using web-based learning portfolios is a challenging task. Data mining attempts to obtain valuable knowledge from data stored in large repositories; the strategy has been considered an appropriate method of knowledge discovery to excavate implicit information (Duan & Da Xu, 2012; Dunham, 2002). In the field of education, data mining is concerned with developing methods to explore the unique types of data that describe learners, and applies these methods to provide a better understanding of the learners, thus, data mining may reveal information that will benefit the learners. Recently, researchers used data mining techniques to analyze learning portfolios, make predictions, and to construct models of student learning performance (Macfadyen & Dawson, 2010). Dringus and Ellis (2005) argued that, from a systems view of LMS, given the textual nature of most asynchronous data held in systems, assessment and the acquisition of valuable information is hindered by limitations of the query and reporting toolset

* Corresponding author. Tel.: +886 2 2621 5656x2874; fax: +886 2 2620 9737.
E-mail addresses: yahan.hu@mis.ccu.edu.tw (Y.-H. Hu), allenlo.tw@gmail.com (C.-L. Lo), sbao@mail.tku.edu.tw (S.-P. Shih).
[1] Tel.: +886 5 272 0411; fax: +886 5 272 1501.

http://dx.doi.org/10.1016/j.chb.2014.04.002
0747-5632/

provided by the systems. LMS with data mining technique can successfully integrate online learning systems to further improve student academic performances (Hanna, 2004; Valsamidis et al., 2012; Wu, 2013).

Previous studies investigated the learning portfolios recorded by LMS systems to understand learner behavior (Gaudioso & Talavera, 2006), determine learning system effectiveness (Mostow et al., 2005; Yu, Jannasch-Pennell, Digangi, & Wasson, 1998; Yu & Wu, 2013), predict academically at-risk students (Essa & Ayad, 2012), or develop an early warning system to provide decision support system for instructors (Kotsiantis, 2012; Macfadyen & Dawson, 2010). However, most of the research dataset was obtained from traditional classroom settings, not from online learning courses; there is not much integrated LMS data that describes all activities undertaken by students during courses, which is available for analysis of academic performances (Hwang, Hsiao, & Tseng, 2003). To our knowledge, no study has predicted student learning performance using learning portfolio datasets when a fully online course is in progress. Therefore, time-dependent variables and other activities undertaken by students during the semester were included in our study, to predict student learning performance using data mining techniques. To identify *time-dependent variables* regarding the students' learning behaviors is a critical task in the development of an LMS. In this study, *time-dependent variables* denote those variables that varied during the learning activity processes (i.e., change with time). In this manner, the study addressed the following research questions:

- *How can data mining techniques accurately predict student learning performance based on activities in a fully online course?*
- *With the inclusion of time-dependent variables, how early in the semester can the early warning system accurately predict student learning performance?*
- *Which data mining technique offers superior predictive power regarding learning performance, when a fully online class is in progress?*

We analyzed a fully online undergraduate information literacy course offered at a national university in Taiwan during 2009. The analysis included 330 student learning portfolios. A "fully online" course refers to one in which all content delivery, communication and assessment is carried out via the LMS (Macfadyen & Dawson, 2010). We evaluated learning portfolio data, obtained while the class was in progress, as our dataset, and built a prediction model to predict at-risk students. To understand the effects of time-dependent variables on academic performance, we collected online course time-dependent variables to build reliable prediction models for an early warning system using C4.5 (Quinlan, 1993), classification and regression tree (CART) (Breiman, Friedman, Stone, & Olshen, 1984), logistic regression (LGR) (Sumner, Frank, & Hall, 2005), and adaptive boosting (AdaBoost) (Freund & Schapire, 1996) as part of our data mining strategy.

The remainder of this study is organized as follows. In Section 2, we review previous studies on the development of LMS and EDM research. Section 3 explains the proposed classification techniques used to build the early warning model and data collection procedure. Section 4 describes the experimental results of classification systems. Section 5 presents the system development and evaluation. Section 6 concludes the study.

2. Literature review

2.1. Learning portfolio and LMS

Learning portfolios are a collection of the events and learning activities undertaken by learners. Building a portfolio is a flexible, evidence-based process that combines reflection with documentation, and encourages student engagement in ongoing and collaborative analysis of learning to provide purposeful and selective outcomes for both improving learning outcomes and assessing the learning process (Zubizarreta & Millis, 2009). Traditionally, a learning portfolio relies on manual data collection and a writing-centered learning process. Difficulties in data storage, searching, and management are obstacles to the development and implementation of learning portfolio evaluation. By contrast, a web-based learning portfolio can be automatically collected, stored, and managed by LMS when learners interact with an e-Learning management system. Consequently, there has been a significant research effort into learning performance assessment using a web-based learning portfolio approach (Rasmussen, Northrup, & Lee, 1997). Previous studies revealed how a learning portfolio can assist instructors in correlating learner behaviors with their learning performance (Agrawal & Srikant, 1994; Hanna, 2004; Macfadyen & Dawson, 2010; Sadler-Smith, 2001). Drawing on the work, and building on a study conducted by Wang and Newlin (2002), researchers identified a strong relationship between LMS usage and learning performance (Campbell & Oblinger, 2007; Goldstein & Katz, 2005). Campbell and Oblinger (2007) further proposed that educators can directly benefit from the analysis of LMS data and development of an early warning system, by identifying at-risk students and implementing early intervention strategies. Wang, Newlin, and Tucker (2001) investigated student behaviors using LMS, and argued that online learning activities can provide an early warning index of student academic achievement. However, among these studies, there are few studies into the effectiveness of LMS-based early warning systems.

2.2. Educational data mining research

The analysis of student usage activity recorded in LMS is becoming increasingly important to EDM (Baker, 2010; Baker & Yacef, 2009). The objectives of these EDM studies include: understanding learner behavior (Chang, Kao, Chu, & Chiu, 2009; Gaudioso & Talavera, 2006), determining the effectiveness of learning systems (Mostow et al., 2005), identifying academically at-risk students (Essa & Ayad, 2012), and developing an early warning system and decision support system for instructors (Gaudioso, Montero, & Hernandez-Del-Olmo, 2012; Kotsiantis, 2012; Macfadyen & Dawson, 2010).

Castro, Vellido, Nebot, and Mugica (2007) reviewed EDM research conducted from 1999 to 2006 and concluded that most EDM studies dealt with Classification and Clustering problems associated with online platforms. The techniques used by the reviewed studies varied with research hypothesis and data characteristics. We reviewed EDM reports related to our study, which were published during the past decade (as shown in Table 1).

