MODELING ACADEMIC ACHIEVEMENT WITH RESPECT TO SELF-READINESS OF PUBLIC UNIVERSITIES GRADUATES

JIWA NORIS BIN HAMID

UNIVERSITI UTARA MALAYSIA
2008
MODELING ACADEMIC ACHIEVEMENT WITH RESPECT TO SELF-READINESS OF PUBLIC UNIVERSITIES GRADUATES

A thesis submitted to the Centre for Graduate Studies in partial fulfillment of the requirements for the degree Master of Science (Information Technology) Universiti Utara Malaysia

By

Jiwa Noris bin Hamid

© Jiwa Noris bin Hamid, 2008. All rights reserved.
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)

JIWA NORIS BIN HAMID

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

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

MODELING ACADEMIC ACHIEVEMENT WITH RESPECT TO  
SELF – READINESS OF PUBLIC UNIVERSITIES GRADUATES

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)  : [Signature]

Tarikh  
(Date)  : 1/6/08
PERMISSION TO USE

In presenting this project in partial fulfillment of the requirements for a postgraduate degree from the 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 scholar purposes may be granted by my supervisor(s) or, in their absent, by the Dean of Centre for Graduate Studies. 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 Universiti Utara Malaysia for any scholar use which may be made of any material from my thesis.

Request for permission to copy or make other use of the material in this thesis in whole or part should be addressed to:

Dean
Centre for Graduate Studies
Universiti Utara Malaysia
06010 Sintok
Kedah Darul Aman
ABSTRACT (ENGLISH)

Students’ performances in higher education institution are evaluated based on their academic achievement. Despite of having technical skills obtained by the graduates during their study in university, it is crucial for them to have other additional skills that help them in decision making and problem solving. This study has presented a model of academic achievement with respect to self-readiness of the public universities graduates. Two data mining methods were used in this study such as logistic regression and neural network to obtain the model with the highest accuracy. The selection of data mining approaches was based on the ability of data mining as a powerful tool for academic analysis purposes. In higher educational institution, data mining can be used for the process of uncovering hidden trends and patterns that help them forecast the students’ achievement. In this study, a dataset comprises of public higher educational institution graduates demographics and self-readiness information was analyzed. The results show that descriptive and self readiness produce higher accurate percentage compared to the self readiness alone. The result also find that neural network is the best model to be developed compared to logistic regression while field of study, citizenship, ready to face working world and challenges, problem solving, decision making and group working are the best predictor for academic achievement.
ABSTRAK

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my supervisor, Associate Professor Fadzilah Siraj for her invaluable guidance, encouragement and knowledge-sharing in completing my thesis work.

I am especially grateful to all my family members especially my lovely wife, Hawa Mohd. Ekhsan and my son, Farhan Iskandar for their support and patience. For those who are not mentioned here, I would like to take this opportunity to thank everyone.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Permission to Use</td>
<td>i</td>
</tr>
<tr>
<td>Abstract (English)</td>
<td>ii</td>
</tr>
<tr>
<td>Abstract (Malay)</td>
<td>iii</td>
</tr>
<tr>
<td>Acknowledgement</td>
<td>iv</td>
</tr>
<tr>
<td>List of Tables</td>
<td>viii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>x</td>
</tr>
<tr>
<td>List of Appendices</td>
<td>xii</td>
</tr>
</tbody>
</table>

## CHAPTER 1: INTRODUCTION

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Overview</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Problem Statement</td>
<td>4</td>
</tr>
<tr>
<td>1.3 Objectives</td>
<td>5</td>
</tr>
<tr>
<td>1.4 Scope</td>
<td>6</td>
</tr>
<tr>
<td>1.5 Thesis Organization</td>
<td>7</td>
</tr>
</tbody>
</table>

## CHAPTER 2: LITERATURE REVIEW

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Academic Achievement</td>
<td>8</td>
</tr>
<tr>
<td>2.2 Self-readiness</td>
<td>11</td>
</tr>
<tr>
<td>2.3 Data Mining</td>
<td>13</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 2.1  Data mining questions in the private sector and their higher education equivalents  16
Table 3.1  Example of demographic questions  25
Table 3.2  Example of self-readiness attributes  25
Table 3.3  CGPA classification  28
Table 3.4  Respondent from various public universities  30
Table 3.5  The selectes attributes before converted to numbers  31
Table 3.6  The selectes attributes after converted to numbers  32
Table 3.7  Data allocation for training, validation, test  36
Table 4.0  Total number of respondent and related attributes  43
Table 4.1  Total number of the respondent based on public universities  44
Table 4.3  Total number of respondent based on mode of study  45
Table 4.4  List of demographic attributes selected for the study  45
Table 4.5  List of self-readiness attributes selected for the study  52
Table 4.6  Respondents’ perception towards self-readiness items with academic achievement  53
Table 4.7  Case processing summary  57
Table 4.8  Model fitting information for self-readiness only  58
Table 4.9  Classification table for self-readiness only  59
Table 4.10  Model fitting information for self-readiness with demographic  59
Table 4.11 Classification table for self-readiness with demographic

