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)

ZULHANA BINTI ZULKIFLE
(86005)

calon untuk Ijazah
(candidate for the degree of) **MSc. (Intelligent System)**

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

HYBRID BLEU ALGORITHM FOR STRUCTURED EXAM MANAGEMENT SYSTEM

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): **PROF. MADYA FADZILLAH SIRAJ**

Tandatangan
(Signature) : _____ Tarikh (Date) : 10/11/08

Nama Penyelia Kedua
(Name of 2nd Supervisor): _____

Tandatangan
(Signature) : _____ Tarikh (Date) : _____

HYBRID BLEU ALGORITHM FOR STRUCTURED EXAM MANAGEMENT SYSTEM

A thesis submitted to the Faculty of Information Technology in partial Fulfillment of the requirement for the degree Master of Science (Intelligent System) Universiti Utara Malaysia

By
ZULHANA BINTI ZULKIFLE

PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirement 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 purposes may be granted by my supervisor or, in her absence, by the dean of the Graduate School. It is also understood that due recognition shall be given to me and Universiti Utara Malaysia for any scholarly use which may be made of any material from my thesis.

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

Dean of Graduate School,
Universiti Utara Malaysia
06010 UUM Sintok
Kedah Darul Aman
Malaysia

ABSTRAK

Pengajaran dan pembelajaran (P & P) merupakan perkara penting dalam sesebuah pendidikan. Dalam menilai sesuatu P & P, pelbagai teknik penilaian boleh dilakukan seperti kuiz, ujian, amali dan sebagainya. Walaubagaimanapun, apabila sesuatu penilaian telah dibuat, pensyarah2 perlu meruangkan masa bagi menyemak penilaian peralatan tulis ketika metersebut, terutama sekali bagi penilaian berbentuk kuiz dan ujian bertulis. Selain itu juga kekangan kerja sampingan turut menyumbang kepada faktor kesibukan untuk menyemak penilaian-penilaian tersebut. Oleh sebab itu satu sistem yang dinamakan "Hybrid BLEU algorithm for Structured Exam Management System" dibangunkan bagi membantu pensyarah-pensyarah dalam menjalankan penilaian yang berbentuk struktur (kuiz atau ujian). Sistem ini mengaplikasikan algorithm BLEU dalam sistem pakar bagi menyemak jawapan dan memberi markah bagi setiap penilaian yang dibuat. Selain itu juga sistem ini menyediakan sistem pakar bagi menyemak kesalahan ejaan ketika memasukkan soalan-soalan ke dalam bank soalan oleh pensyarah dan ketika menjawab penilaian oleh pelajar. Sistem ini telah diujilari bagi modul Cyberpreneurship di Kolej Komuniti Bandar Darulaman. Seramai dua buah kelas yang berkapasiti 30 orang pelajar telah mencuba untuk menjawab ujian menggunakan sistem ini. Berdasarkan ujilari yang telah dibuat, mereka (pelajar) berpendapat sistem ini sangat mudah digunakan. Sistem ini mempunyai antaramuka yang menarik dan mudah difahami oleh para pelajar. Pelajar-pelajar juga tidak perlu membawa sebarang peralatan tulis ketika mengambil peperiksaan. Walaubagaimanapun, mereka berpendapat sistem ini boleh dibuat dalam dwibahasa memandangkan latar belakang Bahasa Inggeris mereka yang kurang memuaskan. Namun dengan adanya sistem pakar bagi menyemak kesalahan ejaan telah banyak membantu mereka dalam memperbetulkan kesilapan ejaan ketika menjawab penilaian tersebut. Para pensyarah pula berpendapat sistem ini boleh dikemaskinikan lagi dengan merujuk format yang disediakan oleh pihak pengurusan. Sistem ini juga membantu para pensyarah menyemak penilaian yang dibuat dengan lebih cepat dan menjimatkan masa. Secara keseluruhannya, sistem ini berjaya bagi membantu sistem penilaian dalam pengajaran dan pembelajaran di Kolej Komuniti.

ABSTRACT

Teaching and learning process is very essential in education. To evaluate teaching and learning process, many techniques can be applied for instance quiz, test, practical and so forth. However, after the assessment has been fulfilled, the lecturers need to spend time to check the assessment especially quiz and written test. Apart from that, other workloads do contribute to this scenario. Due to this problem, "Hybrid BLEU algorithm for Structured Exam Management System" is develop to aid the lecturers during assessment in the construction of quiz and test. This system incorporate algorithm BLEU into the expert system to check the answer and evaluate marks for every assessment. Apart from that, this system consist an expert system to check the spelling error when inserting the question into question bank by the lecturer whilst respond to the student assessment. This system has been conducted for Cyberpreneurship module at Community College Bandar Darulaman. Two classes with thirty students have tested to answer the test using this system. Based on the experimentation, the user (student) concludes that this system is user friendly. It has an attractive interface and easy to be comprehended by students. The students are not required to bring stationary whilst answering the question. Nevertheless, they concluded that this system should be in bilingual in order to take into consideration of the students' proficiency in English language. With the expert system to check the spelling error, can aid them in correcting the error. The lectures believed that this system can be upgraded according to the administration format. As a conclusion, this system has effectively aid the teaching and learning process at Community College.