The research issues broadly divide into two categories. The first category includes reports that make predictions about online test performances, for example, Anozie and Junker (2006) provided an online examination platform that could timely and effectively predict an examination result. Kotsiantis et al. (2004) predicted student performances under a distance learning system by employing ML techniques. Guruler, Istanbullu, and Karahasan (2010) used classification techniques to identify individual student characteristics and their association with future success. The second category includes studies that make predictions about student learning performance through various feature sets. For example, Muehlenbrock (2005) used decision tree (DT) techniques to help students using e-Learning systems to further improve their e-Learning performance. Etchells, Nebot, Vellido, Lisboa, and Mugica (2006) used Fuzzy Inductive Reasoning (FIR) and Orthogonal Search-Based Rule Extraction (OSRE) techniques to construct a

Table 1
Recent studies on learning performance prediction.

	# samples	# attributes	Time attributes					Non-time attributes					Classification techniques	Cross validation
			Demographic	Assignment	Quiz	Forum	Course material	Demographic	Assignment	Quiz	Forum	Course material		
Minaei-Bidgoli et al. (2003)	227	10		V	V		V				V		DT, Bayes, *k*NN, Parzen-window, ANN	V
Macfadyen and Dawson (2010)	118	22	V	V					V		V	V	LR	
Chen, Chen and Li (2007)	183	8				V	V			V	V	V	Fuzzy AR	
Calvo-Flores, Galindo, Jiménez and Pérez (2006)	240	10	V					V	V	V	V		ANN	V
Muehlenbrock (2005)	70	12	V					V		V		V	DT	
Kotsiantis et al. (2003)	354	15						V	V	V	V		Bayes, *k*NN, Ripper, DT	V
Kotsiantis and Pintelas (2005)	354	15						V	V				M5, ANN, LR, SVM	
Kotsiantis et al. (2004)	354	15						V	V	V	V		DT, ANN, Bayes, *k*NN, LR, SVM	V
Minaei-Bidgoli, Tan and Punch (2004)	200	11		V	V			V		V		V	AR	
Etchells et al. (2006)	722	16						V	V	V	V	V	FIR, OSRE	V
Gaudioso et al. (2012)	300	n/a	V									V	Ripper, PART, DT, Bayes	
Kotsiantis (2012)	354	16						V	V			V	M5, ANN, LR, SVM	

DT: Decision Tree; *k*NN: *k*-Nearest Neighbors; AR: Association Rule; ANN: Artificial Neural Network; LR: Linear Regression; Ripper: Repeated Incremental Pruning to Produce Error Reduction; SVM: Support Vector Machine; M5: M5-Rules; FIR: Fuzzy Inductive Reasoning; OSRE: Orthogonal Search-based Rule Extraction.

course grade prediction model. Guruler et al. (2010) used personal characteristics of college freshmen to predict their learning performance. Minaei-Bidgoli, Kashy, Kortmeyer, and Punch (2003) applied *k*NN, ANN, and DT to establish classifiers, and make predictions of student final grades. Kotsiantis, Pierrakeas, and Pintelas (2003) used several machine-learning techniques to predict dropouts among students enrolled in a distance learning course, using demographic and learning performance variables.

It is worth noting that most of these studies adopted non-time-dependent variables in the construction of classifiers. One of the main reasons for this is that the learning activities in the associated studies were conducted in traditional classrooms where the LMS only played as an assisting role. Some studies report that well-developed classifiers can provide early warnings for both educators and students before the students fail a class (Essa & Ayad, 2012; Kotsiantis, Pierrakeas, Zaharakis, & Pintelas, 2003; Kotsiantis & Pintelas, 2005; Macfadyen & Dawson, 2010; Wang & Newlin, 2002). However, these studies summarized their information using complete semester student learning profiles to generate the classification model (Anozie & Junker, 2006; Gaudioso et al., 2012; Wang et al., 2001); thus, the learning information was captured at the end of the courses, not during the courses.

Therefore, we selected a fully online course to determine the complete time-dependent variables of the student learning activities from the LMS, and to develop an early warning system. During experimental evaluation, we collected learning activity data from three different periods during the semester. Consequently, we built three datasets that summarized student learning behaviors at different points to determine how early the early warning system can accurately predict student learning performance. Comparison of the evaluation results from the three datasets verified the effectiveness of the early warning system and confirmed the significance influence of time-dependent variables on student learning performance.

3. Data and classification techniques

3.1. Data preparation

This study collected complete learning portfolio data for the undergraduate students who had taken the Information Literacy and Information Ethics fully online course in a national university from September 2009 to June 2010. The selected course was delivered over two semesters (i.e., fall 2009 and spring 2010) by the same lecturer. According to the course requirement as outlined by the syllabus, students were required to attend online classes and watch online videos in specific time periods.

The LMS used in this study employed a user authentication process. Detailed learning activities for each student were recorded in the LMS database including e-Learning system login and logout events, course material opening and closing events, assignment download and upload events, and discussion read and post events. All student activities were stored in log file format, meaning that a record was generated in the LMS database as an event occurred. Therefore, a series of data preprocessing tasks must be performed, such as session identification, data integration, and data aggregation. Thus, fourteen variables were generated in this study, and their descriptions are shown in Table 2.

Each student record includes four types of input variables, these are variables regarding: login behavior, the use of online course materials, assignment status, and discussion status in the forum. To build a classification model for the early warning system, each student in the dataset is assigned a particular label, i.e., pass or fail the course. A student passes the course if the average score for midterm and final examinations is greater than, or equal to, 60 points, otherwise, the student is classified as fail. The study recruited 300 students, of which 284 students passed the course, and 16 students failed the course.

Table 2
Summary of variables descriptions.

Variables	Descriptions
Course_LoginAllCount	# Login
Course_LoginTime	Total time online (sec)
Course_LoginAVGTime	Average time per session (sec)
ReadTime	Total time material viewed (sec)
ReadTimeAllCount	# Course material viewed
ReadTimeAVG	Average time material viewed (sec)
ReadTimeDOCCount	# Course material viewed (by material category) (sec)
ReadTimeDOCAVG	Total time course material viewed (by material category) (sec)
ReadDocRate	ReadDocRate = ReadTimeDOCCount/# Course material released to date
Hwk_Delay	# Assignment delay
Hwk_NonPayment	# Non-delivery assignment
Forum_JoinRate	Participation rate of forum (# Reply discussion messages /total discussion messages)
Forum_Reply	# Reply discussion messages
IF_PASS	Pass or fail the course

The aim of an LMS-based early warning system is to monitor the learning progress of students, identify at-risk students, and help teachers to develop improvement strategies (Campbell & Oblinger, 2007). Although educators and institutions may wish the system to provide accurate forecasts as soon as learning begins, it is difficult to develop an accurate classification model by analyzing students' previous learning behavior. Thus, in this study we summarize the LMS data, and generate three datasets based on different periods of study. The three datasets T4W, T8W, and T13W denote the sets of learning behavior summary statistics for the first four, eight, and thirteen weeks of the course (Table 3). Section 5 describes comparative analyses of the effects of using LMS data for different stages in the course.

