

Analyzing Teaching Performance of Instructors Using Data Mining Techniques

Sona MARDİKYAN, Bertan BADUR

*Department of Management Information Systems, Bogazici University
Istanbul, Turkey
e-mail: mardikya@boun.edu.tr, badur@boun.edu.tr*

Received: July 2011

Abstract. Student evaluations to measure the teaching effectiveness of instructor's are very frequently applied in higher education for many years. This study investigates the factors associated with the assessment of instructors teaching performance using two different data mining techniques; stepwise regression and decision trees. The data collected anonymously from students' evaluations of Management Information Systems department at Bogazici University. Additionally, variables related to other instructor and course characteristics are also included in the study. The results show that, a factor summarizing the instructor related questions in the evaluation form, the employment status of the instructor, the workload of the course, the attendance of the students, and the percentage of the students filling the form are significant dimensions of instructor's teaching performance.

Keywords: students' evaluation of teaching, educational data mining, stepwise regression, decision trees.

1. Introduction

In recent years, there has been a widespread interest in the student evaluations of teaching (SET) to measure teaching effectiveness of instructors' and quality of teaching. In its early applications, universities and colleges used the result of evaluations to monitor teaching quality and to help teachers improve their teaching effectiveness. However, in recent years SET results are also utilized for other purposes by providing information to different parties in educational institutions (Kulik, 2001). Administrators use ratings in hiring new instructors, in promotion and tenure decisions, in selecting faculty and graduate students for teaching awards, and in assigning teachers to courses. Instructors use SET results to improve their teaching effectiveness and in monitoring the performance of their graduate student assistants. Students use the ratings in selecting courses and selecting teachers for awards.

Although various research studies are conducted to support the validity, reliability, and usefulness of student ratings, it is still subject to long lasting debates as no general agreement has been reached about these issues. One contentious issue is the relationship between the evaluations and learning. Clayson (2009) makes a review of the literature by

performing a meta-analysis. The findings show that a small average relationship exists between learning and the evaluations but the association is situational and not applicable to all teachers, academic disciplines, or levels of instruction. It is concluded that the more objectively learning is measured, the less likely it is related to the evaluations.

Recent researches have explored the effect of a variety of factors on ratings. In those researches there is a general agreement that student ratings are multidimensional but there is disagreement about how many or which dimensions should be used. Marsh and Roche (2000) claimed that SETs are multidimensional, reliable, relatively valid, useful for teaching improvement, and relatively unaffected by biasing factors such as class size, grading tolerance, workload. Cashin (1995) attempted to summarize the conclusions of the major reviews of the SET research covering 67 studies. Referring to several factor analytic studies (Feldman, 1976; Marsh and Dunkin, 1992; Centra, 1993; Braskamp and Ory, 1994) Cashin concluded that student rating forms are multidimensional; they measure several different aspects of teaching. Both Centra (1993) and Braskamp and Ory (1994) identified six factors found in student rating forms as; course organization and planning, clarity and communicating skills, teacher student interaction, course difficulty and workload, grading and examinations, students self-rated learning. Marsh (1984) SEEQ (Student Evaluation of Educational Quality) form has nine dimensions; learning/value, enthusiasm, organization, group interaction, individual rapport, breadth of coverage, exams/grades, assignments, and workload. However, Feldman (1976) identified 28 different dimensions.

A study by Jacobs (2002) classified the factors under four broad categories as course, student, instructor, and administration characteristics. For each of these groups Jacobs identifies the relevant and irrelevant factors to explain the teaching performance of instructors. The following paragraphs present studies related to each of these main categories.