ACKNOWLEDGEMENTS

First of all, I would like to express my appreciation to Allah, the Most Merciful who has granted me the ability and will power to start and complete this project.

I would also like to express my gratitude to my supervisor, Associate Professor, Fadzilah Siraj, Computer Science Department, Faculty Information Technology at the Northern University of Malaysia for her guidance and advice throughout the completion of this project.

Many thanks to the my friends as a expert for this project from Bandar Darulaman Community College, En Idhwan bin Mohamad, Cik Asmarul Shazila, Pn. Nor Analiza and my diploma students for their helps and supports.

To all my friends that have been helping and supporting me throughout the entire duration of this semester, thank you for your kindness.

And lastly, especially for my beloved husband Mazrul Nizam Mustaffa, my daughters Siti Nur Mardhiah and Siti Nur Fatihah and all my family for your patience and understanding of my responsibilities in completing my post graduate study.

Only God knows everything!

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

This chapter briefly explains the background of the project that mainly involves the modified BLUE Algorithm with expert system to manage the structure management exam implemented at Bandar Darulaman Community College. The problem statement, objectives, significance of the project and scopes are also presented in this chapter.

1.1 BACKGROUND

Teaching and Learning are the most crucial part in the education process. However, the assessment is one of the most common way to evaluate student performance. Most of the existing education assessments such as objective testing exercises, structured question, practical exam and other. However, in order to fully assess the students' learning progress, these should be complemented with open-ended questions (Whittington & Hunt, 1999). It has been noted that assessment based only in multiple -choice, fill-in-the-blank or yes/no questions is not accurate enough to measure the amount of knowledge the students have acquired, or whether they have understood the subject, Therefore, the field called Computer-Assisted Assessment

The contents of the thesis is for internal user only

- [illegible] Processing (NLP) techniques with the system

  [illegible] ability of the system in managing textual answers.

REFERENCES

Abney, S. (1996), Part-of-Speech Tagging and Partial Parsing, in Ken Church, Steve Young, and G[illegible], eds. Corpus-Based Methods in Language and Speech, Kluw[illegible] Publishers, Dordrecht.

Bloom, B. (1956), Taxonomy of educational objectives: The classification of educational goals. Handbook I, cognitive domain. New York ; Toronto: Longmans, Green.

Brill, E., Mooney, [illegible]w of Empirical Natural Language Processing, *AI Magazine*. [illegible]3-24, 1997.

Burstein, J. and M[illegible], Toward using text summarization for essay-based feedback, [illegible] *2000*.

Burstein, J., Kukich, [illegible], S., Lu, C., Chodorow, M., Bradenharder, L. and Harris, M. D[illegible] scoring using a hybrid feature identification technique. *In [illegible]e Annual Meeting of the Association of Computation[illegible]*.

Burstein. J, Leacock, [illegible], (2001). Automated evaluation of essays and short answe[illegible] *Proceedings of the International CAA Conference*.

Callan, J., Croft, [illegible], (1995). Trec and tipster experiments with inquery. Inf[illegible]l Management.

Carro, R. M., Pul[illegible]ez, P. (1999). Dynamic generation of adaptive internet[illegible] *of Network and Computer Applications, 22 (4),* 249-257.

Chung, G. an[illegible]thodological approaches to online scoring of essays. [illegible] *UCLA. National Center for Research on Evaluation,* [illegible] *and Testing*.

Conlon, G. [illegible]es of writing ability, New York: Longman, chapter [illegible]s and strategies, pp. 109–125.

Cowie, J. and [illegible]mation extraction. *In Communications of the ACM* 3[illegible]

Cristea, D. Cris[illegible] and I. Pistol. (2005). Summarisation through discour[illegible] *of CICLING 2005*.

D. Callear, J. [illegible] & V. [illegible]h. (2001). CAA of short non-MCQ answers. *Proceed[illegible] [illegible]nter[illegible]nal CAA conference*, Loughborough, UK.

