Int. J. Engng Ed. Vol. 26, No. 1, pp. 87–95, 2010
Printed in Great Britain.

0949-149X/91 $3.00+0.00

A Decision Tree Approach to Predictive Modeling of Student Performance in Engineering Dynamics*

NING FANG,[1] JINGUI LU[2]
[1] *Department of Engineering and Technology Education, College of Engineering, Utah State University, Logan, UT 84322, USA. Email: nfang@engineering.usu.edu*
[2] *CAD Center, Nanjing University of Technology, Nanjing 210009, Jiangsu Province, P. R. China.*

A decision tree model has been developed to predict student performance in Engineering Dynamics based on 750 data records collected from 125 students in two semesters. The predictor variables include a student's cumulative GPA and scores in four prerequisite courses: Engineering Statics, Calculus I, Calculus II, and Physics. The model generates nine decision rules and shows that a student's performance in Statics and cumulative GPA play the two most significant roles in governing the student's performance in Dynamics. The prediction accuracy of the model is more than 80%, which is at least 14% higher than that of the traditional multivariate regression model.

Keywords: decision trees; predictive modeling; multivariate regression; engineering dynamics

1. INTRODUCTION

1.1 Predictive modeling of student performance in Engineering Dynamics

ALMOST EVERY MECHANICAL or civil engineering student is required to take Engineering Dynamics—a high-enrollment, high-impact, and core engineering course. This course is widely regarded by students as one of the most difficult courses, and many students fail because it covers a broad spectrum of foundational engineering concepts and principles [1–3], such as motion, force and acceleration, work and energy, impulse and momentum, and vibrations for both a particle and a rigid body. Prediction of student performance in Engineering Dynamics is important because not only does it help the instructor develop effective course curriculum and teaching strategies, but also facilitates students' increased understanding and development of effective learning strategies [4–6].

1.2 Decision trees

A wide variety of mathematical techniques has been employed in predictive modeling for various applications, including both traditional statistical techniques [7] (such as regression and correlation analysis [8]) and modern data mining techniques [9–11], for example, rough sets [12], fuzzy sets [13], neural networks [14], Bayesian networks [15], genetic algorithm [16], and decision trees [17–19]. As an advanced data mining technique for multiple variables analysis, decision trees have recently received growing attention in research in many disciplines including engineering, business, medicine, education, and so on.

Decision trees are generated by splitting a collection of data records into branch-like segments using a sequence of decision rules [17–19]. These segments generate an inverted decision tree that originates with a root node at the top of the tree. Each internal node denotes a test on an attribute. Each branch represents an outcome of the test. Leaf nodes represent class distribution. The algorithms that are commonly employed to generate decision trees include ID3 (Iterative Dichotomiser 3 [20]), C4.5/C5.0 (the expanded ID3 algorithm [21]), CART (Classification and Regression Tree [22]), and CHAID (Chi-squared Automatic Interaction Detector [23]), and so on. Different splitting criteria are also employed to determine the order in which attributes must be chosen to split the data, i.e. to generate tree branches. For classification trees, the splitting criteria commonly employed include entropy reduction (information gain), GINI index, and chi-square. For regression trees, the splitting criteria commonly employed include variance reduction and F-tests [17–19].

Decision trees have several advantages over traditional statistical techniques [17–19]. For example, a decision tree structure provides an explicit set of "if-then" rules (rather than abstract mathematical equations), making the results easy to interpret and use. Decision tree algorithms process both continuous and categorical data. By finding a strong or weak relationship between input values and target values in a group of observations that form a dataset, decision tree algorithms can also identify the relative importance of each factor investigated.

* Accepted 28 October 2009.

1.3 Research questions and the scope of the present study

The overall goal of the present study is to develop a decision tree model to predict student performance in Engineering Dynamics. The objective is to answer the following three research questions:

1) What specific "if-then" rules can be generated to predict student performance in Engineering Dynamics?
2) To what extent a student's cumulative GPA and performance in prerequisite courses affect student performance in Engineering Dynamics?
3) Are decision tree predictions more accurate than the predictions from the traditional regression-based statistical approach?

The decision tree model developed in the present study includes one outcome/dependent variable (i.e. a student's score in Dynamics) and five predictor/independent variables including the student's cumulative GPA and scores in four prerequisite courses: Engineering Statics, Calculus I, Calculus II, and Physics. Among these five predictor variables, GPA is a comprehensive measurement of a student's cognitive level and problem-solving skills. Statics is the immediate prerequisite course for Dynamics, and numerous concepts of Statics (such as Free Body Diagram) are employed in Dynamics. Calculus I and II measure a student's advanced mathematical skills. Physics measures a student's fundamental understanding of physical principles behind various phenomena.

