ANALYZING ACADEMIC ACHIEVEMENT OF CAS's STUDENTS USING DATA MINING

A thesis submitted to the College of Arts and Sciences in partial fulfillment of the requirements for the degree of Master of Science (Information Technology) Universiti Utara Malaysia

by

Nor Asiah binti Abdul Rahman
KOLEJ SASTERA DAN SAINS  
(College of Arts and Sciences)  
Universiti Utara Malaysia

PERAKUAN KERJA KERTAS PROJEK  
(Certificate of Project Paper)

Saya, yang bertandatangan, memperakukan bahawa  
(I, the undersigned, certify that)  

NOR ASIAH ABDUL RAHMAN  
(89450)

calon untuk ijazah  
(candidate for the degree of)  
MSc. (Information Technology)

telah mengemukakan kertas projek yang bertajuk  
(has presented his/her project paper of the following title)

ANALYZING ACADEMIC ACHIEVEMENT OF  
CAS’S STUDENT USING DATA MINING

seperti yang tercatat di muka surat tajuk dan kulit kertas projek  
(as it appears on the title page and front cover of project paper)

bahawa kertas projek tersebut boleh diterima dari segi bentuk serta kandungan  
dan meliputi bidang ilmu dengan memuaskan.  
(that the project paper acceptable in form and content, and that a satisfactory  
knowledge of the field is covered by the project paper).

Nama Penyelia Utama  
(Name of Main Supervisor): ASSOC. PROF. FADZILAH SIRAJ

Tandatangan  
(Signature):

Tarikh  
(Date): 10-05-2009
PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a postgraduate degree from Universiti Utara Malaysia, I agree that the University Library may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purpose may be granted by my supervisor(s) or, in their absence by the Dean of Academic Affairs. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to Universiti Utara Malaysia for any scholarly use which may be made of any material from my thesis.

Requests for permission to copy or to make other use of materials in this thesis, in whole or in part, should be addressed to

Division of Applied Sciences
College of Arts and Sciences
Universiti Utara Malaysia
06010 UUM Sintok
Kedah Darul Aman.
ABSTRACT (BAHASA MALAYSIA)

Massive information can be collected from students' data in order to produce knowledge. The educational institutions collect students' data such as academic information, demographic, and personal traits. The data collected based on these variables used to predict the students' academic achievement. On this study, the respondents are students who have graduated within the period of six months in the year 2006, 2007 and 2008. Two data mining techniques for analyzing and building the classification model for students' achievement in College of Arts and Sciences (CAS), Universiti Utara Malaysia (UUM) are presented. Initially, the relationship and correlation between students' cumulative grade point average (CGPA) with academic background, demographic, entry qualification, sponsorship and interpersonal skills, students' achievement are analyzed. For model building purposes, final CGPA has been used as a target. The analysis conducted using Multinomial Logistic Regression and Neural Network found that, gender, entry qualification, language qualification (Bahasa Malaysia and English), family income, sponsorship, analytical and analysis skill as well as teamwork are all the best predictors contributed to students' performance. The result obtained through this study can be used to help the management of CAS to make certain decisions and to predict the outcome of current and future students.
ACKNOWLEDGEMENTS

Thanks to ALLAH S.W.T with His bless and mercy for giving me strength to accomplish my thesis and simultaneously completed my masters degree. With this opportunity, I would like to express my grateful to Universiti Utara Malaysia for the valuable chance by funding my education in master degree level.

I would like to express my gratitude to particular individuals who have contributed to the completion of this thesis. First of all, I would like to acknowledge my supervisor who has inspired me in this research area, Associate Professor Fadzilah Siraj and for her remarkable guidance, integrity, generous support and comments as well as her genuineness knowledge and experience throughout the process of writing this thesis.

I would also like to thank my evaluator, Associate Professor Azizi Zakaria, for the impressive comments, suggestions and ideas in evaluating my presentation and thesis.

To a very special person in my life, my beloved husband, Mr. Shamsul Amri Abdul Manap thank you very much for the boundless encouragement, strength and love that came along side towards finishing this thesis.