In the course characteristics main group, class size, discipline, reason for taking course, course level and difficulty level of the course are presented as related factors. Whereas, time of the lecture is found to be an irrelevant factor (Jacobs, 2002). Williams and Ory (1992) found a negative but small correlation between class size and various rating items. The negative correlation indicates that smaller classes are likely to receive higher ratings, but the small magnitude of the correlation indicates that class size is not a very important factor affecting the validity of ratings. In a recent study, Bedard and Kuhn (2008) found a statistically significant nonlinear negative impact of class size on students' evaluation of instructors' effectiveness which is highly robust to the course and instructor fixed effects. Ratings show a substantial variation with respect to the instructor's discipline. Cashin and Sixbury (1992) have shown that the highest ratings are given to courses in the arts and humanities, followed in descending order by biological and social sciences, business and computer science, mathematics, engineering, and physical sciences. Courses in the major field of students or electives tend to get slightly higher ratings compared to the required core courses that are aimed to gain a background to students. Feldman (1978) found a small positive relationship between class ratings and the students' average intrinsic interest in the subject area. Thus, core courses may receive lower ratings as students are less interested in them. Ratings in higher level courses are

likely to be higher than in lower level courses. Within a discipline, the courses that are more difficult or have greater workloads tend to receive higher ratings from students. Contrary to popular opinion, easy professors do not necessarily get high ratings (Cashin and Sixbury, 1992).

In the student characteristics main group expected grade, motivation, major, and gender are related factors. Whereas academic ability, age, class level, general point average, and personality are not related factors to ratings (Jacobs, 2002). The research findings about the relation between expected grade and SET are controversial. Positive but low correlations have been reported between expected grade and student ratings in some studies; however it is reported that students who receive better grades give better evaluations (Marsh and Dunkin, 1992; Krautmann and Sander, 1999; Clayson, 2004). Students can distinguish instructors on the basis of how much they have learned. If students are motivated because of a prior interest in the subject matter or because they chose the class as an elective, then they give more likely higher ratings (Cashin, 1988). Early research indicates that there was no relationship between gender and ratings. However, later research by Lueck *et al.* (1993) found evidence that students tend to give same-gender instructors slightly higher ratings.

In the instructor characteristics main group faculty rank, personality, and research productivity are related factors. Not related factors are age of instructor, years of teaching experience, and gender of the instructor. According to Jacobs (2002), regular faculty tends to receive higher ratings than teaching assistants. Certain personality traits of an instructor may be related to students' overall ratings. Research shows that students appreciate instructors who are knowledgeable, warm, outgoing, and enthusiastic (Murray *et al.*, 1990). Research productivity measured by the number of publications is positively but only slightly correlated with student ratings.

Factors related to administration characteristics are instructor's presence in room during the evaluation, time of evaluation, and student anonymity (Jacobs, 2002).

This study is aimed at identifying the factors associated with the teaching performance of instructors of the Management Information Systems (MIS) Department of Bogazici University. Knowledge Discovery in Databases (KDD) methodology is followed throughout the study. Relevant data is collected anonymously from students, on all the courses offered by the MIS department during the period 2004–2009. The evaluation form in use distinguishes two groups of questions: related to the course and related to the instructor. First, factor analysis is applied to the questionnaire data to uncover independent factors affecting the instructor's overall teaching performance. Together with the factors derived from the questionnaire, variables related to other instructor and course characteristics are also included in the study. Although there are numerous studies conducted about SET, it is not encountered any study applying data mining techniques to this problem. In our study, two well known data mining techniques; stepwise regression and decision tree techniques are applied to the data. The regression results are compared with the outputs obtained from two decision tree algorithms: CHAID (Chi-Squared Automatic Interaction Detection) and CART (Classification and Regression Trees). As a multidisciplinary field, MIS education has some distinguishing characteristics: It offers a wide range of

qualitative and quantitative courses and the practically based courses are offered by part time practitioners. The empirical findings of this study emphasize underlying factors of various aspects of the MIS education based on students' evaluations.

Rest of the study is organized as follows. In Section 2 KDD methodology, and data mining techniques used in the study are provided. In Section 3, the data set is described; preparation and preprocessing steps are explained. Section 4 presents the results of stepwise regression and decision trees. Finally Section 5 concludes with further research directions.