D. Callear, J. [illegible]h, [illegible] V. Soh. (2001). "CAA of short non-MCQ answers". [illegible]ng [illegible] *the 5th International CAA conference*, Loughbor[illegible]

Dagan, I., & [illegible] [illegible] P. [illegible]5). Committee-based sampling for training probab[illegible] *In p[illegible]dings of the Twelfth International Conference on Maci[illegible]* p. 1[illegible]57 San Francisco, CA. Morgan Kaufman.

Delmonte, P. [illegible] [illegible]g with an LFG-based lexicon and conceptual repres[illegible] [illegible] *Humanities, 24* (5-6), 461-488.

Delmonte, R. [illegible] [illegible] Determining essential properties of linguistic objec[illegible] [illegible]aphora resolution. In *proceedings of the Worksh[illegible] in [illegible]rse* (pp. 10-24). Pisa.

Delmonte, P. [illegible] [illegible]t, E. (1992). GETA_RUN—A general text analyz[illegible] [illegible]ding. *In proceedings of the 3rd conference on applied [illegible] [illegible]processing—Association for Computational Linguist[illegible] [illegible]ons*, Trento, 9-10: ACL.

Delmonte. P [illegible] 19[illegible] [illegible]alogues from texts: How to generate answers from [illegible]. In [illegible]'onvegno Nazionale Associazione Italiana di *Intellige[illegible] AI*[illegible]* p. 139-143). Padova: AI*IA

Denton, P. [illegible] [illegible] 'electronic feedback' marking assistant and analysis [illegible] [illegible]letection facility. *In proceedings of the 7th Compu[illegible] [illegible]'onference.*

Dessus, P., [illegible] Ver[illegible] A. (2000). Free text assessment in a virtual campus. [illegible] *7th International Conference on Human System Learni[illegible]*

Dessus, P., [illegible] [illegible] A. (2000). Free text assessment in a virtual campus. [illegible] *'d International Conference on Human System Learni[illegible]*

Dessus, P., [illegible] V[illegible] A. (2000). Free text assessment in a virtual campu[illegible] [illegible] *of [illegible]'d International Conference on Human System Learni[illegible]*

Diana Pérez & [illegible] [illegible] [illegible]005). Application of the BLEU Algorithm for Recog[illegible] [illegible]nts. *In proceedings of the First PASCAL Recog[illegible] [illegible] [illegible]'hallenge.*

Doddington, G. ... mati... luation of Machine Translation Quality Using N-gram ... e Sta... s. *Paper presented at the Human Language Technol... ..., Ma... 24-27, 2002, San Diego, CA, USA.

Foltz, P., La... Lan... T. (1999). The intelligent essay assessor: Applica... tion:... hnology. *Interactive Multimedia Electronic Journal... ...har... earning.*

Foltz, P.W.. ... Lar... T.K. (1999) Automated Essay Scoring : Applic... ation:... chnology. *Proceedings of ED-MEDIA '99 Conferenc... ...arlott... le, USA.*

G. Puscasu ... ewo... temporal resolution. *In proceedings of the Langu... ...d E... on Conference (LREC-2004), 2004.*

Graesser, P. ... gs, K... mer- Hastings, D. Harter, Tutoring Research Group. ... (20... ing latent semantic analysis to evaluate the contrib... nts i... tutor. *Interactive Learning Environments,* 8(2):12...

Hearst, M. ... ate... omated essay grading, In: IEEE Intelligent System... ...-37.

K. Papineni, ... Ward ... Zhu (2001). BLEU: a method for automatic evalu... ... trans... *technical Report RC22176 (W0109- 022),* IBM ... ...n, Th... J. Watson Research Center (2001).

Kanejiya, D. ... ... Pra... (2003). Automatic Evaluation of Students' Answ... ...tical... anced LSA. *Human Language Technology Confer... ...NAACL ...03) Workshop on Building Educational Appli... ...Ed... Canada.*

Kintsch, D. ... ...ahl, ... LSA Research Group. (2000). Developing summa... ... throu... he use of lsa-based feedback. *Interactive Learni... ... 8(... ...09.*

Kishore Pa... ...ukos... Ward, & Wei-Jing Zhu. (2002). BLEU: A Meth... ... Eva... of Machine Translation. *In proceedings of the 40... ...g of...*

Knowles, J. ... ...et 3... ing and teaching on the web: Computer Assis...

Larkey, L. S. [illegible] grading using text categorization techniques. *In Proc[illegible] 21st [illegible] International ACM SIGIR Conference on Resear[illegible] [illegible]mation Retrieval'*, pp. 90–95.

Lewis, D. [illegible] J. ([illegible] Heterogeneous uncertainty sampling for supervi[illegible] *[illegible] of the Eleventh International Conference on Mach[illegible]* pp. 1-[illegible] San Francisco, CA. Morgan Kaufman.