I am deeply grateful to my family and family in law for their support, pray and understanding. And last but not least, I want to acknowledge my friends and co-workers for their trust and concerned.
# TABLE OF CONTENTS

Permission to use ........................................................................................................ i
Abstract (Bahasa Malaysia) ....................................................................................... ii
Abstract (English) ..................................................................................................... iii
Acknowledgement ...................................................................................................... iv
List of Tables ............................................................................................................... vii
List of Figures ............................................................................................................ x
Appendixes ................................................................................................................ xi
Abbreviations ............................................................................................................. xii

## CHAPTER 1: INTRODUCTION

1.1 Research Background ......................................................................................... 1
1.2 Overview of College of Arts and Sciences (CAS) .............................................. 5
1.3 Problem Statement ............................................................................................. 7
1.4 Research’s Objectives ....................................................................................... 8
1.5 Scope .................................................................................................................. 8
1.6 Research Questions ........................................................................................... 9
1.7 Significance of Study ....................................................................................... 9

## CHAPTER 2: LITERATURE REVIEW

2.1 Data Mining in Education ............................................................................... 11
2.2 Academic Achievement .................................................................................. 15
2.3 Logistic Regression ......................................................................................... 26
2.4 Artificial Neural Network (ANN) .................................................................... 30

## CHAPTER 3: METHODOLOGY

3.1 Business Understanding .................................................................................. 35
3.2 Data Understanding ....................................................................................... 35
3.3 Data Preparation .............................................................................................. 37
  3.3.1 Data Description ...................................................................................... 38
  3.3.2 Data Ready to be Used ........................................................................... 39
3.4 Modeling ........................................................................................................... 39
  3.4.1 Multinomial Logistic Regression ................................................................ 40
  3.4.2 Neural Network ...................................................................................... 41
3.5 Evaluation ....................................................................................................... 46
3.6 Deployment ....................................................................................................... 46
CHAPTER 4: DATA ANALYSIS AND FINDINGS

4.1 Descriptive Statistics for UUM Graduates ........................................... 47
4.2 Descriptive Statistics for CAS Students ............................................... 50
  4.2.1 Demographic Characteristics ......................................................... 50
  4.2.2 Academic Information ................................................................. 53
  4.2.3 Employment After Graduation ....................................................... 58
4.3 Multinomial Logistic Regression .......................................................... 62
  4.3.1 Modelling for CAS ................................................................. 63
  4.3.2 Modelling for FTM ................................................................. 69
  4.3.3 Modelling for FE ................................................................. 72
  4.3.4 Modelling for FPPH ................................................................. 75
  4.3.5 Modelling for FSKP ................................................................. 78
  4.3.6 Modelling for FKBK ................................................................. 82
  4.3.7 Modelling for FSPK ................................................................. 85
  4.3.8 Modelling for FSK ................................................................. 88
4.4 Neural Network .................................................................................. 92
  4.4.1 Modelling for CAS (6 targets) ..................................................... 92
  4.4.2 Modelling for FTM ................................................................. 99
  4.4.3 Modelling for FE ................................................................. 104
  4.4.4 Modelling for FPPH ................................................................. 109
  4.4.5 Modelling for FSKP ................................................................. 113
  4.4.6 Modelling for FKBK ................................................................. 118
  4.4.7 Modelling for FSPK ................................................................. 123
  4.4.8 Modelling for FSK ................................................................. 128
  4.4.9 Modelling for CAS (3 targets) ................................................... 132

CHAPTER 5: CONCLUSION

5.1 Discussion ..................................................................................... 139
5.2 Conclusion .................................................................................... 140

REFERENCES .................................................................................. 142
APPENDIX A .......................................................... 148
APPENDIX B .......................................................... 151
APPENDIX C .......................................................... 159
LIST OF TABLES