3.2. Classification techniques

To build an early warning system and predict student performances, we applied three well-known single classification techniques, C4.5, CART, and LGR.

DT is a well-known and powerful supervised learning technique (Quinlan, 1993). It comprises a hierarchical structure comprising nodes and branches; an internal node represents an input variable, the branch of an internal node represents a subset of the values of the corresponding input variable, and a leaf node is associated with a value (or a class label) of the output variable. C4.5 and CART are the two most commonly used DT-based learning techniques. Although both algorithms use similar tree construction processes, they have the following differences: (1) the variable selection criterion of C4.5 is a gain ratio, while that in CART is the Gini index (Breiman et al., 1984); (2) C4.5 allows multi-way splits, while CART only allows binary splits; and (3) C4.5 applies an error-based pruning method, while CART uses a cost-complexity pruning method. The main advantage of DT is that because of its tree-like format, the generated rules are easily observed and interpreted, which reduce the probability of errors arising in a complex problem.

LGR is a widely used statistical technique for modeling an output variable by using a linear combination of one or more input variables. LGR aims to predict the probability of the occurrence of an event by fitting data into a logistic function, thereby allowing inputs with any values to be transformed and confined to a value between 0 and 1. Each regression coefficient represents the corresponding variable's degree of contribution. A positive regression coefficient increases the probability of the output, while a negative value decreases the probability of the output.

In addition to the comparing the performances of C4.5, CART, and LGR, this study used classifier ensembles to enhance the predictive power of the three classification techniques. AdaBoost (Freund & Schapire, 1996) iteratively applies a selected classification algorithm, and evaluates each instance in the training dataset. The weights of instances incorrectly classified by the current classifier are increased for the next round of learning. Thus, AdaBoost encourages a new classifier to learn from instances misclassified by earlier iterations by assigning a larger weight to those instances. After a sequence of classifiers is built, AdaBoost applies a weighted majority vote to make forecasts. Although the concept of AdaBoost is simple, previous studies have shown that several classification algorithms used in conjunction with AdaBoost achieve greater classification accuracy than individual base classifiers do.

4. Experimental results of classification systems

4.1. Design of classification systems and experimental setup

This study used Weka 3.7.3 open-source data mining software (www.cs.waikato.ac.nz/ml/weka), to construct the classification models. Several parameters were chosen to control the quality of the experiments. The J48 (C4.5 in Weka) confidence factor and the minimum number of instances per leaf were defined as 0.25 and 2, respectively. CART parameters (SimpleCart in Weka) defined that a tree should stop growing when the number of instances in a

Table 3
Summary statistics of variables for the three datasets.

Variables	Descriptive statistics		
	T4W	T8W	T13W
Course_LoginAllCount	μ = 6.77, σ = 5.31	μ = 15.46, σ = 10.43	μ = 27.08, σ = 15.99
Course_LoginTime[a]	μ = 25,647, σ = 106885	μ = 58,443, σ = 253580	μ = 97,540, σ = 344941
Course_LoginAVGTime[a]	μ = 5915, σ = 29,956	μ = 6059, σ = 18,031	μ = 6556, σ = 14,799
ReadDocRate	μ = 0.66, σ = 0.32	μ = 0.65, σ = 0.20	μ = 0.70, σ = 0.18
ReadTime[a]	μ = 22,108, σ = 41,400	μ = 49,506, σ = 102606	μ = 78,793, σ = 153203
ReadTimeAllCount	μ = 10.55, σ = 5.68	μ = 22.05, σ = 11.64	μ = 32.79, σ = 16.42
ReadTimeAVG[a]	μ = 3609, σ = 10,002	μ = 4195, σ = 11,007	μ = 4452, σ = 10,550
ReadTimeDOCCount	μ = 3.09, σ = 0.63	μ = 6.72, σ = 1.36	μ = 10.58, σ = 2.11
ReadTimeDOCAVG	μ = 8082, σ = 15,999	μ = 8852, σ = 19,409	μ = 9231, σ = 19,480
Hwk_Delay	μ = 0, σ = 0	μ = 0, σ = 0	μ = 0.16, σ = 0.51
Hwk_NonPayment	μ = 2.19, σ = 0.4	μ = 2.07, σ = 0.36	μ = 0.184, σ = 0.44
Forum_JoinRate	μ = 0.3, σ = 0.19	μ = 0.39, σ = 0.21	μ = 0.38, σ = 0.2
Forum_Reply	μ = 3.74, σ = 2.39	μ = 9.01, σ = 4.82	μ = 12.65, σ = 6.62
IF_PASS	Yes = 284, No = 16		

Note: μ: mean; σ: standard deviations.
[a] Time-dependent variables.

node is less than 2. AdaBoost (AdaBoostM1 in Weka) parameters defining the number of iterations and the weight threshold for pruning, were set as 10 and 100, respectively.

In the collected dataset, 5.33% of students failed the course, resulting in a serious class imbalance problem. Tan, Kumar, and Steinbach (2006) showed that the adjustment of the ratio of two class samples can improve the machine's learning performance. Therefore, we applied a Weka resample module to modify the distribution of instances of the two classes by oversampling the "Fail" class and undersampling the "Pass" class; consequently, the distributions within each class were modified to become almost identical. Some useful instances in the "Pass" class may not be chosen by the resample method, resulting in the loss of valuable classification information. Therefore, we applied a random resampling technique for thirty cycles using different random seeds to construct the datasets. For each generated dataset, a 10-fold cross-validation was then applied to all the experimental evaluations, in which each dataset was partitioned into ten complementary subsets, wherein, any nine were used for model training, and the remaining subset was used for model testing.

4.2. Performance measurement

To evaluate and compare the performance of the classification models, we considered the following three metrics, *prediction accuracy*, *Type I error, and Type II error*. Using a confusion matrix (Fig. 1), these metrics can be calculated as:

$$Prediction\ accuracy = \frac{TP + TN}{TP + FP + FN + TN},$$

$$Type\ I\ error = \frac{FP}{FP + TN},$$

and

$$Type\ II\ error = \frac{FN}{TP + FN}.$$

Note we generated thirty datasets for each summarized dataset (T4W, T8W, and T13W), resulting in generating thirty results for each experiment. All of the experimental evaluations reported in Section 5 are the averages of the results of thirty trials. In addition, the cross-validated paired t-test was applied to compare performances of each pair of classifiers (i.e., both single and ensemble classifiers).

