

PREDICTIVE MODELING AND ANALYSIS OF STUDENT ACADEMIC
PERFORMANCE IN AN ENGINEERING DYNAMICS COURSE

by

Shaobo Huang

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Engineering Education

Approved:

Ning Fang, Ph.D.
Major Professor

Kurt Becker, Ph.D.
Committee Member

Oenardi Lawanto, Ph.D.
Committee Member

Edward Reeve, Ph.D.
Committee Member

Wenbin Yu, Ph.D.
Committee Member

Mark R. McLellan, Ph.D.
Vice President for Research and
Dean of the School of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2011

Copyright © Shaobo Huang 2011

All Rights Reserved

ABSTRACT

Predictive Modeling and Analysis of Student Academic Performance
in an Engineering Dynamics Course

by

Shaobo Huang, Doctor of Philosophy

Utah State University, 2011

Major Professor: Ning Fang, Ph.D.
Department: Engineering and Technology Education

Engineering dynamics is a fundamental sophomore-level course that is required for nearly all engineering students. As one of the most challenging courses for undergraduates, many students perform poorly or even fail because the dynamics course requires students to have not only solid mathematical skills but also a good understanding of fundamental concepts and principles in the field. A valid model for predicting student academic performance in engineering dynamics is helpful in designing and implementing pedagogical and instructional interventions to enhance teaching and learning in this critical course.

The goal of this study was to develop a validated set of mathematical models to predict student academic performance in engineering dynamics. Data were collected from a total of 323 students enrolled in ENGR 2030 Engineering Dynamics at Utah State University for a period of four semesters. Six combinations of predictor variables that

represent students' prior achievement, prior domain knowledge, and learning progression were employed in modeling efforts. The predictor variables include X_1 (cumulative GPA), $X_2 \sim X_5$ (three prerequisite courses), $X_6 \sim X_8$ (scores of three dynamics mid-term exams). Four mathematical modeling techniques, including multiple linear regression (MLR), multilayer perceptron (MLP) network, radial basis function (RBF) network, and support vector machine (SVM), were employed to develop 24 predictive models. The average prediction accuracy and the percentage of accurate predictions were employed as two criteria to evaluate and compare the prediction accuracy of the 24 models.

The results from this study show that no matter which modeling techniques are used, those using $X_1 \sim X_6$, $X_1 \sim X_7$, and $X_1 \sim X_8$ as predictor variables are always ranked as the top three best-performing models. However, the models using $X_1 \sim X_6$ as predictor variables are the most useful because they not only yield accurate prediction accuracy, but also leave sufficient time for the instructor to implement educational interventions. The results from this study also show that RBF network models and support vector machine models have better generalizability than MLR models and MLP network models. The implications of the research findings, the limitation of this research, and the future work are discussed at the end of this dissertation.

(135 pages)

PUBLIC ABSTRACT

Predictive Modeling and Analysis of Student Academic Performance
in an Engineering Dynamics Course

by

Shaobo Huang, Doctor of Philosophy

Engineering dynamics is a fundamental sophomore-level course required for many engineering students. This course is also one of the most challenging courses in which many students fail because it requires students to have not only solid mathematical skills but also a good understanding of dynamics concepts and principles.

The overall goal of this study was to develop a validated set of mathematical models to predict student academic performance in an engineering dynamics course taught in the College of Engineering at Utah State University. The predictive models will help the instructor to understand how well or how poorly the students in his/her class will perform, and hence the instructor can choose proper pedagogical and instructional interventions to enhance student learning outcomes.

In this study, 24 predictive models are developed by using four mathematical modeling techniques and a variety of combinations of eight predictor variables. The eight predictor variables include students' cumulative GPA, grades in four prerequisite courses, and scores in three dynamics mid-term exams. The results and analysis show that each of the four mathematical modeling techniques have an average prediction accuracy of more than 80%, and that the models with the first six predictor variables yield high prediction accuracy and leave sufficient time for the instructor to implement educational interventions.

ACKNOWLEDGMENTS

I would like to express my special appreciation to my advisor, Dr. Ning Fang, for his mentoring, encouragement, and inspiration during my PhD study. Dr. Fang has not only helped me with this research but also provided many suggestions for the writing of this dissertation. He has always been ready to lend a kind hand whenever I need help in my daily life as well. I am also grateful to my committee members, Dr. Kurt Becker, Dr. Oenardi Lawanto, Dr. Edward M. Reeve, and Dr. Wenbin Yu, for their valuable suggestions on my research as well as the help that they have provided during my years of studying abroad. Finally, I would like to thank my family for their unconditional support during my studying abroad. They have always been my source of energy.

Shaobo Huang

CONTENTS

	Page
ABSTRACT.....	iii
PUBLIC ABSTRACT.....	v
ACKNOWLEDGMENTS	vi
LIST OF TABLES	xi
LIST OF FIGURES	xii
CHAPTER	
I. INTRODUCTION	1
Problem Statement.....	4
Research Goals and Objectives.....	4
Research Questions.....	5
Scope of This Research.....	6
Uniqueness of This Research.....	6
II. LITERATURE REVIEW	7
Predictive Modeling of Student Academic Performance.....	7
Statistical and Data Mining Modeling Techniques	23
Chapter Summary	37
III. RESEARCH DESIGN	39
Overall Framework.....	40
Criteria Used for Assessing Prediction Accuracy	50
Determining the Appropriate Sample Size for Predictive Model Development.....	52
Predictive Modeling.....	53
Comparison of the Predictive Models.....	64
IV. RESULTS AND ANALYSIS	65
Descriptive Analysis of the Normalized Data.....	65
Identification of Outliers in the Collected Data.....	70

	Page
Testing of Multiple Collinearity.....	70
Correlation Analysis.....	73
Determining the Appropriate Sample Size	73
Internal and External Validations.....	82
Comparison of Different Modeling Techniques	90
Identifying Academically At-Risk Students.....	91
 V. DISCUSSIONS AND CONCLUSIONS	 96
Summary of This Research.....	96
Answers to the Research Questions	96
Discussion of the Results	99
Implications of the Research Findings.....	101
Limitations of This Research	102
Recommendations for Future Studies.....	104
 REFERENCES	 106
 APPENDIX.....	 117
 CURRICULUM VITAE	 122

LIST OF TABLES

Table	Page
1. Studies on the Relationship Between Class Size and Achievement	11
2. The Effects of Student Prior Knowledge on Academic Performance.....	16
3. Studies That Included the Use of Noncognitive Predictors	22
4. Conversion of Letter Grades.....	48
5. Normalization of the Raw Data	48
6. Descriptive Analysis of the Normalized Data.....	66
7. Collinearity Analysis of the Data Collected in Semester #1	72
8. Correlation Coefficients in Semester #1	74
9. Correlation Coefficients in Semester #2	75
10. Correlation Coefficients in Semester #3	76
11. Correlation Coefficients in Semester #4	77
12. Comparison of Different Sample Sizes.....	80
13. Internal and External Validations of MLR Models.....	84
14. Standardized Coefficients of MLR Models	85
15. Internal and External Validations of MLP Network Models	86
16. Internal and External Validations of RBF Models	87
17. Internal and External Validations of SVM Models.....	88
18. An Example of Prediction: The Dynamics Final Exam Score Was 90 (out of 100) for a Student in Semester #4	92
19. Academically At-Risk Students Correctly Identified by MLR Models.....	93

Table	Page
20. Academically At-Risk Students Correctly Identified by MLP Models	93
21. Academically At-Risk Students Correctly Identified by RBF Models.....	93
22. Academically At-Risk Students Correctly Identified by SVM Models.....	94
23. An Example of Identifying Academically At-Risk Students	95
24. Prediction Accuracy of Models 1 to 6	98

LIST OF FIGURES

Figure	Page
1. Schematic graph of a MLP neural network	30
2. The \mathcal{E} – insensitive loss function.....	33
3. The modeling framework of this study.....	41
4. Student demographics.....	43
5. Effects of C and σ^2 on the average prediction accuracy of the SVM models in Semester #1.....	59
6. Effects of C and σ^2 on the percentage of accurate prediction of the SVM models in Semester #1	59
7. Effects of C and σ^2 on the average prediction accuracy of the SVM models in Semester #2.....	60
8. Effects of C and σ^2 on the percentage of accurate prediction of the SVM models in Semester #2	60
9. Flow chart of genetic algorithms	61
10. Overall framework of genetic algorithm (GA) and SVM.....	62
11. Histogram of students' normalized dynamics final exam scores in Semester #1 ($n = 128$).....	67
12. Histogram of students' normalized dynamics final exam scores in Semester #2 ($n = 58$).....	67
13. Histogram of students' normalized dynamics final exam scores in Semester #3 ($n = 53$).....	68
14. Histogram of students' normalized dynamics final exam scores in Semester #4 ($n = 84$).....	68
15. Scatter plots of the dependent variable Y against the predictor variables X_i	69

Figure	Page
16. Assessing the leverage of the data collected in Semester #1 ($n = 128$).....	71
17. Assessing the discrepancy of the data collected in Semester #1 ($n = 128$).....	71
18. Assessing DFFIT of the data collected in Semester #1 ($n = 128$)	72
19. The minimum sample size determined by Soper's (2004) statistics calculator.	78
20. A sample structure of a MLP network	81

CHAPTER I

INTRODUCTION

Engineering dynamics is a fundamental sophomore-level course that nearly all engineering students majoring in aerospace, mechanics, and civil engineering are required to take (Ibrahim, 2004; Rubin & Altus, 2000; Zhu, Aung, & Zhou, 2010). The course cultivates students' ability to "visualize the interactions of forces and moments, etc., with the physical world" (Muthu & Glass, 1999). It is an essential basis for many advanced engineering courses such as advanced dynamics, machine design, and system dynamics and control (Biggers, Orr, & Benson, 2010; Huang & Fang, 2010).

However, engineering dynamics is also regarded as one of the most challenging courses for undergraduates (Self, Wood, & Hansen, 2004). The course requires students to have solid mathematical skills and a good understanding of fundamental concepts and principles of the field. Many students perform poorly in or even fail this course. The mean score of the final comprehensive exam in the dynamics class is below 70 out of 100 at Utah State University in 2009. On average, only 53% of the engineering dynamics questions were answered correctly in the Fundamentals of Engineering (FE) Examination in U.S. in 2009 (Barrett et al., 2010).

Pedagogical and instructional interventions can improve student academic performance by building up a more solid foundation and enhancing students' learning of engineering concepts and principles (Etkina, Mestre, & O'Donnell, 2005). For example, interventional process of constructing knowledge can help students to relate (and, later, integrate) new information to prior knowledge and achieve complex learning goals

(Etkina et al., 2005; Royer, 1986). Students may be able to construct a hierarchical structure of knowledge and gain better understanding of the principles after training (Dufresne, Gerace, Hardiman, & Mestre, 1992).

To achieve better learning outcomes, the choice of instructional interventions must take into account the diverse academic backgrounds and varied performance of students in relevant courses because each student will have a different reaction to them. For example, a study conducted by Palincsar and Brown (1984) showed that implicit instructions could help average students to achieve greater understanding and success in class, whilst the same teaching method would hinder the learning process of lower-performance students.

Many education researchers and instructors have made extensive efforts in constructing effective models to predict student academic performance in a class (Emerson & Taylor, 2004; Holland, James, & Richards, 1966; Kotsiantis, Pierrakeas, & Pintelas, 2003; Lewis & Castley, 2008; Pittman, 2008). The results of these predictive models can help the instructor determine whether or not a pedagogical and instructional intervention is needed. For example, the instructor can determine how well, or how poorly, students may perform in the class. Then, appropriate pedagogical and instructional interventions (for example, designing an innovative and effective teaching and learning plan) can be developed and implemented to help these academically at-risk students.

Variables such as students' prior knowledge and prior achievement contribute significantly to the prediction accuracy of the model that predicts student academic

performance (Fletcher, 1998). Thompson and Zamboanga (2003) concluded that prior knowledge and prior achievement (such as GPA) are significant predictors of student academic performance in a class and represented 40% to 60% of variance in learning new information (Dochy, 1992; Tobias, 1994). However, if prior knowledge is insufficient or even incorrect, learning and understanding of new information will be hindered (Dochy, Segers, & Buehl, 1999).

Psychological variables, such as goals, are controversial predictors for academic achievement. Some studies found that psychological variables were significant predictors (Cassidy & Eachus, 2000) and increased the amount of variance explained for academic achievement (Allen, Robbins, & Sawyer, 2010). However, other studies discovered that the change in explained variance was not significant when psychological variables were included (French, Immekus, & Oakes, 2005). It has been suggested that the variables have different effects on different learning subjects (Marsh, Vandehey, & Diekhoff, 2008).

Identifying and choosing effective modeling approaches is also vital in developing predictive models. Various mathematical techniques, such as regression and neural networks, have been employed in constructing predictive models. These mathematical techniques all have advantages and disadvantages. For example regression, one of the most commonly used approaches to constructing predictive models, is easy to understand and provides explicit mathematical equations. However, regression should not be used to estimate complex relationships and is susceptible to outliers because the mean is included in regression formulas. On the other hand, neural networks can fit any linear or nonlinear function without specifying an explicit mathematical model for the

relationship between inputs and output; thereby, it is relatively difficult to interpret the results.

In a recent work by Fang and Lu (2010), a decision-tree approach was employed to predict student academic achievement in an engineering dynamics course. Their model (Fang & Lu, 2010) only generates a set of “if-then” rules regarding a student’s overall performance in engineering dynamics. This research focused on developing a set of mathematical models that may predict the numerical scores that a student will achieve on the dynamics final comprehensive exam.

Problem Statement

As stated previously, student low academic performance in the engineering dynamics course has been a long-standing problem. Before designing and implementing any pedagogical and instructional interventions to improve student learning in engineering dynamics, it is important to develop an effective model to predict student academic performance in this course so the instructor can know how well or how poorly the students in the class will perform. This study focused on developing and validating mathematical models that can be employed to predict student academic performance in engineering dynamics.

Research Goals and Objectives

The goal of this study is to develop a validated set of mathematical models to predict student academic performance in engineering dynamics, which will be used to

identify the academically-at-risk students. The predicted results were compared to the actual values to evaluate the accuracy of the models.

The three objectives of the proposed research are as follows:

1. Identify and select appropriate mathematical (i.e., statistical and data mining) techniques for developing predictive models.
2. Identify and select appropriate predictor variables/independent variables that can be used as the inputs of predictive models.
3. Validate the developed models using the data collected in four semesters and identify academically-at-risk students.

Research Questions

Three research questions have been designed to address each research objective of the study. These three research questions include:

1. How accurate will predictions be if different statistical/data mining techniques such as multiple linear regression (MLR), multilayer perceptron (MLP) networks, radial basis function (RBF) networks, and support vector machine (SVM) are used?
2. What combination of predictor/independent variables yields the highest prediction accuracy?
3. What is the percentage of academically at-risk students that can be correctly identified by the model?

Scope of This Research

Student academic performance is affected by numerous factors. The scope of the research is limited to the investigation of the effects of a student's prior achievement, domain-specific prior knowledge, and learning progression on their academic performance in the engineering dynamics course. Psychological factors, such as self-efficacy, achievement goals, and interest, were not included in constructing predictive models.

In the future study, psychological factors will be considered for developing the predictive models and further interviews will be conducted to confirm the identified academically at-risk students and diagnose if those students have psychology-related issues and problems in addition to having academic problems. How to effectively apply the predictive models will also be examined in the future study.

Uniqueness of This Research

A variety of commonly used literature databases were examined, including the Education Resources Information Center, Science Citation Index, Social Science Citation Index, Engineering Citation Index, Academic Search Premier, the ASEE annual conference proceedings (1995-2011), and the ASEE/IEEE Frontier in Education conference proceedings (1995-2011). The only paper on predictive modeling of student academic performance in the engineering dynamics course is done by Fang and Lu (2010). However, not only did their work use only one modeling approach (a decision tree approach), but their work took into account only student prior domain knowledge.

CHAPTER II

LITERATURE REVIEW

This chapter includes two sessions. The first session reviews studies concerning the teaching and learning of engineering dynamics as well as the prediction of student academic performance. Features of engineering dynamics, factors that influence the prediction accuracy, and variables used for developing predictive models in various disciplines are discussed. The second session introduces the statistical and data mining modeling techniques used in this research, including MLR, MLP network, RBF network, and SVM.

Predictive Modeling of Student Academic Performance

Engineering Dynamics

Engineering dynamics is a foundational sophomore-level course required for many engineering students. This course is essential for engineering students because it teaches numerous foundational engineering concepts and principles including motion, force and acceleration, work and energy, impulse and momentum, and vibrations. The course encompasses many fundamental building blocks essential for advanced studies in subsequent engineering courses such as machine design, advanced structural design, and advanced dynamics (North Carolina State University, 2011; Utah State University, 2011).

Most dynamics textbooks used in engineering schools in the U.S. have similar contents (Ibrahim, 2004). Take the popular textbook authored by Hibbeler (2010) as an example. The textbook has 11 chapters covering the following topics on kinematics and

kinetics of particles and rigid bodies:

1. Kinematics of a Particle
2. Kinetics of a Particle: Force and Acceleration
3. Kinetics of a Particle: Work and Energy
4. Kinetics of a Particle: Impulse and Momentum
5. Planar Kinematics of a Rigid Body
6. Planar Kinetics of a Rigid Body: Force and Acceleration
7. Planar Kinetics of a Rigid body: Work and Energy
8. Planar Kinetics of a Rigid Body: Impulse and Momentum
9. Three-Dimensional Kinematics of a Rigid Body
10. Three-Dimensional Kinetics of a Rigid Body
11. Vibrations

Assessment of student academic performance. A student's academic performance is typically assessed by homework, quizzes, and exams. The textbook often includes many dynamics problems that can be used as students' homework assignments. Many homework problems often require students to select and correctly apply dynamics concepts and principles. Quizzes and exams can be of any format that the instructor chooses, such as multiple choice, true or false, matching, and free-response questions. The assessment of a student's performance may also include the student's level of participation in class discussions. However, it is the final comprehensive exam that generally makes up the largest percentage of a student's final grade.

Difficulties in learning dynamics. Engineering dynamics is "one of the most

difficult courses that engineering students encounter during their undergraduate study” (Magill, 1997, p. 15). There are at least three reasons for this. First, solving engineering dynamics problems requires students to have a solid understanding of many fundamental engineering concepts and principles. Students must have the ability to visualize the interactions of forces and moments (Muthu & Glass, 1999) and apply Newton’s Laws, the Principle of Work and Energy, and the Principle of Impulse and Momentum for a particle or for a rigid body. However, some dynamics problems can be solved using different approaches. For example, one can use the Conservation of Energy, Newton’s Second Law, or the Principle of Impulse and Momentum to solve a problem that involves the motion of a bouncing ball (Ellis & Turner, 2003).

Second, solving dynamics problems requires students to have solid mathematical skills. For example, knowledge about cross multiplication, differential equations, and integral equations are required to solve dynamics problems that involve angular impulse and momentum.

Since dynamics brings together “basic Newtonian physics and an array of mathematical concepts” (Self & Redfield, 2001, p. 7465), the prerequisites for engineering dynamics include calculus, physics, and engineering statics. Calculus prepares students with mathematical fundamentals such as differential equations. Physics and statics equip students with a necessary familiarity with such concepts as kinematics, Newton’s Laws, and impulse and momentum.

Third, a large class size increases the challenge level of learning dynamics because it is difficult for the instructor to pay sufficient attention to each individual in a

large class (Ehrenberg, Brewer, Gamoran, & Willms, 2001). Class size refers to the ratio of the number of students to the number of instructors teaching the class during a particular class period. Class size is generally defined as “small” if the student-to-instructor ratio is lower than 30:1 and “large” if the ratio is higher than 70:1 (Kopeika, 1992). Engineering dynamics is often taught in classes with a large number of students. At USU, 50 to 60 students take the class in a fall semester and more than 100 students take it in a spring semester.

Table 1 summarizes seven studies that focused on the relationship between class size and student achievement. Three of them (Nos. 1-3) focused on the effect of class size on achievement for elementary school students. One (No. 4) studied the data collected from elementary school through high school. Three (Nos. 5-7) examined the effect of class size on undergraduate students. These studies, published between 1979 and 2002, yielded mixed results. Two studies (Nos. 3, 5) reported a nonsignificant effect, while the other four (Nos. 1, 2, 4, 6, 7) suggested a negative relationship between class size and student achievement.

Predicting Student Academic Performance

Need for predicting student academic performance. Prediction of student academic performance has long been regarded as an essential research topic in many academic disciplines for a number of reasons. First, predictive models can help the instructor predict student academic performance and then take some proactive measures (Veenstra, Dey, & Herrin, 2008; Ware & Galassi, 2006). With a validated predictive model, an instructor can identify academically at-risk students. The instructor may

Table 1

Studies on the Relationship Between Class Size and Student Achievement

No.	Researcher & year	Participants	Research method	Relationship
1	Cahen & Filby, 1979	Elementary	Qualitative	Negative
2	Angrist & Lavy, 1999	Elementary	Quantitative	Negative
3	Hoxby, 2000	Elementary	Quantitative	N/A
4	Levin, 2001	Elementary to high	Quantitative	Negative
5	Kennedy & Siegfried, 1997	Economics undergraduate	Quantitative	N/A
6	Kopeika, 1992	Engineering undergraduate	Quantitative	Negative
7	Dillon, Kokkelenberg, & Christy, 2002	Undergraduate	Qualitative	Negative

consider adopting specific instructional strategies for those academically at-risk students. For example, if a model predicts that a student will receive a final exam score below 50 (out of 100), he or she will be identified as potentially academically at-risk. The student might first be interviewed, followed by the observation of his/her classroom performance. This will help the instructor to develop a clear understanding of that student's learning skills and difficulties. Based on the instructor's judgment, additional instructional interventions may be implemented on that student. A detailed discussion of these instructional interventions is beyond the scope of this research; however, some examples of additional instructional interventions may include one-on-one tutoring and review of important concepts and principles after class, assigning more representative technical problems for additional student, providing remedial lessons to improve the student's mathematical skill, and asking the student to review previously learned concepts in relevant courses. Computer simulations and visualization of dynamics problems can also

help the student understand the processes on a deeper level.

Additionally, the results of predictive models can help the instructor to develop an effective intervention strategy to reduce the dropout rate of students from relevant courses or programs (Lowis & Castley, 2008). In Lowis and Castley's 2-year study, a questionnaire based on "Seven Principles of Good Undergraduate Teaching" was employed to predict student learning progression and academic achievement. In the first phase of their study, approximately 200 psychology students were surveyed during a scheduled class of their first year at a university in the East Midlands. The results showed that the students who eventually withdrew from the class before the mid-term of their first year had low scores in the questionnaire. In the second phase of their study, 116 psychology freshmen responded to the questionnaire after Week 7. Twenty-eight students were predicted to withdraw. Fifteen of the students were included in the intervention group and were asked to explain reasons for their answers to the questionnaire and to analyze their strengths/weaknesses. The other 13 students were placed in the control group. At the end of the first year, four students in the control group withdrew; however, no student in the intervention group withdrew.

A third positive effect of predictive modeling is that the instructor can employ the predicted results to modify existing course curriculum, such as the redesign of cooperative learning activities like group work. Although cooperative learning is reported to have a positive effect on student academic achievement (Brush, 1997), studies show that the group with ability-matched members would gain higher achievement than the group with one member that performs significantly better than the other members

(Nihalani, Wilson, Thomas, & Robinson, 2010; Onwuegbuzie, Collins, & Elbedour, 2003). Predictive models allow the instructor to identify a student's academic skills. According to the predicted results, the students with compatible skills can be grouped together to maximize the success of cooperative learning for all students involved.

Finally, students themselves can also use the predicted results to develop the learning strategies that are most effective for them personally. A predictive model helps students to develop a good understanding of how well or how poorly they would perform in a course. From the predicted results, academically at-risk students may rethink the way in which they have been studying. Ultimately, with help from the instructor, these students may design a better learning strategy to improve their success in the course.