It must be pointed out that the scope of the present exploratory study is limited in the investigation of the effects of five cognitive factors (i.e. the five above-stated predictor variables) on student performance in Dynamics. The effects of a student's non-cognitive factors (such as learning style, motivation and interest, time devoted to learning, family background, race, and many others [24–27]), the instructor's teaching effectiveness and preparation [28], as well as teaching and learning environment (including the use of new instructional technologies) [29] on student performance is beyond the scope of the present exploratory study and will be dealt with in future work.

1.4 Novelty and significance of the present study

The present study is innovative because no prior literature exists that applies the decision tree approach (or any other data mining and statistical approaches) to model student performance in the Engineering Dynamics course. In prior literature, decision trees were used on other topics, such as business management, production planning and control, student behavior on web-based online courses, student satisfaction to teaching, student persistence to a science or engineering degree [4, 17-23, 30-32]. Selected examples from our extensive literature review are provided in the following paragraphs.

Minaei-Bidgoli et al. [4] employed a variety of classification approaches, including the decision tree approach, to classify students in order to predict their final grade based on features extracted from logged data in an education web-based system. The data they collected included when, for how long, and how many times students access web sources, the number of correct responses students gave on assigned problems, the pattern of correct and incorrect responses, etc.

Thomas and Galambos [30] employed the chi-squared automatic interaction detector (CHAID) algorithm for decision tree analysis to investigate how students' characteristics and experiences affect satisfaction. They concluded that, compared to regression analysis, "decision tree analysis contributed a different perspective by identifying different predictive variables and differences within the student body that shed new light on the heterogeneity of college students."

Nghe et al. [31] applied decision tree and Bayesian network algorithms to predict a student's overall performance in the third year using student records and GPA at the end of the second year at two universities. They concluded that decision tree was consistently 3–12% more accurate than Bayesian network.

Most recently, Mendez et al. [32] applied decision trees to study the factors associated with a student's persistence to earn a science or engineering degree. Their research shows that high school and freshmen GPAs have highest importance for predicting persistence, and other factors such as the number of science and engineering courses taken freshmen year are important for subgroups of the student population.

However, no literature was found on applying decision trees to predict student performance in Engineering Dynamics. By answering the three research questions in subsequent sections of this paper, the present study makes a series of new research findings that were not reported in prior literature. For example, the present study reveals that a student's Statics performance and cumulative GPA play the two most significant roles in governing the student's performance in Dynamics. The major research findings will be summarized at the end of this paper.

The results of the predictions from the present study can be used to help develop an effective teaching strategy to improve student learning of Dynamics. For example, an instructor can use the predictions to identify the number of students who will perform well, averagely, or poorly in the Dynamics class. If the predictions indicate that the number of "academically at-risk" students is significant, an appropriate instructional strategy must be adopted to accommodate the particular needs of those students. The predictions may frustrate those students and challenge the way they have to learn engineering. In this latter case, the instructor must pay particular attention to those students in order to help them succeed in the Dynamics course.

1.5 Structure and contents of this paper

This paper includes three major sections as follows. 1) The "Data Collection" section that describes how data were collected for this research. 2) The "Decision Tree Modeling and Validation" section that answers the first and second research questions. 3) The "Comparison of the Decision Tree Approach with a Traditional Multivariate Statistical Approach" section that answers the third research question.

2. DATA COLLECTION

A validated total of 750 data records (without missing or conflicting data) were collected from 125 students in Semester A (45 students) and Semester B (80 students). Each student was associated with six data records including cumulative GPA (numerical values 0.0–4.0) and scores (i.e. letter grades A, A–, B+, B, B–, C+, C, C–, D+, D, or F) in five course: Dynamics, Statics, Calculus I, Calculus II, and Physics. The Dynamics course was taught by different instructors in Semester A and Semester B.

Table 1 shows the demographics of the 125 students in terms of their sex (male or female) and majors (mechanical engineering, civil engineering, or other majors). In two semesters, 113 (90.4%) students were males, and 101 students (80.8%) were either mechanical or civil engineering majors.

3. DECISION TREE MODELING AND VALIDATION

3.1 Pre-processing of the collected data

The decision tree approach processes only categorical (ordinal or nominal) data, not scale data with numerical values such as 1.35 and 2.78. Therefore, a student's cumulative GPA (numerical values ranging between 0.0 and 4.0) was first converted into one of the 11 letter grades (A, A–, B+, B, B–, C+, C, C–, D+, D, or F). See the first and second rows in Table 2.