2. Methodology

In order to obtain useful knowledge from data, KDD methodology is used. The process of KDD has been defined by Fayyad *et al.* (1996) as "the non-trivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data". KDD process covers mainly goal identification, data selection, data cleaning, data integration, data transformation, data mining, pattern evaluation and knowledge presentation steps.

Data mining as being the heart of these steps is about extracting a pattern with relevant features from large amounts of data. Data mining tasks are generally classified into two broad categories as descriptive and predictive (Han and Kamber, 2006). Two types of predictive data mining tasks can be performed; classification and prediction. Classification is the process of finding a model that describes and distinguishes data classes for the purpose of predicting the discrete target variable whose class label is unknown. However, in prediction the target variable is continuous. Regression analysis, time series are examples for prediction techniques. Decision trees, neural networks and support vector machines can be used for both classification and prediction tasks. Han and Kamber (2006) defined decision tree as a flowchart-like tree structure. The widespread use of decision trees can be accounted for their explanation capability, easy trainability and free of statistical assumptions. The rules extracted from decision trees can easily be understood by even non specialists. They can also be incorporated into decision support systems.

In this study, data are available for a large number of variables as suggested by the literature. To reduce the data and to obtain a concise model for explaining the teaching performance two data mining techniques; regression analysis and decision trees are utilized. For regression analysis the stepwise regression method and for decision trees CHAID and CART algorithms are applied. The same set of candidate independent variables are used as inputs in both methods. Stepwise regression procedure is one of the variable selection methods in regression analysis. It is an iterative procedure to identify which independent variables provide the best model (Kleinbaum *et al.*, 2008). The CHAID (Kass, 1980) algorithm uses statistical Pearson chi-square for nominal outputs, likelihood-ratio for ordinal outputs and F test for continuous outputs, but it does not perform any pruning method. The CART (Breiman *et al.*, 1984) algorithm generates binary trees using Gini and twoing splitting criteria and cost-complexity pruning.

3. Data

Student course evaluations are conducted during the last two weeks of each semester by the registration office in Bogazici University. According to the university policy, the undergraduate courses including less than 8 students and the graduate courses including less than 6 students are not evaluated. Moreover, for a course when the number of student evaluators falls below 8 (6 for graduate courses), the evaluation results about this course is not reported.

3.1. Data Description and Preparation

The data set used in this study includes all reported MIS graduate and undergraduate course sections offered from Fall 2004 to Summer Term 2009 except project, seminar, and graduate foundation courses. After this exclusion the final data set contains totally 259 course sections offered by both fulltime and part-time instructors of MIS department.

The data include two basic categories of variables. The first group consists of the data obtained from evaluation questionnaires where students response anonymously 13 questions about course (Q1–Q6) and instructor (Q7–Q13) characteristics. The 5-point Likert scale is used in this questionnaire. We include all the questions except Question 6 which is intended to measure the performance of the teaching assistant. Questions of the evaluation form are given in Table 1.

Table 1
Questions of the evaluation form

COURSE	
Q1:	What is your current Grade Point Average (GPA) (5:3.50-4.00, 4:3.00-3.49, 3:2.50-2.99, 2:2.00-2.49, 1:00-1.99)
Q2:	What proportion of classes did you attend in this course? (5: all, 1: none)
Q3:	Were the materials, e.g. textbooks/handouts prescribed satisfactory and useful? (5:yes, 1:no)
Q4:	Compared to other courses you have taken at BU how much did you have to work for this course? (5:very much, 1:very little)
Q5:	Compared to other courses you have taken at BU how much did you learn from this course? (5:very much, 1:very little)
Q6:	Did the laboratory work meet the needs of the course? (5:yes, 1:no)
THE INSTRUCTOR	
Q7:	Apart from classroom contact, was available to give outside help. (5:yes, 1:no)
Q8:	Graded exams and other materials fairly. (5:yes, 1:no)
Q9:	Graded exams and other materials on time. (5:yes, 1:no)
Q10:	Had a well-prepared course outline (syllabus). (5:yes, 1:no)
Q11:	Spoke English correctly and understandably/clearly. (5:yes, 1:no)
Q12:	How would you rate the overall teaching effectiveness of your instructor? (5: excellent, 1:poor)
Q13:	Would you choose to take another course with the same instructor? (5:yes, 1:no)