		Predicted class	
		Fail	Pass
Actual class	Fail	*TP*	*FN*
	Pass	*FP*	*TN*

Fig. 1. Confusion matrix.

4.3. Results

The evaluation results are presented in Table 4. We compared the prediction results of the three single classifiers. The overall accuracy experimental results (T4W, T8W, and T13W) showed that both C4.5 and CART provide an overall accuracy of greater than 93%, significantly better than the accuracy of the LGR method ($p = 0.05$) ($t = 23.066$, $p < 0.000$ for C4.5; and $t = 23.630$, $p < 0.000$ for CART). Additionally, although CART exhibited slightly greater accuracy than C4.5 did, it lacks of significance for $p = 0.05$ ($t = -1.886$, $p < 0.000$). These findings show that compared to LGR, the DT-based early warning system provides greater accuracy in characterizing student learning performances.

In classification problems, a prediction model with high accuracy does not necessarily exhibit the same behavior in both Type I and Type II errors. A Type I error means that the probabilities that a student with good learning performance is misclassified as one with poor learning performance. If Type I error is too high, then the school will waste resources on students who do not need special attention. A Type II error represents the probability that a student with poor learning performance is misclassified as performing well. If Type II error is too high, then the early warning system is unable to correctly identify at-risk students, and therefore cannot provide advance warning of low performance.

The C4.5 and CART Type I error performance was better than that of LGR. Type I error for C4.5 and CART was less than 5% over three groups of experiments, while LGR produced significantly greater Type I errors for all three datasets. Both C4.5 and CART produced less than 10% Type II errors, and these methods produced significantly better predictions than those of LGR. CART produced slightly less Type I and Type II errors than C4.5 for both T4W and T8W, although the error rate for CART was greater than that of C4.5 for T13W. Our results show that as expected, the accuracy of predicting student learning performance increases as the semester progresses. Because LGR failed to perform satisfactorily at identifying students with poor learning achievements, the LGR classifier was excluded from all subsequent experiments.

To further investigate the influence of time-dependent variables in early warning systems, we generate three datasets by removing all time-dependent variables from the original datasets. Specifically, the three datasets, NT4W, NT8W, and NT13W, were generated by removing the ReadTime, ReadTimeAVG, Course_LoginTime, and Course_LoginAVGTime variables from the respective groups T4W, T8W, and T13W. The results are shown in Table 5. The accuracies of C4.5 in NT4W, NT8W, and NT13W were 0.941, 0.944, and 0.913, respectively. Compared to the corresponding results in T4W, T8W, and T13W, we found that considering time-dependent variables produced significantly more accurate forecasts at the 0.05 level ($t = -23.948$, $p < 0.000$). On the other hand, the CART experimental results were similar to those of C4.5. Compared to datasets without time-dependent variables, applying CART to T4W, T8W, and T13W increased accuracy by approximately 1–4%, which is statistically significant at the 0.05 level ($t = -15.811$, $p < 0.000$). Our experimental results are in agreement with literature reports, and show the existence of correlations between time-dependent variables and online

Table 4
Comparative results of the three datasets.

	T4W			T8W			T13W		
	C4.5	LGR	CART	C4.5	LGR	CART	C4.5	LGR	CART
Accuracy	0.934	0.769	0.950	0.951	0.754	0.952	0.953	0.867	0.949
Type I error	0.004	0.219	0.002	0.002	0.229	0.002	0.003	0.107	0.002
Type II error	0.128	0.244	0.103	0.099	0.261	0.096	0.093	0.162	0.103

Table 5
Comparative influence of time attributes on prediction.

	C4.5		CART	
	T4W	NT4W	T4W	NT4W
Accuracy	0.934	0.941	0.950	0.937
Type I error	0.004	0.000	0.002	0.000
Type II error	0.128	0.123	0.103	0.130
	T8W	NT8W	T8W	NT8W
Accuracy	0.951	0.944	0.952	0.940
Type I error	0.002	0.001	0.002	0.000
Type II error	0.099	0.114	0.096	0.122
	T13W	NT13W	T13W	NT13W
Accuracy	0.953	0.913	0.949	0.907
Type I error	0.003	0.033	0.002	0.034
Type II error	0.093	0.112	0.103	0.125

Table 7
Variable selection rating for the 3 groups.

Variables	T4W	T8W	T13W	Average ranking
Course_LoginTime[a]	6.24	6.10	6.54	1
ReadTimeDOCCount	6.70	4.62	7.26	2
Course_LoginAVGTime[a]	5.73	5.59	6.67	2
ReadTimeAllCount	3.77	4.94	6.13	5
ReadDocRate	3.46	4.24	7.10	7
Course_LoginAllCount	3.58	5.58	5.57	5
ReadTime[a]	4.19	4.97	5.34	4
ReadTimeAVG[a]	4.12	3.99	5.30	8
ReadTimeDOCAVG	4.34	3.80	5.09	8
Forum_JoinRate	1.15	3.09	4.05	10
Forum_Reply	0.80	3.30	3.73	11
Hwk_NonPayment	0.12	0.12	2.78	12
Hwk_Delay	0.00	0.00	0.00	13

[a] *Note:* indicates time-dependent variables.

learning performance. Our results also confirm that, regardless of classification techniques, we can increase the accuracy of an early warning system in identifying at-risk students by considering time-dependent variables.

We employed classifier ensembles to further enhance the predictive power of the applied classification techniques. Adaboost, a popular classifier ensemble, can integrate with several other supervised learning algorithms. Here, we consider two classifier ensembles, AdaBoost + C4.5 and AdaBoost + CART. The results are shown in Table 6.