Table 4.1  Case processing summary for CAS ........................................... 63
Table 4.2  Distribution of target ................................................................ 63
Table 4.3  Classification table for CAS ...................................................... 64
Table 4.4  Model Fitting Information (CAS) ................................................. 64
Table 4.5  Pseudo R-Square (CAS) .............................................................. 65
Table 4.6  Likelihood Ratio Tests (CAS) ...................................................... 66
Table 4.7  List of Variables Included in students' achievement dataset (CAS) ........................................................................................................... 68
Table 4.8  Classification Table (FTM) .......................................................... 69
Table 4.9  Model Fitting Information (FTM) ................................................ 69
Table 4.10 Pseudo R-Square (FTM) ............................................................ 70
Table 4.11 Likelihood Ratio Tests (FTM) .................................................... 70
Table 4.12 List of Variables Included in students' achievement dataset (FTM) ........................................................................................................... 71
Table 4.13 Classification Table (FE) ............................................................ 72
Table 4.14 Model Fitting Information (FE) ................................................ 72
Table 4.15 Pseudo R-Square (FE) ............................................................... 73
Table 4.16 Likelihood Ratio Tests (FE) ....................................................... 73
Table 4.17 List of Variables Included in students' achievement dataset (FE) ........................................................................................................... 74
Table 4.18 Classification Table (FPPH) ....................................................... 75
Table 4.19 Model Fitting Information (FPPH) .............................................. 75
Table 4.20 Pseudo R-Square (FPPH) .......................................................... 76
Table 4.21 Likelihood Ratio Tests (FPPH) .................................................. 76
Table 4.22 List of Variables Included in students' achievement dataset (FPPH) ........................................................................................................ 77
Table 4.23 Classification Table (FSKP) ........................................................ 78
Table 4.24 Model Fitting Information (FSKP) ............................................. 78
Table 4.25 Pseudo R-Square (FSKP) ........................................................... 79
Table 4.26 Likelihood Ratio Tests (FSKP) ................................................... 79
Table 4.27 List of Variables Included in students' achievement dataset (FSKP) ........................................................................................................ 81
Table 4.28 Classification Table (FKBM) ...................................................... 82
Table 4.29 Model Fitting Information (FKBM) .......................................... 82
Table 4.30 Pseudo R-Square (FKBM) ........................................................ 83
Table 4.31 Likelihood Ratio Tests (FKBM) ................................................. 83
Table 4.32 List of Variables Included in students' achievement dataset (FKBM) ........................................................................................................ 84
Table 4.33 Classification Table (FPSM) ...................................................... 85
Table 4.34 Model Fitting Information (FPSM) .......................................... 85
Table 4.35 Pseudo R-Square (FPSM) ........................................................ 86
Table 4.36 Likelihood Ratio Tests (FPSM) ................................................. 86
Table 4.37 List of Variables Included in students' achievement dataset (FPSM) ........................................................................................................ 87
Table 4.38 Classification Table (FSK). ........................................... 88
Table 4.39 Model Fitting Information (FSK). .................................. 88
Table 4.40 Pseudo R-Square (FSK). .............................................. 89
Table 4.41 Likelihood Ratio Tests (FSK) .......................................... 89
Table 4.42 List of Variables Included in students’ achievement dataset
(FSK). .................................................................................. 90
