UNIVERSITI PUTRA MALAYSIA

ESTIMATION OF EXAMINEES' ABILITY THROUGH COMPUTER ADAPTIVE TESTING BASED ON NEURAL NETWORK APPROACH

AZAM KAZEMI

FSKTM 2010 13

ESTIMATION OF EXAMINEES' ABILITY THROUGH COMPUTER ADAPTIVE TESTING BASED ON NEURAL NETWORK APPROACH

By

AZAM KAZEMI

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Master of Science

December 2010

DEDICATION

To

My Dear husband for his encouragement,

My parents who devoted their life to their children,

My lovely daughter, Bahar

And to all my friends in University Putra Malaysia

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfillment of the requirement for the degree of Master of Science

ESTIMATION OF EXAMINEES' ABILITY THROUGH COMPUTER ADAPTIVE TESTING BASED ON NEURAL NETWORK APPROACH

By

AZAM KAZEMI

December 2010

Chairman: Masrah Azrifah Azmi Murad, PhD

Faculty: Computer Science and Information Technology

Examinee's knowledge is measured through exams. A key purpose of using an exam is to determine the proficiency level of each examinee based on his/her responses to the administered test. A main problem of traditional test is that the asked questions are not match to actual ability of examinees and doesn't measure examinee's proficiency accurately.

Computer Adaptive Testing (CAT) has been developed to address this issue. In CAT, each examinee has to answer the questions that are tailored to his/her ability level. Some of the features such as test security, immediate score reporting, improved efficiency and measurement precision have increased popularity of CAT. It uses models of proficiency estimation such as Item Response Theory (IRT). It is a statistical method with theoretical foundation that is being widely used in the field of modern educational testing technology and psychological testing. However, this model has some drawbacks.

IRT model relates the response of an examinee to a specific item to his/her ability level and characteristics of the item. But relationship between items characteristics and person's skill are very complex and nonlinear. In addition, it relies on strict assumption and need a large amount of data to precise measurement. These limitations are the motivation behind this research to use other adaptive approach to estimate the proficiency level in the CAT.

In this thesis, we proposed a novel solution based on Artificial Neural Network (ANN) to address the above mentioned limitations. The ANN with adaptive features is a suitable scheme for solving complex non-linear problems. In addition, it has the ability to learn and generalize. These strong potentials make it an appropriate method to measure proficiency level of examinees in CAT systems.

This work has been organized in two phases. In the first phase, we use 3-PL (three parameter logistic) dichotomous and polytomous model of IRT to estimate examinees' ability in adaptive testing. Statistical approaches such as maximum likelihood estimation method and Bayesian approach are used for this purpose. In the second phase, estimation of examinees' ability has been obtained with multi-layer feed forward neural network with back propagation algorithm. Experiments have been repeated under different scenarios and results indicate the advantages of the proposed scheme by obtaining better accuracy in performance.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia
sebagai memenuhi keperluan untuk ijazah Master Sains

PENGANGGARAN KEBOLEHAN CALON PEPERIKSAAN MELALUI PENGUJIAN MUDAH SUAI BERKOMPUTER BERDASARKAN PENDEKATAN RANGKAIAN NEURAL

Oleh

AZAM KAZEMI

Disember 2010

Pengerusi: Masrah Azrifah Azmi Murad, PhD

Fakulti: Sains Komputer dan Teknologi Maklumat

Pengetahuan seseorang calon selalunya diuji menerusi peperiksaan. Tujuan utama peperiksaan adalah untuk menentukan tahap kemahiran setiap calon berdasarkan jawapan-jawapan yang diberi dalam sesuatu ujian. Masalah utama yang dihadapi oleh ujian tradisional adalah soalan-soalan yang ditanya selalunya tidak berpadanan dengan keupayaan sebenar calon dan tidak dapat mengukur tahap kemahiran calon dengan tepat.