AdaBoost + C4.5 exhibited accuracies of 0.972, 0.977, and 0.979 in T4W, T8W, and T13W, respectively, producing significantly better performance than C4.5 at the 0.05 level ($t = 17.945$, $p < 0.000$). For the same three datasets, AdaBoost + CART exhibited accuracies of 0.972, 0.978, and 0.980, respectively, significantly better than the experimental results for CART alone (as shown in Table 5), which were also statistically significant at the 0.05 level ($t = 15.133$, $p < 0.000$). We compared Adaboost + CART and Adaboost + C4.5. The accuracy of AdaBoost + CART is not significantly better than that of AdaBoost + C4.5 at the 0.05 level ($t = -0.597$, $p = 0.540$), however, on the important Type II error, the AdaBoost + CART produced lower error rates than Adaboost + C4.5 did for all three datasets, demonstrating that Adaboost + CART's classification performance is the more stable. The classification performance comparisons between the selected classifiers demonstrated that the incorporation of the AdaBoost ensemble technique improves early-warning system performance; the best classifier identified by this study was provided by combining CART with AdaBoost. In addition to the comparisons of prediction performance of these classifiers, we evaluated the importance of each input variable. For each generated dataset, we used Weka 3.7.3 to calculate scores for all input variables based on the gain ratio for each input variable and the associated output variable. Table 7 lists the scores averages for thirty generated datasets from T4W, T8W, and T13W. The most significant input variables for the construction of the early warning system are Course_LoginTime, ReadTimeDOCCount, Course_LoginAVGTime, and ReadTime. With the exception of ReadTimeAVG, all the time-dependent variables shown in Table 3 are essential to the identification of poor learning. Therefore, we propose that early warning systems incorporate these variables to gain more accurate predictions.

Table 6
Experimental results of combining AdaBoost techniques.

Datasets	Evaluation	AdaBoost + C4.5	AdaBoost + CART
T4W	Accuracy	0.972	0.972
	Type I error	0.007	0.009
	Type II error	0.049	0.048
T8W	Accuracy	0.977	0.978
	Type I error	0.000	0.000
	Type II error	0.047	0.045
T13W	Accuracy	0.979	0.980
	Type I error	0.000	0.000
	Type II error	0.042	0.041

5. System development and evaluation

5.1. Development of early warning system

Based on the evaluation findings for various datasets, an early warning prototype system was developed. Interactions between the early warning system and the LMS are shown in Fig. 2. The proposed system comprised three main modules: a data mining engine, a knowledge base, and an inference engine. The data mining engine was responsible for generating early warning rules; as an example, decision rules induced by the Adaboost + CART algorithm at Weeks 4, 8, and 13 are summarized in Table 8. Only certain critical rules are reported in this paper because of space limitations. The knowledge base is a repository for the knowledge that has been verified by domain experts. The knowledge base is

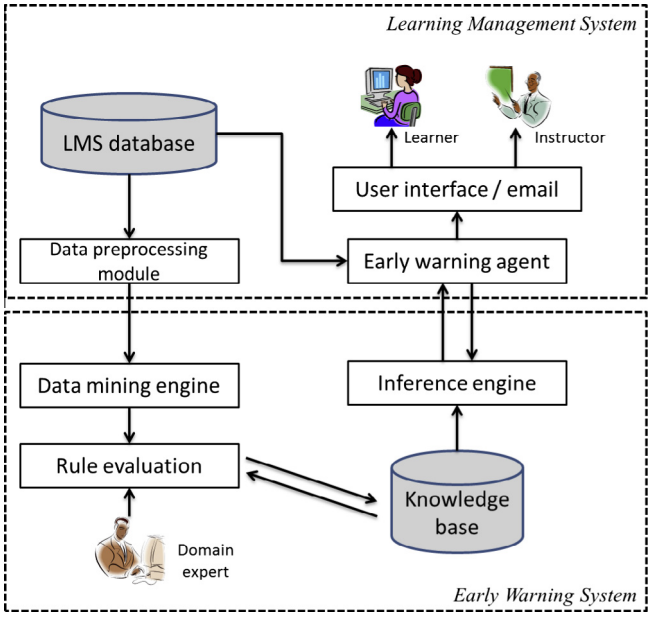

Fig. 2. System structure.

Table 8
The partial decision rules extracted by the decision tree technique.

Dataset	No.	Decision rule
T4W	1	**If** ReadTimeDOCCount ⩾ 2.5 & ReadTime < 25 **then** Class = fail
	2	**If** ReadTimeDOCCount < 2.5 & ReadTimeAllCount < 4.5 & ReadDocRate < 0.275 **then** Class = fail
	3	**If** ReadTimeDOCCount ⩾ 2.5 & ReadTimeAllCount < 4.5 & ReadDocRate < 0.275 **then** Class = fail
T8W	4	**If** Course_LoginCountAll ⩽ 10 & ReadTimeAVG < 4978 & Forum_JoinRate < 0.595 & ReadTimeDOCAVG < 1858 **then** Class = fail
	5	**If** 10 < ReadTimeAllCount ⩽ 16 & ReadTimeAllCount ⩽ 9 **then** Class = fail
	6	**If** ReadTimeAllCount > 16 & ReadDocRate ⩽ 0.14 **then** Class = fail
T13W	7	**If** ReadDocRate ⩽ 0.33 **then** Class = fail
	8	**If** 0.33 < ReadDocRate ⩽ 0.56 & ReadTimeAVG ⩽ 323 then Class = fail
	9	**If** 0.33 < ReadDocRate ⩽ 0.56 & ReadTimeAVG > 323 & ReadTimeAllCount > 26 **then** Class = fail

easily built, because the knowledge extracted by the decision tree takes the form of if-then rules. The inference engine, also known as the rule interpreter, is a computer program for determining learning outcomes (i.e., pass or fail) based on a student's learning portfolio. To this end, in the LMS, an early warning agent was developed. Each time a student's learning portfolio is updated in the LMS, current learning records are summarized by the early warning agent, and the student's information are transmitted to the inference engine to evaluate learner performance. The inference engine analyzes the data, and promptly identifies poor learners, when their learning records match the corresponding decision rules. In the meanwhile, the early warning agent automatically notifies both poor learner and instructor through the email and the user interface in LMS. For example, in Week 13, if the ReadDocRate of a learner was between 0.33 and 0.56, and their average time of material viewed (ReadTimeAVG) was lower than 323 s, a "fail the course" alert would be sent to both the instructor and the learner (according to decision rule No. 8 in Table 8).

5.2. Mechanism for learning improvement

Fig. 3 illustrates the learning improvement mechanism. The learning improvement mechanism focuses on several strategies from the inference engine demonstrated in the earning warning system. During the semester, the earning warning system automatically judges at-risk students and sends warning alert through emails. At the same time, instructor also gets the warning alert email showing those at-risk students. According to the early-warning forecasts generated by the system, instructors can timely adjust their teaching methods, or adopt adaptive teaching approaches to meet the needs of students with poor learning performance as the course proceeds. Both poor performance learners and the instructor can book office hour to have teaching assistant counseling or instructor tutoring to assist students in need of additional support and guidance. After the guidance by teaching assistant or instructor, if the at-risk students have mental issues or other problems cannot be worked out by either teaching assistant or instructor, the earning warning system then inform the counseling center and notify the parents of at-risk students to have further supports from professional counseling psychologist and parents. Through the mechanisms described above, the overall teaching quality and learning performance can be improved greatly and reduce dropouts.