Table 2
The second group variables and scales

Variables	Scale
PARTTIME	Nominal (Binary)
FIRST	Nominal (Binary)
SIZE	Ratio
FILLOFSIZE	Ratio (0 to 1)
YEAR12	Nominal (Binary)
YEAR34	Nominal (Binary)
GRAD	Nominal (Binary)
ELECTIVE	Nominal (Binary)
SUMMER	Nominal (Binary)
ANALYTICAL	Nominal (Binary)
BEHAVIORAL	Nominal (Binary)
PROGRAMMING	Nominal (Binary)
DESIGN	Nominal (Binary)

The second group of variables that are chosen mainly based on the SET literature contains information about the additional characteristics of the courses and instructors. There are two variables about the instructor: whether or not the instructor is a part time and giving the course for the first time. The remaining set of variables contains information about the nature of the course; size, the ratio of the number of students filled the questionnaire to the class size, the level of the class (first or second year, third or fourth year, graduate), whether or not the course is elective, and offered in summer term. As a multidisciplinary program, MIS curriculum includes managerial, quantitative, programming, and system development and design type courses. According to Cashin (1992) course evaluations show variations with respect to the field of the course. For this reason, four binary variables (analytical, behavioral, programming, design) representing the major underlying fields of the course are created. These variables are not mutually exclusive. The second group variables with their measurement scales are presented in Table 2.

The student responses to each question were averaged for each course section so as to form 12 different aggregate variables. Since the data does not include missing values, no missing value handling techniques were used. Moreover, neither the Likert scale variables nor the second group variables contain outliers.

3.2. Data Reduction and Transformations

The average of Q12 and Q13 (named as AVERQ1213) is used as the dependent variable, since among all the variables; these are the only two variables that measure the overall teaching performance of the instructor. All the other variables from the questionnaire are treated as candidate independent variables of the models. To reduce the number of independent variables and to understand the patterns of relationship among them the factor analysis was applied. According to the Barlett test and Kaiser–Meyer–Olkin (KMO)

Table 3
Rotated component matrix

	Component		
	1	2	3
Q1	-.028	-.143	.881
Q2	.211	.358	.583
Q3	.837	.147	.031
Q4	-.030	.853	.011
Q5	.836	.263	.152
Q7	.776	.138	.212
Q8	.871	-.172	-.094
Q9	.836	-.110	-.033
Q10	.898	.223	.042
Q11	.731	.292	.192

measure the dataset is adequate for the factor analysis. KMO value is 0.861 (> 0.5) and Berlett test p-value is 0.00. The three components are identified with eigen values exceeding 1.0 which explain 68.65% of the variability. Varimax rotation method was used to obtain the rotated factor loadings. According to Rotated Component Matrix (Table 3) results Q3, Q5, Q7, Q8, Q9, Q10 and Q11 load to the first component, whereas Q4 loads to the second and, Q1 and Q2 load to the third. These results are meaningful since the first component includes all the questions measuring the attitudes of the instructor, the second component primarily loaded by a single question related to the course and, finally the third component reflects student characteristics. The representative variable of the first component (named as COMP1) is formed by taking the average of the questions Q3, Q5, Q7, Q8, Q9, Q10 and Q11. Since only Q4 loads to the second component, it is taken as another separate variable. Although Q1 and Q2 load to the third component, these variables are taken as separate variables because they measure different attitudes of the students.

3.3. Prediction Models and Results

In this study, first we constructed a regression model to predict a measure for instructors' teaching performance using student evaluations and other explanatory variables. As a second prediction model two types of decision tree algorithms were applied to the same data set. The stepwise regression model was estimated using the SPSS 19.0 and the decision trees were built using the AnswerTree 3.1.

4. Regression Analysis

Statistical estimates were obtained by ordinary least squares estimation of a linear model. In addition to the four variables obtained from factor analysis, variables related to other

Table 4
Stepwise regression model

Model	Unstandardized coefficients		t	Sig.	Collinearity statistics	
	B	Std. error			Tolerance	VIF
(Constant)	-.780	.202	-3.851	.000		
COMP1 *	1.213	.036	34.067	.000	.893	1.120
PARTTIME	.257	.044	5.872	.000	.884	1.132
Q4	-.203	.035	-5.782	.000	.856	1.169
FILLOFSIZE	.192	.079	2.450	.015	.980	1.020
Q2	.086	.038	2.258	.025	.883	1.133
R ²	.854					
Adjusted R ²	.851					
SE Estimate	.285					

instructor and course characteristics were also used as independent variables in a regression model to predict the performance of the instructor.

There exist significant correlations among some of the variables that may cause the multicollinearity problem in the linear regression model. To avoid this problem stepwise regression method was applied. The candidate independent variables of the model are as follows: Q1, Q2, Q4, COMP1, PARTTIME, FIRST, SIZE, FILLOFIZE, YEAR12, GRAD, ELECTIVE, SUMMER, ANALYTICAL, BEHAVIORAL, PROGRAMMING, and DESIGN.

The summary statistics of stepwise regression results are presented in Table 4. The variable COMP1 is the first variable entered into the model strongly explaining the performance of the instructor, and then PARTTIME, Q4, FILLOFSIZE, Q2 enter the model sequentially. R² value of the final model is 0.854. The F-test ($F_{258} = 296.432, p = 0.000$) result shows that the model is significant at a level of 0.05 significance level. Therefore, these five variables explain the 85% variability of the instructor teaching performance. All the t-test results of the independent variables are significant for 0.05 significance level. All the coefficients of the variables are positive, except the coefficient of Q4. The attitudes of the instructor (COMP1), the attendance of the student (Q2), the ratio of the number of students filled the questionnaire to the class size (FILLOFSIZE) and being a part time instructor (PARTTIME) positively affect the instructor's performance whereas the workload of the course (Q4) negatively affects the dependent variable.

In this model there is no multicollinearity since stepwise method eliminates this problem. Furthermore, VIF values of the independent variables entered into the final model are all around 1 which also indicates that there is almost no collinearity between these variables (Kleinbaum *et al.*, 2008). To check heteroscedasticity of the residuals, a non-parametric test Spearman Rank Correlation test was applied. Results indicate that there is no significant Spearman's rank correlation coefficient between the absolute value of the residuals of the model and the independent variables. Therefore our model satisfies homoscedasticity assumption. The normality of the residuals was tested applying

Kolmogorov–Smirnov test. According to the results of the test (Kolmogorov–Smirnov $Z = 0.969$, Asymp. Sig. = 0.305) the distribution of residuals is appropriate to the normal distribution. As a result of the diagnostic checks, the model obtained from stepwise method is satisfying all the assumptions of the linear regression model, and can be used to explain the instructor’s teaching performance.

4.1. Decision Tree Analysis

CHAID and CART algorithms were applied by choosing AVERQ1213 as the dependent variable as in the regression analysis. The same accuracy estimation methods; holdout and 10-fold cross validation (Han and Kamber, 2006) were used in the two algorithms. In holdout accuracy estimation method sampling with a ratio of 80% training and 20% test samples were implemented. In 10-fold cross validation method, the data set was randomly partitioned into 10 folds. Overall accuracy of the models was calculated by averaging the test sample error of each fold. In order to assess the goodness of fit, a risk measure was monitored. This risk estimate was based on calculating within-node variance about the mean of each node.

CHAID decision tree models were built for both holdout and 10-fold cross validation with the following parameters; maximum tree depth as 3, minimum number of cases for parent as 10, and minimum number of cases for child as 2. These parameters were used as stopping rules for the decision tree growth. Minimum number of cases for the parent node parameter allows the user to specify minimum number of observations in a parent node. Minimum number of cases for a child node enables to control the minimum number of cases in child nodes. The splitting a node will not be performed if the number of cases in any child node is less than this threshold.