Pengujian mudah suai berkomputer (CAT) telah dibangunkan bagi menyelesaikan isu ini. Dengan menggunakan CAT, setiap calon peperiksaan hanya perlu menjawab soalan-soalan yang berpadanan dengan tahap keupayaannya. Ciri-ciri seperti keselamatan ujian, laporan markah secara segera, peningkatan kecekapan dan ketepatan penilaian telah menambah popularitи CAT.

CAT menggunakan model-model penganggaran kemahiran seperti *Item Response Theory* (IRT). Ia merupakan satu kaedah statistik dengan teori asas yang telah digunakan secara meluas di dalam bidang teknologi pengujian pendidikan moden dan pengujian psikologi. Walau bagaimanapun, model ini mempunyai beberapa kelemahan.

Model IRT akan mengaitkan jawapan daripada calon dengan satu item khusus dari tahap keupayaannya dan ciri-ciri item tersebut. Akan tetapi, hubungan antara ciri-ciri item dan kemahiran seseorang adalah sangat kompleks dan tidak linear. Tambahan lagi, ia bergantung kepada andaian yang sempit dan memerlukan jumlah data yang besar bagi pengukuran yang tepat. Batasan-batasan ini merupakan motivasi di sebalik penyelidikan ini untuk menggunakan pendekatan mudah suai lain bagi menganggar tahap kemahiran di dalam CAT.

Dalam tesis ini, kami mencadangkan satu kaedah novel berdasarkan rangkaian neural buatan (ANN) bagi menyelesaikan batasan-batasan yang dinyatakan. ANN dengan ciri-ciri mudah suai merupakan satu skim yang sesuai bagi menyelesaikan masalah-masalah kompleks dan tidak linear. Selain itu, ia mempunyai keupayaan untuk belajar dan membuat kesimpulan umum. Potensi-potensi ini menjadikannya kaedah yang sesuai bagi mengukur tahap kemahiran calon-calon peperiksaan dalam sistem CAT.

Kajian ini terbahagi kepada dua fasa. Dalam fasa pertama, kami menggunakan 3-PL *(three parameter logistic)* model IRT dikotomi dan politomi bagi menganggar keupayaan calon-calon peperiksaan dalam ujian mudah suai. Pendekatan statistik seperti Maximum Likelihood Estimation dan kaedah Bayesian digunakan untuk tujuan ini.

Manakala dalam fasa kedua, anggaran keupayaan calon-calon peperiksaan telah diperolehi menggunakan rangkaian neural *multi-layer feed forward* dengan algoritma rambatan belakang. Eksperimen-eksperimen telah diulang di bawah senario yang berbeza dan keputusan menunjukkan kelebihan dalam skim yang dicadangkan dengan mencapai kejituan yang lebih baik dalam prestasi.

ACKNOWLEDGEMENTS

In the name of GOD, the most gracious and merciful. I thank GOD for giving me the patient and strength to complete this research work. After that, I would like to thank my advisor, Dr. Masrah Azrifah Azmi Murad for kindly providing suggestions which helped me in all the time of research and writing of this thesis. She showed great patience in improving the quality of my work. Also, I would like to thank my supervisory committee member, Dr. Rodziah Atan. Her precious time and feed-backs are greatly appreciated.

I would like to acknowledge my family. My husband, Mohammad helped me to overcome the many difficulties on the way to completing this research work. I will not forget his supports.

And finally I also give my thanks to my Dear mother and my Dear father. Again, thanks to all of the people who supported me to finish my study at Universiti Putra Malaysia.

I certify that an Examination Committee has met on 24 December 2010 to conduct the final examination of Azam Kazemi on her Master of Science thesis "Estimation of Examinees' Ability through Computer Adaptive Testing based on Neural Network Approach" in accordance with Universiti Pertanian Malaysia (Higher Degree) Act 1980 and Universiti Pertanian Malaysia (Higher Degree) Regulations 1981. The Committee recommends that the candidate be awarded the relevant degree. Members of the Examination Committee are as follows:

Shamala K. Subramaniam, PhD
Senior Lecturer
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairman)

Norwati Mustapha, PhD
Senior Lecturer
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Internal Examiner)