Table 4.43a Results to determine the best epoch .............................. 92
Table 4.43b Results to determine the best data allocation ............... 93
Table 4.43c Results to determine the best average hidden unit ....... 94
Table 4.43d Results to determine the best learning rate ............... 95
Table 4.43e Results to determine the best momentum rate ........... 96
Table 4.44 Parameter values for neural network model with 6 targets 96
Table 4.45a Results to determine the best epoch .............................. 99
Table 4.45b Results to determine the best data allocation ............... 100
Table 4.45c Results to determine the best average hidden unit ..... 100
Table 4.45d Results to determine the best learning rate ............... 101
Table 4.45e Results to determine the best average learning rate .... 102
Table 4.45f Results to determine the best learning rate ............... 102
Table 4.45g Results to determine the best average momentum rate 103
Table 4.46 Parameter values for FTM model ................................. 103
Table 4.47a Results to determine the best epoch .............................. 104
Table 4.47b Results to determine the best data allocation ............... 105
Table 4.47c Results to determine the best average hidden unit ....... 106
Table 4.47d Results to determine the best learning rate ............... 107
Table 4.47e Results to determine the best average learning rate .... 107
Table 4.47f Results to determine the best learning rate ............... 108
Table 4.48 Parameter values for FE model ................................. 108
Table 4.49a Results to determine the best epoch .............................. 109
Table 4.49b Results to determine the best data allocation ............... 110
Table 4.49c Results to determine the best average hidden unit ....... 111
Table 4.49d Results to determine the best learning rate ............... 111
Table 4.49e Results to determine the best average learning rate .... 112
Table 4.49f Results to determine the best momentum rate ........... 112
Table 4.50 Parameter values for FPPH model ............................... 114
Table 4.51a Results to determine the best epoch .............................. 114
Table 4.51b Results to determine the best data allocation ............... 114
Table 4.51c Results to determine the best average hidden unit ....... 115
Table 4.51d Results to determine the best learning rate ............... 116
Table 4.51e Results to determine the best average learning rate .... 116
Table 4.51f Results to determine the best momentum rate ........... 117
Table 4.51g Results to determine the best average momentum rate 117
Table 4.52 Parameter values for FSKP model ............................... 118
Table 4.53a Results to determine the best epoch .............................. 119
Table 4.53b Results to determine the best data allocation ............... 119
Table 4.53c Results to determine the best average hidden unit ....... 120
Table 4.53d Results to determine the best learning rate ................ 121
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.53e</td>
<td>Results to determine the best momentum rate</td>
<td>122</td>
</tr>
<tr>
<td>4.53f</td>
<td>Results to determine the best average momentum rate</td>
<td>122</td>
</tr>
<tr>
<td>4.54</td>
<td>Parameter values for FKBM model</td>
<td>123</td>
</tr>
<tr>
<td>4.55a</td>
<td>Results to determine the best epoch</td>
<td>123</td>
</tr>
<tr>
<td>4.55b</td>
<td>Results to determine the best data allocation</td>
<td>124</td>
</tr>
<tr>
<td>4.55c</td>
<td>Results to determine the best average hidden unit</td>
<td>125</td>
</tr>
<tr>
<td>4.55d</td>
<td>Results to determine the best learning rate</td>
<td>126</td>
</tr>
<tr>
<td>4.55e</td>
<td>Results to determine the best average learning rate</td>
<td>126</td>
</tr>
<tr>
<td>4.55f</td>
<td>Results to determine the best momentum rate</td>
<td>127</td>
</tr>
<tr>
<td>4.55g</td>
<td>Results to determine the best average momentum rate</td>
<td>127</td>
</tr>
<tr>
<td>4.56</td>
<td>Parameter values for FPSM model</td>
<td>128</td>
</tr>
<tr>
<td>4.57a</td>
<td>Results to determine the best epoch</td>
<td>128</td>
</tr>
<tr>
<td>4.57b</td>
<td>Results to determine the best data allocation</td>
<td>129</td>
</tr>
<tr>
<td>4.57c</td>
<td>Results to determine the best average hidden unit</td>
<td>130</td>
</tr>
<tr>
<td>4.57d</td>
<td>Results to determine the best learning rate</td>
<td>131</td>
</tr>
<tr>
<td>4.57e</td>
<td>Results to determine the best momentum rate</td>
<td>131</td>
</tr>
<tr>
<td>4.58</td>
<td>Parameter values for FSK model</td>
<td>132</td>
</tr>
<tr>
<td>4.59</td>