Validation of the predictive models. Validation of the predictive models includes internal and external validation and reflects the differences between predicted values and actual values (Das et al., 2003; Bleeker et al., 2003). Internal validation is the “estimation of the prediction accuracy of a model in the same study used to develop the model” (Glossary Letter I, 2011, para. 51). External validation is the process of validating the developed models “using truly independent data external to the study used to develop the models” (Glossary Letter E, 2011, para. 69). Das et al. (2003) employed prediction accuracy to assess the internal and external validation of the predictive models. Artificial neural network and multiple-logistic-regression models were developed to predict outcome of lower-gastrointestinal haemorrhage. Data from 190 patients in one institution were used to train and internally validate the predictive models. The predictive models were externally validated by using data from 142 patients in another institution.

Prediction accuracy was calculated by the ratio of the correct predictions to total predictions. Results showed that neural network models had similar prediction accuracy to multiple-logistic-regression models in internal validation, but were, however, superior to multiple-logistic-regression models in external validation.

Another study conducted by Bleeker et al. (2003) suggested that external validation, which was assessed by prediction accuracy, was necessary in prediction research. In total, 376 datasets were used to develop and internally validate a predictive model and 179 datasets were used to externally validate the model. The ROC area was employed to measure prediction accuracy, and dropped from 0.825 in internal validation to 0.57 in external validation. The poor external validation indicated necessary of refitting the predictive model. The ROC area of refitted model was 0.70.

Factors that influence the prediction accuracy of predictive models. The prediction accuracy of a predictive model is affected by at least two factors: (1) the selection of predictors and (2) the mathematical techniques that are used to develop the predictive model. On the one hand, the prediction accuracy of a predictive model changes with different predictors. Lykourantzou, Giannoukos, and Mpardis (2009) compared the mean absolute error of prediction accuracy generated by different predictors. In their study, data of 27 students or 85% of a class in a 2006 semester were used to train the model, and data of five students or 15% in the same semester were used as the internal validation dataset. Another dataset of 25 students in a 2007 semester were used for external validation. Students took four multiple-choice tests: mc1, mc2, mc3, and mc4. Three predictive models developed using neural network were compared: model #1 used

mc1 and mc2 as input variables; model #2 used mc1, mc2, and mc3 tests; and model #3 used all four tests. While keeping all other conditions the same but with different predictors, the mean absolute error of prediction accuracy was 0.74 for model #1, 1.30 for model #2, and 0.63 for model #3.

On the other hand, the mathematical techniques used to develop a predictive model also affect the accuracy of prediction. In the same study (Lykourantzou et al., 2009), two modeling techniques—neural network and multiple linear regression—were compared. In terms of the mean absolute error, predictions from all the neural network models were more accurate than those of MLR models. The mean absolute error of the prediction accuracy of neural network models was only 50% of that of the corresponding MLR models. Another comparison was made by Vandamme, Meskens, and Superby (2007) which predicted students' academic success early in the first academic year. In total, 533 students from three universities were classified into three achievement categories: low-risk, medium-risk, and high-risk students. The mathematical techniques used in the Vandamme et al. (2007) study included decision trees, neural networks, and linear discriminant analysis. Their results showed that linear discriminant analysis had the highest rate of correct classifications based on the collected samples. However, none of the three models had a high rate of correct classification. They found that a larger sample size was needed to increase the rate of correct classification for each model.

Factors that affect student academic performance. The following paragraphs introduce the factors that affect student academic performance.

Prior domain knowledge. Domain knowledge is an individual's knowledge of a

particular content area, such as mathematics (Alexander, 1992; Dochy, 1992). Prior domain knowledge is defined as the knowledge that is available before a certain learning task and contains conceptual and meta-cognitive knowledge components (Dochy, De Rijdt, & Dyck, 2002). Prior domain knowledge is often measured by the grades earned in diagnostic exams or pretests (see Table 2). In this research, prior domain knowledge refers to the mathematical and physical knowledge students learned in the prerequisite courses.

Table 2

The Effects of Student Prior Knowledge on Academic Performance

Researcher & year	Participants		Sample size	Major/class	Variables examined
	Freshman	Higher			
Danko- McGhee & Duke, 1992	100%		892	Intermediate Accounting	Overall GPA, related course grades, diagnostic exam
O'Donnell & Dansereau, 2000	100%		108	Education and psychology	Prior knowledge of the ANS and PT ^a
Hicks & Richardson, 1984		100%	312	Intermediate Accounting	Diagnostic test, overall GPA, principles GPA
Thompson & Zamboanga, 2004	85%	25%	353	Psychology	ACT, pretest
Hailikari et al., 2008	67%	33%	139	Mathematics	Math tasks, GPA

^aANS: autonomic nervous system; PT: probability theory

A number of studies, such as those shown in Table 2, have investigated the effect of prior domain knowledge on student academic performance. Two of these studies (Hailikari, Nevgi, & Komulainen, 2008; Thompson & Zamboanga, 2004) focused on the impact of prior domain knowledge on student academic achievement at the college level. Hailikari and colleagues' (2008) study indicated that compared to prior academic success and self-belief, a student's prior domain knowledge was the strongest variable that contributed to his/her academic achievement in related classes ($\beta = .42, p < .001$). Thompson and Zamboanga (2004) designed a study to investigate the effect of prior domain knowledge on course achievement for freshmen psychology students. Their prior domain knowledge was measured by using two pretests, one to determine academic knowledge of psychology and another to gauge familiarity with popular psychology. The results of this study showed that for both pretests, psychological knowledge ($r = .37$) and popular psychology ($r = .20$), were significantly ($p < .01$) correlated with new learning. However, only the pretest of scholarly knowledge was identified as the most significant predictor for student academic performance.

Other similar studies have been conducted with students from different academic backgrounds including Hicks and Richardson (1984) and Danko-McGhee and Duke (1992) who used diagnostic tests to investigate the effect of students' prior domain knowledge on new learning. Hicks and Richardson (1984) found that a high correlation existed between diagnostic scores and course scores that students earned in an intermediate accounting class ($r = .57, p < .001$). A 2-year study was conducted by Danko-McGhee and Duke (1992) to explore the variables related to students' grades in an

accounting course. These research findings supported Hicks and Richardson's (1984) conclusion that the diagnostic examination, which was related to prerequisite courses, shared a relatively high variance with course performance ($R^2 = .19$).

However, it must be noted that the quality of students' prior domain knowledge is a significant factor. In other words, prior knowledge that contains inaccuracies and misconceptions may also hinder new learning (Hailikari et al., 2008; O'Donnell & Dansereau, 2000; Thompson & Zamboanga, 2004). Fisher, Wandersee, and Moody (2000) found that prior knowledge profoundly interacted with learning and resulted in a diverse set of outcomes. New learning may be seriously distorted if prior knowledge contains significant misconceptions or inaccuracies of a subject matter.

Extensive literature review shows that prior domain knowledge is generally a reliable predictor of student academic performance in a variety of courses. Approximately 95% of studies in different academic fields support the claim that students' prior knowledge, especially domain knowledge, has a significant positive impact on student academic performance (Dochy et al., 2002). Nevertheless, the impact varies according to the amount, completeness, and correctness of students' prior knowledge. As Dochy et al. (2002, p. 279) concluded, "the amount and quality of prior knowledge substantially and positively influence gains in new knowledge and are closely linked to a capacity to apply higher order cognitive problem-solving skills."

Prior achievement. In this study, prior achievement refers to a student's cumulative GPA, not the grade the student earned in a particular course.

On the one hand, prior achievement is correlated with prior knowledge and affects

academic performance. Hicks and Richardson (1984) studied the impact of prior knowledge and prior achievement on the academic performance of accounting students. The descriptive analysis they performed showed that a moderate correlation ($r = .31$) existed between a student's overall GPA (prior achievement) and diagnostic score (prior knowledge) in a particular class.

On the other hand, some studies in a variety of academic disciplines confirmed that GPA (prior achievement) has a significant direct effect on student achievement. In the same study mentioned above, Hicks and Richardson (1984) also found a strong correlation ($r = .52$) between a student's overall GPA and his/her final grade in an accounting course. A simple linear regression was employed based on students' overall GPAs and course grades. The results showed that overall GPA shared 27.3% variance of a student's final grade. Based on the data collected from 471 students who had been recruited from four sections in an introductory psychology course, Harachiewicz, Barron, Tauer, and Elliot (2002) found that student high school performance was a positive predictor of their short-term and long-term academic success. Similar results have also been found in economics (Emerson & Taylor, 2004), mathematics (Hailikari, Nevgi, & Ylanne, 2007), agriculture (Johnson, 1991), chemistry (Ayan & Garcia, 2008), and engineering (Fletcher, 1998; Wilson, 1983) disciplines.

Some studies investigated the impact of prior achievement on academic success without specifying students' majors. For example, Hoffman and Lowitzki (2005) collected a set of data from 522 "non-major students" at a private Lutheran university to study the effect of students' characteristics on their academic success. The results

revealed that the impact of high school grades varied with a student's ethnicity and race. Prior achievement was a significant and strong predictor of academic performance for white students and students of color, but not for non-Lutherans. Although the sample was very similar to the overall population at the university level, the research findings may not be generalizable because of the strong religion influence in Hoffman and Lowitzki's (2005) study.

Standardized tests. The Scholastic Aptitude Test (SAT) and the American College Test (ACT) are two standardized tests widely used to measure students' academic skills in the U.S. (Harachiewics et al., 2002). Some studies suggested that SAT/ACT scores were significant predictors of academic performance, but SAT/ACT scores were not as precise an indicator as was prior achievement (Camara & Echternacht, 2000; Fleming & Garcia, 1998; Hoffman, 2002). Some other studies found no relationship between SAT scores and achievement in a group of students (Emerson & Taylor, 2004).

The predictive validity of standardized test scores may be affected by some factors such as race. Fleming (2002) conducted a study to compare the impact of standardized test scores on students of different races. His results indicated that, on average, standardized test scores had a correlation of 0.456 with student academic success. However, SAT has higher predictive validity for Black freshmen who attended Black colleges ($R^2 = .158$) than for White freshmen attending primarily White colleges ($R^2 = .092$).

Students' grades may also affect the predictive validity of standardized test scores. In the above-mentioned article (Felmig, 2002) that studied prediction of student

academic performance from standardized test scores, SAT/ACT scores were found to be significant predictors in the first year of college. However, SAT/ACT scores had a weak or even nonsignificant relationship with academic performance as a student's academic career progressed. It is therefore reasonable to conclude that standardized tests, which are generally taken by students in high school, have significant and high correlation coefficients for student academic performance in the first year in college, but have a weak and low correlation with student academic performance beyond the first year.

Other influencing factors. Some research considered noncognitive variables, such as personality traits like leadership and self-efficacy, as predictors of student academic performance (see Table 3). It was found that the effects of noncognitive variables on student academic achievement differ according to the target groups and purpose of the predictive model. For example, in Ting's (2001) study, different predictors were identified for different target groups. For all students, SAT total score, positive self-concept, leadership experiences, and preference of long-term goals were identified as significant predictors. In predicting GPA for all male students, leadership experience did not contribute much and was excluded from the model. In predicting GPA for all female students, preference of long-term goals was excluded from the model.

In Lovegreen's (2003) study, all noncognitive variables had little contribution in predicting academic success of female engineering students in their first year of college. Although Lovegreen (2003) included similar noncognitive predictors, as did Ting (2001) and other researchers, different conclusions were made. The participants in Lovegreen's (2003) study (100 female first-year engineering students in a research-extensive

Table 3

Studies That Included the Use of Noncognitive Predictors

Study	Survey	Participants	Key cognitive predictors	Key noncognitive predictors
Ting, 2001	NCO	2800 first-year engineering students at North Carolina State University in fall of 1996	<ul style="list-style-type: none"> • SAT including mathematics, verbal, and total; • GPA in fall and spring 	<ul style="list-style-type: none"> • Positive self-concept; • Self-appraisal system; • Preference of long-term goals; • Leadership experience; • Demonstrated community service
Imbrie et al., 2006	Study Process Questionnaire;	1595 first-year engineering students in 2004, 1814 in 2005, and 1838 in 2006 at a large midwestern university	<ul style="list-style-type: none"> • Learning effectiveness; • Meta-cognition; 	<ul style="list-style-type: none"> • Motivation; • Self-efficacy; • Leadership; • Expectancy-value; • Team vs. individual orientation
Veenstra et al., 2009	PFEAS	2004-2005 engineering and general college freshman classes at the University of Michigan	<ul style="list-style-type: none"> • High school academic achievement; • Quantitative skills 	<ul style="list-style-type: none"> • Commitment to career and educational goals; • Confidence in quantitative skills; • Study habits; • Strong support person
Lovegreen, 2003	Noncognitive questionnaire	100 female first-year engineering students at a large research-intensive university	<ul style="list-style-type: none"> • SAT verbal and math 	<ul style="list-style-type: none"> • Self-assessment; • Self-concept; • Understanding of racism; • Long-range goals; • Strong support person; • Leadership experience; • Demonstrated community service

university) were different from those in other studies. The conflicting results from Lovegreen's (2003) study and other studies, such as Ting (2001), indicated that the contribution of noncognitive variables varies with target student groups and the purpose of the model.

As the first step for predicting student academic performance in engineering dynamics, this study focuses on the effects of a student's prior achievement and prior domain knowledge. The effects of noncognitive variables on student performance in engineering dynamics will be the focus of more studies in the future.

Statistical and Data Mining Modeling Techniques

Data mining is also called knowledge discovery in database (Han & Kamber, 2001). It integrates statistics, database technology, machine learning, pattern recognition, artificial intelligence, and visualization (Pittman, 2008). Data mining analyzes the observational datasets to summarize "unsuspected relationships" between data elements (Hand, Mannila, & Smyth, 2001). It has two functions:(a) to explore regularities in data, and (b) to identify relationships among data and predict the unknowns or future values. For the purpose of this research, three data mining techniques (MLP network, RBF network, and SVM) and one statistical technique, which are all commonly used for predictive modeling, are described.

Multiple Regression

Multiple regression takes into account the effect of multiple independent variables on a dependent variable and determines the quantitative relationships between them. If

the relationship between independent variables and a dependent variable is linear, a MLR may be employed. MLR is a “logical extension” of simple linear regression based on the least square principle (Field, 2005). It establishes quantitative linear relationships among these variables by using

$$\hat{y}_i = b_0 + b_1x_{i1} + b_2x_{i2} + \cdots + b_nx_{in}$$

where \hat{y}_i is the predicted value of a dependent variable;

x_i is the predictor, also called the predictor variable or the independent variable;

b_0 is the predicted intercept of y_i ;

b_i is the regression coefficient.

In the least-square estimation process, parameters for the multiple regression model, which can minimize the sum of squared residuals between the observed value and the predicted value, are calculated as (Everitt, 2009)

$$\hat{b} = (X'X)^{-1}X'y$$

where $y = [y_1, y_2, \cdots, y_m]'$

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1n} \\ 1 & x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix}$$

However, if the relationship between independent variables and the dependent

variable is nonlinear, three approaches are commonly used to estimate the nonlinear relationship in multiple regression: polynomial regression, nonlinear transformation (also called intrinsically nonlinear), and nonlinear regression (Cohen, Cohen, West, & Aiken, 2003).

Polynomial regression can approximate any unknown nonlinear relationships among the variables using additively exponential functions (Criddle, 2004)

$$\hat{Y} = b_0 + b_1X_1 + b_2X_1^2 + b_3X_1^3 + \dots + b_nX_n^n + \dots$$

The highest order (e.g., X^3 is of order 3) in polynomial regression determines the shape (the number of bends) of regression. For example, the quadratic equation

$\hat{Y} = b_0 + b_1X + b_2X^2$ generates one bend (a parabola) in regression. The cubic equation

$\hat{Y} = b_0 + b_1X + b_2X^2 + b_3X^3$ causes two bends (an S-shape) in regression.

By introducing variables X_i^2 , X_i^3 , etc., nonlinear relationships between X_i and Y can be determined. Regression equation 3 is “linear in the parameters” and can be analyzed with multiple regression (Cohen et al., 2003).

However, the variables X_i ($i=1,2,\dots,n$) need to be centered before employing polynomial regression because the equation is meaningful only if the variables X_i have meaningful zeros (Cohen et al., 2003). The full function for polynomial regression is:

$$\hat{Y} = b_0 + b_1(X_1 - \bar{X}_1) + b_2(X_1 - \bar{X}_1)^2 + b_3(X_1 - \bar{X}_1)^3 + \dots + b_n(X_n - \bar{X}_n)^n + \dots$$

Nonlinear transformation can change the relationship between the predictors X_i

and the dependent variable Y by changing the scale or units of the variables, such as changing $X (Y)$ to $\log X (\log Y)$, $a^X (a^Y)$, or $\sqrt{X} (\sqrt{Y})$. Nonlinear transformation can help simplify the relationships between X_i and Y by eliminating heteroscedasticity, and normalizing residuals (Cohen et al., 2003).

Three elements must be considered before choosing between the transformed variables and the original variables. First, one must consider whether the transformation is supported by relevant theories. Some psychophysical theories require nonlinear transformation to estimate the parameters of a model. The second aspect is the interpretation of the model. The final factor is the improvement of fit. Nonlinear transformation can substantially improve the overall fit of the model through simplifying the relationships between predictors and the dependent variable (Cohen et al., 2003).

Nonlinear regression is used to estimate the parameters of a nonlinear model which cannot be linearized by nonlinear transformation. A particular nonlinear equation must be specified to conduct nonlinear regression based on theory or the appropriateness of the relationships between predictors and the dependent variable, for example,

$$Y = c(e^{dX}) + \varepsilon_i \text{ (Cohen et al., 2003).}$$

Selection of predictor/independent variables. Four approaches are typically used to select appropriate predictor/independent variables from a list of candidate variables: forward selection, backward selection, stepwise regression, and the enter approach. With the forward selection approach, candidate independent variables are entered one by one into the initial model, which is a constant. The candidate variables that do not have a statistically significant contribution to the mean value of the predicted

value are excluded.

In the backward selection approach, all candidate independent variables are first included in the model. Then, candidate variables are successively removed until all remaining variables in the model cause a statistically significant change in the mean value of the predicted value if eliminated.

The stepwise regression method is a combination of both forward and backward selection. The initial model for the stepwise regression approach is a constant. Candidate independent variables are added to the model one by one. If a candidate variable makes a significant change to the mean of the predicted value, the variable will be temporarily kept in the model. If a candidate variable does not contribute significantly, the variables which were kept in the model earlier are removed from the model one by one to see if any more significant contributions will be generated by discarding one of the candidate variables.

With the enter approach, all candidate variables must be included in the model at first, with no regard to sequencing. Significant levels and theoretical hypotheses can assist a researcher in deciding which variables should be retained. Generally, the enter approach is the default method of variable entry in many commercial software packages, for example, SPSS.

Factors that affect the prediction accuracy of multiple regression. In theory, the best model should be achieved through any one of the three automatic selecting approaches (forward selection, backward elimination, and stepwise regression). However, an inferior model might be selected if, for example, two candidate independent variables

(such as X_1 and X_2) are highly correlated with each other. If this is the case, then, at least one candidate independent variable must be excluded from the model. Assume an automatic variable selection approach, such as stepwise, retains X_1 . It is possible that the model with X_2 is equal to or even better than the model containing X_1 . It is suggested that a healthy degree of skepticism be maintained when approaching the multiple regression model with automatic selection methods (Everitt, 2009).

Applications of multiple regression models. The multiple regression models have been widely employed for predicting student academic performance in a variety of disciplines. Delauretis and Molnar (1972) used stepwise regression to predict eight semesters of grade-point averages (GPA) for the 1966 freshman in engineering class at Purdue University. Precollege indicators, including high school rank, SAT score, ACT score, and cumulative college GPA, were incorporated into the predictor set. Based on a large sample size, Delauretis and Molnar (1972) found that college GPA was an effective predictor. Prediction accuracy ranged from 0.54 to 0.68 ($p < .01$) when precollege measurements and college GPA were used as predictors; however, prediction accuracy declined to 0.26 when using precollege measurements only. Delauretis and Molnar (1972) concluded that “it is overly simplistic to investigate GPA solely” and that further study was needed to construct a comprehensive model.

Marsh et al. (2008) developed multiple regression models to predict student academic performance (measured by GPA) in an introductory psychology course. Student information such as age, gender, classification, ACT, SAT, and general psychology exam scores collected from 257 students were used as predictors. Their results showed that

general psychology exam scores were an effective variable to predict GPA ($R^2_{exam1-5} = .46$), and general psychology exam scores had equal or greater predictive power than did SAT or ACT scores ($R^2_{SAT} = .06$, $R^2_{ACT} = .14$). Therefore, Marsh et al. (2008) suggested that scores in other required courses be used to predict student academic performance.

Neural Networks

Neural networks refer to a set of interconnected units/neurons that function in parallel to complete a global task. Two types of neural networks most commonly used include MLP and RBF networks. These two types of neural network models are introduced in the following paragraphs.

MLP network. MLP network, also known as multilayer feed forward neural network, is the neural network model that has been most widely studied and used (Maimon, 2008). It has a promising capability for prediction because of its ability regarding “functional mapping problems” in which one needs to identify how input variables affect output variables (Cripps, 1996; Maimon, 2008). Error back propagation is one of its key learning methods.

The schematic diagram graph of a multilayer perception neural network is shown in Figure 1. An MLP network contains an input layer, one or more hidden layers, and an output layer. Each layer consists of a set of interconnected neurons. The neurons, which include nonlinear activation functions, learn from experience without an explicit mathematical model about the relationship between inputs and outputs (Cripps, 1996). Sample data enter the network via the input layer, and exit from the output layer after

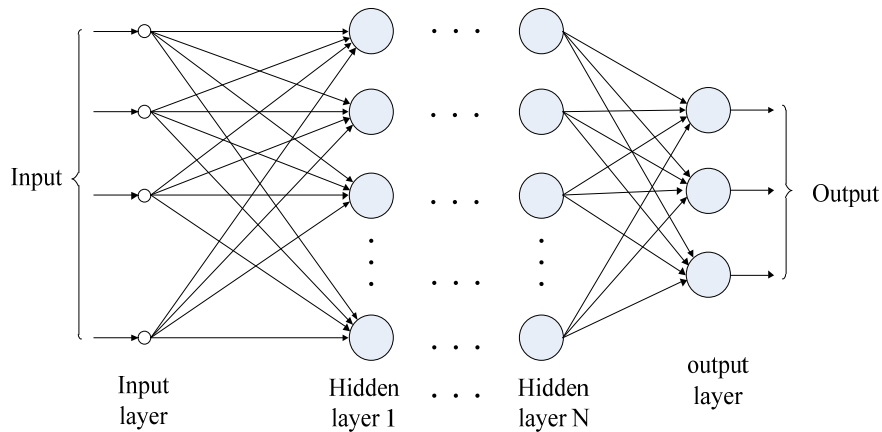


Figure 1. Schematic graph of a MLP neural network.

being processed by each hidden layer. Each layer can only influence the one next to it. If the output layer does not yield the expected results, the errors go backward and distribute to the neurons. Then the network adjusts weights to minimize errors.

Several factors may influence the accuracy of MLP, such as the number of layers, units in the hidden layers, activation function, weight, and learning rate. Increasing the number of layers and units may improve the prediction accuracy of the MLP network; however, it also increases complications and training time. Initial weight determines whether the network can reach a global minimum. The learning rate determines how much the weight is changed each time.

RBF network. RBF network is a three-layer feed-forward network. It takes the RBF function as the activation function in the hidden layer, and a linear function as the activation function in the output layer (Maimon, 2008). This RBF network approach can estimate any continuous function, including nonlinear functions, and has a good generalization capability.

The prediction accuracy of the RBF network is mainly affected by the number of units in the hidden layer. If the number is too small, the network is too simple to reflect the objective; however, if the number is too large, over-fit may occur and the generalization capability of the network would decline.

Factors that affect the prediction accuracy of neural network models.

