

**PENENTUAN KERELEVANAN DOKUMEN MENGGUNAKAN
RANGKAIAN RAMBATAN BALIK**

Tesis ini dikemukakan kepada Sekolah Siswazah bagi
memenuhi keperluan Ijazah Sarjana Sains (Teknologi Maklumat),
Universiti Utara Malaysia

oleh

Fadhilah Binti Mat Yamin

KEBENARAN PENGGUNAAN

Tujuan utama penghasilan tesis ini adalah bagi memenuhi syarat pengijazahan Ijazah Sarjana Universiti Utara Malaysia. Dengan ini, saya bersetuju untuk membenarkan tesis saya ini dibuat penelitian atau kajian oleh pihak Perpustakaan Sultanah Bahiyah Universiti Utara Malaysia. Dengan persetujuan ini juga saya bersetuju untuk membenarkan tesis saya ini dibuat salinan dalam apa jua bentuk sekalipun, secara keseluruhan ataupun secara sebahagian sahaja bagi tujuan akademik semata-mata tetapi dengan syarat mestilah disusuli dengan mendapatkan terlebih dahulu persetujuan daripada penasihat tesis ini dan sekiranya tiada, juga boleh diperolehi daripada Dekan Sekolah Siswazah. Adalah difahamkan bahawa segala bentuk salinan atau cetakan semula tesis ini bagi tujuan mengaut keuntungan semata-mata adalah tidak dibenarkan sama sekali tanpa mendapatkan persetujuan bertulis daripada saya terlebih dahulu. Adalah difahamkan juga, segala bentuk pengiktirafan yang bakal diterima daripada tesis ini berhak diberi kepada saya dan juga Universiti Utara Malaysia tidak kira sama ada ia melibatkan penggunaan hanya sebahagian sahaja atau keseluruhan daripada isi kandungan tesis ini.

Segala bentuk permohonan untuk mendapatkan keizinan menyalin secara keseluruhan atau sebahagian maklumat yang terkandung di dalam tesis ini haruslah dialamatkan kepada :

Dekan Sekolah Siswazah
Universiti Utara Malaysia
06010 UUM Sintok
Kedah Darul Aman

ABSTRAK

Sistem capaian maklumat ialah salah satu daripada cabang Sains Komputer yang berkaitan dengan proses pencapaian dokumen yang relevan daripada pangkalan data. Banyak enjin carian yang telah dibangunkan untuk membantu pengguna mendapatkan maklumat yang relevan dari Internet. Walau bagaimanapun, keupayaan enjin carian maklumat yang sedia ada masih lagi terbatas ekoran daripada masalah timbunan maklumat serta pemulangan dokumen yang kurang relevan dengan keperluan pengguna. Oleh itu, penyelidikan ini bertujuan untuk mengkaji salah satu daripada teknik Kepintaran Buatan iaitu Rangkaian Neural bagi mengukur kerelevanan sesuatu dokumen berbanding kata kunci pengguna. Pembelajaran Rangkaian Neural yang digunakan dalam kajian ini ialah berdasarkan kepada algoritma pembelajaran rambatan balik. