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

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) : Fadzilah

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

Prestasi graduan di peringkat pengajian tinggi biasanya dinilai berdasarkan kepada pencapaian akademik. Selain daripada kemahiran teknikal yang diperolehi semasa menuntut di universiti, adalah penting untuk para graduan mempunyai kemahiran tambahan yang lain bagi membantu mereka membuat keputusan dan menyelesaikan masalah. Kajian ini telah mempersembahkan model pencapaian akademik berhubung kesediaan diri graduan di universiti awam tempatan. Dua teknik perlombongan data iaitu rangkaian neural dan regresi logistik telah digunakan untuk mendapatkan model dengan peratusan ketepatan yang paling tinggi. Pemilihan pendekatan perlombongan data adalah berdasarkan kepada keupayaannya sebagai alat yang sangat berkuasa untuk tujuan analisa akademik. Di peringkat pengajian tinggi, perlombongan data boleh digunakan untuk proses membongkar trend dan paten tersembunyi dalam membantu meramal pencapaian graduan. Dalam kajian ini, set data yang mengandungi maklumat demografik dan kesediaan diri graduan telah dianalisa. Hasil kajian mendapati demografik dan kesediaan diri menghasilkan peratus ketepatan yang lebih tinggi berbanding kesediaan diri semata-mata. Hasil kajian juga mendapati bahawa rangkaian neural menghasilkan model yang lebih baik berbanding logistik regresi sementara bidang pengajian, kewarganegaraan, kesediaan diri untuk menghadapi dunia dan cabaran pekerjaan, kebolehan menyelesaikan masalah, kebolehan membuat keputusan dan kebolehan bekerja sebagai satu pasukan merupakan peramal yang baik untuk pencapaian akademik graduan.

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

Permission to Use i

Abstract (English) ii

Abstract (Malay) iii

Acknowledgement iv

List of Tables viii

List of Figures x

List of Appendices..... xii

CHAPTER 1: INTRODUCTION

1.1 Overview 1

1.2 Problem Statement 4

1.3 Objectives 5

1.4 Scope 6

1.5 Thesis Organization 7

CHAPTER 2: LITERATURE REVIEW

2.1 Academic Achievement 8

2.2 Self -readiness..... 11

2.3 Data Mining 13

2.4 Data Mining application in Higher Education..... 15

2.4.1 Predictive..... 16

2.4.1.1 Neural Networks..... 19

2.4.1.2 Regression..... 20

CHAPTER 3: RESEARCH METHODOLOGY

3.1 Survey..... 24

3.1.1 Instrument..... 26

3.1.2 Respondents data..... 27

3.1.3 Target..... 27

3.2 Cross Industry Standard Process 28

3.3 Information System Development Research Process..... 40

CHAPTER 4: FINDING AND RESULTS

4.1 Results 42

4.1.1 Descriptive 45

4.1.2 Logistic Regression..... 57

4.1.3 Neural Network..... 64

CHAPTER 5: CONCLUSION AND RECOMMENDATION 70

REFERENCES 71

APPENDIX A..... 78

APPENDIX B..... 85

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	60
Table 4.12	The accuracy of logistic regression	61
Table 4.13	Significant values for self- readiness items and demographic	62
Table 4.14	The results of Chi Square and significant value for self- readiness and demographic items	63
Table 4.15	Parameter values initialization	64
Table 4.16	Result of neural networks accuracy for initial experiment	65
Table 4.17	The effects of hidden unit on neural networks results	66
Table 4.18	Training and test results for different hidden units and different weight seeds	67
Table 4.19	The effect of data allocation on neural networks result	67
Table 4.20	The accuracy percentage for self- readiness only dataset	68
Table 4.21	The comparison of the accuracy percentage for demographic with self- readiness and self- readiness only dataset	68
Table 4.22	Parameter values for neural network model	69

LIST OF FIGURES

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

Figure 4.7	Percentage of the respondents by mode of study	50
Figure 4.8	Percentage of the respondents by entry qualification	51
Figure 4.9	Percentage of the respondents by family income	51
Figure 4.10	Snapshot of self- readiness attributes	52
Figure 4.11	The percentage of "more sensitive towards current affairs" among respondents	54
Figure 4.12	The percentage of "creative and critical thinking" among respondents	55
Figure 4.13	The percentage of "develop self confidence" among respondents	56

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

Abdullah concerned about varsity rankings. (2008, January 23). *NST Online*. Retrieve on January 15, 2008 from http://www.nst.com.my.