Although neural networks are good at learning and modeling, one possible shortcoming of neural networks is over fitting, which cannot be overlooked. When over fitting occurs, the predictive capability of the neural network model will be decreased (Fulcher, 2008). This means that the model is highly accurate only when the training dataset is used, but prediction falters if other dataset is included.

To avoid the over fitting phenomenon, it is necessary to prune the model, that is, separate the data that are used for building the predictive model into the training and testing datasets, and use the testing dataset to modify the model to prevent over fitting. In this way, the prediction accuracy of the neural network model can be improved when dealing with different datasets (Linoff & Berry, 2011).

Applications of neural network models. Although neural networks do not yield an explicit set of mathematical equations as does the MLR approach, it is popular in the educational research community because of its outstanding performance compared to traditional techniques such as multiple regression. Lykourantzou et al. (2009) used neural network models to predict student achievement in an e-learning class. Scores of four multiple-choice tests in an e-learning class in the 2006 semester (mc1, mc2, mc3, and mc4) were used as predictors. Data from 27 students or 85% of the class were used to

train the model, and data from five students or 15% of the class in the same semester were used as the internal validation dataset. Another set of data from 25 students in 2007 was used as the external validation dataset. Three neural network models were compared: NN1 model using mc1 and mc2 as inputs; NN2 model using mc1, mc2, and mc3 as inputs; and NN3 model using all mc tests as inputs. With different inputs, the mean absolute error of NN1, NN2, and NN3 was 0.74, 1.30, and 0.63, respectively. The neural network models were also compared with MLR models. A comparison of the mean absolute errors showed that all neural network models performed much better than the regression models. The prediction error of neural network models was approximately 50% compared to the corresponding regression models.

Support Vector Machine

SVM is a learning system developed by Vapnick (1995) based on the structural risk minimization (SRM) principle. Compared to the traditional empirical risk minimization (ERM) principle, which minimizes the errors in training data, SRM minimizes an upper bound on the expected risk. This feature enables SVM to be more accurate in generalization.

The SVM method was first used to handle classification problems (pattern recognition) by mapping nonlinear functions into linear functions “in a high dimensional feature space” (Cristianini & Taylor, 2000). However, by introducing a loss function, a SVM model can also be applied to regression problems as well (Gunn, 1998). For regression purposes, ϵ -insensitive loss function is often used (Deng & Tian, 2004; Stitson, Weston, Gammerman, & Vapnik, 1996). ϵ is the number that is so small that

smaller than which the predictive error (difference between the predicted value $f(x)$ and the actual value y) can be ignored. In general, ε is set as a small positive number or zero, for example, 0.001. Equation 1 and Figure 2 illustrate the ε -insensitive loss function.

$$L_\varepsilon = |y - f(x, \omega)|_\varepsilon = \begin{cases} 0 & \text{for } |y - f(x, \omega)| \leq \varepsilon \\ |y - f(x, \omega)| - \varepsilon & \text{otherwise} \end{cases} \quad (1)$$

where ω is the parameter to identify

ε is a user-defined precision parameter

Given a set of data $\{x_i, y_i\}, i=1, \dots, n, x_i \in R^d, y_i \in R$, where R^d is a Euclidean space, the linear regression function commonly used is shown in Equation 2 (Smola & Scholkopf, 2004):

$$f(x) = (w \cdot x) + b \quad (2)$$

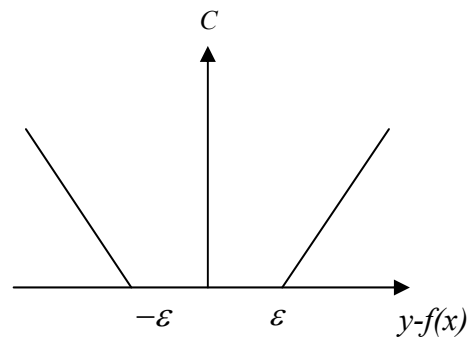


Figure 2. The ε -insensitive loss function.

The objective of regression is to find a function in the form of Equation 2 to yield minimal loss-function. Therefore, the initial constrained optimization problem is

$$\begin{aligned} & \min_{\omega \in R^n, b \in R} \quad \frac{1}{2} \|\omega\|^2 \\ \text{subject to} \quad & ((\omega \cdot x_i) + b) - y_i \leq \varepsilon, \quad i = 1, \dots, l \\ & y_i - ((\omega \cdot x_i) + b) \leq \varepsilon, \quad i = 1, \dots, l \end{aligned}$$

Considering the fitting error, two slack variables $\xi_i \geq 0$ and $\xi_i^* \geq 0$ are introduced. To minimize the ε -insensitive loss function $\|\omega\|^2/2 + C \sum_{i=1}^l (\xi_i + \xi_i^*)$, the equivalent primal optimization problem becomes

$$\begin{aligned} & \min_{\omega \in R^n, b \in R} \quad \|\omega\|^2/2 + C \sum_{i=1}^l (\xi_i + \xi_i^*) \\ \text{subject to} \quad & \begin{cases} y_i - \omega \cdot x_i - b \leq \varepsilon + \xi_i \\ \omega \cdot x_i + b - y_i \leq \varepsilon + \xi_i^* \end{cases} \quad i = 1, 2, \dots, l \end{aligned}$$

where constant $C > 0$. Constant C measures “the trade-off between complexity and losses” (Cristianini & Taylor, 2000) and stands for the penalty on the sample data which has a larger error than ε . To solve this quadratic optimization problem, Lagrange multipliers $\alpha_i, \alpha_i^*, \eta_i, \eta_i^*$ are introduced as (Cristianini & Taylor, 2000)

$$\begin{aligned} L(\omega, b, \xi, \xi^*) = & \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^l (\xi_i + \xi_i^*) - \sum_{i=1}^l \alpha_i (\varepsilon + \xi_i - y_i + \omega \cdot x_i + b) \\ & - \sum_{i=1}^l \alpha_i^* (\varepsilon + \xi_i^* + y_i - \omega \cdot x_i - b) - \sum_{i=1}^l (\eta_i \xi_i + \eta_i^* \xi_i^*) \end{aligned}$$

Then we have

$$\left\{ \begin{array}{l} \frac{\partial L}{\partial \omega} = \omega - \sum_{i=1}^l (\alpha_i - \alpha_i^*) \cdot x_i = 0 \\ \frac{\partial L}{\partial b} = \sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0 \\ \frac{\partial L}{\partial \xi_i} = C - \alpha_i - \eta_i = 0 \\ \frac{\partial L}{\partial \xi_i^*} = C - \alpha_i^* - \eta_i^* = 0 \end{array} \right.$$

The Lagrangian dual problem of the primary problem is defined as follows:

$$\begin{aligned} \max : W(\alpha, \alpha^*) &= \sum_{i=1}^l y_i (\alpha_i - \alpha_i^*) - \varepsilon \sum_{i=1}^l (\alpha_i + \alpha_i^*) \\ &\quad - \frac{1}{2} \sum_{i,j=1}^l (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) (x_i \cdot x_j) \\ \text{subject to} \quad &\sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0 \quad i = 1, 2, \dots, l \\ &0 \leq \alpha_i, \alpha_i^* \leq C \end{aligned}$$

The regression function at a given point is determined as

$$f(x) = (\omega \cdot x) + b = \sum_{i=1}^l (\alpha_i - \alpha_i^*) (x_i \cdot x) + b$$

where $(x_i \cdot x)$ is the dot product of vector x_i and vector x .

Nonlinear regression problems in a low-dimensional space can be mapped into linear regression problems in a high-dimensional space. The mapping process can be undertaken by SVM through using the kernel function $k(\cdot)$ to replace the dot product of vectors (Collobert & Bengio, 2001). Polynomial kernel, Gaussian kernel, and hyperbolic tangent kernel are often used. They are expressed as (Hong & Hwang, 2003)

Polynomial kernel $K(x, y) = (\langle x, y \rangle + 1)^P$

Gaussian kernel $K(x, y) = e^{-\frac{\|x-y\|^2}{2\sigma^2}}$

Hyperbolic tangent kernel $K(x, y) = \tanh(k \langle x, y \rangle + \theta)$

The optimization problem is thus defined as

$$\max : W(\alpha, \alpha^*) = \sum_{i=1}^l y_i (\alpha_i - \alpha_i^*) - \varepsilon \sum_{i=1}^l (\alpha_i + \alpha_i^*) - \frac{1}{2} \sum_{i,j=1}^l (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) K(x_i \cdot x_j)$$

$$\text{subject to} \quad \sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0 \quad i = 1, 2, \dots, l$$

$$0 \leq \alpha_i, \alpha_i^* \leq C$$

The regression function is

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) K(x_i \cdot x) + b$$

Factors that affect the prediction accuracy of SVM models. The prediction accuracy of SVM is mainly affected by two parameters: the penalty factor C and the kernel parameter. The penalty factor C determines penalty for the data whose deviations are larger than precision ε . They affect the prediction accuracy and the SVM model's ability to generalize. The kernel parameter affects the generalization ability of the SVM model. However, there is no standard method for optimizing the two parameters. The method most often used is the grid method (Chen, Wang, & Lee, 2004; Friedrichs & Igel,

2005).

Applications of SVM models. SVM has been used for many applications, such as pattern identification and image processing (Romon & Christodoulou, 2006). In recent years, SVM has also been applied in control engineering (Mohandes, Halawani, & Rehman, 2004). However, SVM has not yet been widely applied in educational research. One study using SVM to predict the dropout rate of new students was conducted by Kotsiantis et al. (2003). Data were collected from four written assignments, face-to-face consulting meeting with tutors, and final examinations. Various techniques were employed to identify dropout-prone students by using the collected data as well as other information including sex, age, and parental occupation. The results showed that SVM performed better than neural networks after the third training phase, which included both the data used for the second step and the data from the first written assignment. Only ordinal data were included in the study of Kotsiantis et al. (2003). However, a study has not yet been conducted to investigate the prediction accuracy of SVM in educational research that involves the use of continuous data.

Chapter Summary

In this chapter, studies of predicting student academic performance as well as four modeling techniques that can be used for developing predictive models were reviewed. It is shown that (a) academic performance of sophomore and junior students can be predicted by prior achievement and prior domain knowledge; and (b) modeling techniques, including multiple regression, MLP network, RBF network, and SVM may

influence the prediction accuracy of the models. Prediction accuracy can be employed to assess the internal and external validation of the predictive models.

CHAPTER III

RESEARCH DESIGN

The goal of this study was to develop a validated set of statistical and data mining models to predict student academic performance in an engineering dynamics course. This chapter describes how the predictive models were developed using six combinations of predictors and four modeling techniques (MLR, MLP network, RBF network, and SVM). The models were developed and validated based on the quantitative data of student academic performance collected during four semesters from 2008 to 2011. The criteria used to evaluate and compare the models are also defined.

The three objectives of this research were as follows:

1. Identify and select appropriate mathematical (i.e., statistical and data mining) techniques for constructing predictive models.
2. Identify and select appropriate predictor variables (i.e., independent variables) that can be used as inputs for predictive models.
3. Validate the developed models using the data collected during multiple semesters to identify academically-at-risk students.

Three research questions were designed to address each research objective:

1. How accurate will predictions be if different statistical and data mining modeling techniques such as traditional multiple linear regression, MLP networks, RBF networks, and SVM are used?
2. What particular combination of predictor variables will yield the highest prediction accuracy?

3. What is the percentage of academically-at-risk students that can be correctly identified by the models?

Overall Framework

Cabena, Hadjinian, Stadler, Verhees, and Zanasi (1997) created a five-stage model for data mining processes, including the determination of business objectives, data preparation, data mining, results analysis, and knowledge assimilation. Feelders, Daniels, and Holsheimer (2000) illustrated six stages of the data mining process, including defining the problem definition, acquiring background information, selection and preprocessing of data, analyzing and interpreting, as well as reporting acquired data. Pittman (2008) proposed a data mining process model for education, which includes determining a dataset based on student retention rates, domain knowledge, and data availability. The next steps would be extracting data from a data warehouse, generating instances, calculating derived variables, and assigning outcome variables. The last step would entail generating descriptive and exploratory statistics for the dataset and eliminating highly correlated variables and normalizing numeric data elements.

The modeling framework of this study was based on the data mining process models described above. Figure 3 shows the modeling framework.

Data Collection

Students who were enrolled in ENGR 2030 Engineering Dynamics in the College of Engineering at Utah State University in Fall 2008-Spring 2011 participated in this study (see the Appendix for a copy of the IRB approval letter). Approximately 120

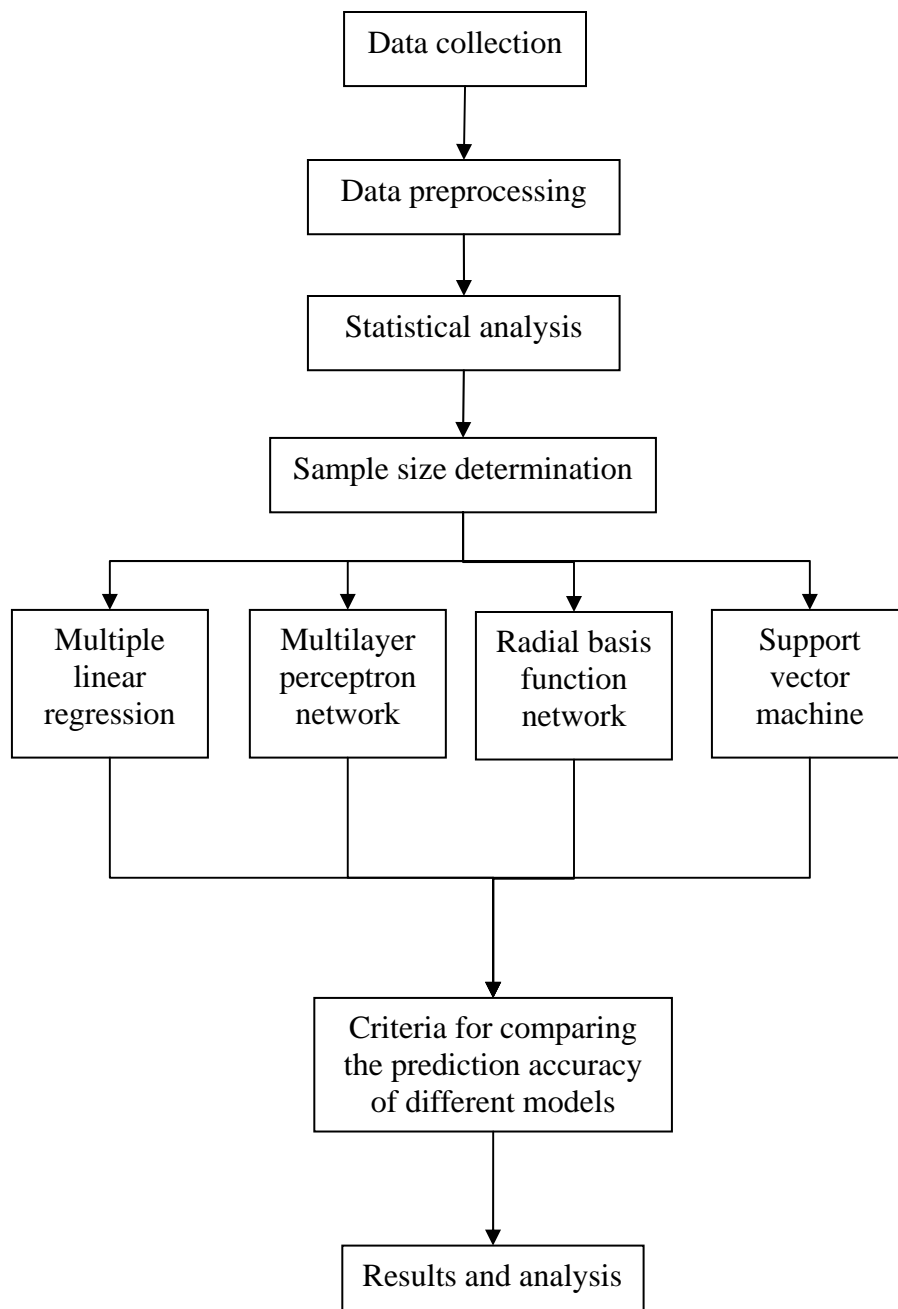
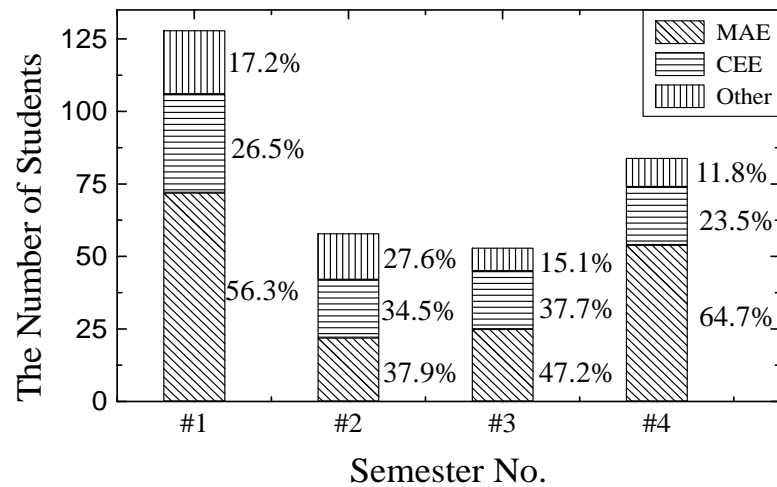


Figure 3. The modeling framework of this study.

students enrolled in the engineering dynamics course in spring semester, and 60 students enrolled in this course in fall semester.

Information regarding student academic performance was collected from a total of 324 students in four semesters: 128 students in Semester #1 (Spring 2009), 58 students in Semester #2 (Fall 2008), 53 students in Semester #3 (Fall 2009), and 85 students in Semester #4 (Spring 2011). The reason for assigning Spring 2009 as Semester #1 was the largest number of students enrolled in that semester; therefore, the data collected in Spring 2009 were more representative. Figure 4 shows student demographics. As seen in Figure 4, the majority of the 324 students were either mechanical and aerospace engineering majors (174, or 53.7%) or civil and environmental engineering majors (94, or 29%).

Candidate variables to be used as predictors. Based on extensive literature review and the experience in teaching engineering dynamics, data regarding students' prior achievement, domain-specific prior knowledge, and learning progression were collected. Eight variables (X_1, X_2, \dots, X_8) were selected as the candidate predictor/independent variables of the predictive models. X_1 (cumulative GPA) indicates prior achievement. $X_2 \sim X_5$ (grades earned in the prerequisite courses for engineering dynamics) indicate prior domain knowledge. $X_6 \sim X_8$ (grades earned from three engineering dynamics mid-term exams) indicate learning progression in this particular course. Data collected from four semesters in Fall 2008-Spring 2011 were used to develop and validate the models.



*MAE: Mechanical and aerospace engineering

*CEE: Civil and environmental engineering

*Other: Biological engineering, general engineering, pre-engineering, undeclared, or nonengineering majors

Figure 4. Student demographics.

The reasons for selecting these particular variables are discussed below.

- X_1 (cumulative GPA) was included because it is a comprehensive measurement of a student's overall cognitive level.
- X_2 (statics grade) was included because numerous concepts of statics (such as free-body diagram, force equilibrium, and moment equilibrium) are employed throughout the dynamics course.
- X_3 and X_4 (calculus I and II grades) are an accurate measurement of a student's mathematical skills needed to solve calculus-based dynamics problems.
- X_5 (physics grade) was used to measure a student's basic understanding of physical concepts and principles behind various dynamics phenomena.
- X_6 (score of dynamics mid-term exam #1) measures student problem-solving

skills concerning “kinematics of a particle” and “kinetics of a particle: force and acceleration.”

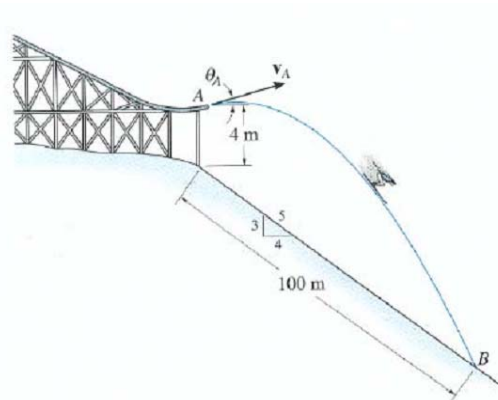
- X_7 (score of dynamics mid-term exam #2) measured student problem-solving skills concerning “kinetics of a particle: work and energy” and “kinetics of a particle: impulse and momentum.”
- X_8 (score of dynamics mid-term exam #3) is a measurement of student problem-solving skills on “planar kinetics of a rigid body” and “planar kinetics of a rigid body: force and acceleration.”

The following examples explain three representative dynamics problems used to prepare students for the three dynamics mid-term and final exams. Knowledge of projectile motion, impulse and momentum, and general plane motion are tested in examples 1, 2, and 3, respectively.

Example 1:

Given: Skier leaves the ramp at $\theta_A = 25^\circ$ and hits the slope at B .

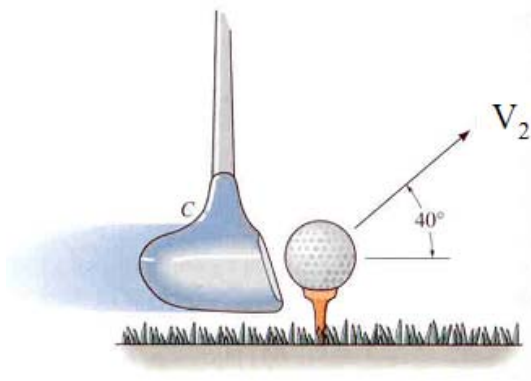
Find: The skier's initial speed v_A



Example 2:

Given: A 40 g golf ball is hit over a time interval of 3 ms by a driver. The ball leaves with a velocity of 35 m/s, at an angle of 40° . Neglect the ball's weight while it is struck.

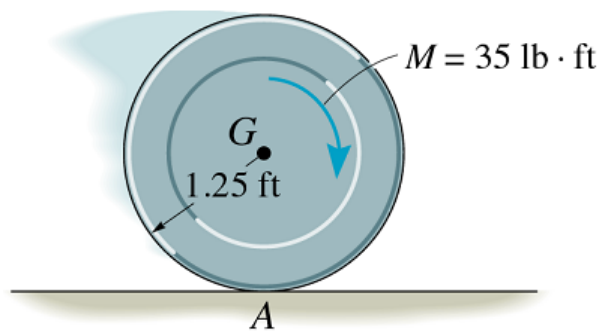
Find: The average impulsive force exerted on the ball.



Example 3:

Given: A 50 lb driving-wheel has a radius of gyration $k_G = 0.7$ ft. While rolling, the wheel slips with $\mu_K = 0.25$.

Find: The acceleration a_G of the mass center G .

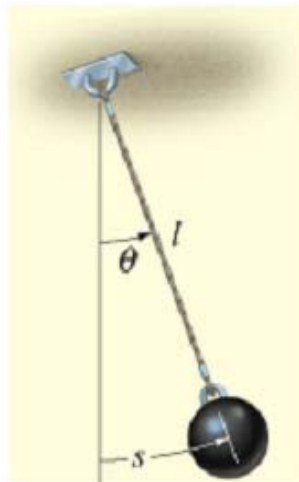


Independent variables. The dynamics final exam (the output Y) is comprehensive and covers all the above-listed dynamics topics as well as three additional topics that students learned after mid-term exam #3. The three additional topics included “planar kinetics of a rigid body: work and energy,” “planar kinetics of a rigid body: impulse and momentum,” and “vibration.” The following is one more example of the type of questions found on the final exam. This quotation (example 4) examines a student’s problem-solving skills in dealing with undamped free vibration.

Example 4:

Given: The bob has a mass m and is attached to a cord of length l . Neglect the size of the bob.

Find: The period of vibration τ for the pendulum.



For each student, nine data points, including eight predictor variables and one dependent variable, were collected: X_1 (cumulative GPA), X_2 (statics grade), X_3 and X_4 (calculus I and II grades), X_5 (physics grade), X_6 (score of dynamics mid-term exam #1),

X_7 (score of dynamics mid-term exam #2), X_8 (score of dynamics mid-term exam #3), and Y (dynamics final exam grade).