The decision tree approach produces a set of explicitly stated decision rules. However, if the number of decision rules produced from data is too large, it is difficult for a user to interpret them and to use them directly for prediction purposes. Therefore, in order to reduce the number of decision rules generated from the decision tree approach, the above 11 letter grades of a student's cumulative GPA were represented by four letter grades (A, B, C, DF), as shown in the third row of Table 2. Students' scores (one of the 11 letter grades) in Dynamics, Statics, Calculus I, Calculus II, and Physics were also converted in four letter grades A, B, C, and DF using Table 2.

3.2 Training dataset and validation dataset

The pre-processed data were divided into two groups: training dataset and validation dataset. In Semester A, the training dataset contained the records from the first 15 students (in alphabetical order of their last names in the class), and the records from the remaining 30 students were used as the validation dataset. In Semester B, the training dataset had the records from the last 16 students (in alphabetical order of their last names in the class), and the records from the remaining 64 students were used as the validation dataset. Note that the training dataset was not subjectively selected based upon student performance. Table 3 shows the training dataset using Semester A as a representative example.

3.3 Results and analysis

Based on the training dataset of each semester, a decision tree was constructed for each semester by using the popular ID3 (Iterative Dichotomiser 3) alrogithm [20] that employed the method of top-down induction of decision trees. The outcome variable (i.e. dependent variable) is a student's Dynamics grade. The predictor variables (i.e., independent variables) include the student's cumulative GPA and grades in Statics, Calculus I,

Table 1. Student demographics (the total number of students = 125)

	Male	Female	Mechanical engineering major	Civil engineering major	Other majors
Semester A	42	3	23	13	9
Semester B	71	9	38	27	15
Two semesters	113	12	61	40	24

Table 2. Conversion of cumulative GPA to letter grades

Numerical value	**3.835 – 4.0**	**3.5 – 3.835**	**3.165 – 3.5**	**2.835 – 3.165**	**2.5 – 2.835**	**2.165 – 2.5**	**1.835 – 2.165**	**1.5 – 1.835**	**1.165 – 1.5**	**0.835 – 1.165**	**0 – 0.835**
Letter grade	A	A–	B+	B	B–	C+	C	C–	D+	D	F
Grade use for modeling	A		B			C			DF		

Table 3. Training dataset (semester A as an example)

Student No. 1–15	Dynamics	GPA*	Statics	Calculus I	Calculus II	Physics
1	C	B	B	C	C	B
2	B	B	B	B	C	A
3	B	B	C	C	C	B
4	B	B	C	C	B	B
5	B	B	C	B	C	B
6	DF	C	C	B	C	B
7	B	B	B	C	C	B
8	B	A	B	A	A	B
9	B	B	C	A	C	A
10	B	B	B	B	B	B
11	DF	C	DF	A	C	B
12	C	B	B	B	C	B
13	C	B	B	C	C	C
14	DF	B	C	A	B	B
15	A	A	A	B	A	A

*Conversion of the cumulative GPA from numerical values (0.0-4.0) to letter grades: 3.5–4.0 = A; 2.5–3.5 = B; 1.5–2.5 = C; below 1.5 = DF. See Table 2 also.

Calculus II, and Physics. To prevent the decision tree's growing too large and, therefore, too difficult to be understood, the maximum depth of the tree was five, and the maximum number of branches from a node was three.

Entropy reduction (information gain) [20] was employed as the splitting criterion to determine the order in which attributes were chosen to split the data. Entropy is a concept from Information Theory that measures the homogeneity of a dataset, or how disordered a dataset is [17–19]. The higher the entropy (or uncertainty) of a dataset, the more information is required to describe the dataset completely. The decision tree algorithm aims to decrease the entropy of the dataset until leaf nodes are reached at which point the subset has zero entropy and represents instances all of one class. The entropy of a dataset S is mathematically expressed as

$$\text{Entropy}(S) = \sum_{i-1}^{c} p_i \log_2 p_i \qquad (1)$$

where p_i is the proportion of instances in the dataset that take the *i*th value of the target attribute.

Information gain (entropy reduction between two levels) is then calculated as

$$\text{Gain}(S, A) = \text{Entropy}(S) - \sum_{v \in A} \frac{|S_v|}{|S|} \text{Entropy}(S_v) \qquad (2)$$

where v is a value of an attribute A, |Sv| is the subset of instances of S where A takes the value v, and |S| is the number of instances. On the right-hand side of Equation (2), the first term is the entropy of the original dataset S, and the second term is the expected value of the entropy after S is partitioned using attribute A.