Figs. 4 and 5 demonstrate two screenshots of the early warning system for learners and instructors, respectively. As shown in Fig. 4, when a student logs into the LMS, his/her current learning status can be displayed by the dashboard (in the left of the screen). The key performance indicators (KPIs) on the dashboard are the input variables with high gain ratio, which were automatically calculated by data mining engine. When a learner clicks on one of the KPIs on the dashboard, the learning history will display in the right of the screen. A learner can compare his or her learning history

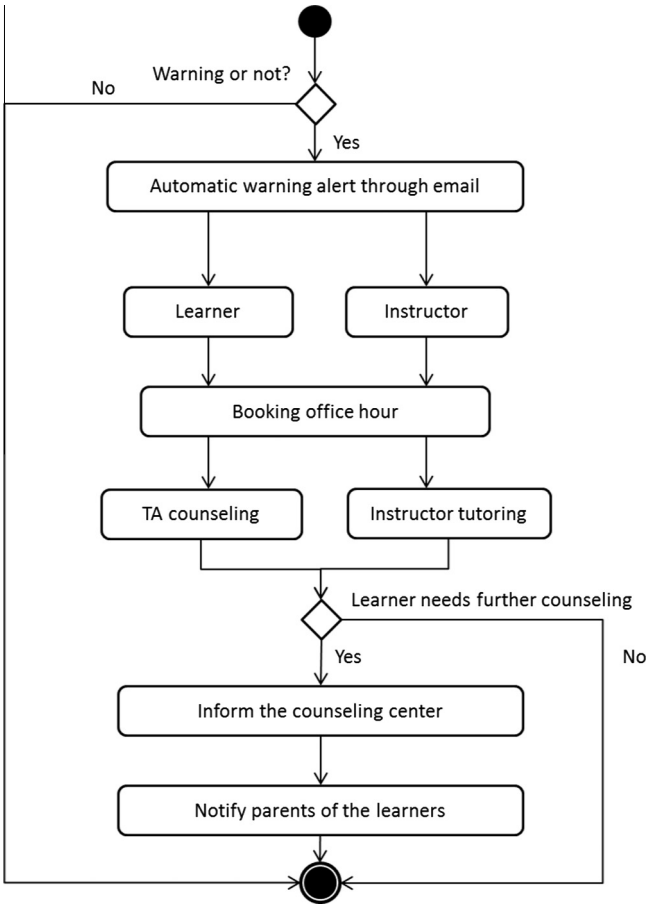

Fig. 3. Mechanism for learning improvement.

with class-wide statistics regarding the best, the average, and the worst values of the selected KPI. This will help students to self-monitor their own learning performance. Consequently, the system can facilitate the identification of at-risk students at any time point, considerably minimizing the negative effects of poor learning. In addition, as shown in Fig. 5, the current learning status of the class is listed at the top of the screen. After selecting a KPI in the table, the instructor can view the learning history in a box-and-whisker plot, revealing the range and the quartile information regarding the selected KPI in the class.

5.3. System usability survey

To evaluate the usability of the system developed in this study, we designed a subjective questionnaire (Aparicio, De Buenaga, Rubio, & Hernando, 2012). Seven items were developed with a

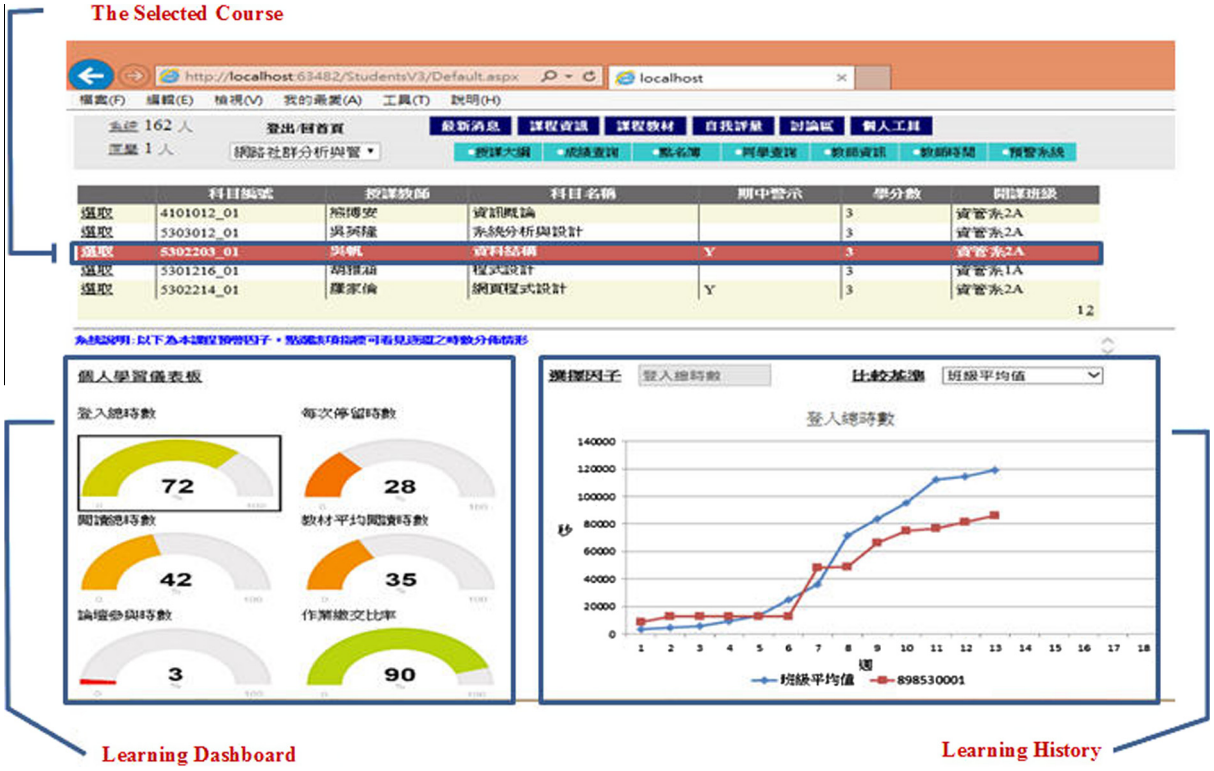

Fig. 4. Screenshot of early warning system for learners.