Abu, B., Johan, O. M., Mansor, S. M. S. S. & Jaafar, H. (2007). Kepelbagaian Gaya Pembelajaran Dan Kemahiran Belajar Pelajar Universiti Di Fakulti Pendidikan, UTM Johor . Research Vote No: 71881. Jabatan Asas Pendidikan, UTM.

Ahmad, A., Idrus, S. M., Malik, N. N. N.A., Murad, N. A., Ngajikin, N. H. & Esa, R. M. (2006). Pengaruh Gaya Hidup Pelajar Kejuruteraan di Kolej Kediaman UTM Terhadap Pencapaian Akademik *National Student Development Conference (NASDEC)* 2006, 8-9 August 2006, Kuala Lumpur, Malaysia.

Adams, J. (2000). Work Experience as a Predictor of MBA Performance. *College Student Journal.* June, 2000.

A step in the right direction. (2007, November 7). *The Star Online*. Retrieved on January 15, 2008 from http://www.thestar.com.

Behrouz, M., Kashy, D. A., Kortemeyer, G. & Punch, W. F. (2003). Predicting Student Performance: An Application of Data Mining with the Educational Web-based System Lon-Capa. *33rd ASEE/IEEE Frontiers in Education Conference*. November 5-8, Boulder.

Belanger, F., Lewis, T., Kasper G. M., Smith, W. J., & Harrington, K. V. (2007). Are Computing Student Different? An Analysis of Coping Strategies and Emotional Intelligence. *IEEE Transactions on Education*, 50 (3), August 2007.

Be wary of corruption. (2007, December 7). *The Star Online*. Retrieved on January 15, 2008 from http://www.thestar.com

Byrne, P. & Lyons, G. (2001). The effects of students attributes on success in programming. *ACM SIGCSE Bulletin*, Vol (33)- 3, September 2001, pg 49 – 52.

Camara, W. J. & Echternacht, G. (2000). *The SAT 1 and High School Grades: Utility in Predicting Success in College*. College Boarg Research Report No. RN-10. New York: Research and Development.

Carrol, B. R. (2006). *The effects of school violence and crime on academic achievement*. Master dissertation. Davidson College.

Close, W. & Solberg, S. (2007). Predicting Achievement, distress, and retention among lower-income Latino youth. *Journal of Vocational Bahavior*. Retreived January 20, 2008 from Science Direct database.

Chamorro-Premuzic, T. & Furnham, A. (2003). Personality Predicts Academic Performance: Evidence from two longitudinal university samples. *Research in Personality*, 37, 319 – 338.

Cripps, A. (1996). Using Artificial Neural Nets to Predict Academic Performance.

Darmayanti, T. (1994). Readines for Self-Directed Learning and Achievement of the Students of Universitas Terbuka (The Indonesian Open Learning University). *Published Master Thesis*. Michigan: UMI Dissertation Services.

DeBerard, M. S, Spielmans, G. I. & Julka, D. L. (2004). *Predictus of academic achievement and retention among college freshmen: A longitudinal study*. College Student Journal, 38 (1), 66 – 80.

Ditcher, A. & Tetley, J (1999). Factors influencing university students' academic success: what do students and academics think?. *HERDSA Annual International Conference*, Melbourne, 12-15 July.