Table 4 summarizes the major results of decision tree modeling for the two semesters. Fig. 1 and Table 5 show, respectively, the structure of the developed decision tree and the associated decision rules, both using Semester A as a representative example. In Fig. 1, the letters (A, B, C, and DF) in bold frames are Dynamics grades; and the letters in italics (*A*, *B*, *C*, and *DF*) are grades of other relevant courses.

The following observations and analysis can be made from Tables 4 and 5 and Figure 1.

Table 4. Summary of the major results of decision tree modeling

	Training dataset	Validation dataset	Number of decision rules generated*	Relative importance of predictor variables*	Prediction accuracy of the model using the validation dataset
Semester A	15	30	9	GPA: 74.5 Statics: 100 Calculus I: 43.2 Calculus II: 43.2 Physics: 51.6	83.3%
Semester B	16	64	9	GPA: 68.1 Statics: 100 Calculus I: 24.7 Calculus II: 56.4 Physics: 39.1	85.9%

*The data in these two columns were generated from the training dataset, not including the validation dataset.

Table 5. Associated decision rules (Semester A as an example)

Rule	IF	THEN
#1	Statics = A	Dynamics = A
#2	Statics = B AND Calculus II = A	Dynamics = B
#3	Statics = B AND Calculus II = B	Dynamics = B
#4	Statics = B AND Calculus II = C AND Physics = A	Dynamics = B
#5	Statics = B AND Calculus II = C AND Physics = B	Dynamics = C
#6	Statics = B AND Calculus II = C AND Physics = C	Dynamics = C
#7	Statics = C AND GPA = B	Dynamics = B
#8	Statics = C AND GPA = C	Dynamics = DF
#9	Statics = DF	Dynamics = DF

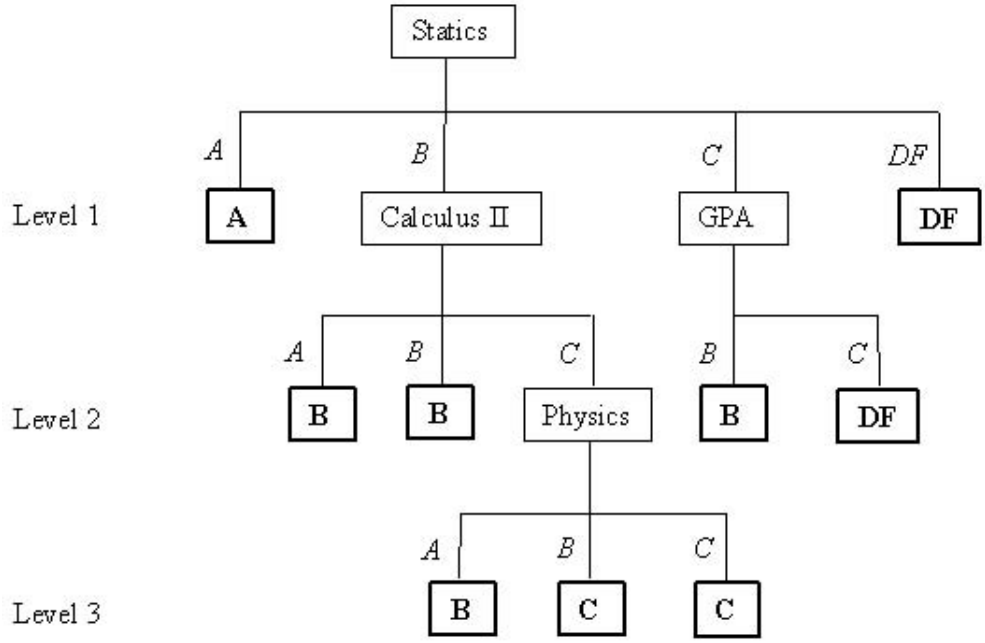

Fig. 1. Structure of developed decision tree (Semester A as an example).

1) Validation of the developed decision tree models.

   As shown in Table 4, the prediction accuracy of the developed decision tree models is more than 80% for both semesters: 83.3% for Semester A and 85.9% for Semester B. Note that the Dynamics course was taught by two different instructors in Semester A and Semester B. Students were also different in these two semesters. Therefore, more than 80% of prediction accuracy in both semesters validates the developed decision tree models.

2) Relative importance of the predictor variables.

   As shown in Table 4, student performance in Statics and cumulative GPA are the two most important variables governing student performance in Dynamics in both semesters. The relative importance of Statics and cumulative GPA is: Statics 100 and GPA 74.5 for Semester A; Statics 100 and GPA 68.1 for Semester B. The specific numerical value of relative importance of predictor variables varies in different semesters due to the use of different training datasets.