Fig. 5. Screenshot of early warning system for instructors.

five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The survey results are shown in Table 9. Finally, a total of 40 participants were invited to complete the survey. The sample included 30 undergraduate students and 10 instructors. Of the students who participated in the system usability survey, 16 (53.33%) were female and 14 (46.67%) were male; 5

Table 9
Survey results.

Statements		Answers				
		Strongly disagree	Disagree	Neutral	Agree	Strongly agree
Learning	The system has helped student/instructor to extract useful learning status of the class	0(0)	0(0)	6(1)	17(4)	7(5)
	The system has helped student/instructor to reduce the time needed to understand the learning status, as compared with no early warning system	0(0)	0(0)	2(2)	18(4)	10(4)
Usability	The interface is user-friendly	0(0)	0(0)	5(2)	14(3)	11(5)
	It's easy to find the information I need	0(0)	0(0)	6(1)	15(5)	9(4)
	I feel comfortable using the system	0(0)	0(0)	6(1)	15(4)	9(5)
	The system speed is reasonable	0(0)	0(0)	5(2)	18(4)	7(4)
	It's easy to learn its use	0(0)	0(0)	9(1)	14(3)	7(6)

Note: Numbers in parentheses represent the survey results of instructors.

instructors were female, and 5 instructors were male. None of the participants indicated that the system was difficult to learn or use. Most students and instructors agreed that the system was able to extract useful information in a timely manner regarding learning status. Furthermore, most students and instructors agreed that the system interface was user-friendly, information was easy to find, and the system was comfortable, relatively fast, and easy to use. Specifically, 82.5% felt that the interface was user-friendly (40% strongly agree, 42.5% agree), 82.5% believed that information was easy to find (32.5% strongly agree, 50% agree), 82.5% felt comfortable using the system (35% strongly agree, 47.5% agree), 82.5% perceived that the system speed was reasonable (27.5% strongly agree, 55% agree), and 75% of participants found the system easy to use (32.5% strongly agree, 42.5% agree).

6. Conclusion

E-Learning has become a popular study method in recent years. Previous studies have proposed the concept of an early warning system through analyzing various samples. For example, some used physical classes activities as samples (e.g. Hwang et al., 2003), others used demographic attributes and performance attributes as samples (Kotsiantis et al., 2003), and still others used LMS data after the end of the course (e.g. Macfadyen & Dawson, 2010). There have been few reports of the influences of learning behaviors and time-dependent variables on the learning performance prediction. Thus, these studies are unable to conduct timely assessments and verifications using actual online curricular data.

In this study, we selected a purely online course as a research data set and compared several data mining techniques, including C4.5, LGR, and CART to develop an early warning system for an e-Learning environment. Our overall results showed that time-dependent variables are important in identifying student learning performance. Systems that employed time-dependent variables achieved greater prediction accuracy than the LGR system, which did not consider time-related events; both C4.5 and CART methods significantly outperformed conventional LGR. Furthermore, combination with AdaBoost significantly improved the predictive power of both C4.5 and CART. After a short learning period, our early warning system can provide accurate forecasts by analysis of student learning portfolio data as a course of study progresses. Our results demonstrated that the inclusion of data mining techniques is useful in the construction of an early warning system.

Our findings confirm that the LMS platform can provide timely and automated prediction by integrating the prediction model proposed by this study. The early warning system can automatically record student learning behaviors and progress while a course of study is underway. According to the early-warning forecasts generated by the system, instructors can quickly adjust their teaching speed, methods, or adopt adaptive teaching approaches to meet the needs of students with poor learning performance as the course proceeds. Course instructors can also initiate student counseling programs, to assist students in need of additional support and guidance. Therefore, the overall learning quality and learning performance can be improved greatly and reduce dropouts.

There are several limitations that may affect the overall generalizability of this study. First, because our data were derived from a fully online course within a single institution, more curriculum data could be considered to verify the effectiveness of the early warning system. Second, student demographics and previous learning performance data may affect the accuracy of learning performance predictions. Future studies might integrate such features into early warning systems to provide more robust learning performance predictions. Third, we did not practically implement an early warning system for further investigation. Future studies can develop early warning systems with evaluation mechanisms that allow educators to monitor learner performance, and provide prompt feedback to learners. Finally, more advanced supervised learning techniques might be considered to construct more accurate classifiers.

References

Agrawal, R., & Srikant, R. (1994). *Fast algorithms for mining association rules*. Paper presented at the 20th international conference on very large data bases, Santiago de Chile, Chile.

Anozie, N., & Junker, B. W. (2006). *Predicting end-of-year accountability assessment scores from monthly student records in an online tutoring system*. Paper presented at the AAAI workshop on educational data mining, Menlo Park, CA.

Aparicio, F., De Buenaga, M., Rubio, M., & Hernando, A. (2012). An intelligent information access system assisting a case based learning methodology evaluated in higher education with medical students. *Computers & Education, 58*(4), 1282–1295.

Baker, R. (2010). Data mining for education. In P. Peterson, E. Baker, & B. McGaw (Eds.). *International encyclopedia of education* (Vol. 7, pp. 112–118). Oxford, UK: Elsevier.

Baker, R., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining, 1*(1), 3–17.

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. Boca Raton, FL, USA: Chapman & Hall/CRC.

Calvo-Flores, M. D., Galindo, E. G., Jiménez, M. C. P., & Pérez, O. P. (2006). *Predicting students' marks from moodle logs using neural network model*. Paper presented at the 4th international conference on multimedia and information and communication technologies in education, Seville, Spain.

Campbell, J., & Oblinger, D. (2007). Academic analytics. *Educause Center for Applied Research, 42*(4), 40–57.

Castro, F., Vellido, A., Nebot, A., & Mugica, F. (2007). Applying data mining techniques to e-learning problems. In R. A. Tedman & D. K. Tedman (Eds.). *Evolution of teaching and learning paradigms in intelligent environment* (Vol. 62, pp. 183–221). Berlin, Heidelberg: Springer-Verlag.

Chang, Y. C., Kao, W. Y., Chu, C. P., & Chiu, C. H. (2009). A learning style classification mechanism for e-learning. *Computers & Education, 53*(2), 273–285.

Chen, C. M., Chen, M. C., & Li, Y. L. (2007). *Mining key formative assessment rules based on learner profiles for web-based learning systems*. Paper presented at the IEEE International conference on advanced learning technologies, Niigata, Japan.