To solve the problems shown in examples 1, 2, 3, and 4, knowledge of the prerequisite courses including statics (X_2), calculus ($X_3 \sim X_4$), and physics (X_5) are required, such as scalars and vectors, the free-body diagram, moment of a force, integral and differential equations, kinematics in two dimensions, impulse and momentum, and tense force.

Data Preprocessing

The collected data ($Y, X_1, X_2, X_3, \dots, X_8$) were initially in different scales of measurement: X_1 varies from 0.00 to 4.00, while $X_2, X_3, X_4,$ and X_5 are letter grades from A to F; X_6 and X_8 vary from 0.00 to 15.00; X_7 from 0.00 to 16.00; and Y from 0.00 to 100.00. Before they could be of any use in mathematical models, these raw data must be preprocessed, which is described in the following paragraphs.

First, to establish a standard unit for all variables and make models easier to construct, all letter grades in $X_2, X_3, X_4,$ and X_5 were converted into the corresponding numerical values using Table 4.

Second, the numerical values of all data were normalized, so each datum varied within the same scale from 0 to 1, as shown in Table 5. There were two purposes for applying normalization. The first one was to avoid the cases in which one variable received a higher or lower weight for its coefficient due to its initial low or high scale of measurements. The second purpose was to decrease data processing time. The normalized value of data was calculated through dividing the initial value of the data by

Table 4

Conversion of Letter Grades

Letter grade	A	A-	B+	B	B-	C+	C	C-	D+	D	F
Numerical value	4.00	3.67	3.33	3.00	2.67	2.33	2.00	1.67	1.33	1.00	0.00

Table 5

Normalization of the Raw Data

Variables	Initial value of data	Normalized value of data
X_1 Cumulative GPA	0.00 - 4.00 (numerical value)	Initial value/4
X_2 Engineering Statics	Letter grade A, A-, B+, B, etc.	Initial value/4
X_3 Calculus I	Letter grade A, A-, B+, B, etc.	Initial value/4
X_4 Calculus II	Letter grade A, A-, B+, B, etc.	Initial value/4
X_5 Physics	Letter grade A, A-, B+, B, etc.	Initial value/4
X_6 Mid-Exam #1	0.00 - 15.00 (numerical value)	Initial value/15
X_7 Mid-Exam #2	0.00 - 16.00 (numerical value)	Initial value/16
X_8 Mid-Exam #3	0.00 - 15.00 (numerical value)	Initial value/15
Y Final Exam	0.00 - 100.00 (numerical value)	Initial value/100

its range. For instance, the range of GPA that a student could receive was 0.00-4.00.

Suppose that one student earned a GPA of 3.55, then that student's normalized GPA would be $3.55 \div 4.00 = 0.8875$.

The following five steps were performed before the predictive models were constructed:

First, in the case of missing student data, averages of all other records for the student were filled in to utilize the model to its full extent. For example, assuming the collected data for one student are $X_1 \sim X_7$, the missing value for X_8 would be estimated using Equation 3.

$$X_8 = \frac{X_1 + X_2 + X_3 + X_4 + X_5 + X_6 + X_7}{7} \quad (3)$$

However, the student would be excluded from the study if three or more data points were missing because glaring error may be introduced to the models if replacing these missing data points with average value of the student. Two cases in Semester #1, eight cases in Semester #2, four cases in Semester #3, and five cases in Semester #4 missed one data point, respectively. One case in Semester #2 missed two data points. One case in Semester #4 missed four data points that had to be excluded from the sample. Finally, the valid samples collected from the four semesters were as follows: 128 data sets in Semester #1, 56 data sets in Semester #2, 58 data sets in Semester #3, and 84 data sets in Semester #4. A total of 323 students, or in other words $323 \times 9 = 2,907$ data points from all four semesters, were collected.

A second challenge was to identify the outliers, which may be generated by measurement errors and rare cases. Outliers may significantly affect the correlation between independent and dependent variables by changing slope coefficients and standard error deviation. However, not all outliers deserve attention. Leverage, discrepancy, and influence were employed to identify the problematic outliers.

Third, descriptive statistics of the normalized data were employed. Information about the mean and standard deviation of the variables was generated. Histograms and scatter plots were employed to present the distribution of the data, including normality and the relationships between predictors and dependent variables.

Fourth, multiple collinearity was tested, which may occur when two or more

independent variables share too much variance. If adding one variable makes another variable flip the sign in regression, or the sign of one variable differs from theoretical expectations, there might be collinearity problems. Diagnostic statistical analysis was performed to detect collinearity. The variance inflation factor and tolerance redundancy were determined to assess the degree of collinearity (Cohen et al., 2003).

Finally, the correlation matrix was developed. Pearson's correlation, a number ranging from -1 to +1 that measures the degree and direction of the correlation between two continuous variables, was employed to demonstrate the correlation between eight independent variables and one dependent variable. The positive value for a correlation coefficient implies that the two variables trend in the same direction, while the negative value for a correlation coefficient implies the two variables trending in the opposite direction. The higher the absolute value of a correlation coefficient, the stronger the relationship between the two variables. Pearson's correlation coefficient r is computed as

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{(n-1)S_X S_Y}$$

where X and Y are two variables, and S_X and S_Y are the sample standard deviations of X and Y (Howell, 2010).

Criteria Used for Assessing Prediction Accuracy

Data collected from the first semester were employed for internal validation of the predictive models. Data collected from the other three semesters were used to test for

external validation. Data for both internal and external validations were from the same population because the participants learned the dynamic course with the same instructor at the same university.

The prediction accuracy of each model was examined by using the following two criteria:

1. Average prediction accuracy (APA) indicates, on average, how well the model predicts the final exam scores of students in the dynamics course. The average prediction accuracy for the final exam scores was calculated as

$$\text{APA} = 1 - \frac{1}{n} \cdot \sum_{i=1}^n \left| \frac{P_i - A_i}{A_i} \right| \times 100\%$$

where n is the total number of cases, P_i is the predicted final exam score of the i^{th} student in the class ($i = [1, n]$), and A_i is the actual final exam score of the i^{th} student. The higher the average prediction accuracy, the better the model.

2. Percentage of accurate predictions (PAP). The percentage of accurate predictions among all predictions was calculated as the number of accurate predictions divided by the total number of predictions. In this study, an accurate prediction was defined as the prediction in which the predicted value is within 90-110% of the actual value (namely, the prediction error is $\pm 10\%$). The higher the percentage of accurate predictions, the better the model.

Determining the Appropriate Sample Size for Predictive Model Development

Statistical analysis was performed to determine the minimum sample size for developing effective predictive models. Different sample sizes were tested to determine the appropriate sample size to be used in the training of predictive models.

To determine the minimum sample size for developing predictive models, a desired power needs to be set. The power is defined as the probability that a null hypothesis will be rejected when the null hypothesis is false (Bezeau & Graves, 2001; Cohen, 1962). The cost of committing type II error when compared to the cost of gathering research data determines which power to choose. Generally, a quite large power is 0.95 or higher, and a small power is around 0.60 (Cohen et al., 2003). Most studies choose a power value from 0.70 to 0.90. The power value of 0.80, which falls between 0.70 and 0.90, is a reasonable one to choose (Cohen, 1988) and was used in this study.

Power analysis concerns the relationships among power, sample size, significance criterion (α), and the effect size (ES) f^2 . The necessary sample size can be determined if the ES, desired power, and α are available. Generally, the more predictors included, the larger the sample size needed. To estimate the minimum sample size to develop all predictive models, the number of predictors was set at eight, which is the maximum value in this study. An online statistics calculator (Soper, 2004) was used to estimate the sample size at the given desired power of 0.8, the alpha level of 0.05, and the number of predictors of eight. A medium effect size was employed as the anticipated effect size.

To confirm sample size for training the predictive models, the MLR technique

was employed to develop a set of models using different sample sizes. MLR was selected to determine sample size because it has been a traditional statistical technique in educational research and was easy to use

Dataset for Semester #1 (Spring 2009) was randomly split into a training dataset and a testing dataset using various combinations as follows:

- 30% of the full dataset as the training dataset and the remaining 70% as the testing dataset
- 40% of the full dataset as the training dataset and the remaining 60% as the testing dataset
- 50% of the full dataset as the training dataset and the remaining 50% as the testing dataset
- 60% of the full dataset as the training dataset and the remaining 40% as the testing dataset
- 100% of the full dataset as the training dataset.

Five MLR models were generated. Datasets collected during Semester #1 (Spring 2009) were used for internal validation while data sets collected in Semester #2 (Fall 2008) were used for external validation. APA and PAP were employed to compare the prediction accuracy of the five models generated by using five different sample sizes.

Predictive Modeling

Training data were finally selected based on the appropriate sample size, which was using the method described in the section above. Four statistical and data mining

techniques were used to develop the predictive models by using six combinations of predictor variables listed below.

- I. X_1 used as predictor
- II. $X_1, X_2, X_3, X_4,$ and X_5 used as predictors
- III. X_6 as the predictor
- IV. $X_1, X_2, X_3, X_4, X_5,$ and X_6 used as predictors
- V. $X_1, X_2, X_3, X_4, X_5, X_6,$ and X_7 used as predictors
- VI. $X_1, X_2, X_3, X_4, X_5, X_6, X_7,$ and X_8 used as predictors

Combination I and II only consider a student's prior achievement and prior knowledge before taking the dynamics course. Combination III only considers a student's early performance in the dynamics class by including results on Exam #1. Combination IV considers not only prior achievement, but also a student's early performance (the first dynamics mid-term exam) in class. Combination V takes into consideration a student's prior achievement and the performance in the first and second dynamics mid-term exams. Combination VI includes a student's prior achievement and the performance in all three dynamics mid-term exams (i.e., Exams #1, #2, and #3).

The predictive models developed with the first combination of predictors can be applied before the dynamics course starts. Thus, it would be possible for the instructor to design a specific course curriculum and choose proper learning aids according to the predicted results at the beginning of the semester. The predictive models with combinations III-VI can only be used as the dynamics course proceeds. X_6 would not become incorporated until the end of the first quarter of the semester, and X_7 not until the

middle of the semester, while X_8 would not come into play until the last quarter of the semester. The instructor may choose different combinations during different periods in a semester according to the needs of each class.

The predictive models were developed by using four statistical and data mining techniques, including MLR, MLP network, RBF network, and SVM, as well as the six combinations of predictors.

The commercial software package SPSS 18 was employed for constructing multiple regression, MLP, and RBF models. MATLAB was used to develop the SVM models. All candidate predictors in various combinations were adopted as inputs for MLP, RBF, and SVM models regardless of the statistical significance of the candidate predictors.

MLR Models

The MLR models were developed using the “enter” mode. The statistical significance threshold of 0.05 was adopted, which is the most commonly used threshold for predicting student academic performance $p < .05$ (Marsh et al., 2008; Thompson & Zamboanga, 2004; Ting, 2001).

However, all the inputs were kept in the regression models regardless of their significance level. The reason is justified as follows. When different modeling techniques are used to create a new model, the contribution of each predictor varies with the techniques. For example, cumulative GPA was the most important predictor for one MLR model, while it was the second most important predictor in another MLP model. The results would be biased if the regression models used only significant predictors while the

other models used that all input predictors regardless of significance.

Explicit mathematical equations were generated in the following form:

$$\hat{Y} = a_0 + [b_i \ \cdots \ b_n][X_i \ \cdots \ X_n]^T$$

where matrix X_i represents one of the six combinations of predictors; and the matrix of b_i represents corresponding regression coefficients.

Neural Network Models

An arbitrary value was set for MLP/RBF models using a random number generator. A small testing sample, generally smaller than the training sample, is able to train the neural network more efficiently. Eighty percent data were used as the training sample, while the other twenty percent were used as the testing sample to trace errors during training to prevent overtraining. The default value of relevant parameters in SPSS, such as the minimum relative change in training error, the minimum relative change in training error ratio, and the maximum training epochs, were adopted and optimized automatically with specific criteria and algorithms.

SVM Models

M files in MATLAB were employed to construct the SVM models. The RBF kernel, one of the kernels most commonly seen in SVM regression, was used (Chapelle & Vapnik, 2000; Hong & Hwang, 2003; Thissen, van Brakel, de Weijer, Melssen, & Buydens, 2003; Trafalis & Ince, 2000). The basic idea of SVM regression is to map the data into a high dimensional feature space via a nonlinear map (Chapelle, Vapnik,

Bousquet, & Mukherjee, 2002; Hearst, 1998). As described in the second session in Chapter II, the following dual-Lagrangian problem is solved when constructing an SVM model:

$$\max : W(\alpha, \alpha^*) = \sum_{i=1}^l y_i (\alpha_i - \alpha_i^*) - \varepsilon \sum_{i=1}^l (\alpha_i + \alpha_i^*) - \frac{1}{2} \sum_{i,j=1}^l (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)(x_i \cdot x_j)$$

$$\text{subject to} \quad \sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0 \quad i = 1, 2, \dots, l$$

$$0 \leq \alpha_i, \alpha_i^* \leq C$$

The regression function is

$$f(x) = (\omega \cdot x) + b = \sum_{i=1}^l (\alpha_i - \alpha_i^*) (x_i \cdot x) + b$$

Two parameters, penalty factor C and the width of kernel σ^2 , affect the prediction accuracy of an SVM model. Eight points from a prediction accuracy curve were selected to show how the two parameters C and σ^2 affect prediction accuracy.

The grid method is often used to optimize C and σ^2 for SVM models (Cherkassky & Ma, 2004; Momma & Bennett, 2002; Staelin, 2002). In the defined range and minimum unit, the grid method searches by increasing one unit of one variable at a time. For example, assume C is in the range $[a, b]$, with every m as a unit; while σ^2 is in the range $[w, z]$, with every n as a unit. The grid method first generates results with $[C, \sigma^2] = [a, w]$; then results with $[C, \sigma^2] = [a+m, w+n]$; and finally results with $[C, \sigma^2] = [b, z]$.

However, the extent to which C and σ^2 can be optimized depends on the pace of the grid method (Staelin, 2002), that is the m and n values mentioned above. If m and n are large, the optimization results may not be accurate enough because the relationship between the prediction accuracy of SVM and the number of parameters is not linear. If m and n are small, the grid method will be time consuming. For example, in SVM models, data collected in Semester #1 and #2 were used to demonstrate how the penalty factor C and the width of kernel σ^2 affect the prediction accuracy of SVM models when $X_1 \sim X_8$ are used as predictors. The full dataset collected in Semester #1 was used to train and internally validate the SVM models. Data collected in Semester #2 were used for external validation. Figures 5-8 show how the change of the two parameters C and σ^2 affect the average prediction accuracy and the percentage of accurate predictions using internal and external validations. The results show that the penalty factor C and the width of kernel σ^2 affect the prediction accuracy of the SVM model in a nonlinear way.

Genetic algorithms were employed to overcome the shortcomings of the grid method and optimize parameters C and σ^2 (Pai & Hong, 2005). In this study, genetic algorithms select the fittest member and pass the genetic information from one generation to the next. Selection, crossover, and mutation are three main processes associated within genetic algorithms. The flow chart of genetic algorithms is shown in Figure 9.

The relevant parameters of genetic algorithms were set as follows:

The maximum number of generations (max gen) = 200

The size of the population (sizepop) = 20

The probability of crossover (pcrossover) = 0.4

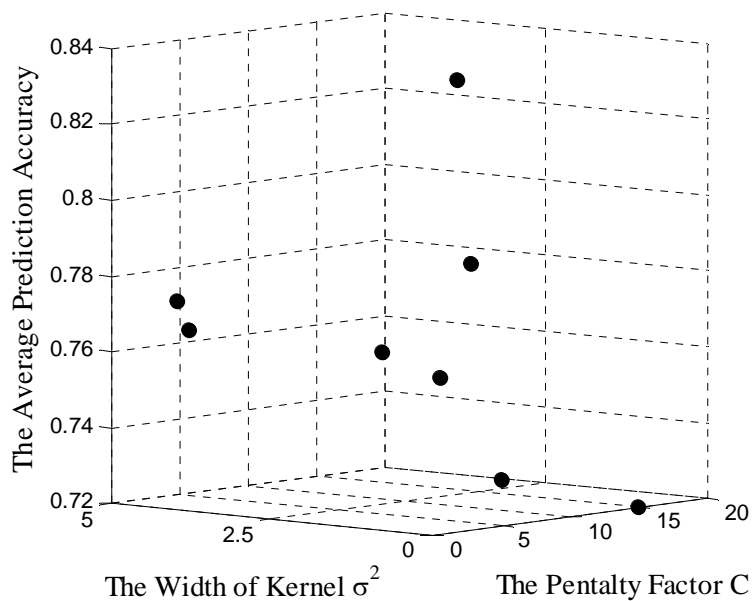


Figure 5. Effects of C and σ^2 on the average prediction accuracy of the SVM model in Semester #1.

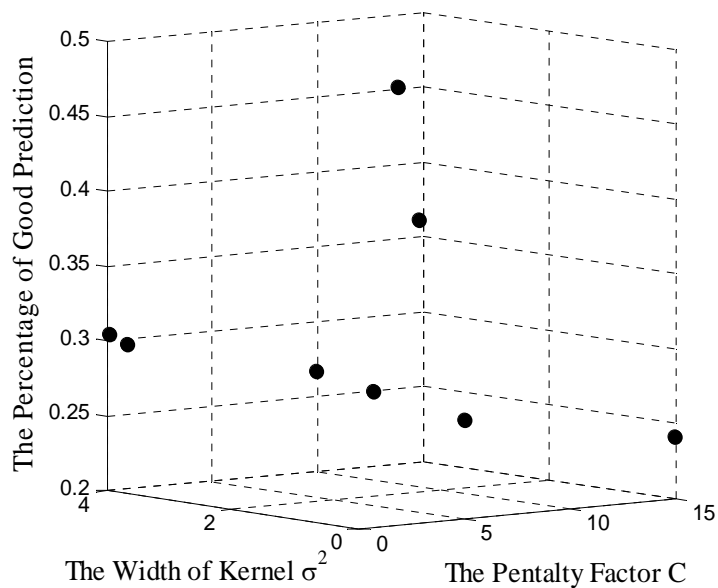


Figure 6. Effects of C and σ^2 on the percentage of accurate prediction of the SVM model in Semester #1.

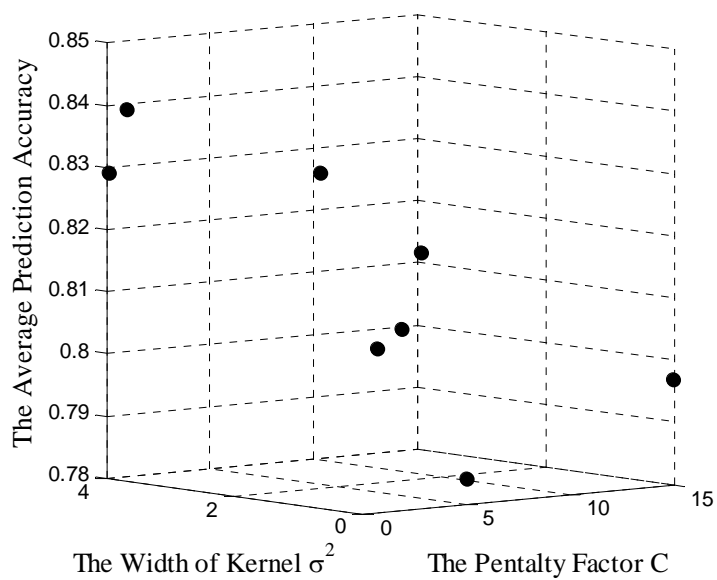


Figure 7. Effects of C and σ^2 on the average prediction accuracy of the SVM model in Semester #2.

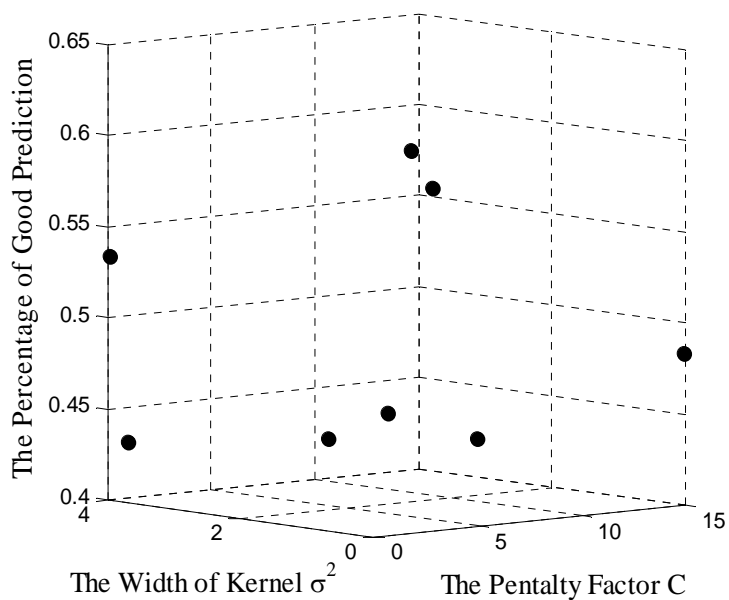


Figure 8. Effects of C and σ^2 on the percentage of accurate prediction of the SVM model in Semester #2.

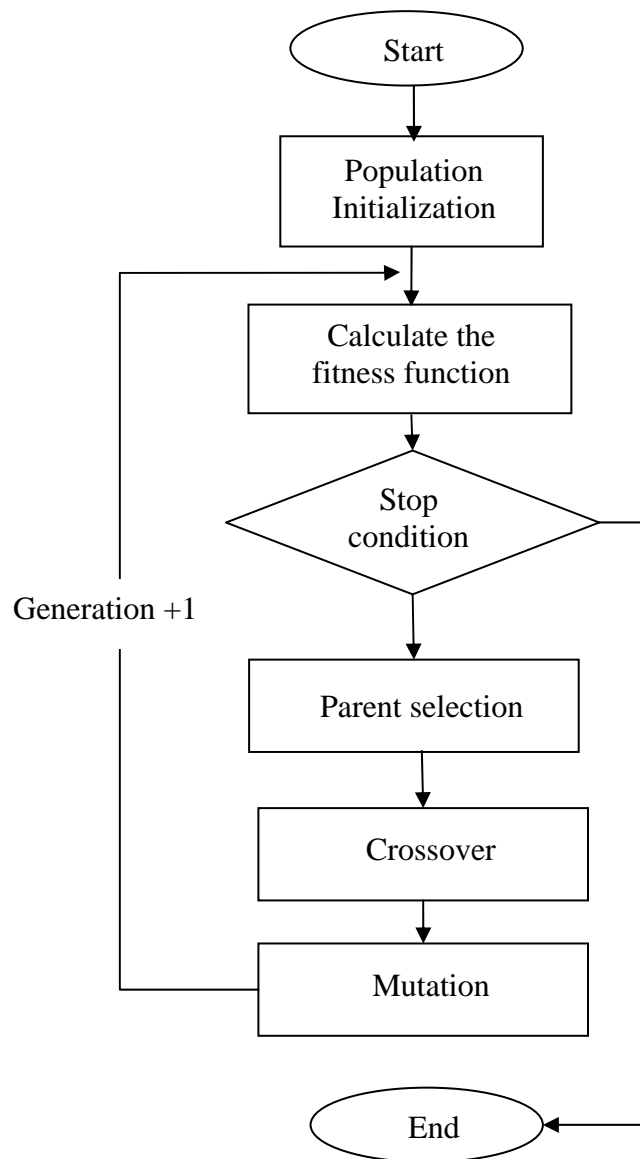


Figure 9. Flow chart of genetic algorithms.

The probability of mutation (pmutation) = 0.01

The range of penalty factor C (cbound) = [0.01, 400]

The range of width of the kernel σ^2 (gbound) = [0.001, 1000].

The overall framework of genetic algorithm and SVM is demonstrated in Figure 10.