   The importance of Statics can also be seen clearly from Fig. 1. Refer to Level 1 of the developed decision tree for Semester A. A student's Dynamics grade decreases as the student's Statics grade reduces from A to DF. If the student's Statics grade is B or C, other variables (Calculus II and GPA) take effect, and the decision tree evolves to Level 2. For a student whose Statics grade is B and Calculus II grade is C, another variable (Physics) takes effect, and the decision tree evolves to Level 3. A set of decision tree rules was also generated from the decision tree structure shown in Fig. 1, which is described as follows.

3) The generated decision rules

   As shown in Tables 4 and 5, nine decision rules were generated for each semester. Each decision rule (Table 5) provides a simple, literal explanation to the graphic representation of the decision tree structure (Fig. 1). For example, if a student has high performance (grade A) in Statics, the student also has high performance in Dynamics, as indicated by decision rule #1. On the contrary, if a student has poor performance (grade DF, or grade C with low GPA C) in Statics, the student also has poor performance in Dynamics, as indicated by decision

rules #8 and #9. However, if a student gets a middle grade B in Static, student performance in Dynamics will be affected by other variables, as indicated by decision rules #2–#6.

Based on many years of teaching experience of the authors of this paper, all these decision rules reflect real-world situations at a typical university learning environment. Therefore, these decision rules can be used as general guidelines to predict student performance in Dynamics. However, on the other hand, it also must be pointed out that these decision rules were generated from the training dataset employed in the present study. The decision rules change if using a different training dataset. When applying the decision tree approach for predictive modeling of student performance in Dynamics at another institution of higher learning, it is suggested that the training dataset be collected from that particular institution so as to best and most accurately represent teaching and learning at that particular institution.

4. COMPARISON OF THE DECISION TREE APPROACH WITH A TRADITIONAL MULTIVARIATE STATISTICAL APPROACH

A variety of traditional multivariable statistical approaches, such as regression analysis, has been employed for predictive modeling that involves multiple outcome and predictor variables. Regression analysis includes logistic regression and linear/nonlinear regression. Logistic regression only applies to an outcome variable that is a categorical dichotomy (the predictor variables can be either continuous or categorical) [33, 34]. In the present study, the outcome variable (i.e. a student's Dynamics grade) is not binary and has more than two values (A, B, C, and DF). Therefore, logistic regression does not apply to the present study. The following study is conducted to compare the prediction accuracy of the decision tree approach and the multivariate linear regression approach—one of the most commonly used approaches in educational research. Through the data-driven comparison of the two approaches, we can also develop a deeper understanding of the advantages and disadvantages of each approach.

4.1 Multivariate linear regression

To make the results comparable, the same training dataset (see Table 3, for example) and validation dataset that were used in the decision tree modeling were employed again in multiple linear regression analysis for each semester. All letter grades that a student earned in Dynamics, Statics, Calculus I, Calculus II, and Physics were converted to numerical values using standard criteria: A = 4.0, A– = 3.67; B+ = 3.33; B = 3.0 , B– = 2.67; C+ = 2.33; C = 2.0; C– = 1.67; D+ = 1.33; D = 1.0; and F = 0.0. A student's cumulative GPA took its actual numerical value.

The following regression equations were generated: Equation (3) for Semester A and Equation (4) for Semester B.

$$\text{Dynamics score} = -2.039 + 1.654 \times \text{GPA} - 0.253 \times \text{Statics score} - 0.500 \times \text{Calculus I score} + 0.239 \times \text{Calculus II score} + 0.329 \times \text{Physics score} \quad (3)$$

$$\text{Dynamics score} = -2.391 + 1.256 \times \text{GPA} + 0.474 \times \text{Statics score} - 0.333 \times \text{Calculus I score} + 0.172 \times \text{Calculus II score} + 0.035 \times \text{Physics score} \quad (4)$$

The coefficient of determination, R^2, is 0.781 for Equation (3) and 0.894 for Equation (4). The R^2 values imply that the five predictor variables account for 78.1% of the variability in a student's Dynamics performance in Semester and for 89.4% of the variability in a student's Dynamics performance in Semester B.

4.2 Comparison of prediction accuracy

The prediction accuracy of the decision tree approach and the multivariate linear regression approach is compared in Table 6. As seen from Table 6, the decision tree approach improves the prediction accuracy by 16.6% in Semester A and by 14.0% in Semester B.