Dringus, L. P., & Ellis, T. (2005). Using data mining as a strategy for assessing asynchronous discussion forums. *Computers & Education, 45*(1), 141–160.
Duan, L., & Da Xu, L. (2012). Business intelligence for enterprise systems: A survey. *Industrial Informatics, IEEE Transactions on, 8*(3), 679–687.
Dunham, M. H. (2002). *Data mining: Introductory and advanced topics*. New Jersey, USA: Prentice Hall.
Essa, A., & Ayad, H. (2012). *Student success system: Risk analytics and data visualization using ensembles of predictive models*. Paper presented at the 2nd international conference on learning analytics and knowledge, Vancouver.
Etchells, T., Nebot, A., Vellido, A., Lisboa, P., & Mugica, F. (2006). *Learning what is important: Feature selection and rule extraction in a virtual course*. Paper presented at the 14th European symposium on artificial neural networks, Bruges, Belgium.
Freund, Y., & Schapire, R. (1996). *Experiments with a new boosting algorithm*. Paper presented at the 13th international conference on machine learning, Bari, Italy.
Gaudioso, E., & Talavera, L. (2006). Data mining to support tutoring in virtual learning communities: Experiences and challenges. In R. C. & V. S. (Eds.), *Data mining in learning* (pp. 207–225). Southampton, Boston: WIT Press.
Gaudioso, E., Montero, M., & Hernandez-Del-Olmo, F. (2012). Supporting teachers in adaptive educational systems through predictive models: A proof of concept. *Expert Systems with Applications, 39*(1), 621–625.
Goldstein, P. J., & Katz, R. N. (2005). *Academic analytics: The uses of management information and technology in higher education*. Louisville: Educause.
Guruler, H., Istanbullu, A., & Karahasan, M. (2010). A new student performance analysing system using knowledge discovery in higher educational databases. *Computers & Education, 55*(1), 247–254.
Hanna, M. (2004). Data mining in the e-learning domain. *Campus Wide Information Systems, 21*(1), 29–34.
Hwang, G. J., Hsiao, C. L., & Tseng, J. C. R. (2003). A computer-assisted approach to diagnosing student learning problems in science courses. *Journal of Information Science and Engineering, 19*(2), 229–248.
Kotsiantis, S. B. (2012). Use of machine learning techniques for educational proposes: A decision support system for forecasting students' grades. *Artificial Intelligence Review, 37*(4), 331–344.
Kotsiantis, S. B., & Pintelas, P. E. (2005). *Predicting students marks in hellenic open university*. Paper presented at the 5th IEEE international conference on advanced learning technologies, Kaosiung, Taiwan.
Kotsiantis, S. B., Pierrakeas, C. J., Zaharakis, I. D., & Pintelas, P. E. (2003). *Efficiency of machine learning techniques in predicting studentsā performance in distance learning systems*. Paper presented at the Symposium on recent advances in mechanics, Athens, Greece.
Kotsiantis, S. B., Pierrakeas, C. J., & Pintelas, P. E. (2003). Preventing student dropout in distance learning using machine learning techniques. *Lecture Notes in Computer Science, 2774*, 267–274.
Kotsiantis, S. B., Pierrakeas, C. J., & Pintelas, P. E. (2004). Predicting students' performance in distance learning using machine learning techniques. *Applied Artificial Intelligence, 18*(5), 411–426.
Lust, G., Elen, J., & Clarebout, G. (2013). Students' tool-use within a web enhanced course: Explanatory mechanisms of students' tool-use pattern. *Computers in Human Behavior, 29*(5), 2013–2021.
Macfadyen, P. L., & Dawson, S. (2010). Mining lms data to develop an "early warning system" for educators: A proof of concept. *Computers & Education, 54*(2), 588–599.
Minaei-Bidgoli, B., Kashy, D. A., Kortmeyer, G., & Punch, W. F. (2003). *Predicting student performance: An application of data mining methods with an educational web-based system*. Paper presented at the 33rd annual frontiers in education conference, Westminster.
Minaei-Bidgoli, B., Tan, P. N., & Punch, W. F. (2004). *Mining interesting contrast rules for a web-based educational system*. Paper presented at the International conference on machine learning and applications, Louisville, KY, USA.
Mostow, J., Beck, J., Cen, H., Cuneo, A., Gouvea, E., & Heiner, C. (2005). *An educational data mining tool to browse tutor-student interactions: Time will tell*. Paper presented at the Educational data mining, Muehlenbrock.
Muehlenbrock, M. (2005). *Automatic action analysis in an interactive learning environment*. Paper presented at the international conference on artificial intelligence in education, Amsterdam.
Quinlan, J. R. (1993). *C4. 5: Programs for machine learning*. San Francisco, CA, US: Morgan kaufmann.
Rasmussen, K., Northrup, P., & Lee, R. (1997). Implementing web-based instruction. In B. H. Khan (Ed.), *Web-based instruction* (pp. 341–346). Englewood Cliffs: Educational Technology Publications.
Romero, C., Ventura, S., & García, E. (2008). Data mining in course management systems: Moodle case study and tutorial. *Computers & Education, 51*(1), 368–384.
Sadler-Smith, E. (2001). The relationship between learning style and cognitive style. *Personality and Individual Differences, 30*(4), 609–616.
Sumner, M., Frank, E., & Hall, M. (2005). *Speeding up logistic model tree induction*. Paper presented at the 9th European conference on principles and practice of knowledge discovery in databases, Porto, Portugal.
Tan, P., Kumar, V., & Steinbach, M. (2006). *Introduction to data mining*. Boston, MA, USA: Pearson Addison Wesley.
Valsamidis, S., Kontogiannis, S., Kazanidis, I., Theodosiou, T., & Karakos, A. (2012). A clustering methodology of web log data for learning management systems. *Journal of Educational Technology & Society, 15*(2), 154–167.
Wang, A., & Newlin, M. (2002). Predictors of performance in the virtual classroom: Identifying and helping at-risk cyber-students. *Transforming Education through Technology Journal, 29*(10), 21–25.
Wang, A. Y., Newlin, M. H., & Tucker, T. L. (2001). A discourse analysis of online classroom chats: Predictors of cyber-student performance. *Teaching of Psychology, 28*(3), 222–226.
Wu, H. (2013). Examining students' online interaction in a live video streaming environment using data mining and text mining. *Computers in Human Behavior, 29*(1), 90–102.
Yu, C., Jannasch-Pennell, A., Digangi, S., & Wasson, B. (1998). Using sas/intrnet for evaluating web-based courses. *Educational Media International, 35*(3), 157–161.
Yu, F.-Y., & Wu, C.-P. (2013). Predictive effects of online peer feedback types on performance quality. *Journal of Educational Technology & Society, 16*(1), 332–341.
Zubizarreta, J., & Millis, B. J. (2009). *The learning portfolio: Reflective practice for improving student learning*. San Francisco, CA: Jossey-Bass.