Table 4.12 The accuracy of logistic regression

Table 4.13 Significant values for self-readiness items and demographic

Table 4.14 The results of Chi Square and significant value for self-readiness and demographic items

Table 4.15 Parameter values initialization

Table 4.16 Result of neural networks accuracy for initial experiment

Table 4.17 The effects of hidden unit on neural networks results

Table 4.18 Training and test results for different hidden units and different weight seeds

Table 4.19 The effect of data allocation on neural networks result

Table 4.20 The accuracy percentage for self-readiness only dataset

Table 4.21 The comparison of the accuracy percentage for demographic with self-readiness and self-readiness only dataset

Table 4.22 Parameter values for neural network model
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>The process flow of the study</td>
<td>24</td>
</tr>
<tr>
<td>3.2</td>
<td>Snapshot of demographic questions</td>
<td>26</td>
</tr>
<tr>
<td>3.3</td>
<td>Snapshot of self-readiness questions</td>
<td>27</td>
</tr>
<tr>
<td>3.4</td>
<td>Phases of CRISP-DM Reference Model</td>
<td>29</td>
</tr>
<tr>
<td>3.5</td>
<td>Sample of data set before converted using SPSS 12.0</td>
<td>33</td>
</tr>
<tr>
<td>3.6</td>
<td>Process of data conversion using SPSS 12.0</td>
<td>33</td>
</tr>
<tr>
<td>3.7</td>
<td>Data set after converted in SPSS 12.0</td>
<td>34</td>
</tr>
<tr>
<td>3.8</td>
<td>Network architecture of the study</td>
<td>35</td>
</tr>
<tr>
<td>3.9</td>
<td>Neural Connection interface</td>
<td>36</td>
</tr>
<tr>
<td>3.10</td>
<td>Data loaded into Neural Connection</td>
<td>37</td>
</tr>
<tr>
<td>3.11</td>
<td>Parameter setting for MLP network</td>
<td>37</td>
</tr>
<tr>
<td>3.12</td>
<td>Parameter setting for MLP training</td>
<td>38</td>
</tr>
<tr>
<td>3.13</td>
<td>Multinominal Logistic Regression using SPSS 12.0</td>
<td>39</td>
</tr>
<tr>
<td>3.14</td>
<td>Prototype system structure diagram</td>
<td>40</td>
</tr>
<tr>
<td>4.0</td>
<td>Snapshot of demographic attributes</td>
<td>46</td>
</tr>
<tr>
<td>4.1</td>
<td>Percentage of the respondents by gender</td>
<td>46</td>
</tr>
<tr>
<td>4.2</td>
<td>Percentage of the respondents by age</td>
<td>47</td>
</tr>
<tr>
<td>4.3</td>
<td>Percentage of the respondents by citizenship</td>
<td>47</td>
</tr>
<tr>
<td>4.4</td>
<td>Percentage of the respondents with disabilities</td>
<td>48</td>
</tr>
<tr>
<td>4.5</td>
<td>Percentage of the respondents by CGPA</td>
<td>49</td>
</tr>
<tr>
<td>4.6</td>
<td>Percentage of the respondents by field of study</td>
<td>49</td>
</tr>
</tbody>
</table>

x
Figure 4.7  Percentage of the respondents by mode of study  

Figure 4.8  Percentage of the respondents by entry qualification  

Figure 4.9  Percentage of the respondents by family income  

Figure 4.10  Snapshot of self-readiness attributes  

Figure 4.11  The percentage of “more sensitive towards current affairs” among respondents  

Figure 4.12  The percentage of “creative and critical thinking” among respondents  

Figure 4.13  The percentage of “develop self confidence” among respondents
CHAPTER 1

INTRODUCTION

This chapter discusses the overview of the study implemented in this study. This chapter consists of research background, problem statement, objectives, significance and scope of the study conducted. At the end of this chapter, thesis organization is presented to describe the structure of this report.

1.1 Overview of the study

Higher education creates opportunities and helps to equip individuals with the appropriate knowledge and skills. In Malaysia, higher educational institutions can be divided into two, public higher educational institutions, which are established and maintained by the Government and private higher educational institutions, which are privately funded institutions and have been registered with the Ministry of Education.
The contents of the thesis is for internal user only
REFERENCES


Luan, J. (2004). Data Mining Applications in higher education. SPSS Inc.

Luan, J. (2006). Data Mining Applications in Higher Education. SPSS Inc.


Engineering. 34th ASEE/IEEE Frontiers in Education Conference, October 20-23, 2004, Savannah, GA.