Using Semester A as an example, Table 7 lists the detailed, item-to-item comparison of predictions by the two approaches. Bold letters in Table 7 are wrong predictions. Out of 30 validation data, the decision tree approach made five wrong predictions; however, the linear regression approach made 10 wrong predictions. All these results show that the decision tree approach is better than the linear regression approach.

Furthermore, the developed regression equations (1) and (2) contain negative coefficients for some predictor variables, which do not represent real-world situations. For example, the coefficient of the "Statics" variable in Equation (1) is –0.253, which means a student will have higher performance in Dynamics if the student performs poorly in Statics. The coefficient of the "Calculus I" variable in Equation (2) is –0.333, which means a student will have higher performance in Dynamics if the student performs poorly in Calculus I. These purely mathematical predictions do not make physical sense in a real-world teaching and learning environment.

Table 6. Comparison of prediction accuracy

	Liner regression	Decision tree	Improvement by using decision tree
Semester A	66.7%	83.3%	16.6%
Semester B	71.9%	85.9%	14.0%

Table 7. Detailed comparison of predictions (Semester A as an example)

Student No.	Predictor variables			Dynamics score			Predicted by decision tree	Predicted by regression
	GPA	Statics	Cal. I	Cal. II	Physics	Actual		
16	C	C	B	B	B	DF	DF	**C**
17	B	B	B	B	A	B	B	B
18	B	C	C	C	C	C	**B**	C
19	A	A	A	A	A	A	A	A
20	B	C	B	C	C	C	**B**	C
21	B	B	B	B	C	B	B	**C**
22	A	A	B	B	A	A	A	A
23	B	C	C	C	C	DF	**B**	**C**
24	A	A	A	A	A	A	A	A
25	A	C	A	B	B	B	B	B
26	C	C	C	C	B	C	**DF**	C
27	B	B	B	B	B	B	B	B
28	B	C	B	B	A	B	B	**A**
29	B	C	C	C	B	B	B	B
30	A	B	B	B	B	B	B	B
31	B	C	C	C	A	B	B	B
32	B	B	C	B	A	B	B	B
33	B	B	B	A	B	B	B	B
34	B	B	A	C	B	C	C	C
35	A	C	B	B	B	B	B	B
36	B	C	B	C	A	B	B	B
37	B	B	B	A	B	B	B	**A**
38	A	B	B	B	B	B	B	B
39	B	C	B	C	A	B	B	B
40	A	A	A	A	B	A	A	**B**
41	C	C	B	C	A	C	**DF**	C
42	B	C	B	B	B	B	B	**C**
43	B	B	A	C	B	C	C	**DF**
44	B	C	B	C	B	B	B	**C**
45	B	A	B	B	B	A	A	**C**

4.3 Relative importance of the predictor variables

Following the American Psychological Association guidelines for reporting multiple regression, Tables 8 and 9 provide the details of regression results based on the training dataset in Semester A and Semester B, respectively.

Table 8. Regression results based on the training dataset in Semester A

	Unstandardized coefficients		Standardized coefficients β
	B	Standard error	
Constant	–2.039	1.182	
Cumulative GPA	1.654	0.630	0.760
Statics	–0.253	0.304	–0.206
Calculus I	–0.500	0.241	–0.436
Calculus I	0.239	0.259	0.194
Physics	0.329	0.403	0.186

Table 9. Regression results based on the training dataset in Semester B

	Unstandardized coefficients		Standardized coefficients β
	B	Standard error	
Constant	–2.391	0.712	
Cumulative GPA	1.256	0.334	0.666
Statics	0.474	0.235	0.349
Calculus I	–0.333	0.247	–0.234
Calculus I	0.172	0.229	0.141
Physics	0.035	0.178	0.026

The magnitude (the absolute value) of standardized coefficients β listed in Tables 8 and 9 can be used to estimate the relative importance of the predictor variables in affecting a student's Dynamics performance. In Semester A, the order of relative importance is GPA > Calculus I > Statics > Calculus II > Physics. In Semester B, the order of relative importance is GPA > Statics > Calculus I > Calculus II > Physics. Comparing these results with the data shown in Table 4, one can find that the order of relative importance changes if using different modeling approaches. However, both the decision tree approach and the regression approach highlight the predominant importance of GPA in affecting a student's performance in Dynamics. This seems reasonable because GPA represents the comprehensive problem-solving skills of a student. Therefore, although with various shortcomings, the traditional regression approach still has its value in providing supplemental, supporting insights in interpreting the prediction results from the decision tree approach.