The SVM package LibSVM (Chang & Lin, 2001) was the method of preference for regression calculation in this study. LibSVM enables users to easily apply SVM as a tool (Chang & Lin, 2001). The Matlab main code is as follows:

```
% Load the training data (data collected in Semester #1) and the external
validation data (data collected in Semesters #2, #3, and #4)
```

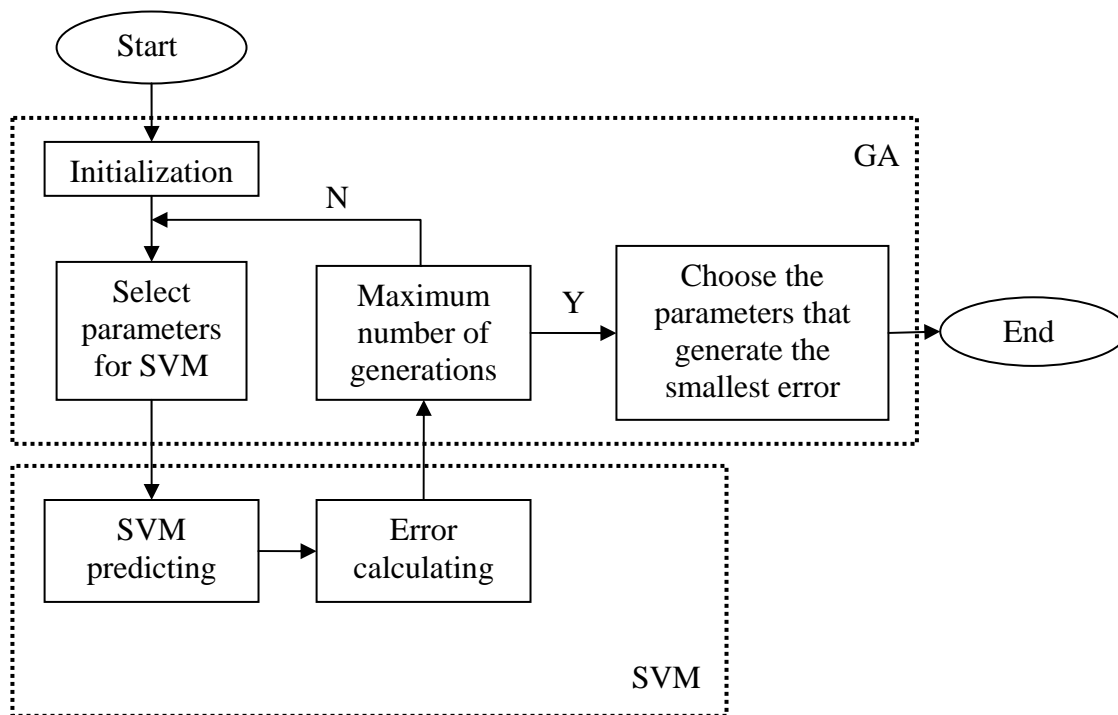


Figure 10. Overall framework of genetic algorithm (GA) and SVM.

```

load train_in;

load train_out;

load vali_in_Sem2;

load vali_out_Sem2;

load vali_in_Sem3;

load vali_out_Sem3;

load vali_in_Sem4;

load vali_out_Sem4;

% Search for the best parameters by using a genetic algorithm

[bestCVmse,bestC,bestG] = GA_SVM (train_in,train_out);

% Train the predictive model with the best parameters, where '-c' sets the
penalty factor C of  $\epsilon$ -loss function; '-g' sets the width of the kernel; and '-s 3' sets the
loss function for regression as  $\epsilon$ -loss function.

cmd = ['-c ',num2str(bestC),' -g ',num2str(bestG),' -s 3 '];

model = svmtrain (train_out,train_in,cmd);

% Apply the developed model to the data collected from Semesters #2, #3, and #4.

[ptrain1,train_mse2] = svmpredict(vali_out_Sem2, vali_in_Sem2,model);

[ptrain2,train_mse3] = svmpredict(vali_out_Sem3, vali_in_Sem3,model);

[ptrain3,train_mse4] = svmpredict(vali_out_Sem4, vali_in_Sem4,model);

```

Comparison of the Predictive Models

The predictive models developed by using the training dataset were applied to the full datasets collected during Semesters #2, #3, and #4. Because each semester presented a new set of students, the datasets collected in Semesters #2, #3, and #4 can be used to assess external validity of the developed models and examine the generalizability of the developed models.

Moreover, to investigate which combination, among the 24 combinations of candidate predictors and mathematical techniques, yields the most accurate prediction, the predicted results using the data collected in Semesters #2, #3, and #4 were compared. Two criteria were adopted: the average percentage of predictive accuracy and the percentage of accurate prediction. Prediction accuracy measures the degree of proximity of the predicted results to actual values. The percentage of accurate prediction represents the percentage of cases whose predicted values are within 90-110% of the actual values (namely, the prediction error is $\pm 10\%$).

CHAPTER IV

RESULTS AND ANALYSIS

Presented in this chapter are the results of the preprocessing of data, the selection of sample size, the effects of relevant parameters of the predictive models, and internal and external validations of those predictive models.

Descriptive Analysis of the Normalized Data

Table 6 shows the results of descriptive analysis of the normalized data collected during the four semesters. As seen in Table 6, most variables of X_1 - X_8 and Y in Semesters #2 and #3 had lower means and higher standard deviations, and some variables in Semester #4 had higher means and lower standard deviations. For example, compared to students in Semester #1 as a whole, students in Semesters #2 and #3 had lower cumulative GPAs, lower statics scores, lower dynamics mid-exam #3 scores, and higher standard deviations in GPAs, statics, and dynamics mid-exam #3 scores. Meanwhile, students in Semester #4 had higher cumulative GPAs, higher statics scores, higher physics scores, and lower standard deviations in GPAs, statics, and physics scores.

The above research findings imply that students in Semesters #2 and #3 did not perform as well as students in Semester #1, and that students in Semesters #2, #3, and #4 were more diverse in their academic performance. Figures 11-14 further show the histograms of students' normalized final exam scores in the dynamics course throughout the four semesters. The distribution of the final exam scores comes closest to a normal distribution during Semesters #2 and #4, and to a bimodal distribution in Semester #3.

Table 6

Descriptive Analysis of the Normalized Data

Variable	Semester #1 Students ($n = 128$)		Semester #2 Students ($n = 58$)		Semester #3 Students ($n = 53$)		Semester #4 Students ($n = 84$)	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
X_1 Cumulative GPA	0.8586	0.09569	0.8110	0.11207	0.8379	0.10613	0.8593	0.08774
X_2 Engineering Statics	0.8076	0.18898	0.6725	0.20628	0.7738	0.24276	0.8411	0.17349
X_3 Calculus I	0.7580	0.18555	0.7642	0.19330	0.7223	0.19369	0.7304	0.18296
X_4 Calculus II	0.7813	0.18336	0.7284	0.20030	0.7145	0.20884	0.7727	0.19129
X_5 Physics	0.7925	0.15960	0.7356	0.18682	0.7479	0.16748	0.8618	0.11892
X_6 Mid-exam #1	0.7870	0.15764	0.7109	0.18474	0.7255	0.15164	0.7808	0.12842
X_7 Mid-exam #2	0.7778	0.13716	0.7813	0.14446	0.7276	0.15226	0.7176	0.14317
X_8 Mid-exam #3	0.8477	0.12407	0.8080	0.14989	0.7709	0.15200	0.8257	0.13430
Y Final exam	0.7175	0.16683	0.6916	0.15754	0.6647	0.17726	0.7077	0.16474

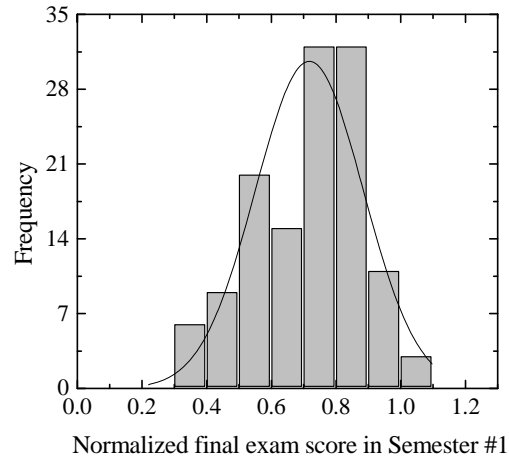


Figure 11. Histogram of students' normalized dynamics final exam scores in Semester #1 ($n = 128$).

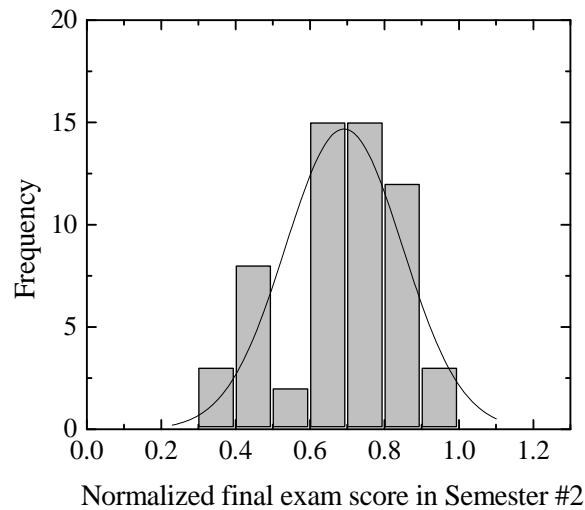


Figure 12. Histogram of students' normalized dynamics final exam scores in Semester #2 ($n = 58$).

Thus, Semesters #2, #3, and #4 provided excellent “external” cases to validate the generalization ability of the predictive models developed from the data collected in Semester #1. Figure 15(a-h) shows the scatter plots of the final exam scores against each

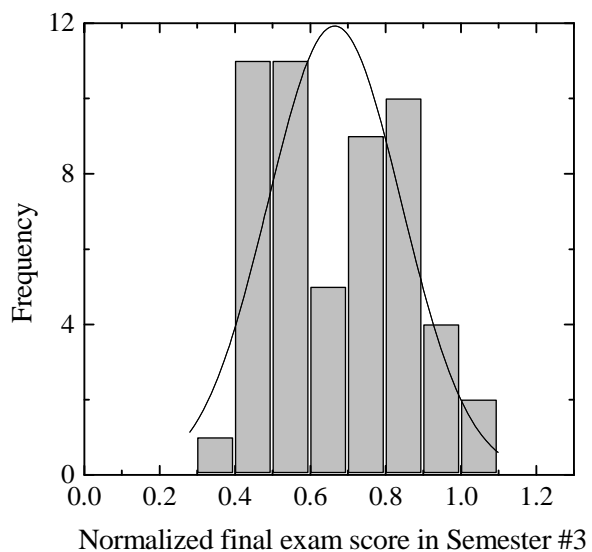


Figure 13. Histogram of students' normalized dynamics final exam scores in Semester #3 ($n = 53$).

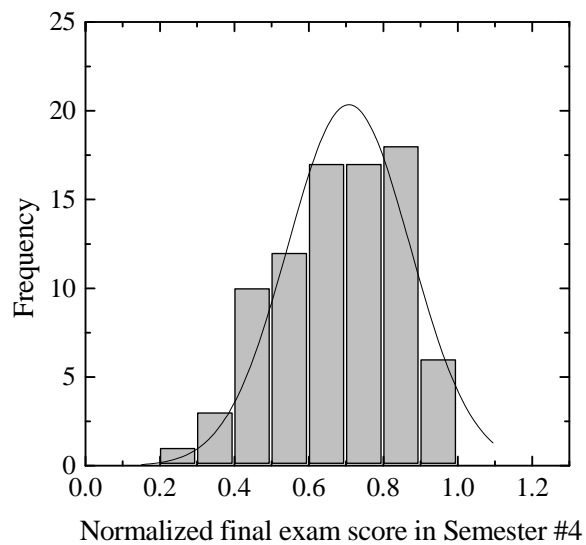


Figure 14. Histogram of students' normalized dynamics final exam scores in Semester #4 ($n = 84$).

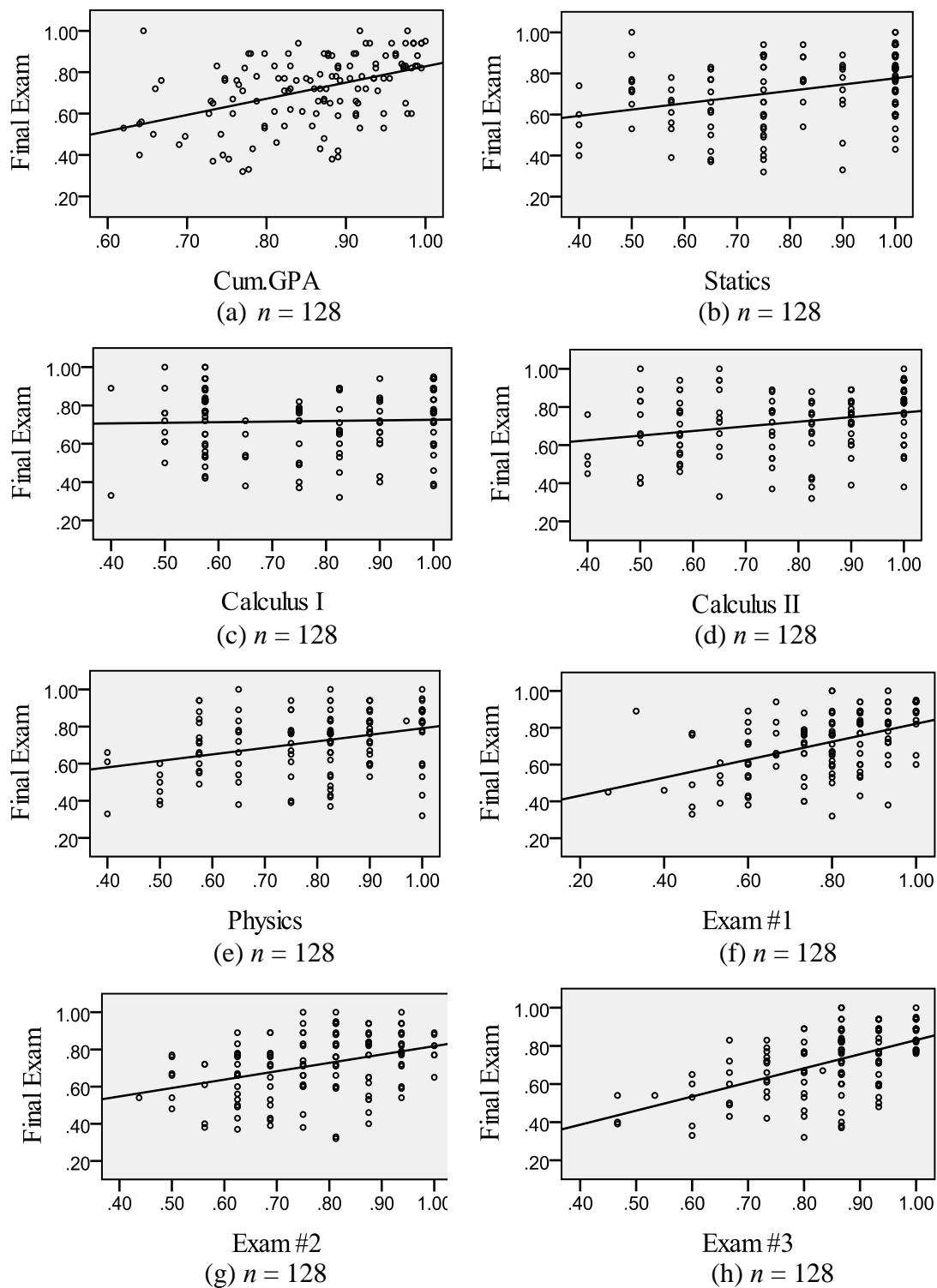


Figure 15. Scatter plots of the dependent variable Y against the predictor variables X_i .

predictor variable in Semester #1 ($n = 128$). Nearly all predictor variables (except X_4 calculus I) had a linear relationship with the dynamics final exam score. Figure 15(c) shows that calculus I (X_4) had nearly no effect on the dynamics final exam score.

Identification of Outliers in the Collected Data

Leverage, discrepancy, DFFIT, and DFBETAS were employed to test if there were outliers in the collected data. Leverage assesses how unusual case i is on the independent variables. Discrepancy measures the difference between the predicted and the actual value. DFFIT assesses the overall impact of case i on the regression results. DFBETAS assess the influence of case i on regression coefficients.

The cutoff value for leverage was $3k/n = 0.19$ (Cohen et al., 2003), where k is the number of predictor variables ($k = 8$), and n is the total number of cases used to develop the models ($n = 128$). The cutoff value was ± 3.5 for discrepancy and ± 1 for influence (Cohen et al., 2003). Figures 16-18 show that no case exceeds the cutoff value.

The influence on a specific regression coefficient was also tested using DFBETAS. No outlier was identified by DFBETAS. Therefore, it can be concluded that there was no outlier in the data collected in Semester #1. This implies that all data collected in this semester can be used to develop predictive models.

Testing of Multiple Collinearity

Table 7 illustrates collinearity analysis, which is used in cases where all eight predictors ($X_1 \sim X_8$) are included in the regression model. All tolerances are higher than 0.2,

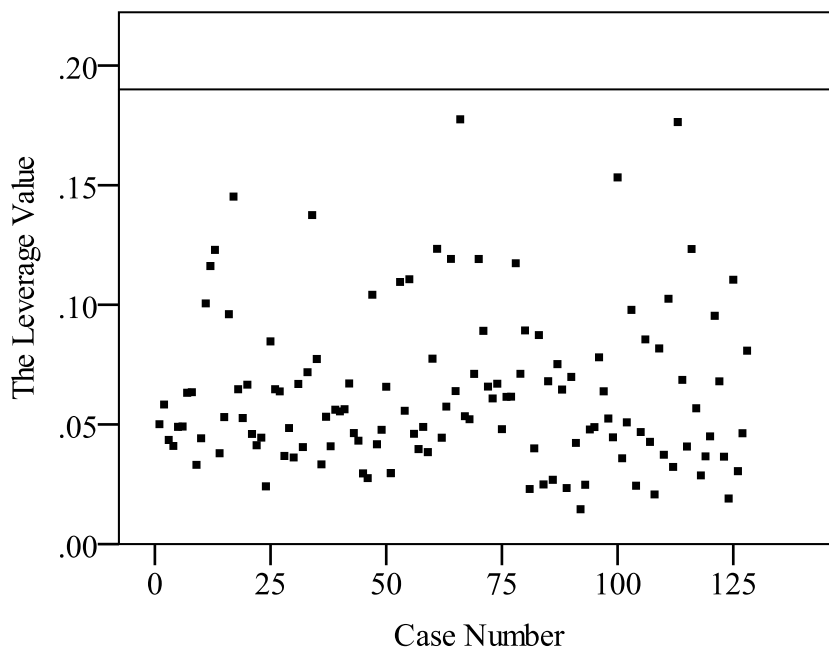


Figure 16. Assessing the leverage of the data collected in Semester #1 ($n = 128$).

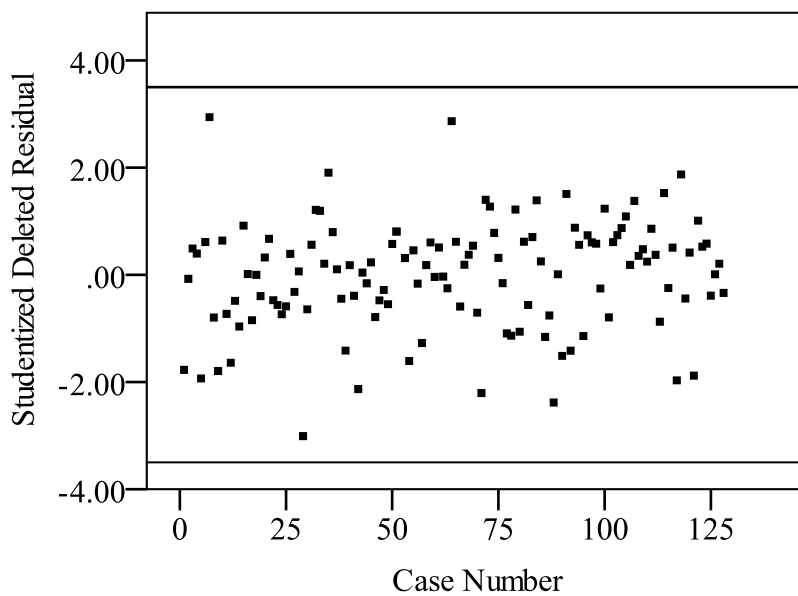


Figure 17. Assessing the discrepancy of the data collected in Semester #1 ($n = 128$).

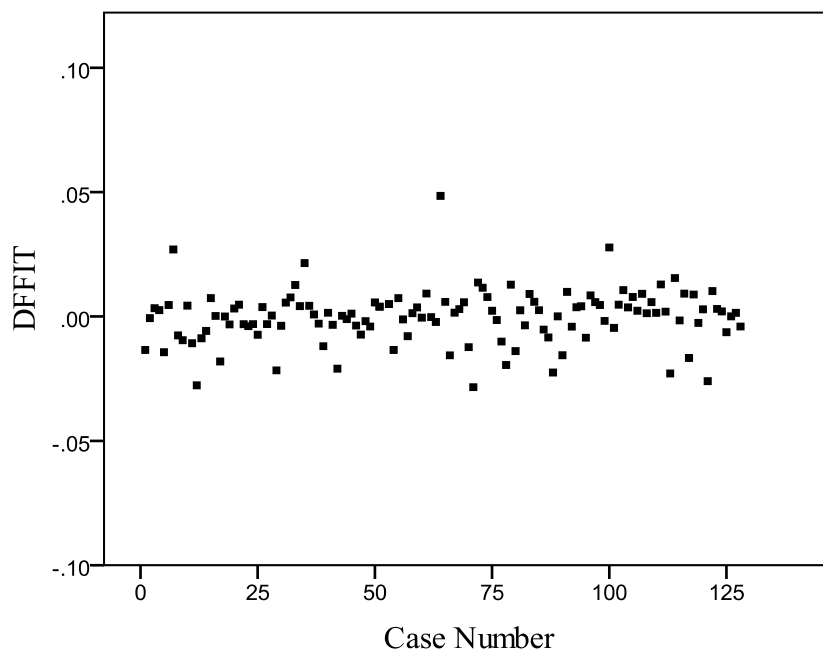


Figure 18. Assessing DFFIT of the data collected in Semester #1 ($n = 128$).

Table 7

Collinearity Analysis of the Data Collected in Semester #1

Model	Collinearity analysis	
	Tolerance	Variance inflation factor
(Constant)		
X_1 Cumulative GPA	0.331	3.025
X_2 Statics	0.480	2.082
X_3 Calculus I	0.900	1.111
X_4 Calculus II	0.531	1.882
X_5 Physics	0.781	1.280
X_6 Dynamics mid-term exam #1	0.674	1.484
X_7 Dynamics mid-term exam #2	0.656	1.523
X_8 Dynamics mid-term exam #3	0.739	1.353

and the variance inflation factors are less than five. The results indicate that collinearity is not an issue that needs to be considered in predictive modeling in this study.

Correlation Analysis

As seen from Tables 8 to 11, a statistically significant corelationship ($p < 0.01$ or $p < 0.05$) exists between the dynamics final exam score and each of the eight predictor variables for all four semesters with only one exception: the corelationship between the dynamics final exam score and the Calculus I grade. This latter corelationship is not statistically significant in Semesters #1 and #4 ($p > 0.05$) but is statistically significant in Semesters #2 ($r = 0.270, p < 0.05$) and #3 ($r = 0.301, p < 0.05$). This result is consistent with the research findings shown in Figure 15(c) that in Semester #1, the effect of Calculus I on the dynamics final exam score was small. However, to generate a general predictive model to cover as many cases as possible, it was decided to include the Calculus I grade as a predictor variable in the predictive models.