Early checks on student options. (2008, January 14). *The Star Online*. Retrieved on January 15, 2008 from http://www.thestar.com

Exploit Your Potential, Grads told. (2007, December 4). *The Star Online*. Retrieved on January 15, 2008 from http://www.thestar.com

Facey-Shaw, L. & Golding, P. (2005). Effects of peer tutoring and attitude on academic performance of first year introductory programming students. 35th ASEE/IEEE Frontiers in Education Conference, October 19 – 22, Indianapolis, IN.

Golding, P. & Donaldson, O. (2006). Predicting Academic Performance. *36th ASEE/IEEE Frontiers n Education Conference*, October 28-31, 2006, San Diego, CA.

Golding, P. & McNamarah, S. (2005). Predicting Academic Performance in the School of Computing & Information Technology (SCIT). *35th ASEE/IEEE Frontiers in Education Conference*. October 19-22, Indianapolis.

Graduate Careers Council of Australia. Retrieved on March 3, 2008 from http://www.graduatecareers.com.au/content/view/full/3137

Han, J. & Chang, K. C. (2002). Data Mining for web Intelligence. Computer, vol. 35 (11), 64 – 70.

Harding, T. S., Passow, H. J., Carpenter, D. D. & Finelli, C. J. (2003). An examination of the relationship between academic dishonesty and professional behavior. *33rd ASEE/IEEE Frontiers in Education Conference*, November 5-8, Boulder, CO.

Hendrich, U. G. (2007). *The prediction of first-year academic achievement at a South African University of Technology: The importance of English Language Proficiency*. Unpublished paper.

Herzog, S. (2006). Estimating Student Retention and Degree-Completion Time: Decision Tree and Neural Networks Vis-à-vis Regression. *New Direction of Institutional Research*, Fall 2006.San Francisco, CA: Jossey: Bass.

Hii, C. (2007, December 23). Distance Learning a boon for busy working adults. *The Star Online*.

Ho, D. Y. F. & Spinks, J. A. (2004). *Multivariate prediction of academic performance by Hong Kong University students*. Contemporary Educational Psychology, 10 (3), 249 -259.

Hoefer, P. & Gould, J. (2000). *Assessment of admission criteria for predicting students' academic performance. Journal of Education for Business*, 75(4), 225.

Johnson, M. D. (2005). *Academic Performance of Transfer versus "Native" students in natural resources and sciences*. College Student Journal, 39 (3), 570 – 579.

Junemann, M. A. P., Lagos, P. A. S., & Arriagada, R. C. (2007). Neural Networks to predict Schooling Failure/ Success. In Junemann, M. A. P., Lagos, P. A. S., & Arriagada, R. C., *Nature Inspired problem- Solving Methods n Knowledge Engineering* (pp. 571 – 579). Berlin/ Heidelberg: Springer.

Kalles, D. & Pierrakeas, C. (2006). Analyzing Student Performance in Distance Learning with Genetic Algorithms and Decision Trees. *Applied Artificial Intelligence,* 20: 655 – 674.

Kamini, R. S. (2008, January 23). Universities: New Courses to Boost Student Intake. *NST Online*. Retrieved on January 15, 2008 from http://www.nst.com.my.

Leslie, K. (2007). *The impact of emotional intelligence on the academic performance of at-risk high school students*. Ph.D. Dissertation. University of the Incarnate Word, Texas, US.

Linacre, J. M. (2001). Data Mining and Rasch Measurement. *Rasch Measurement Transactions*, Fall 2001, 15 (2), 826-7.

Ling, K. S. (2008, January 23). Schools: An Alternative Path. *NST Online*. Retrieved on January 15, 2008 from http://www.nst.com.my.

Luan, J. (2002). Data Mining and Knowledge Management in Higher Education – Potential Applications. *Preceeding at AIR 2002*, Toronto, Canada.

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

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

Lufi, D., Parish-Plass, & Cohen, A. (2003). *Persistence in Higher Education and its relationship to other personality variables*. College Student Journal, 37 (1), 50 - 59.