5. CONCLUSIONS

As the first exploration to apply the decision tree approach for predictive modeling of student performance in Engineering Dynamics, the present study makes scientific contributions to the body of

knowledge by answering the three research questions that were stated before in the "Introduction" section of this paper. Based on the modeling effort and analysis on a validated total of 750 student data records in two semesters, our research findings are summarized in the following paragraphs.

Nine specific "if-then" rules were generated to predict student performance in Engineering Dynamics for each semester. For example, if a student has a high performance (grade A) in Statics, the student also has high performance in Dynamics. If a student has a poor performance (grade DF, or grade C with low GPA C) in Statics, the student also has a poor performance in Dynamics. However, if a student gets a middle grade B in Statics, student performance in Dynamics will be affected by other variables.

These decision rules generally reflect real-world situations at a typical institution of higher learning and can be used as general guidelines to predict student performance in Dynamics. When extending the decision tree approach to a particular institution of higher learning, the suggestions is made to collect the training dataset from that particular institution so as to best and most accurately represent teaching and learning at that particular institution.

The decision tree analysis reveals that, in both Semester A and Semester B, a student's Statics performance and cumulative GPA play the two most significant roles in governing the student's performance in Dynamics. The prediction accuracy of the developed decision tree models is more than 80% for both semesters: 83.3% for Semester A and 85.9% for Semester B. Compared to the traditional multivariate linear regression approach, the decision tree approach improves the prediction accuracy by 16.6% in Semester A and by 14.0% in Semester B. A major shortcoming of the traditional multivariate linear regression approach is that the developed regression equations contain negative coefficients for some predictor variables, which are difficult to interpret because they do not make physical sense in a real-world teaching and learning environment.

REFERENCES

1. R. C. Hibbeler, *Engineering Mechanics Dynamics* (11th edition), Pearson Prentice Hall, Upper Saddle River, NJ. 2007.
2. M. B. Rubin and E. Altus, An alternative method for teaching dynamics. *Int. J. Eng. Educ.*, **16**, 2000, pp. 447–456.
3. R. Kumar and M. Plummer, Using contemporary tools to teach dynamics in engineering technology. *Int. J. Eng. Educ.* **13**, 1997, pp. 407–411.
4. B. Minaei-Bidgoli, D. A. Kashy, G. Kortemeyer and W. F. Punch, Predicting student performance: an application of data mining methods with an educational web-based system. *Proceedings of the 33rd ASEE/IEEE Frontiers in Education Conferences*, November 5–8, Boulder, CO. 2003.
5. R. M. Felder, K. D. Forrest, L. Baker-Ward, E. J. Dietz and P. H. Mohr, A longitudinal study of engineering student performance and retention. I. success and failure in the introductory course. *Journal of Engineering Education,* **82**, 1993, pp. 15–21.
6. R. J. Delauretis and G. E. Molnar, Academic prediction: paradigm research or let's see what happens? *IEEE Transactions on Education*, **E-15**, 1972, pp. 32–36.
7. L. Cohen, L. Manion and K. Morrison, *Research Methods in Education* (6th edition), Routledge, Oxon, UK. 2007.
8. J. Cohen, P. Cohen, S. G. West and L. S. Aiken, *Applied Multiple Regression/correlation Analysis for the Behavioral Sciences* (3rd ed.), Lawrence Erlbaum Associates, Inc., Publishers, Mahwah, NJ. 2003.
9. J. Han and M. Kamber, *Data Mining: Concepts and Techniques*, Morgan Kaufmann publishers, San Fransisco, CA (2007).
10. K. J. Cios, W. Pedrycz and R. M. Swiniarski, *Data Mining Methods for Knowledge Discovery*, Kluwer, Boston, MA. 1998.
11. C. Romero, S. Ventura and E. Garcia, Data mining in course management system: Moodle case study and tutorial. *Computers and Education*, **51**, 2008, pp. 368–384.
12. S. Pawlak, Decision analysis using rough sets. *International Transactions in Operational Research*, **1**, 1994, pp. 107–114.
13. J. Ma and D. Zhou, Fuzzy set approach to the assessment of student-centered learning. *IEEE Transactions on Education*, **43**, 2000, pp. 237–241.
14. R. Marti, Artificial neural networks for prediction. In J. Wang (Ed.), *Encyclopedia of Data Warehousing and Mining,* Idea Group Inc, Hershey, PA. 2006.
15. A. Bashir, L, Khan, and M. Awad, Bayesian networks. In J. Wang (Ed.), *Encyclopedia of data warehousing and mining,* Idea Group Inc, Hershey, PA. 2006.
16. A. Kamrani, W. Rong and E. Gonzalez, A genetic algorithm methodology for data mining and intelligent knowledge acquisition. *Computers & Industrial Engineering*, **40**, 2001, pp. 361–377.
17. A. Abu-Hanna and N. de Keizer, Integrating classification trees with logistic regression in intensive care prognosis. *Artificial Intelligence in Medicine,* **29**, 2003, pp. 5–23.
18. L. Breiman, J. H. Friedman, R. A. Olshen and C. J. Stone, *Classification and Regression Trees*, Chapman & Hall, New York, NY. 1984.
19. L. Rokach and O. Maimon, *Data Mining with Decision Trees: Theory and Applications (Machine Perception and Artificial Intelligence)*, World Scientific Publishing, Hackensack, NJ. 2008.
20. J. R. Quinlan, Induction of decision trees. *Machine Learning*, **1**, 1986, pp. 81–106.