Determining the Appropriate Sample Size

Soper's (2004) statistical calculator was used to determine the minimum sample size in this study. The effect size, power, number of predictors, and probability level were also factors in the determination. The effect size was anticipated by the squared multiple correlation and the power level was set as 0.8, as discussed in Chapter III. Figure 19 shows that the minimum sample size was 46 for the development of predictive modeling in this study.

Table 8

Correlation Coefficients in Semester #1

Variable	X_1 Cumulative GPA	X_2 Engineering Statics	X_3 Calculus I	X_4 Calculus II	X_5 Physics	X_6 Dynamics mid-exam #1	X_7 Dynamics mid-exam #2	X_8 Dynamics mid-exam #3	Y Dynamics final exam
X_1	1.000								
X_2	0.695**	1.000							
X_3	0.108	1.000	1.000						
X_4	0.194*	0.668**	0.477**	1.000					
X_5	0.416**	0.360**	0.190*	0.375**	1.000				
X_6	0.475**	0.446**	0.023	0.365**	0.246**	1.000			
X_7	0.468**	0.418**	0.020	0.266**	0.234**	0.437**	1.000		
X_8	0.298**	0.347**	-0.123	0.186*	0.207*	0.358**	0.421**	1.000	
Y	0.448**	0.346**	0.035	0.267**	0.335**	0.461**	0.370**	0.550**	1.000

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

Table 9

Correlation Coefficients in Semester #2

Variable	X_1 Cumulative GPA	X_2 Engineering Statics	X_3 Calculus I	X_4 Calculus II	X_5 Physics	X_6 Dynamics mid-exam #1	X_7 Dynamics mid-exam #2	X_8 Dynamics mid-exam #3	Y Dynamics final exam
X_1	1.000								
X_2	0.628**	1.000							
X_3	0.329*	0.284*	1.000						
X_4	0.619**	0.424**	0.411**	1.000					
X_5	0.512**	0.326*	0.371**	0.448**	1.000				
X_6	0.569**	0.502**	0.233	0.416**	0.358**	1.000			
X_7	0.578**	0.555**	0.379**	0.396**	0.297*	0.421**	1.000		
X_8	0.598**	0.603**	0.288*	0.568**	0.425**	0.530**	0.430**	1.000	
Y	0.636**	0.730**	0.270*	0.408**	0.377**	0.582**	0.672**	0.666**	1.000

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

Table 10

Correlation Coefficients in Semester #3

Variable	X_1 Cumulative GPA	X_2 Engineering Statics	X_3 Calculus I	X_4 Calculus II	X_5 Physics	X_6 Dynamics mid-exam #1	X_7 Dynamics mid-exam #2	X_8 Dynamics mid-exam #3	Y Dynamics final exam
X_1	1.000	0.660**	0.401**	0.483**	0.435**	0.408**	0.596**	0.607**	0.569**
X_2		1.000	0.298*	0.362**	0.391**	0.455**	0.434**	0.498**	0.466**
X_3			1.000	0.354**	0.466**	0.330*	0.329*	0.372**	0.301*
X_4				1.000	0.437**	0.321*	0.128	0.136	0.473**
X_5					1.000	0.337*	0.233	0.220	0.375**
X_6						1.000	0.554**	0.546**	0.493**
X_7							1.000	0.727**	0.328*
X_8								1.000	0.458**
Y									1.000

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

Table 11

Correlation Coefficients in Semester #4

Variable	X_1 Cumulative GPA	X_2 Engineering Statics	X_3 Calculus I	X_4 Calculus II	X_5 Physics	X_6 Dynamics mid- exam #1	X_7 Dynamics mid- exam #2	X_8 Dynamics mid- exam #3	Y Dynamics final exam
X_1	1.000	0.580**	0.179	0.457**	0.630**	0.423**	0.456**	0.492**	0.479**
X_2		1.000	0.081	0.299**	0.487**	0.367**	0.430**	0.234**	0.343**
X_3			1.000	0.214	0.089	0.062	0.017	0.106	-0.020
X_4				1.000	0.373**	0.305**	0.178	0.328**	0.408**
X_5					1.000	0.342**	0.555**	0.251*	0.390**
X_6						1.000	0.417	0.418**	0.569**
X_7							1.000	0.309**	0.466**
X_8								1.000	0.527**
Y									1.000

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

Anticipated effect size (f^2):

Desired statistical power level:

Number of predictors:

Probability level:

Minimum required sample size: 46

Figure 19. The minimum sample size determined by Soper's (2004) statistics calculator.

In theory, the larger the sample size, the better the prediction accuracy of a model. Five regression models with different sample sizes were developed using a variety of combinations of predictors $X_1 \sim X_8$. The mathematical formula of each regression model (I-V) is expressed below.

Model I:

$$Y_1 = -0.429 + 0.567X_1 - 0.233X_2 - 0.040X_3 + 0.050X_4 + 0.281X_5 \\ + 0.258X_6 + 0.122X_7 + 0.334X_8$$

Model II:

$$Y_2 = -0.380 + 0.520X_1 - 0.006X_2 + 0.213X_3 + 0.051X_4 + 0.079X_5 \\ + 0.084X_6 - 0.055X_7 + 0.585X_8$$

Model III:

$$Y_3 = -0.309 + 0.556X_1 - 0.194X_2 + 0.002X_3 - 0.028X_4 + 0.102X_5 \\ + 0.251X_6 - 0.070X_7 + 0.591X_8$$

Model IV:

$$Y_4 = -0.334 + 0.500X_1 - 0.201X_2 - 0.021X_3 - 0.057X_4 + 0.154X_5 \\ + 0.281X_6 + 0.053X_7 + 0.540X_8$$

Model V:

$$Y_5 = -0.369 + 0.515X_1 - 0.097X_2 + 0.024X_3 - 0.085X_4 + 0.149X_5 \\ + 0.233X_6 - 0.001X_7 + 0.556X_8$$

To confirm that the minimum sample size is 46, a sample size of 39 (in Model I) was also studied. Data from Semesters #1 and #2 were used for internal and external validations, respectively. Table 12 shows the results from these internal and external validations.

As illustrated in Table 12, in general, the prediction accuracy of the developed regression models was found to reduce in external validation by up to 1.1% (for Model I). However, the percentage of accurate prediction was reduced by up to 12.7% (for Model II). Based on the results of both internal and external validations, it can be concluded that the developed regression models have excellent predictability with 87%-91% of the average prediction accuracy, but they have only moderate ability (46%-66%) to generate accurate predictions (again, an accurate prediction is defined as the prediction with $\pm 10\%$ of prediction error).

The percentage of accurate prediction for Model II (a sample size larger than 46) was higher than that for Model I (a sample size smaller than 46) in both internal and external validations. However, when sample size increases from 30% to 40%, the average prediction accuracy decreases only slightly.

Three larger sample sizes for training the model were tested, including 64 (50% of the data collected in Semester #1), 77 (60% of the data collected in Semester #1), and 128

Table 12

Comparison of Different Sample Sizes

Regression model	Sample size (training dataset / full dataset)	Average prediction accuracy (%)		Percentage of accurate predictions (%)	
		Internal validation (Semester #1)	External validation (Semester #2)	Internal validation (Semester #1)	External validation (Semester #2)
I	39 (30%)	89.2	88.1	51.7	46.6
II	51 (40%)	87.7	87.3	61.0	48.3
III	64 (50%)	90.7	89.8	65.6	56.9
IV	77 (60%)	89.4	90.1	60.8	56.9
V	128(100%)	90.3	90.5	65.6	56.9

(100% of the data collected from Semester #1). In term of the percentage of accurate prediction, Models III, IV, and V outperformed Models I and II. In general, the average prediction accuracy increased with a larger sample size. Therefore, in this study, the full dataset of Semester #1 ($n = 128$) was employed as the training dataset for developing all types of predictive models (MLR, MLP, RBF, and SVM).

Six MLR models were developed using the full dataset ($n = 128$) collected from Semester #1 and six combinations of predictors. These MLR models have explicit mathematical equations as follows:

MLR Model 1:

$$Y = 0.047 + 0.781X_1$$

MLR Model 2:

$$Y = 0.022 + 0.715X_1 + 0.034X_2 - 0.063X_3 - 0.077X_4 + 0.204X_5$$

MLR Model 3:

$$Y = 0.334 + 0.487X_6$$

MLR Model 4:

$$Y = -0.053 + 0.567X_1 - 0.025X_2 - 0.041X_3 - 0.101X_4 + 0.191X_5 + 0.334X_6$$

MLR Model 5:

$$Y = -0.079 + 0.502X_1 - 0.036X_2 - 0.036X_3 - 0.090X_4 + 0.186X_5 + 0.303X_6 + 0.138X_7$$

MLR Model 6:

$$Y = -0.369 + 0.515X_1 - 0.097X_2 + 0.024X_3 - 0.085X_4 + 0.149X_5 + 0.233X_6 - 0.001X_7 + 0.556X_8$$

However, there were no simple mathematical equations for other types (MLP, RBF, and SVM) of predictive models. MLP and RBF networks have multiple layers and neurons. For example, Figure 20 shows a simple architecture for a MLP network that has

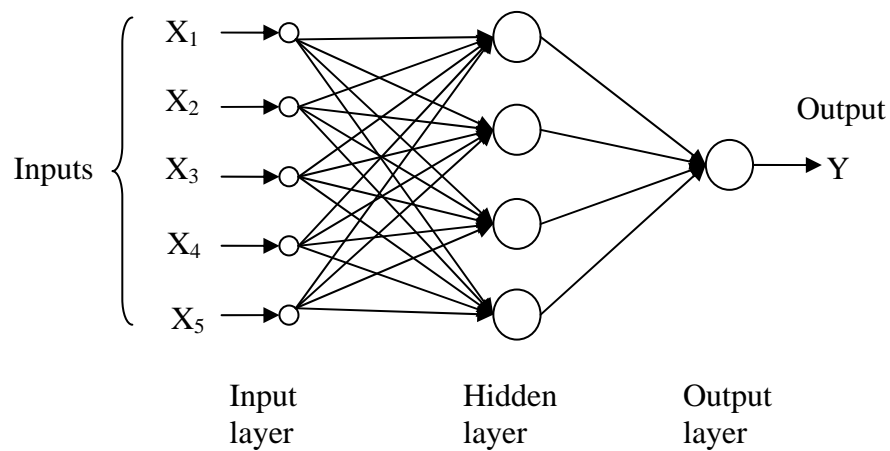


Figure 20. A sample structure of a MLP network.

five inputs, one hidden layer, four neurons in the hidden layer, and one output. Equation 4 shows the mathematical output of the j^{th} neuron in the hidden layer h_j and the output Y :

$$\begin{aligned} h_j &= f\left(\sum_{i=1}^5 w_{ij} X_i\right) \\ Y &= f\left(\sum_{j=1}^4 w_j h_j\right) \end{aligned} \quad (4)$$

where h_j is the output of the j^{th} neuron in the hidden layer, w_{ij} is the weight between the i^{th} input and the j^{th} neuron, and $f(\bullet)$ is the activation function.

Relevant parameters for MLP, RBF, and SVM models were adjusted to ensure these types have the highest possible prediction accuracy. For example, the penalty factor

C was 2.23 and the width of kernel σ^2 was 0.06 for the SVM model using $X_1 \sim X_8$ as predictors. The penalty factor C was 0.28 and the width of kernel σ^2 was 0.63 for the SVM model using $X_1 \sim X_5$ as predictors.

Internal and External Validations

Results of Internal Validation (Using Data from Semester #1)

Tables 13-16 compare different combinations of predictors and show that the models using $X_1 \sim X_8$ as predictors generate the best prediction, except for one case in which the percentage of accurate predictions from model #24 is 3.1% lower than that of model #23.

Comparison of different modeling techniques: The average APA is 88.1% for MLR models (Table 13), 89.0% for MLP models (Table 15), 88.4% for RBF models

(Table 16), and 84.0% for SVM models during Semester #1 (Table 17). The average PAP is 57.3% for all MLR models (Table 13), 55.9% for all MLP models (Table 15), 51.7% for all RBF models (Table 16), and 60.6% for all SVM models (Table 17) in Semester #1.

In terms of the average prediction accuracy (APA) and the percentage of accurate prediction (PAP), all types of models yield accurate predictive results. SVM models have relatively low APA, but relatively high PAP. Among the four types of models, RBF models yield the lowest average PAP.

Results of External Validation (Using Data from Semesters #2, #3, and #4)

MLR. Comparison of different combinations of predictors: Table 13 shows that the average APA varies from 88.4% to 89.5%, and the average PAP varies from 47.1% to 59.9% among the six different combinations of predictors. In terms of APA and PAP, the top three best-performing MLR models are #6, #5, and #4 and the worst-performing is model #3.

Comparison of model performance in different semesters: Table 13 shows that on average, the MLR models generate the lowest APA (87.9%) and PAP (50.8%) in Semester #3. The APA and PAP for external validation are 1.7% and 6.5%, respectively, lower than those for internal validation. In Semester #2, the MLR models generate the highest APA (90.2%) and PAP (59.2%).

Table 14 further shows the *R*-square and standardized coefficients β of each model. It is shown that the MLR models explain 20.1% - 44.7% of student academic performance in the engineering dynamics course. If all eight predictor variables are

Table 13

Internal and External Validations of MLR Models

MLR model no.	Predictor	Semester #1 students ($n = 128$) (internal validation)		Semester #2 students ($n = 58$) (external validation)		Semester #3 students ($n = 53$) (external validation)		Semester #4 students ($n = 84$) (external validation)	
		Average prediction accuracy (%)	Percentage of accurate predictions (%)	Average prediction accuracy (%)	Percentage of accurate predictions (%)	Average prediction accuracy (%)	Percentage of accurate predictions (%)	Average prediction accuracy (%)	Percentage of accurate predictions (%)
1	X_1	88.3	50.0	90.2	56.9	87.7	54.7	88.1	53.6
2	$X_1 \sim X_5$	88.6	54.7	90.1	62.1	87.8	49.1	88.1	53.6
3	X_6	88.2	54.7	89.9	53.4	86.7	40.4	88.5	47.6
4	$X_1 \sim X_6$	89.5	58.6	89.9	62.1	88.4	50.9	89.0	53.6
5	$X_1 \sim X_7$	89.5	60.2	90.7	63.8	88.4	50.9	89.3	52.4
6	$X_1 \sim X_8$	90.3	65.6	90.5	56.9	88.2	58.5	89.8	64.3

Table 14

Standardized Coefficients of MLR Models

MLR model no.	Predictor	R^2	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8
1	X_1	0.201	0.448 ^{**}	-	-	-	-	-	-	-
2	$X_1 \sim X_5$	0.238	0.410 ^{**}	0.039	-0.07	-0.085	0.196 [*]	-	-	-
3	X_6	0.212	-	-	-	-	-	0.461 ^{**}	-	-
4	$X_1 \sim X_6$	0.311	0.325 [*]	-0.028	-0.045	-0.111	0.183 [*]	0.315 ^{**}	-	-
5	$X_1 \sim X_7$	0.320	0.288 [*]	-0.041	-0.040	-0.099	0.178 [*]	0.286 ^{**}	0.113 [*]	-
6	$X_1 \sim X_8$	0.447	0.295 [*]	-0.110	0.027	-0.093	0.142 [*]	0.220 ^{**}	-0.001	0.413 ^{**}

^{**} Correlation is significant at the 0.01 level (2-tailed).

^{*} Correlation is significant at the 0.05 level (2-tailed).

Table 15

Internal and External Validations of MLP Network Models

MLP model no.	Predictor	Semester #1 students ($n = 128$) (internal validation)		Semester #2 students ($n = 58$) (external validation)		Semester #3 students ($n = 53$) (external validation)		Semester #4 students ($n = 84$) (external validation)	
		Average prediction accuracy (%)	Percentage of accurate prediction (%)	Average prediction accuracy (%)	Percentage of accurate prediction (%)	Average prediction accuracy (%)	Percentage of accurate prediction (%)	Average prediction accuracy (%)	Percentage of accurate prediction (%)
7	X_1	88.1	48.4	90.1	58.6	87.7	52.8	88.1	53.6
8	$X_1 \sim X_5$	88.6	54.7	90.1	60.3	87.6	45.3	88.0	51.2
9	X_6	88.2	51.6	90.0	56.9	86.8	39.6	87.8	47.6
10	$X_1 \sim X_6$	89.4	57.0	90.1	60.3	88.4	49.1	89.2	54.8
11	$X_1 \sim X_7$	89.5	57.0	90.5	60.3	88.4	50.9	88.7	53.6
12	$X_1 \sim X_8$	90.3	66.4	90.6	56.9	88.0	52.8	89.6	61.9

Table 16

Internal and External Validations of RBF Network Models

RBF model no.	Predictor	Semester #1 students ($n = 128$) (internal validation)		Semester #2 students ($n = 58$) (external validation)		Semester #3 students ($n = 53$) (external validation)		Semester #4 students ($n = 84$) (external validation)	
		Average prediction accuracy (%)	Percentage of accurate prediction (%)	Average prediction accuracy (%)	Percentage of accurate prediction (%)	Average prediction accuracy (%)	Percentage of accurate prediction (%)	Average prediction accuracy (%)	Percentage of accurate prediction (%)
13	X_1	87.8	50.0	89.8	62.1	87.8	50.9	87.7	52.4
14	$X_1 \sim X_5$	88.1	50.8	90.0	62.1	87.4	49.1	86.7	51.2
15	X_6	88.0	51.6	90.0	58.6	87.0	43.4	88.3	52.4
16	$X_1 \sim X_6$	88.4	50.0	90.6	63.8	88.3	52.8	89.0	57.1
17	$X_1 \sim X_7$	88.7	50.8	90.9	69.0	88.4	52.8	88.7	51.2
18	$X_1 \sim X_8$	89.1	57.0	91.0	72.4	89.0	56.6	90.4	65.5

Table 17

Internal and External Validations of SVM Models

SVM model no.	Predictor	Semester #1 students ($n = 128$) (internal validation)		Semester #2 students ($n = 58$) (external validation)		Semester #3 students ($n = 53$) (external validation)		Semester #4 students ($n = 84$) (external validation)	
		Average prediction accuracy (%)	Percentage of accurate prediction (%)	Average prediction accuracy (%)	Percentage of accurate prediction (%)	Average prediction accuracy (%)	Percentage of accurate prediction (%)	Average prediction accuracy (%)	Percentage of accurate prediction (%)
19	X_1	80.5	51.6	90.4	62.1	87.3	50.9	87.9	50.0
20	$X_1 \sim X_5$	80.4	56.3	90.6	65.5	87.3	47.2	87.9	47.6
21	X_6	85.7	59.4	88.8	56.9	87.2	42.5	88.2	51.2
22	$X_1 \sim X_6$	82.0	64.8	90.3	63.8	88.5	52.8	89.1	54.8
23	$X_1 \sim X_7$	85.1	67.2	91.1	65.5	88.9	52.8	89.5	59.5
24	$X_1 \sim X_8$	90.2	64.1	90.9	69.0	89.0	62.3	90.3	60.7

included in the model, the most important predictor variables that affect prediction accuracy are: dynamics mid-term exam #3 ($\beta_8 = 0.413$), cumulative GPA ($\beta_1 = 0.295$), dynamics mid-term exam #1 ($\beta_6 = 0.220$), and physics ($\beta_5 = 0.142$).

MLP network. Comparison of different combinations of predictors: When the predictors change in MLP models, the average APA varies from 88.2% to 89.4%, and the average PAP varies from 48.0% to 57.2% as illustrated in Table 15. In terms of APA and PAP, the top three best-performing MLP models are #12, #11, and #10 and the worst-performing is model #9.

Comparison of model performance in different semesters: Table 15 shows that on average, the MLP models generate the lowest APA (87.8%) and PAP (48.4%) in Semester #3. The APA and PAP determined from external validation are 1.2% and 7.5%, respectively, which is lower than those determined from internal validation. For Semester #2, the MLP models generate the highest APA (90.2%) and PAP (58.9%), which are higher than those for internal validation (APA = 89.0% and PAP = 55.9%).

RBF. Comparison of different combinations of predictors: Table 16 shows that the average APA is no lower than 88.0% and the average PAP is no lower than 51.5% in RBF models with different combinations of predictors. In terms of APA and PAP, the top three best-performing RBF models are #18, #16, and #17 and the worst-performing is model #15.

Table 16, which compares model performance during different semesters, shows that the RBF models also have low prediction accuracy in Semester #3 when the average of APA is 88.0% and the average of PAP is 50.9%. The APA and PAP from external

validation are almost the same as those from internal validation. In Semester #2, the RBF models generate the highest APA (90.4%) and PAP (64.7%), which are higher than those from internal validation (APA = 88.4% and PAP = 51.7%).

SVM. Table 17, which compares different combinations of predictors, shows the average APA varies from 88.1% to 90.1%, while the average PAP varies from 50.2% to 64.0% among the six different combinations of predictors. In terms of APA and PAP, the top three best-performing SVM models are models #24, #23, and #22 and the worst-performing model is model #21.

Table 17, which compares model performances in different semesters, shows that on average, the PAP for external validation in Semesters #2 and #3 are 9.2% and 6.6%, respectively, which is lower than those from internal validation. In Semester #2, the SVM models generate the highest APA (90.4%) and PAP (63.8%), which are higher than those from internal validation (APA = 84.0% and PAP = 60.6%).

Comparison of Different Modeling Techniques

From Tables 13-16, the following observations are made:

1. In internal validation, SVM models have relatively low APA, but relatively high PAP.
2. RBF models yield the lowest average PAP among the four types of models in internal validation.
3. Although MLP models generate good APA in external validation, RBF and SVM models outperform MLP models in terms of PAP. RBF and SVM

models have the nearly the same level of performance in terms of APA and PAP. The MLP models have the lowest performance among the four types of models based on the data collected in this study.

Table 18 shows an example of prediction with different modeling techniques and different combinations of predictors.

Identifying Academically At-Risk Students

One of the purposes of this study is to identify academically at-risk students. Tables 18-21 show the percentage of academically at-risk students that have been correctly identified by the four types of predictive models. A cell in the table is called a “good cell” if the value in it is larger than 50, which means that more than 50% of academically at-risk students are correctly identified by the model. In Tables 18- 21, there are a total of 19 “good cells” which are highlighted in bold.

Comparison of different combinations of predictors: The models with $X_1 \sim X_8$ as predictors yield nine good cells. The models with $X_1 \sim X_7$ and $X_1 \sim X_6$ as predictors have four good cells. The average percentage of academically at-risk students correctly identified in Semesters #2-#4 (external validation) is 58.8% for models using $X_1 \sim X_8$ as predictors, 41.2% for models using $X_1 \sim X_7$ as predictors, and 40.9% for models using $X_1 \sim X_6$ as predictors.

Comparison of different modeling techniques: Both RBF and SVM models generate seven good cells. However, SVM Model #19 fails to correctly identify any academically at-risk student in Semester #4. On average, RBF models correctly identify 64.1% of

academically at-risk students in Semester #2, 46.7% of those students in Semester #3, and 28.1% in Semester #4. SVM models identify 64.1% of those students in Semester #2, 44.7% in Semester #3, and 10.5% in Semester #4. Table 23 shows an example of identifying academically at-risk students.