CHAID decision tree model selected COMP1 variables as the most relevant variable to instructor teaching performance with holdout sampling. Average instructor performance is positively related with this variable. The second related variable is PART-TIME which appears in the second level of the tree. Part time instructors’ evaluations are slightly higher than that of the full time faculty members’. The estimated risk measures are summarized in Table 5 for the hold out and 10-fold cross validation sampling methods. According to the estimated risk statistics on the test samples, it can be claimed that the variables selected by the CHAID decision tree model successfully explain the instructors teaching performance.

CART Decision Tree model creates binary decision trees which are more complex than the trees produced by CHAID. CART algorithm performs cost-complexity pruning based on standard error rule or minimum risk criteria (Breiman *et al.*, 1984). In this study, standard error rule pruning was applied which chooses the smallest subtree whose risk is close to that of the subtree with the minimum risk. In addition to pruning, the tree growth was restricted with the following parameters; maximum tree depth as 5, minimum number of cases for parent as 10, and minimum number of cases for child as 2. As CART makes only binary splits the depths of a CART tree is higher than the depth of the CHAID tree. For this reason, the maximum tree depth parameter of the CART model was set at a higher value than that of the CHAID model. The Gini index was used as splitting criterion.

Table 5
The risk estimates of the models

Tree	Sampling	Risk estimate	Standard error of risk estimate
CHAID	Hold out	0.0569	0.0116
	Cross validation	0.0661	0.0073
CART	Hold out	0.0613	0.0125
	Cross validation	0.0639	0.0069

CART decision tree model with holdout sampling selected COMP1 as the most relevant variable as in the regression and CHAID models. Moreover DESIGN, PARTTIME, Q4 AND Q2 variables appeared in the sublevels of the CART tree. According to the prediction rules represented by CART model, the teaching performance of instructors increases as COMP1 variable increases.

The goodness of fit statistics test samples for the hold out and 10-fold cross validation sampling methods are given in Table 5. According to the risk estimates of the cross validation methods figures CART model predicted the teaching performance better than the CHAID model.

4.2. Findings

According to the results of stepwise regression and decision tree analyses the instructors' teaching performance is explained with mostly instructor related questions summarized by the variable COMP1. Instructors, who have well prepared course outlines, use satisfactory materials, help the students outside the lectures, grade exams fairly and on time receive higher evaluations. An additional instructor characteristics; the employment status of the instructor that is not included in the questionnaire is found to be significant. This result may support the claim that the MIS students appreciate the industry practice of the part time instructors. Giving the course for the first time is not found to be a significant factor in explaining the teaching performance. It is understood that instructors are well prepared to their courses even they are giving the course for the first time.

Among the variables related to course characteristics only the workload of the course is a significant factor. This finding is contradictory with the study of Cashin and Sixbury (1992). In MIS department, courses having lower workloads tend to receive higher ratings. Although in the literature, a small but negative impact of the class size is indicated in affecting the instructors' performance (Williams and Ory, 1992; Bedard and Kuhn, 2008), in our study this variable is not found to be an important factor. At the beginning of the study, it is hypothesized that the students would rate courses belonging to different underlying fields of MIS education differently, but the results do not support this claim. Another insignificant factor is the level of the course contrary to the literature (Feldman, 1978). Additionally, courses offered in summer terms and elective courses do not get higher ratings.

From the variables related to student characteristics, student attendance positively affects the teaching performance. Thus, instructors would be advised to encourage more students to attend the courses. Another important factor which positively influences the performance is the percentage of the students that fill the questionnaire. Therefore, more students can be promoted to be involved in the evaluation process. As it is reported in the literature (Jacobs, 2002) the general point average of the student is not a significant factor. This result is supported in our study for the MIS education as well.