21. J. R. Quinlan, *C4.5: Programs for Machine Learning*, Morgan Kaufmann Publishers, San Mateo, CA. 1993.
22. L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*, Chapman & Hall (Wadsworth, Inc.), New York, NY. 1984.
23. G. V. Kass, An exploratory technique for investigating large quantities of categorical data. *Journal of Applied Statistics*, **29**, 1980, pp. 119–127.
24. R. Riding and S. Rayner, *Cognitive Styles and Learning Strategies: Understanding Style Differences in Learning and Behavior*, David Fulton Publishers, London, UK. 1998.
25. B. F. French, J. C. Immekus, and W. Oakes, A structural model of engineering students success and persistence. *Proceedings of the 33rd ASEE/IEEE Frontiers in Education Conference*, November 5–8, Boulder, CO. 2003.
26. S. Ransdell, Predicting college success: the importance of ability and non-cognitive variables. *Int. J. Educ. Res.*, **35**, 2001, pp. 357–364.
27. T. J. Tracey and W. E. Sedlacek, Noncognitive variables in predicting academic success by race. *Measurement and Evaluation in Guidance*, **16**, 1984, pp. 171–178.
28. Y. Chen and L. B. Hoshower, Student evaluation of teaching effectiveness: an assessment of student perception and motivation. *Assessment & Evaluation in Higher Education*, **28**, 2003, pp. 71–88.
29. E. Graaff, G. N. Saunders-Smits, and M. R. Nieweg, *Research and Practice of Active Learning in Engineering Education*, Palls Publications, Amsterdam, Holland. 2005.
30. E. H. Thomas and N. Galambos, What satisfies students? mining student-opinion data with regression and decision tree analysis. *Research in Higher Education*, **45**, 2004, pp. 251–269.
31. N. T. Nghe, P. Janecek and P. Haddawy, A Comparative Analysis of Techniques for Predicting Academic Performance. *Proceedings of the 37th ASEE/IEEE Frontiers in Education Conferences*, October 10–13, Milwaukee, WI. 2007.
32. G. Mendez, T. D. Buskirk, S. Lohr and S. Haag, Factors associated with persistence in science and engineering majors: an exploratory study using classification trees and random forests. *J. Eng. Educ.*, **97**, 2008, pp. 57–70.
33. M. P. Driscoll, *Psychology of Learning for Instruction* (2nd ed.), Allyn & Bacon, Needham Heights, MA. 2000.
34. B. H. Cohen, *Explaining Psychological Statistics* (2nd edition), John Wiley & Sons, Inc., New York, NY. 2000.

Ning Fang is an Associate Professor in the College of Engineering at Utah State University, USA. He teaches Engineering Dynamics. His areas of interest include computer-assisted instructional technology, curricular reform in engineering education, the modeling and optimization of manufacturing processes, and lean product design. He earned his Ph.D., MS, and BS degrees all in Mechanical Engineering and is the author of more than 60 technical papers published in refereed international journals and conference proceedings. He is a Senior Member of the Society for Manufacturing Engineering and a member of the American Society of Mechanical Engineers. He is also a member of the American Society for Engineering Education and a member of the American Educational Research Association.

Jingui Lu is a Professor in the CAD Center at Nanjing University of Technology, P. R China. His research areas and expertise include computer-aided design, computational intelligence, and a variety of data mining techniques such as genetic algorithm, neural network, and decision tree. He obtained his Ph.D., MS, and BS degrees all in Mechanical Engineering. He was a Research Fellow of Tokyo Institute of Technology, Japan, from 1997 to 1998, and a Visiting Professor of Utah State University, USA in 2007. He is a Senior Member of the Society of Mechanical Engineers of China.