<td>New target with three (3) classes</td>
<td>132</td>
</tr>
<tr>
<td>4.60a</td>
<td>Results to determine the best epoch</td>
<td>133</td>
</tr>
<tr>
<td>4.60b</td>
<td>Results to determine the best data allocation</td>
<td>134</td>
</tr>
<tr>
<td>4.60c</td>
<td>Results to determine the best hidden unit</td>
<td>135</td>
</tr>
<tr>
<td>4.60d</td>
<td>Results to determine the best learning rate</td>
<td>135</td>
</tr>
<tr>
<td>4.60e</td>
<td>Results to determine the best average learning rate</td>
<td>136</td>
</tr>
<tr>
<td>4.60f</td>
<td>Results to determine the best momentum rate</td>
<td>136</td>
</tr>
<tr>
<td>4.61</td>
<td>Parameter values for neural network model with 3 targets</td>
<td>137</td>
</tr>
<tr>
<td>4.62</td>
<td>Comparison of accuracy between logistic regression and neural network</td>
<td>138</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>A List of CAS’s Programmes Offered by Faculty</td>
<td>5</td>
</tr>
<tr>
<td>Figure 3.1</td>
<td>Phases of CRISP-DM Reference Model</td>
<td>34</td>
</tr>
<tr>
<td>Figure 3.2</td>
<td>Data sample for 2006</td>
<td>36</td>
</tr>
<tr>
<td>Figure 3.3</td>
<td>Data sample for 2007</td>
<td>36</td>
</tr>
<tr>
<td>Figure 3.4</td>
<td>Data sample for 2008</td>
<td>36</td>
</tr>
<tr>
<td>Figure 3.5</td>
<td>The Variables Distribution for Three Different Years</td>
<td>38</td>
</tr>
<tr>
<td>Figure 3.6</td>
<td>Sample Data for Multinomial Logistic Regression</td>
<td>39</td>
</tr>
<tr>
<td>Figure 3.7</td>
<td>Sample Data for Neural Network</td>
<td>39</td>
</tr>
<tr>
<td>Figure 3.8</td>
<td>Multinomial Logistic Regression with SPSS 12.0</td>
<td>40</td>
</tr>
<tr>
<td>Figure 3.9</td>
<td>MLP Icon to Control the Parameter</td>
<td>42</td>
</tr>
<tr>
<td>Figure 3.10</td>
<td>The Architecture of Multilayer Perceptron</td>
<td>42</td>
</tr>
<tr>
<td>Figure 3.11</td>
<td>Data Allocation</td>
<td>43</td>
</tr>
<tr>
<td>Figure 3.12</td>
<td>Setting the Input Layer</td>
<td>44</td>
</tr>
<tr>
<td>Figure 3.13</td>
<td>Multilayer Perceptron illustration</td>
<td>45</td>
</tr>
<tr>
<td>Figure 3.14</td>
<td>Training session</td>
<td>45</td>
</tr>
<tr>
<td>Figure 4.1</td>
<td>The total number of UUM graduates for year 2006 to 2008</td>
<td>48</td>
</tr>
<tr>
<td>Figure 4.2</td>
<td>Undergraduate students for year 2006, 2007 and 2008 by faculties</td>
<td>48</td>
</tr>
<tr>
<td>Figure 4.3</td>
<td>Comparison by gender for year 2006, 2007 and 2008</td>
<td>49</td>
</tr>
<tr>
<td>Figure 4.4</td>
<td>Distribution between races for all UUM graduates</td>
<td>49</td>
</tr>
<tr>
<td>Figure 4.5</td>
<td>The total number of CAS students</td>
<td>50</td>
</tr>
<tr>
<td>Figure 4.6</td>
<td>Comparison of CAS students by year</td>
<td>51</td>
</tr>
<tr>
<td>Figure 4.7</td>
<td>Comparison by gender for CAS students</td>
<td>52</td>
</tr>
<tr>
<td>Figure 4.8</td>
<td>Distribution between gender and races</td>
<td>52</td>
</tr>
<tr>
<td>Figure 4.9</td>
<td>The percentage of programmes enrolled by students at three different years</td>
<td>53</td>
</tr>
<tr>
<td>Table 4.10</td>
<td>Distribution of entry qualification of CAS graduates</td>
<td>54</td>
</tr>
<tr>
<td>Figure 4.11</td>
<td>Distribution of students’ CGPA by class</td>
<td>55</td>
</tr>
<tr>
<td>Figure 4.12</td>
<td>Distribution of CGPA and study programmes</td>
<td>55</td>
</tr>
<tr>
<td>Figure 4.13</td>
<td>CGPA distribution between gender</td>
<td>56</td>
</tr>
<tr>
<td>Figure 4.14</td>
<td>CGPA and entry qualification by gender difference</td>
<td>56</td>
</tr>
<tr>
<td>Figure 4.15</td>
<td>CGPA distribution among races</td>
<td>57</td>
</tr>
<tr>
<td>Figure 4.16</td>
<td>Distribution of students’ CGPA by family income</td>
<td>58</td>
</tr>
<tr>
<td>Figure 4.17</td>
<td>The percentage of graduates by employment</td>
<td>59</td>
</tr>
<tr>
<td>Figure 4.18</td>
<td>Distribution of students employment after graduation</td>
<td>60</td>
</tr>
<tr>
<td>Figure 4.19</td>
<td>The percentage of students’ CGPA and employment</td>
<td>60</td>
</tr>
<tr>
<td>Figure 4.20</td>
<td>Employed students and CGPA achievement</td>
<td>61</td>
</tr>
<tr>
<td>Figure 4.21</td>
<td>Distribution of employed students by study programmes</td>
<td>62</td>
</tr>
<tr>
<td>Figure 4.22</td>
<td>Multilayer Perceptron Architecture for CAS</td>
<td>97</td>
</tr>
<tr>
<td>Figure 4.23</td>
<td>Confusion matrix for testing result</td>
<td>98</td>
</tr>
<tr>
<td>Figure 4.24</td>
<td>Confusion matrix for training result</td>
<td>98</td>
</tr>
<tr>
<td>Figure 4.25</td>
<td>Confusion matrix for testing result</td>
<td>137</td>
</tr>
<tr>
<td>Figure 4.26</td>
<td>Confusion matrix for training result</td>
<td>138</td>
</tr>
</tbody>
</table>
APPENDIXES