5. Conclusions

The student evaluations of their classes and instructors are widespread in higher education. The relationship between student evaluations and the teaching performance of instructors has been debated for many years. This study was conducted to understand the key factors affecting the teaching performance of the instructors in MIS at Bogazici University. KDD methodology was followed through out the study. In modeling step, for the purpose of data reduction and variable selection two widely used data mining techniques; stepwise regression and decision trees (CHAID and CART) were applied.

The stepwise regression results are supported by the decision trees findings. According to these results, the most important factor to explain the instructors' teaching performance is the instructor attitudes that are primarily measured by the evaluation process. This result is not peculiar to the MIS education since the same evaluation form is used in all departments of Bogazici University. The courses offered by part time instructors tend to receive higher ratings in MIS education. Hence, MIS departments should offer a variety of elective courses given by part time instructors. In the course related variables, the workload of the course negatively affects the instructor teaching performance, while the percentage of students filling the questionnaire positively affects the same variable. In addition, the attendance of the student is another important factor that influences positively the performance of the instructor. Hence, the instructors that attract more students to the classes are evaluated more successfully. Although MIS is an interdisciplinary field, the course type (analytical, behavioral, programming and design) considered in this study is not found to be significant. It can be concluded that, students are aware of this specific characteristics of MIS education and do not differentiate the underlying field courses when evaluating the teaching performance of instructors.

This study reveals some important aspects of instructors' teaching performance in MIS and informatics education. The results can be utilized in curriculum design, in assigning teachers to courses, in awarding the instructors. In future, this study can be applied to other disciplines to investigate field specific and general factors explaining teaching performance of the instructors.

References

- Bedard, K., Kuhn, P. (2008). Where class size really matters: class size and student ratings of instructor effectiveness. *Economics of Education Review*, 27(3), 253–265.

- Braskamp, L.A., Ory, J.C. (1994). *Assessing Faculty Work: Enhancing Individual and Institutional Performance*. Jossey-Bass, San Francisco.
- Breiman, L., Friedman, J., Olshen, R.A., Stone, C.J. (1984). *Classification and Regression Trees*. Chapman & Hall, New York.
- Cashin, W.E. (1988). Students ratings of teaching: a summary of the research. *IDEA Paper No. 20*. Manhattan, Center for Faculty Evaluation & Development, Division of Continuing Education, Kansas State University.
- Cashin, W.E. (1995). Student ratings of teaching: the research revisited. *IDEA Paper No. 32*. Manhattan, Center for Faculty Evaluation & Development, Division of Continuing Education, Kansas State University.
- Cashin, W.E., Sixbury, G.R. (1992). Comparative data by academic field. *IDEA Technical Report No. 8*, Kansas State University.
- Centra, J.A. (1993). *Reflective Faculty Evaluation: Enhancing Teaching and Determining Faculty Effectiveness*. Jossey-Bass, San Francisco.
- Clayson, D.E. (2004). A test of reciprocity effect in the student evaluation of instructors in marketing classes. *Marketing Education Review*, 14(2), 11–21.
- Clayson, D.E. (2009). Student evaluations of teaching: are they related to what students learn? A meta-analysis and review of the literature. *Journal of Marketing Education*, 31(1), 16–30.
- Fayyad, U.M., Piatetsky-Shapiro, G., Smyth, P. (1996). From data mining to knowledge discovery: an overview. In: Fayyad, U.M., Piatetsky-Shapiro, G., Smyth, P., Uthurusamy, P.R. (Eds.), *Advances in Knowledge Discovery and Data Mining*. The MIT Press, Massachusetts, 1–34.
- Feldman, K.A. (1976). The superior college teacher from the student's view. *Research in Higher Education*, 5, 243–288.
- Feldman, K.A. (1978). Course characteristics and college students' ratings of their own teachers: what we know and what we don't. *Research in Higher Education*, 9, 199–242.
- Han, J., Kamber, M. (2006). *Data Mining Concepts and Techniques*. 2nd ed., Morgan Kaufmann, San Francisco.
- Jacobs, C.L. (2002). Student ratings of college teaching: what research has to say. <http://www.indiana.edu/~best/multiop/ratings.shtml>.
- Kass, G.V. (1980). An exploratory technique for investigating large quantities of categorical data. *Applied Statistics*, 29(2), 119–127.
- Kleinbaum, D.G., Kupper, L.L., Muller, K.E., Nizam, A. (2008). *Applied Regression Analysis and Other Multivariable Methods*. 4th ed., Thomson Brooks/Cole Publishing, California.
- Krautmann, A.C., Sander, W. (1999). Grades and student evaluations of teachers. *Economics of Education Review*, 18, 56–93.
- Kulik, J.A. (2001). Student ratings: validity, utility, and controversy. In: Theall M., Abrami, P.C. Mets, L.A. (Eds.), *New Directions for Institutional Research*. Jossey-Bass, 109, 9–24.
- Lueck, T.L., Endres, K.L., Caplan, R.E. (1993). The interaction effects of gender on teaching evaluations. *Journalism Educator*, 48(3), 46–54.
- Marsh, H.W. (1984). Students' evaluations of university teaching: dimensionality, reliability, validity, potential biases, and utility. *Journal of Educational Psychology*, 76(5), 707–754.
- Marsh, H.W., Dunkin, M.J. (1992). Students' evaluations of university teaching: a multidimensional perspective. In: Smart, J.C. (Ed.), *Higher Education: Handbook of Theory and Research*. Springer, 8, 143–233.
- Marsh H.W., Roche, L.A. (2000). Effects of grading leniency and low workload on students' evaluations of teaching: popular myth, bias, validity, or innocent bystanders? *Journal of Educational Psychology*, 92(1), 202–228.
- Murrey H.G., Rushton, J.P., Paunonen, S.V. (1990). Teacher personality traits and student instructional ratings in six types of university courses. *Journal of Educational Psychology*, 82(2), 250–261.
- Williams, R., Ory, J.C. (1992). A further look at class size, discipline differences and student ratings. Unpublished manuscript, Office of Instructional Resources, University of Illinois.