Abu Bakar Md Sultan, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Internal Examiner)

Azuraliza Abu Bakar, PhD
Associate Professor
Faculty of Technology and Information Science
Universiti Kebangsaan Malaysia
(External Examiner)

BUJANG KIM HUAT, PhD
Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

This thesis was submitted to the senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Master of Science. The members of the Supervisory Committee were as follows:

Masrah Azrifah Azmi Murad, **PhD**
Senior Lecturer
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairman)

Rodziah Atan, PhD
Senior Lecturer
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

HASANAH MOHD GHAZALI, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

DECLARATION

I declare that the thesis is my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously and is not concurrently submitted for any other degree at Universiti Putra Malaysia or at any other institution.

_____

AZAM KAZEMI

Date: 24 December 2010

TABLE OF CONTENTS

		Page
DEDICATION		ii
ABSTRACT		iii
ABSTRAK		v
ACKNOWLEDGEMENTS		viii
APPROVAL		ix
DECLARATION		xi
LIST OF FIGURES		xiv
LIST OF TABLES		xvi
LIST OF ABBREVIATIONS		xviii
CHAPTER		
1	**INTRODUCTION**	1
	1.1 Background	1
	1.2 Problem Statement	3
	1.3 Significance of Study	5
	1.4 Research Objectives	6
	1.5 Scope of Study	7
	1.6 Thesis Organization	7
2	**LITERATURE REVIEW**	9
	2.1 Computer Adaptive Testing	9
	2.1.1 Advantages of Computer Adaptive Testing	11
	2.1.2 Components of Computer Adaptive Testing	12
	2.2 Item Response Theory (IRT)	14
	2.2.1 Dichotomous Item Response Theory	17
	2.2.2 Polytomous Item Response Theory	24
	2.3 Estimate Examinee‘s Ability through IRT Model	29
	2.3.1 Maximum Likelihood Estimation (MLE)	31
	2.3.2 Bayesian Method	32
	2.4 Weaknesses of IRT Model	35
	2.5 Artificial Neural Network (ANN)	35
	2.5.1 Topology of Neural Networks	39
	2.5.2 Learning of Neural Networks	41
	2.6 Sample Size of IRT Model	51
	2.7 Artificial Intelligence and Computer Adaptive Testing	53
	2.8 Summary	58
3	**METHODOLOGY**	59
	3.1 Simulation Data	60
	3.2 Estimate Examinee’s Ability through Neural Network	65
	3.2.1 Number of Hidden Layers and Hidden Neurons	66

3.2.2 Transfer Function 67
3.2.3 Data Preprocessing 68
3.2.4 Improving Generalization 69
3.2.5 Neural Network Training 70
3.2.6 Neural Network Testing 74
3.3 Estimate Examinee's Ability based on IRT Model 75
3.4 Evaluation Performance 78
3.5 Summary 79

4 **EXPERIMENTAL RESULT AND ANALYSIS** 80
4.1 Estimation of Examinees' Ability through NN for Data Sample I 80
4.2 Estimation of Examinees' Ability through NN for Data Sample II 85
4.3 Estimation of Examinees' Ability through NN for Data Sample III 88
4.4 Estimation of Examinees' Ability through NN for Data Sample IV 92
4.5 Estimation of Examinees' Ability through NN for Data Sample V 96
4.6 Estimation of Examinees' Ability through NN for Data Sample VI 100
4.7 Estimation Ability by MLE and BEM - Data Sample I 104
4.8 Estimation Ability by MLE and BEM - Data Sample II 105
4.9 Estimation Ability by MLE and BEM - Data Sample III 107
4.10 Estimation Ability by MLE and BEM - Data Sample IV 108
4.11 Estimation Ability by MLE and BEM - Data Sample V 109
4.12 Estimation Ability by MLE and BEM - Data Sample VI 109
4.13 Comparison Results 111
4.14 Summary 113

5 **CONCLUSIONS AND FUTURE WORK** 115

REFERENCES 119
BIODATA OF STUDENT 124
PUBLICATIONS 125