Table 18

An Example of Prediction: The Dynamics Final Exam Score was 90 (out of 100) for a Student in Semester #4

Model type	Model no.	Predicted score	Prediction accuracy (%)	Is it an accurate prediction?
MLR	1	75	83.3	N
	2	80	88.9	N
	3	76	84.4	N
	4	80	88.9	N
	5	81	90.0	Y
	6	84	93.3	Y
MLP	7	74	82.2	N
	8	84	93.3	Y
	9	73	81.1	N
	10	83	92.2	Y
	11	79	87.8	N
	12	87	96.7	Y
RBF	13	72	80.0	N
	14	77	85.6	N
	15	76	84.4	N
	16	78	86.7	N
	17	79	87.8	N
	18	81	90.0	Y
SVM	19	77	85.6	N
	20	82	91.1	Y
	21	82	91.1	Y
	22	83	92.2	Y
	23	85	94.4	Y
	24	87	96.7	Y

Table 19

Academically At-Risk Students Correctly Identified by MLR Models

MLR model no.	Predictor variables	Semester #1 (%)	Semester #2 (%)	Semester #3 (%)	Semester #4 (%)
1	X_1	25.0	23.1	22.7	5.3
2	$X_1 \sim X_5$	32.1	38.5	31.8	5.3
3	X_6	28.6	30.8	27.3	26.3
4	$X_1 \sim X_6$	35.7	46.2	36.4	15.8
5	$X_1 \sim X_7$	28.6	46.2	36.4	21.1
6	$X_1 \sim X_8$	39.3	76.9	63.6	47.4

Table 20

Academically At-Risk Students Correctly Identified by MLP Models

MLP model no.	Predictor variables	Semester #1 (%)	Semester #2 (%)	Semester #3 (%)	Semester #4 (%)
7	X_1	7.1	23.1	18.2	10.5
8	$X_1 \sim X_5$	7.1	46.2	40.9	15.8
9	X_6	10.7	38.5	36.4	0.0
10	$X_1 \sim X_6$	7.1	53.8	45.5	10.5
11	$X_1 \sim X_7$	10.7	46.2	13.6	15.8
12	$X_1 \sim X_8$	39.3	76.9	59.1	36.8

Table 21

Academically At-Risk Students Correctly Identified by RBF Models

RBF model no.	Predictor variables	Semester #1(%)	Semester #2(%)	Semester #3(%)	Semester #4(%)
13	X_1	14.3	69.2	27.3	26.3
14	$X_1 \sim X_5$	7.1	38.5	40.9	21.1
15	X_6	14.3	38.5	31.8	36.8
16	$X_1 \sim X_6$	14.3	76.9	54.5	15.8
17	$X_1 \sim X_7$	14.3	76.9	50.0	15.8
18	$X_1 \sim X_8$	21.4	84.6	63.6	52.6

Table 22

Academically At-Risk Students Correctly Identified by SVM Models

SVM model no.	Predictor variables	Semester #1(%)	Semester #2(%)	Semester #3(%)	Semester #4(%)
19	X_1	10.7	30.8	27.3	0.0
20	$X_1 \sim X_5$	10.7	30.8	22.7	5.3
21	X_6	28.6	84.6	54.5	26.3
22	$X_1 \sim X_6$	17.9	76.9	45.5	5.3
23	$X_1 \sim X_7$	21.4	92.3	59.1	10.5
24	$X_1 \sim X_8$	14.3	69.2	59.1	15.8

Table 23

An Example of Identifying Academically At-Risk Students

Model type	Model no.	For the particular student		Average prediction accuracy of the model (%)	Is the student correctly identified as academically at-risk?
		Predicted score	Actual score		
MLR	1	54	50	92.0	Y
	2	56	50	88.0	Y
	3	59	50	82.0	Y
	4	50	50	100.0	Y
	5	48	50	96.0	Y
	6	32	50	64.0	Y
MLP	7	60	50	80.0	N
	8	56	50	88.0	Y
	9	59	50	82.0	Y
	10	52	50	96.0	Y
	11	58	50	84.0	Y
RBF	12	33	50	66.0	Y
	13	49	50	98.0	Y
	14	60	50	80.0	N
	15	58	50	84.0	Y
	16	60	50	80.0	N
	17	50	50	100.0	Y
SVM	18	57	50	86.0	Y
	19	56	50	88.0	Y
	20	64	50	72.0	N
	21	54	50	92.0	Y
	22	52	50	96.0	Y
	23	44	50	88.0	Y
	24	45	50	90.0	Y

CHAPTER V

DISCUSSIONS AND CONCLUSIONS

This chapter summarizes the major research findings and discusses the limitations of this study and possible future work.

Summary of This Research

Student low academic performance in engineering dynamics has been a long-standing problem. A valid predictive model would provide the instructor with a tool to predict how well, or how poorly, the students in the class will perform in this particular course. In this study, a validated set of statistical and data mining models have been developed to predict student academic performance in an engineering dynamics course by using six combinations of predictor variables and four statistical and data mining modeling techniques. Twenty-four predictive models have been developed. The average prediction accuracy and the percentage of accurate predictions have been employed as two criteria to evaluate and compare the prediction accuracy of the 24 models. The following paragraphs summarize the major findings from this research.

Answers to the Research Questions

Research Question #1 : How accurate will predictions be if different statistical and data mining modeling techniques such as traditional MLR, MLP networks, RBF networks, and SVM are used?

A total of 24 predictive models have been developed by using MLR, MLP, RBF,

and SVM techniques. The prediction accuracy of MLP models remains nearly unchanged in spite of the change in relevant parameters, such as the maximum training epochs. The initial value of these parameters does not significantly affect the prediction accuracy of MLP and RBF models. The prediction accuracy of SVM models is affected by changing the penalty factor C and the width of kernel σ^2 . In cases in which all above-mentioned parameters are optimized, and based on the average prediction accuracy and the percentage of accurate predictions, the order of the overall prediction accuracy of the four types of models is:

$$\text{MLP} < \text{MLR} < \begin{matrix} \text{RBF} \\ \text{SVM} \end{matrix}$$

Research Question #2 : What combination of predictor/independent variables yields the highest prediction accuracy?

According to the combinations of predictors, the 24 models are grouped into the following six sets:

1. Models using X_1 as predictors
2. Models using $X_1, X_2, X_3, X_4,$ and X_5 as predictors
3. Models using X_6 as the only predictor
4. Models using $X_1, X_2, X_3, X_4, X_5,$ and X_6 as predictors
5. Models using $X_1, X_2, X_3, X_4, X_5, X_6,$ and X_7 as predictors
6. Models using $X_1, X_2, X_3, X_4, X_5, X_6, X_7,$ and X_8 as predictors

Table 24 summarizes the prediction accuracy of the six sets of models that use different combinations of predictors. The results indicate that the best combination of

Table 24

Prediction Accuracy of Models 1 to 6

Combination of predictor variables	Prediction accuracy	
	APA (%)	PAP (%)
1) X_1	88.6	54.9
2) $X_1 \sim X_5$	88.5	53.7
3) X_6	88.3	49.2
4) $X_1 \sim X_6$	89.2	56.3
5) $X_1 \sim X_7$	89.5	56.9
6) $X_1 \sim X_8$	89.8	61.5

predictors is model 6, which includes all predictors $X_1 \sim X_8$.

Research Question #3: What is the percentage of academically at-risk students that can be correctly identified by the predictive model?

The percentage of academically at-risk students who are correctly identified by the predictive models varies from 0% to 92.3%, depending on the particular combination of predictor variables. The top three predictor combinations that correctly identify the highest percentage of those students are as follows:

$$X_1 \sim X_8 > \begin{matrix} X_1 \sim X_6 \\ X_1 \sim X_7 \end{matrix}$$

RBF and SVM models performed similarly in Semesters #2 and #3 when predictor-combinations $X_1 \sim X_6$, $X_1 \sim X_7$, and $X_1 \sim X_8$ were employed to develop the models. However, RBF models performed much better than the SVM models in Semester #4 in terms of identifying the percentage of academically at-risk students. RBF models correctly identified 64.1% of academically at-risk students in Semester #2, 46.7% in

Semester #3, and 28.1% in Semester #4.

Discussion of the Results

The following points can be deduced from the comparison of the four types of models with different combinations of predictors:

1. No matter what modeling techniques are used, models with $X_1 \sim X_6$, $X_1 \sim X_7$, and $X_1 \sim X_8$ are always ranked as the best-performing models. Including students' in-class performance measurements ($X_6 \sim X_8$) as predictor variables increases the prediction accuracy of the models because they ($X_6 \sim X_8$) represent student achievement throughout the dynamics course.

2. The best combination of predictors that yield the highest prediction accuracy is $X_1 \sim X_8$. This combination works well for all models. However, X_7 and X_8 are the last two dynamics mid-term exams. Including X_7 and X_8 as predictor variables is not beneficial for the instructor because it might be too late for him or her to implement educational interventions to improve student learning. Therefore, the models with $X_1 \sim X_6$ as predictors are the most useful because they not only yield accurate prediction results, but also leave sufficient time for the instructor to implement educational interventions.

3. In general, the prediction accuracy of the models that include X_6 (dynamics mid-term exam #1) as the only predictor is lower than that of the models with X_1 (cumulative GPA) as the only predictor. This is because X_1 is a more comprehensive representation of a student's skills and knowledge than X_6 . However, X_6 has more influence than X_1 on student academic performance in dynamics final exam in MLR

models. This is because partial topics tested in the final exam were covered in mid-term exam #1 (X_6).

4. In general, the prediction accuracy of all models in Semester #3 is lower than in the other three semesters. One possible reason is the distribution of data. Student performance varies from semester to semester. The distribution of the dynamics final exam score is close to a normal distribution in Semesters #1, #2, and #4, but is a bimodal distribution in Semester #3.

5. Compared to MLR and MLP models, RBF and SVM models have lower prediction accuracy in internal validation but higher prediction accuracy in external validation in terms of both APA and PAP. One possible reason is that RBF and SVM models are more robust against disturbance when applying the predictive models to different semesters. In other words, RBF and SVM models have better generalizability.

6. $X_2 \sim X_4$ had non-significant or even negative coefficients in MLR models. This may be caused by the correlation between predictor variables; for example, the correlation coefficient between X_4 and X_1 was 0.668. However, the correlation was not problematic because collinearity was not an issue in this study. To keep the predictor variables consistent with those for MLP, RBF, and SVM models, all predictor variables were kept in the MLR models.

7. MLP models have the lowest prediction accuracy compared to the models developed by the other three types of modeling techniques. On the one hand, the MLP, multilayer feed forward neural networks, has “difficulty in making correct predictions on data that are contradictory to the ones used for their training” (Lykourantzou et al., 2009).

Many factors that are unpredictable, such as students' health and motivation, affect student academic performance. Some students may have high prior achievement and in-class performance (measured by the dynamics mid-term exams), but low achievement in the dynamics final exam. On the other hand, the scaled conjugate gradient algorithm was used to adjust the weight values in the MLP networks. However, that algorithm didn't guarantee that the weight values were globally optimal. The risk of local minima of parameters limits the performance of MLP models.

Implications of the Research Findings

The following is an overview of the research completed in the study: (a) Different combinations of predictors have been identified to predict student academic performance in an engineering dynamics course; (b) various statistical and data mining techniques have been used and compared in developing predictive models; and (c) models have been used to identify academically at-risk students in the engineering dynamics course.

The research findings from this study imply that RBF and SVM models are the best at predicting the "average" academic performance of all students in the dynamics class. The models using X_1 and $X_1 \sim X_5$ as predictors only take into account a student's prior knowledge and prior achievement, and can be used only as an initial attempt to estimate student performance in dynamics. These models can be developed before the course even begins. The positive aspect of these types of models is that the instructor has sufficient time to consider what proactive measures he or she will use to improve performance in the new semester.

However, if the instructor would like to predict student performance more accurately, he/she should not use the models with X_1 , $X_1 \sim X_5$, or X_6 as predictors because those models have low prediction accuracy, especially a low percentage of accurate predictions. The models with $X_1 \sim X_6$ as predictors are recommended because they have moderate predictability to generate accurate predictions and also leave enough time for the instructor to implement educational interventions.

If the main purpose is to identify academically at-risk students, the instructor should use RBF models with $X_1 \sim X_6$ or $X_1 \sim X_7$ as predictors because they represent student prior knowledge, prior achievement, and in-class performance in the dynamics course. RBF models are more robust to the change of data in term of identifying academically-at-risk students.

Finally, although the models including $X_1 \sim X_8$ are the mathematically best among the four types of models, they cannot be used until after the third exam when the semester is almost over and when educational interventions for academically at-risk students are difficult to implement. Therefore, the primary application of the models with $X_1 \sim X_8$ as predictors might be labeled as “interpretation” rather than “prediction,” which means these models can be used to “explain” how each of the eight predictor variables affects a student’s final exam score.

Limitations of This Research

This research has several limitations. First, only some cognitive variables including prior achievement and prior domain knowledge were concerned in this study.

Non-cognitive variables such as motivation, interest, major, and gender were not included in this study. Although the APAs of the predictive models were high, the PAPs were moderate. *R*-square of the MLR models also showed that more than 50% of student academic performance in dynamics was not explained by the cognitive predictor variables used in this study.

Second, the grades that a student earned in prerequisite courses might not truly reflect the student's knowledge of those topics. A student may have taken prerequisite courses years ago. By the time he/she takes dynamics, his/her knowledge of prerequisite courses may have improved. For example, some students took calculus courses more than two semesters before they took dynamics, and got only a C- in the calculus final exam. However, they may have received more practice with calculus problems through some other courses, such as physics, and it is possible that they would now understand calculus at a level higher than their below-average grade would suggest. The prediction accuracy is reduced when the grade earned in calculus is used as a predictor variable.

Third, no differentiation is made between norm-referenced and criterion-referenced scores in the data collected. Different predictor variables might use different criteria. A student who earns 60 (out of 100) in a criterion-referenced system may receive an A in a norm-referenced system (Gronlund & Waugh, 2009). Thus, a student who got an A in a prerequisite class might not truly understand the given topics as well as his/her grade indicates, and may receive a low grade in the dynamics course.

Recommendations for Future Studies

Educational research shows that some psychological factors, such as learning style, self-efficacy (Ransdell, 2001; Riding & Rayner, 1998; Tracey & Sedlacek, 1984), motivation and interest, and teaching and learning environment (Graaff, Saunders-Smiths, & Nieweg, 2005), also play a role in student learning and thus affect student achievement. Therefore, future studies should include psychological variables in the models so as to increase their prediction accuracy (Lin, Imbrie, & Reid, 2009). A longitudinal study could be employed that involves the measurement of student psychological factors as well as other information such as students' majors.

To better assess student prior domain knowledge, a pretest prior to the start of the dynamics course is suggested in future studies. The pretest should cover the topics in statics, calculus, and physics, such as free-body diagrams, integral and differential equations, and impulse and momentum.

In addition to mid-term exams, dynamics homework may also be included as a predictor variable in the predictive models. Student performance in homework assignments reflect student learning progression and problem-solving skills. In the future studies, efforts will be made to investigate whether the prediction accuracy of the models can be increased by including student performance in dynamics homework assignments as an additional predictor variable.

Finally, it must be pointed out that the predictive models developed in this study were based on the data collected at Utah State University. The developed models can be employed as a general tool to predict student academic performance in dynamics course,

so they can benefit both teaching and learning. When extending the modeling techniques to another institution of higher learning, it is recommended to collect the data on student academic performance at that particular institution to develop corresponding predictive models. This will ensure that the corresponding predictive models best reflect the features of teaching and learning at that particular institution.

REFERENCES

- Alexander, P. A. (1992). Domain knowledge: Evolving themes and emerging concerns. *Educational Psychologist, 27*(1), 33-51.
- Allen, J., Robbins, S. B., & Sawyer, R. (2010). Can measuring psychosocial factors promote college success? *Applied Measurement in Education, 23*(1), 1-22.
- Angrist, J. D., & Lavy, V. (1999). Using maimonides' rule to estimate the effect of class size on scholastic achievement. *The Quarterly Journal of Economics, 114*(2), 533-575.
- Ayan, M. N. R., & Garcia, M. T. C. (2008). Prediction of university students' academic achievement by linear and logistic models. *The Spanish Journal of Psychology, 11*(1), 275-288.
- Barrett, S. F., LeFevre, E. W., Steadman, J. W., Tietjen, J. S., White, K. R., & Whitman, D. L. (2010). Using the fundamentals of engineering (FE) examination as an outcomes assessment tool. *National Council of Examiners for Engineering and Surveying*. Retrieved from www.ncees.org/Documents/Public/White%20paper%202010%20FAQ.pdf
- Bezeau, S., & Graves, R. (2001). Statistical power and effect sizes of clinical neuropsychology research. *Journal of Clinical and Experimental Neuropsychology, 23*(3), 399-406.
- Biggers, S., Orr, M., & Benson, L. (2010, June). *Integrated dynamics and statics for first semester sophomores in mechanical engineering*. Paper presented at the American Society for Engineering Education Annual Conference & Exposition, Louisville, KY.
- Bleeker, S. E., Moll, H. A., Steyerberg, E. W., Donders, A. R. T., Derksen-Lubsen, G., Grobbee, D. E., & Moons, K. G. M. (2003). External validation is necessary in prediction research: A clinical example. *Journal of Clinical Epidemiology, 56*, 826-832.
- Brush, T. A. (1997). The effects on student achievement and attitudes when using integrated learning systems with cooperative pairs. *Educational Technology Research and Development, 45*(1), 51-64.
- Cabena, P., Hadjinian, P., Stadler, R., Verhees, J., & Zanasi, A. (1997). *Discovering data mining: From concept to implementation*. Englewood Cliffs, NJ: Prentice Hall.
- Cahen, L. S., & Filby, N. N. (1979). The class size/achievement issue: New evidence and a research plan. *The Phi Delta Kappan, 60*(7), 492-495, 538.

- Camara, W. J., & Echternacht, G. (2000). *The SAT I and high school grades: Utility in predicting success in college* (College Board Report No. RN-10). New York, NY: College Entrance Examination Board.
- Cassidy, S., & Eachus, P. (2000). Learning style, academic belief systems, self-report student proficiency and academic achievement in higher education. *Educational Psychology, 20*(3), 318-322.
- Chang, C. C., & Lin, C. J. (2001). *LibSVM: A library for support vector machines*. Retrieved from <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>
- Chapelle, O., & Vapnik, V. (2000). Model selection for support vector machines. In S. A. Solla, T. K. Leen, & K-R. Müller (Eds.), *Advances in neural information processing systems*, (Vol. 12, pp. 216-222). Cambridge, MA: MIT.
- Chapelle, O., Vapnik, V., Bousquet, O., & Mukherjee, S. (2002). Choosing multiple parameters for support vector machines. *Machine Learning, 46*(1-3), 131-159.
- Chen, P. W., Wang, J. Y., & Lee, H. M. (2004). Model selection of SVMs using GA approach. *Proceedings of the International Joint Conference on Neural Networks, Budapest, Hungary, 2035–2040*.
- Cherkassky, V., & Ma, Y. (2004). Practical selection of SVM parameters and noise estimation for SVM regression. *Neural Networks, 17*, 113-126.
- Cohen, J. (1962). The statistical power of abnormal-social psychological research: A review. *Journal of Abnormal and Social Psychology, 65*(3), 145-153.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Mahwah, NJ: Erlbaum.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). Outliers and multicollinearity: Diagnosing and solving regression problems. In D. Riegert (Ed.), *Applied multiple regression/correlation analysis for the behavioral science* (Vol. II, pp. 391-430). Mahwah, NJ: Erlbaum.
- Collobert, R., & Bengio, S. (2001). SVM torch: Support vector machines for large-scale regression problems. *Journal of Machine Learning Research, 1*, 143-150.
- Criddle, K. R. (2004). Nonlinear regression and autoregression. In *Intermediate statistics and applied regression analysis* (pp. 195-229). Providence, UT: East-West Bridge Publishing House.
- Cripps, A. (1996, February). Using artificial neural nets to predict academic performance. In J. Hightower (Chair), *Proceedings of the Association for Computing Machinery Symposium on Applied Computing*, Philadelphia, PA.

- Cristianini, N., & Taylor, J. S. (2000). *An introduction to support vector machines and other kernel-based learning methods*. New York, NY: Cambridge University.
- Danko-McGhee, K., & Duke, J. C. (1992). Predicting student performance in accounting classes. *Journal of Education for Business*, 67(5), 270-274.
- Das, A., Ben-Menachem, T., Cooper, G. S., Chak, A., Sivak, M. V., Jr., Gonet, J. A., & Wong, R. C. K. (2003). Prediction of outcome in acute lower-gastrointestinal haemorrhage based on an artificial neural network: Internal and external validation of a predictive model. *The Lancet*, 362, 1261-1266.
- Delauretis, R. J., & Molnar, G. E. (1972). Academic prediction: Paradigm research or let's see what happens? *IEEE Transactions on Education*, E15(1), 32-36.
- Deng, N., & Tian, Y. (2004). *A new method of data mining: Support vector machine*. Beijing, China: Science Press.
- Dillon, M., Kokkelenberg, E. C., & Christy, S. M. (2002). The effects of class size on student achievement in higher education: Applying an earnings function. *Cornell Higher Education Research Institute (CHERI)*, 15. Retrieved from <http://digitalcommons.ilr.cornell.edu/cheri/15>
- Dochy, F. (1992). *Assessment of prior knowledge as a determinant for future learning: The use of knowledge state tests and knowledge profiles*. Utrecht, London: Lemma B. V.
- Dochy, F., De Rijdt, C., & Dyck, W. (2002). Cognitive prerequisites and learning. *The Institutes for Learning and Teaching in Higher Education and SAGE Publications*, 3(3), 265-284.
- Dochy, F., Segers, M., & Buehl, M. M. (1999). The relation between assessment practices and outcomes of studies: The case of research on prior knowledge. *Review of Educational Research*, 69(2), 145-186.
- Dufresne, R. J., Gerace, W. J., Hardiman, P. T., & Mestre, J. P. (1992). Constraining novices to perform expertlike problem analyses: Effects on schema acquisition. *The Journal of the Learning Sciences*, 2(3), 307-331.
- Ehrenberg, R. G., Brewer, D. J., Gamoran, A., & Willms, J. D. (2001). Class size and student achievement. *Psychological Science in the Public Interest*, 2(1), 1-30.
- Ellis, G. W., & Turner, W. A. (2003). Helping students organize and retrieve their understanding of dynamics. *Proceedings of the American Society for Engineering Education Annual Conference & Exposition*. Retrieved from <http://www.science.smith.edu/departments/Engin/pdf/ASEE%20Helping%20Student.pdf>

- Emerson, T. L., & Taylor, B. A. (2004). Comparing student achievement across experimental and lecture-oriented sections of a principle of microeconomics course. *Southern Economic Journal*, 70(3), 672-693.
- Etkina, E., Mestre, J. P., & O'Donnell, A. (2005). The impact of the cognitive revolution on science learning and teaching. In J. M. Royer (Ed.), *The cognitive revolution in educational psychology* (pp. 119-164). Greenwich, CT: Information Age.
- Everitt, B. S. (2009). *Multivariable modeling and multivariate analysis for the behavioral sciences*. London, UK: CRC.
- Fang, N., & Lu, J. (2010). A decision tree approach to predictive modeling of student performance in engineering dynamics. *International Journal of Engineering Education*, 26(1), 87-95.
- Feelders, A., Daniels, H., & Holsheimer, M. (2000). Methodological and practical aspects of data mining. *Information & Management*, 37(5), 271-281.
- Field, A. (2005). *Discovering statistics using SPSS*. London, UK: Sage.
- Fisher, K. M., Wandersee, J. H., & Moody, D. E. (2000). Meaningful and mindful learning. In K. Tobin (Ed.), *Mapping biology knowledge* (pp. 77-94). Dordrecht, Netherlands: Kluwer.
- Fleming, J. (2002). Who will succeed in college? When the SAT predicts black students' performance. *The Review of Higher Education*, 25(3), 281-296.
- Fleming, J., & Garcia, N. (1998). Are standardized tests fair to African Americans? Predictive validity of the SAT in Black and White institutions. *Journal of Higher Education*, 69(5), 471-495.
- Fletcher, J. (1998). *A study of the factors affecting advancement and graduation for engineering students* (Doctoral dissertation). Available from Dissertations and Theses at Auburn University database. (Publication No. AAT 9912921)
- French, B. F., Immekus, J. C., & Oakes, W. (2005). An examination of indicators of engineering students' success and persistence. *Journal of Engineering Education*, 94(4), 419-425.
- Friedrichs, F., & Igel, C. (2005). Evolutionary tuning of multiple SVM parameters. *Neurocomputing*, 64, 107-117.
- Fulcher, J. (2008). Artificial neural networks. In J. Flucher, & L. C. Jain (Eds.), *Computational intelligence: A compendium* (pp. 26-30). Berlin, Germany: Springer.