S. Mardikyan received her BS degree in control and computer engineering from Technical University of Istanbul, Turkey, MS degree in industrial engineering from Bogazici University, and PhD in quantitative methods from Istanbul University. She is working in Management Information Systems Department of Bogazici University as an assistant professor. In the meantime, she is the vice director of School of Applied Disciplines. During her 16 years working experience in MIS department, she has taught more than 10 different courses and involved in various research and administrative activities. Her research areas are statistics, operations research and data mining.

B. Badur received his BS degree in chemical engineering from Bogazici University Istanbul, Turkey, MS degrees in industrial engineering and chemical engineering from Bogazici University, MA degree in economics from Bogazici University and PhD in economics from Bogazici University. Currently, he is working in Management Information Systems Department of Bogazici University as an assistant professor. In the MIS department, he has taught economics, programming and data mining courses at both undergraduate and graduate levels and conducted various MA theses. His research interests are data mining and agent based modeling.

Dėstytojo mokymo veiklos analizavimas, naudojant duomenų gavybos metodus

Sona MARDIKYAN, Bertan BADUR

Jau daugelį metų studentų vertinimai, kurie padeda išmatuoti dėstytojo mokymo efektyvumą, naudojami aukštojo mokslo sistemoje. Tiriama veiksniai, susiję su dėstytojų mokymo efektyvumo vertinimu naudojant du skirtingus duomenų gavybos metodus: pažingsninę regresiją ir sprendimų medžius. Straipsnio autoriai anonimiškai rinko duomenis iš Bogazici universiteto Vadybos informacinių sistemų fakulteto studentų. Be to, atsižvelgiama į kintamuosius, kurie susiję su kitais dėstytojais bei kursų charakteristikomis.

Rezultatai rodo, kad dėstytojų įdarbinimo statusas, darbo krūvis, studentų lankomumas yra reikšmingi rodikliai dėstytojo mokymo veiklai įvertinti.