Appendix A  Data Description .................................................. 148
Appendix B  List of tables for data analysis .............................. 151
Appendix C  Modelling and result summary ............................... 159
**ABBREVIATIONS**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN</td>
<td>Artificial Neural Network</td>
</tr>
<tr>
<td>CAS</td>
<td>College of Arts and Sciences</td>
</tr>
<tr>
<td>CGPA</td>
<td>Cumulative Grade Point Average</td>
</tr>
<tr>
<td>COB</td>
<td>College of Business</td>
</tr>
<tr>
<td>COLGIS</td>
<td>College of Law, Government and International Studies</td>
</tr>
<tr>
<td>CS</td>
<td>Computer Science</td>
</tr>
<tr>
<td>CRISP-DM</td>
<td>Cross Industry Standard Process for Data Mining</td>
</tr>
<tr>
<td>DM</td>
<td>Data Mining</td>
</tr>
<tr>
<td>FE</td>
<td>Faculty of Economics</td>
</tr>
<tr>
<td>FKBM</td>
<td>Faculty of Communication and Modern Languages</td>
</tr>
<tr>
<td>FPH</td>
<td>Faculty of Tourism and Hospitality</td>
</tr>
<tr>
<td>FPSM</td>
<td>Faculty of Humanities and Social Development</td>
</tr>
<tr>
<td>FSK</td>
<td>Faculty of Quantitative Sciences</td>
</tr>
<tr>
<td>FSKP</td>
<td>Faculty of Cognitive Sciences and Education</td>
</tr>
<tr>
<td>FTM</td>
<td>Faculty of Information Technology</td>
</tr>
<tr>
<td>GPA</td>
<td>Grade Point Average</td>
</tr>
<tr>
<td>ICT</td>
<td>Information and Communication Technology</td>
</tr>
<tr>
<td>IPTA</td>
<td>Institut Pengajian Tinggi Awam</td>
</tr>
<tr>
<td>KDD</td>
<td>Knowledge Discovery in Databases</td>
</tr>
<tr>
<td>MUET</td>
<td>Malaysian University English Test</td>
</tr>
<tr>
<td>UUM</td>
<td>Universiti Utara Malaysia</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

This chapter consists of a study on the students' performance of College of Arts and Sciences (CAS), Universiti Utara Malaysia (UUM). Performance information is gathered from students' final semester results. Research background, problem statements, project's objectives, scope, research questions and significance of the study are highlighted in this chapter.

1.1 RESEARCH BACKGROUND

Students' performance in academic achievement is the major concern in the universities (Fennolar, Roman, & Cuestas, 2007). The increasing of students attending university has developed the interest in identifying factors to predict academic performance. In higher education, the issues of prediction and explanation of academic performance and a study to identify the key indicators to the academic success and persistence of students are extremely important (Komaraju, Karau & Ramayah, 2007; Ervina & Md Nor, 2005).
The contents of the thesis is for internal user only
REFERENCES


http://www.davidson.edu/academic_economics/Student%E2%80%99s Research%20Papers/Brendan%20Carroll%20paper.pdf


Luan, J. (2001). Data mining as driven by knowledge management in higher education- persistence clustering and prediction. *SPSS Public Conference, UCSF*.