Liere, R., [illegible] ([illegible] ive learning with committees for text categ[illegible] *[illegible] the Fourteenth National Conference on Artifici[illegible]* p. 5[illegible] Providence,RI.

Little, J. (20[illegible] [illegible] of essays. Term Paper, Expert Systems.

M. Glass. [illegible] Systems for Tutorial Applications. *In proces[illegible] [illegible]* CIRCSIM-Tutor Intelligent System. *AAAI 2000 F[illegible]*

Mason, O. [illegible] [illegible]02). Automated free text marking with paperle[illegible] *[illegible] the Sixth International Computer Assisted Asses[illegible]* [illegible] University, Loughborough, UK.

Mikhailov, [illegible] [illegible]fficient Engineering Systems Through Artific[illegible] [illegible]

Ming, P.Y.. [illegible] , T.L. (2000). Intelligent essay marking syste[illegible] *[illegible]ogether,Feb. 2000*, Polytechnic, Singapore.

Ming, Y., M[illegible] [illegible]00). Intelligent essay marking system. *In Learne[illegible]*

Mitchell, T.. [illegible] and N. Aldridge (2002). Towards Robust Comp[illegible] [illegible]xt Responses. *In presented to the CAA Confe[illegible]* [illegible]*rsity.*

Nunamaker, [illegible] [illegible]991). System Development in Information Syste[illegible] [illegible]*gement Information Systems, 7:3,* pp. 89 – 106.

O. Postolac[illegible] [illegible]oreference model on excerpt from a novel. *In pr[illegible]* *Summer School in Logic Language and Inform[illegible]* [illegible]. France.

Palmer, J., [illegible] [illegible]. (200[illegible]), Automated essay grading system applied [illegible] [illegible] subject - how can we do it better?, *In Inform[illegible]* [illegible]*e* – Where Parallels Interesect.

Perez, E. Al... ...2004), Application of the Bleu method for evaluat... ...e-learning environment. *In proceedings of LREC-2...*

Porter, M. (19... ...ix stripping. Program 14(3), 130–137.

Ros'e, C., Ro... ...ehn, K. (2003). 'A hybrid text classification approa... ...ays' Build Educational Applications Using Nature... ...-75.

Rudner, L. a... ...A... ...ated essay scoring using bayes' theorem, in 'Proce... ...of the National Council on Measurement in Educati...

Rudner, L. ... ...ted essay scoring using bayes' theorem. *In proce... ...f the National Council on Measurement in Educa...*

Rudner, L.M... ...ated essay scoring using Bayes' Theorem. *The Jo... ...g and Assessment*, 1(2), 3-21.

S. Valenti, ...elli, An overview of current research on autom... ...of Information Technology Education, 2:319–...

Sealey, C., ...D. (2003), At the coal face. experiences of compu... ...*eding... of the 7th Computer Assisted Assess...*

Streeter, L... ...L. MacCuish (2003). The credible grading machi... ...in the note.

Sukkarieh, ...Pulman N. (2003). .Auto-marking: using compu... ...free text responses. *In paper presented at the ... International Association for Educational Assess...*

Sukkarieh. ... ...Raikes N. (2003). Auto-marking: using compu... ...short, free text responses. *In paper presented at the ... International Association for Educational Asses...*

T. Mitchell. ... ...& N. Aldridge (2002). Towards robust compu... ...ses. *In proceedings of the 6th CAA Conf...*

Thomas, D. [illegible] Petre (2004). E-Assessment using Latent Seman[illegible] Co[illegible] Science Domain: A Pilot Study. *In Proc. of the eLe[illegible] [illegible]nguistics and Computational Linguistics for eLearn[illegible] [illegible]04.* pp 38- 44, Geneva, 2004 [THD04].

Valenti, F. [illegible] 003). An overview of current research on automa[illegible] *l of Information Technology Education,* 2:319–

Vicedo & [illegible] [illegible]e of pronominal anaphora resolution to questi[illegible] *[illegible] of the 38th Annual Meeting of the Associ[illegible] [illegible]stics (ACL),* pages 555–562, 2000.

Whittingdo[illegible] [illegible]aches to the Computerised Assessment of Free-[illegible] *[illegible]rnational Computer Assisted Assessment Confe[illegible] [illegible]e June 1999.*

Wiemer-Hastings, [illegible] [illegible]n (2005) Initial Results and Mixed Direc[illegible] [illegible]NG.

Williams, R. [illegible] [illegible]natically Grading Essays with Markit©. *In proce[illegible] [illegible]nce [illegible]04 Conference*, Rockhampton, Quee[illegible]