Mesir, B.,Yatim, J. M., Aziz, A. A., Jamal, M. H. & Abideen, M. Z.(2006). Gaya Hidup Pelajar Cemerlang dalam Persekitaran Kampus.*National Student Development Conference (NASDEC) 2006*, 8-9 August 2006, Kuala Lumpur, Malaysia.

Mohamed, M. (2007, November 14). Plan to shape varsities of world class. *The Star Online*. Retrieved on January 15, 2008 from http://www.thestar.com

Mustapha not happy with English Programmes at varsities. (2008, January 23). *The Star Online*. Retrieved on January 15, 2008 from http://www.thestar.com

Naik, B. (2004). Using Neural Network to Predict MBA student success. *College Student Journal.* March, 2004.

Nguyen, T. N., Jacenek, P. & Haddawy, P. (2007). A Comparative Analysis of Techniques for Predicting Academic Performance. *37th ASEE/IEEE Frontiers in Education Conference*. October 10-13, Milwaukee.

Nik Anis, M. (2007, November 27). PM: Acquire Knowledge. *The Star Online*. Retrieved on January 15, 2008 from http://www.thestar.com

Nunamaker, J. F. & Chen, M. (1990).System Development in Information System Research. *Proceedings of the 23rd Annual Hawaii International Conference on System Science*.

Ogor, E. N. (2007). Student Academic Performance Monitoring and Evaluation Using Data Mining Techniques. *Electronics, Robotics and Automotive Mechanics Conference (CERMA 2007)*. 354-359.

Peters, K. A. (2007). The academic success of students at an HBCU in Maryland. Ph.D thesis. Morgan State University.

Sam, S. (2007). *Academic Success through group work: Linking social skills, achievement, learning behaviors*. Ph.D. Dissertation. George Mason University, Virginia, US.

Sattar, M. A., Lee, C. J. & Caso, R. (2004). Work in progress: *The Relationship between Objective and /or Subjective Admission Criteria and success Outcome in Undergraduate*

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

Stacy, K. (2007). Academic success: *Are virtual high schools working in Georgia?* Ph. D. Dissertation, Capella University, Minnessota, US.

Superby, J. F. & Vandamme, J. P. (2006). Determination of Factors influencing the achievement of the first-year university students using data mining methods. *Proceedings of the Workshop on Educational Data Mining at the 8th International Conference on Intelligent Tutoring Systems (ITS 2006)*. Jhongli, Taiwan.

Suziena, U. N. (2008, January 23). Higher Education: Incentives for postgraduates. *NST Online*. Retrieved on January 15, 2008 from http://www.nst.com.my.

The Manage Mentor . *Demand Forecating using Qualitative & Quantitative methods*. Retreived January 26, 2008, from

http://www.themanagementor.com/EnlightenMentorAreas/finance/mgr/DemandForecastQualitativeQuantitative.htm

The Spur of Poor Grades. (2008, January 23). *NST Online*. Retrieved on January 15, 2008 from http://www.nst.com.my.

Thomas, E. H. & Galambos, N. (2004). What Satisfies Students? Mining Student-Opinion Data with Regression and Decision Tree Analysis. *Research in Higher Education.*, 45 (3), 251-269.

Ullah, H. & Wilson, M. A. (2007). *Students' Academic Success and Its Association to Student Involvement with Learning and Relationships with Faculty and Peers. College Student Journal*; Dec 2007 Part B, 41 (4), 1192 – 1202.

Yahaya, A., Boon, Y., Hashim, S. & Hamid, W. Z. W. (2002). Hubungan Di Antara Gaya Pembelajaran Dengan Pencapaian Akademik Pelajar. *International Conference on The challenge on Learning and Teaching in Brave New World". Hatyai City College Thailand and School of Cognitive and Education*, UUM,14-16 October 2002.

Yue, Y. & Scutter, M. (2000). Quantitative decision making for cued land reconnaissance: A defense case study. *International Transactions in Operational Research*, 7 (2000), 449 – 463.