- Graaff, E., Saunders-Smiths, G. N., & Nieweg, M. R. (2005). *Research and practice of active learning in engineering education*. Amsterdam, The Netherlands: Palls.
- Glossary Letter E. (2011). In *Journal of Clinical Oncology*. Retrieved from http://jco.ascopubs.org/cgi/glossarylookup?lookup=by_letter&letter=E
- Glossary Letter I. (2011). In *Journal of Clinical Oncology*. Retrieved from http://jco.ascopubs.org/cgi/glossarylookup?lookup=by_letter&letter=I
- Gronlund, N. E., & Waugh, C. K. (2009). *Assessment of student achievement*. Boston, MA: Allyn & Bacon.
- Gunn, S. R. (1998). *Support vector machines for classification and regression*. Retrieved from the Support Vector Machines website: <http://www.svms.org/tutorials/Gunn1998.pdf>
- Hailikari, T., Nevgi, A., & Komulainen, E. (2008). Academic self-beliefs and prior knowledge as predictors of student achievement in mathematics: A structural model. *Educational Psychology, 28*(1), 59-71.
- Hailikari, T., Nevgi, A., & Ylanne, S. L. (2007). Exploring alternative ways of assessing prior knowledge, its components and their relation to student achievement: A mathematics based case study. *Studies in Educational Evaluation, 33*(3-4), 320-337.
- Han, J., & Kamber, M. (2001). Introduction. In D. C. Diane, H. Severson, & B. Breyer (Eds.), *Data mining: Concepts and techniques* (pp. 1-38). San Francisco, CA: Morgan Kaufman.
- Hand, D., Mannila, H., & Smyth, P. (2001). *Principles of data mining*. Cambridge, MA: MIT.
- Harachiewicz, J. M., Barron, K. E., Tauer, J. M., & Elliot, A. J. (2002). Predicting success in college: A longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation. *Journal of Educational Psychology, 94*(3), 562-575.
- Hearst, M. A. (1998). Support vector machines. *IEEE Intelligent Systems and Their Applications, 13*(4), 18-21.
- Hicks, D. W., & Richardson, F. M. (1984). Predicting early success in intermediate accounting: The influence of entry examination and GPA. *Issues in Accounting Education, 2*(1), 61-67.
- Hibbeler, R. C. (2010). *Engineering mechanics dynamics*. Upper Saddle River, NJ: Pearson.

- Hoffman, J. L. (2002). The impact of student curricular involvement on student success: Racial and religious differences. *Journal of College Student Development*, 43(5), 712–739.
- Hoffman, J. L., & Lowitzki, K. E. (2005). Predicting college success with high school grades and test scores: Limitations for minority students. *The Review of Higher Education*, 28(4), 455-474.
- Holland, J. L., James, J. R., & Richards, M. (1966). Academic and non-academic accomplishment in a representative sample taken from a population of 612,000. *American College Testing Program*. Retrieved from <http://eric.ed.gov/PDFS/ED014093.pdf>
- Hong, D. H., & Hwang, C. (2003). Support vector fuzzy regression machines. *Fuzzy Sets and Systems*, 138(2), 271-281
- Howell, D. C. (2010). *Statistical methods for psychology*. Belmont, CA: Wadsworth.
- Hoxby, C. M. (2000). The effects of class size on student achievement: New evidence from population variation. *The Quarterly Journal of Economics*, 115(4), 1239-1285.
- Huang, S., & Fang, N. (2010, June). *Regression models for predicting student academic performance in an engineering dynamics course*. Paper presented at the American Society for Engineering Education Annual Conference & Exposition, Louisville, KY.
- Ibrahim, S. R. (2004). An integrated approach for the engineering dynamics course. *Proceedings of the American Society for Engineering Education Annual Conference & Exposition, Salt Lake City, UT* (pp. 7535-7541). Washington, DC: ASEE.
- Imbrie, P. K., Lin, J. J. J., Reid, K., & Malyscheff, A. (2006). *Using hybrid data to model student success in engineering with artificial neural networks*. Retrieved from http://www.engconfintl.org/8axabstracts/Session%202C/rees08_submission_64.pdf
- Johnson, D. M. (1991). Student achievement and factors related to achievement in a state FFA agricultural mechanics contest. *Journal of Agricultural Education*, 32(3), 23-28.
- Kennedy, P. E., & Siegfried, J. J. (1997). Class size and achievement in introductory economics: Evidence from the TUCE III data. *Economics of Education Review*, 16(4), 385-394.

- Kopeika, N. S. (1992). On the relationship of number of students to academic level. *IEEE Transactions on Education*, 35(4), 294-295.
- Kotsiantis, S. B., Pierrakeas, C. J., & Pintelas, P. E. (2003, September). *Preventing student dropout in distance learning using machine learning techniques*. Paper presented at the AI Techniques in Web-Based Educational Systems at Seventh International Conference on Knowledge-Based Intelligent Information and Engineering Systems, Oxford, England.
- Levin, J. (2001). For whom the reductions count: A quantile regression analysis of class size and peer effects on scholastic achievement. *Empirical Economics*, 26(1), 221-246.
- Lin, J. J. J., Imbrie, P. K., & Reid, K. J. (2009, July). *Student retention modeling: An evaluation of different methods and their impact on prediction results*. Paper presented at the Research in Engineering Education Symposium, Palm Cove, Queensland.
- Linoff, G. S., & Berry, M. J. A. (2011). Artificial neural networks. In R. Elliott, A. O. Tulton, D. Scribner, & P. Lowell (Eds.), *Data mining techniques for marketing, sales, and customer relationship management* (pp. 211-255). Indianapolis, IN: Wiley.
- Lovegreen, T. A. (2003). *Predicting the academic success of female engineering students during the first year of college using the SAT and noncognitive variables* (Master thesis). Virginia Polytechnic Institute and State University. (No. ETD-04252003-184329).
- Lewis, M., & Castley, A. (2008). Factors affecting student progression and achievement: Prediction and intervention. *Innovations in Education and Teaching International*, 45(4), 333-343.
- Lykourantzou, L., Giannoukos, L., & Mpardis, G. (2009). Early and dynamic student achievement prediction in e-learning courses using neural networks. *Journal of the American Society for Information Science and Technology*, 60(2), 372-380.
- Magill, M. A. (1997). Classroom models for illustrating dynamics principles, part I- particle kinematics and kinetics. *Proceeding of American Society of Engineering Education Annual Conference and Exposition, Milwaukee, WI* (pp. 15-18). Washington, DC: ASEE.
- Maimon, O. Z. (2008). Neural network methods. In O. Maimon & L. Rokach (Eds.), *Soft computing for knowledge discovery and data mining* (pp. 17-44). New York, NY: Springer.

- Marsh, C. M., Vandehey, M. A., & Diekhoff, G. M. (2008). A comparison of an introductory course to SAT/ACT scores in predicting student performance. *The Journal of General Education*, 57(4), 244-255.
- Mohandes, M. A., Halawani, T. O., & Rehman, S. (2004). Support vector machines for wind speed prediction. *Renewable Energy*, 29(6), 939-947.
- Momma, M., & Bennett, K. P. (2002). A pattern search method for model selection of support vector regression. *Proceedings of SIAM Conference on Data Mining*. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.101.8309&rep=rep1&type=pdf>
- Muthu, G., & Glass, J. (1999). Engineering a flexible access instruction program in engineering dynamics. *Higher Education Research and Development Society of Australasia Annual International Conference, Melbourne, Australia* (pp. 12-15). Crawley, Western Australia: HERDSA.
- Nihalani, P. K., Wilson, H. E., Thomas, G., & Robinson, D. H. (2010). What determines high- and low-performing groups? The superstar effect. *Journal of Advanced Academics*, 21(3), 500-529.
- North Carolina State University. (2011). *College of Engineering distance education graduate degree programs, degree requirements and course offerings*. Retrieved from <http://engineeringonline.ncsu.edu/onlinecourses/Engineering%20Online%20Graduate%20Degree%20Programs%20and%20Course%20Requirements.pdf>
- O'Donnell, A. M., & Dansereau, D. F. (2000). Interactive effect of prior knowledge and material format on cooperative teaching. *Journal of Experimental Education*, 68(2), 101-118.
- Onwuegbuzie, A. J., Collins, K. M. T., & Elbedour, S. (2003). Aptitude by treatment interactions and matthew effects in graduate-level cooperative-learning groups. *The Journal of Educational Research*, 96(4), 217.
- Pai, P. F., & Hong, W. C. (2005). Forecasting regional electricity load based on recurrent support vector machines with genetic algorithms. *Electric Power Systems Research*, 74(3), 417-425.
- Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, 1(2), 117-175.
- Pittman, K. (2008). *Comparison of data mining techniques used to predict student retention* (Doctoral dissertation). Available from Nova Southeastern University database. (UMI No. 3297573)

- Ransdell, S. (2001). Predicting college success: The importance of ability and non-cognitive variables. *International Journal of Educational Research*, 35(4), 357-364.
- Riding, R., & Rayner, S. (1998). *Cognitive styles and learning strategies: Understanding style differences in learning and behavior*. London, UK: David Fulton.
- Romon, M. M., & Christodoulou, C. (2006). *Support vector machines for antenna array processing and electromagnetics*. San Rafael, CA: Morgan & Claypool.
- Royer, J. M. (1986). Designing instruction to produce understanding: An approach based on cognitive theory. In G. D. Phye & T. Andre (Eds.), *Cognitive classroom learning: Understanding, thinking, and problem solving* (pp. 83-113). Orlando, FL: Academic.
- Rubin, M. B., & Altus, E. (2000). An alternative method for teaching dynamics. *International Journal of Engineering Education*, 16(5), 447-456.
- Self, B., & Redfield, R. (2001). New approaches in teaching undergraduate dynamics. *Proceedings of the American Society for Engineering Education Annual Conference & Exposition, Albuquerque, NM* (pp. 7465-7475). Washington, DC: ASEE.
- Self, B. P., Wood, J. J., & Hansen, D. (2004, July). *Teaching undergraduate kinetics using LEGO mindstorms race car competition*. Paper presented at the American Society for Engineering Education, Salt Lake City, UT.
- Smola, A. J., & Scholkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing*, 14(3), 199-222.
- Soper, D. S. (2004). *Statistics calculators*. Retrieved from <http://www.danielsoper.com/statcalc/>
- Staelin, C. (2002). *Parameter selection for support vector machine* (Report No. HPL-2002-354). Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.4.2355&rep=rep1&type=pdf>
- Stitson, M. O., Weston, J. A. E., Gammerman, A. V., & Vapnik, V. (1996). *Theory of support vector machines* (Report No. CSD-TR-96-17). London, UK: Royal Holloway University.
- Thissen, U., van Brakel, R., de Weijer, A. P., Melssen, W. J., & Buydens, L. M. C. (2003). Using support vector machines for time series prediction. *Chemometrics and Intelligent Laboratory System*, 69(1), 35-49.

- Thompson, R. A., & Zamboanga, B. L. (2003). Prior knowledge and its relevance to student achievement in introduction to psychology. *Teaching of Psychology, 30*(2), 96-101.
- Thompson, R. A., & Zamboanga, B. L. (2004). Academic aptitude and prior knowledge as predictors of student achievement in instruction to psychology. *Journal of Educational Psychology, 96*(4), 778-784.
- Ting, S. R. (2001). Predicting academic success of first-year engineering students from standardized test scores and psychosocial variables. *International Journal of Engineering Education, 17*(1), 75-80.
- Tobias, S. (1994). Interest, prior knowledge, and learning. *Review of Educational Research, 64*(1), 37-54.
- Tracey, T. J., & Sedlacek, W. E. (1984). Noncognitive variables in predicting academic success by race. *Measurement and Evaluation in Guidance, 16*(4), 171-178.
- Trafalis, T. B., & Ince, H. (2000). Support vector machine for regression and applications to financial forecasting. *Proceedings of the IEEE-INNS-ENNS International Joint Conference, 6*, 348-353.
- Utah State University. (2011). *Mechanical engineering, BS*. Retrieved from http://catalog.usu.edu/preview_program.php?catoid=3&poid=1262&returnto=244
- Vandamme, J. P., Meskens, N., & Superby, J. F. (2007). Predicting academic performance by data mining methods. *Education Economics, 15*(4), 405-419.
- Vapnick, V. N. (1995). *The nature of statistical learning theory*. New York, NY: Springer-Verlag.
- Veenstra, C. P., Dey, E. L., & Herrin, G. D. (2008). Is modeling of freshman engineering success different from modeling of non-engineering success? *Journal of Engineering Education, 97*(3), 467-479.
- Veenstra, C. P., Dey, E. L., & Herrin, G. D. (2009). *A model for freshman engineering retention*. Retrieved from <http://advances.asee.org/vol01/issue03/papers/aee-vol01-issue03-p07.pdf>
- Ware, W. B., & Galassi, J. P. (2006). Using correlational and prediction data to enhance student achievement in K-12 schools: A practical application for school counselors. *Professional School Counseling, 9*(5), 344-356.
- Wilson, K. M. (1983). *A review of research on the prediction of academic performance after the freshman year* (College Board Report No. 83-2). New York, NY: College Entrance Examination Board.

Zhu, W., Aung, K., & Zhou, J. (2010, June). *Design improvement and analysis on 3D haptic-augmented learning tools for dynamics course*. Paper presented at the American Society for Engineering Education Annual Conference & Exposition, Louisville, KY.

APPENDIX



USU Assurance: FWA#00003308

Protocol # 2503

1/10/2010

SPO #:
AES #: UTA00

MEMORANDUM

TO: Ning Fang
Shaobo Huang

FROM: Kim Corbin-Lewis, IRB Chair *Kim Corbin-Lewis*
True M. Fox, IRB Administrator *True M. Fox*

SUBJECT: A Study on Student Academic Performance in an Engineering Dynamics
Course: Predictive Modeling and its Applications

Your proposal has been reviewed by the Institutional Review Board and is approved under expedite procedure #7

- There is no more than minimal risk to the subjects.
 There is greater than minimal risk to the subjects.

This approval applies only to the proposal currently on file for the period of one year. If your study extends beyond this approval period, you must contact this office to request an annual review of this research. Any change affecting human subjects must be approved by the Board prior to implementation. Injuries or any unanticipated problems involving risk to subjects or to others must be reported immediately to the Chair of the Institutional Review Board.

Prior to involving human subjects, properly executed informed consent must be obtained from each subject or from an authorized representative, and documentation of informed consent must be kept on file for at least three years after the project ends. Each subject must be furnished with a copy of the informed consent document for their personal records.

The research activities listed below are expedited from IRB review based on the Department of Health and Human Services (DHHS) regulations for the protection of human research subjects, 45 CFR Part 46, as amended to include provisions of the Federal Policy for the Protection of Human Subjects, November 9, 1998.

7. Research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies.



USU IRB Approval: Jan. 10, 2010; Page 1 of 3
IRB Approval Terminates: 01/09/2012
Protocol #2503
IRB Password Protected per IRB Administrator

INFORMED CONSENT

A Study on Student Academic Performance in an Engineering Dynamics Course:
Predictive Modeling and Its Applications

Introduction/ Purpose Professor Ning Fang in the Department of Engineering and Technology Education at Utah State University (USU) is conducting a research study to find out more about student academic performance in the Engineering Dynamics course. You have been asked to take part because you have been enrolled in this course. There will be approximately 200 participants per year at Utah State University College of Engineering. There will be approximately 600-800 total participants in this study.

Your participation in this study is completely optional. Your final grade of the Engineering Dynamics course will not be affected in any negative way if you decide not to participate in this study. If you have not taken the Engineering Dynamics course before, your final grade of the course is calculated by using a computer program that is designed based on the pre-set course syllabus (for example, homework accounts for 15% of the final grade, class attendance 5.5%, keynotes 5%, mid-exams 54%, and final exam 25%). All homework and exam questions are multiple choice questions that are graded independently by either a computer program (on USU Blackboard) or a teaching assistant. If you have taken the Engineering Dynamics course before, the data obtained from you will only be used to develop and refine the initial predictive model. The initial predictive model will then be validated and applied to the students who will be enrolled in the course in new semesters.

Procedures This study consists of two closely-related tasks. Task 1 focuses on the development of a mathematical model to predict students' final exam score in ENGR 2030 Engineering Dynamics course offered at USU College of Engineering. Task 2 focuses on the application of the developed model in ENGR 2030 Engineering Dynamics course. The general procedure is described in detail as follows.

In Task 1, relevant data is first collected from participating students as the inputs of the model. The data to be collected include student GPA, their scores in four pre-requisite courses (Statics, Calculus I, Calculus II, and Physics), and student mid-term and final exam scores in Engineering Dynamics. Then, a variety of statistical and data-mining techniques, such as multi-variable regression, neural network, support vector machine, and/or decision trees, will be used to develop a mathematical model to predict student final exam score in Dynamics. The prediction accuracy of the model will be examined and validated.

In Task 2, the predictive model developed in Task 1 will be applied to participating students in new semesters to predict how well or how poorly they will perform in the Engineering Dynamics course. A set of survey questions will also be administered to seek detailed comments of the students on their particular learning needs.

It is totally optional for you to decide whether or not you are willing to participate in this study. No student has to participate. If you choose to participate in this study, you need to sign on this letter of written consent. Your signature indicates the following:



USU IRB Approval: Jan. 10, 2010; Page 2 of 3
IRB Approval Terminates: 01/09/2012
Protocol #2563
IRB Password Protected per IRB Administrator

INFORMED CONSENT

A Study on Student Academic Performance in an Engineering Dynamics Course: Predictive Modeling and Its Applications

1. You are willing to release your cumulative GPA, grades in four pre-requisite course (Engineering Statics, Calculus I and II, and Physics, mid-term scores and the final exam score in the Dynamics course to Professor Fang for this particular research project.
2. You are willing to complete two surveys to ask about your learning experience and the instructor's teaching quality. Selected participants will also be interviewed to obtain more detailed information about their learning experience.

New Findings During the course of this research study, you will be informed of any significant new findings (either good or bad), such as changes in the risks or benefits resulting from participation in the research, or new alternatives to participation that might cause you to change your mind about continuing in the study. If new information is obtained that is relevant or useful to you, or if the procedures and/or methods change at any time throughout this study, your consent to continue participating in this study will be obtained again.

Risks The risks or discomforts to the participants are minimal.

Benefits This study may benefit not only the students but also the instructor of the Engineering Dynamics course. First, the predictive model can be used to predict how well or how poorly a student will perform in the final exam of the Engineering Dynamics course. If the prediction shows that a student will perform poorly, then that student may consider taking pro-active measures, such as spending more time on learning course materials, re-studying some old course materials that they are supposed to master in relevant pre-requisite courses, and seeking more instructional assistance from the teaching assistants and the instructor after the class to discuss technical questions.

Second, the instructor of the Engineering Dynamics course can also use the predicted results to develop a good understanding of how well or how poorly the students will perform and then take pro-active measures to best suit student problem-solving skills and meet their needs. For example, if the predicted results show that a significant number of students in the class will perform poorly, then the instructor may consider adding recitation sessions and more office hours after the class to help improve student learning.

Explanation & offer to answer questions Ning Fang has explained this research study to you and answered your questions. If you have other questions or research-related problems, you may reach Professor at (435) 797-2948

Voluntary nature of participation and right to withdraw without consequence Participation in research is entirely voluntary. You may refuse to participate or withdraw at any time without consequence or loss of benefits.



USU IRB Approval: Jan. 10, 2010; Page 3 of 3
IRB Approval Terminates: 01/09/2012
Protocol #2503
IRB Password Protected per IRB Administrator

INFORMED CONSENT

A Study on Student Academic Performance in an Engineering Dynamics Course:
Predictive Modeling and Its Applications

Confidentiality The data will be protected in a variety of ways. First, each participating student will be assigned a participant ID number. Each data collection will be associated with that generated ID number. The collected raw data will be stored in a locked file cabinet in a locked room, more specifically, in the USU Industrial Science Building Room 106C. Scanned copies of survey results will be stored in a secured, password-protected computer of the PI. Only the PI has access to the collected raw data. The PI will also create and maintain a coded list that could be used to match names and participant student ID codes to protect privacy. A coding sheet associating ID number with student names will be kept separately and locked in a filing cabinet in the PI's office to maintain confidentiality. Only the PI has access to this coding document. After data has been collected, the document which connects student names and ID numbers will be destroyed immediately. All these measures maximize the safety of all the data and participants' confidentiality.

IRB Approval Statement The Institutional Review Board for the protection of human participants at USU has approved this research study. If you have any pertinent questions or concerns about your rights or a research-related injury, you may contact the IRB Administrator at (435) 797-0567 or email irb@usu.edu. If you have a concern or complaint about the research and you would like to contact someone other than the research team, you may contact the IRB Administrator to obtain information or to offer input.

Copy of consent You have been given two copies of this Informed Consent. Please sign both copies and keep one copy for your files.

Investigator Statement "I certify that the research study has been explained to the individual, by me or my research staff, and that the individual understands the nature and purpose, the possible risks and benefits associated with taking part in this research study. Any questions that have been raised have been answered."

Principal Investigator (Ning Fang)
(435-797-2948)
(ning.fang@usu.edu)

Student Researcher (Shaobo Huang)
(435-797-1796)
(shaobo.huang@aggiemail.usu.edu)

Signature of Participant By signing below, I agree to participate.

Participant's signature

Date

CURRICULUM VITAE

SHAOBO HUANG, PhD

Engineering and Technology Education Department
Utah State University, Logan, UT 84322-6000
Phone: (435)-363-5730
E-mail: shaobo.huang@aggiemail.usu.edu

GPA: 3.86/4.0

Education Background

- Ph.D. Engineering and Technology Education, Utah State University, May 2009~Dec 2011.
Dissertation title: Predictive modeling and analysis of student academic performance in an engineering dynamics course
Advisor: Dr. Ning Fang
- M.S. Control Theory and Control Engineering, Harbin Engineering University, 2006~2009
Thesis title: Intelligent control system of submarine course-keeping
Advisor: Dr. Sheng Liu
- B.S. Automation Engineering, Qingdao University, 2002~2006

Awards and Honors

- Presidential Fellowship, Utah State University, 2009-2010
 - Dissertation Fellowship, Utah State University, 2011-2012
 - The 2nd Place for Oral Presentation, Intermountain Graduate Research Symposium, organized by Utah State University, 2011
 - Scholarship, Qingdao University, 2005-2006
 - Outstanding Student Leadership, Qingdao University, 2004
 - Scholarship, Qingdao University, 2002-2003
-

Teaching and Research Experience

- Teaching Assistant for the Engineering Dynamics course, Utah State University, January 2011 ~ date
- Research Assistant, Utah State University, 2009 ~ date
- A key member of the graduate project “Intelligent Control of Submarine Course System”, Harbin Engineering University, 2007~2008
- Member of the project “Research on Signal Process for Bistatic Sonar System”, Harbin Engineering University, 2008

Refereed Publications

- **Huang, S.**, & Fang, N. (2010). Prediction of Student Academic Performance in an Engineering Dynamics Course: Development and Validation of Multivariate Regression Models. *International Journal of Engineering Education*, 26(4), 1-10.
 - **Huang, S.**, & Fang, N. (2010). *Regression models for predicting student academic performance in an engineering dynamics course*. Proceedings of the American Society for Engineering Education Annual Conference & Exposition, Louisville, KY.
 - **Huang, S.**, & Fang, N. (2011). *WIP: Prediction of students' academic performance in an introductory engineering course*. Proceedings of the 2011 Frontiers in Education Conference (Accepted).
 - Fang, N., & **Huang, S.** (2009). *An electronic classroom response system for active learning and engineering education research*. Proceedings of the International Symposium on Total Engineering Education. Shanghai, China. 91-98.
 - Sheng, L., & **Huang, S.** (2010). Model Predictive Control of Ship Course with LSSVM. *Control Engineering*.
 - Jin, H., Wang, L., Li, D., & **Huang, S.** (2008). Research on Control Strategy for Ship Anti-roll at Zero Speed, *Ship Engineering*.
-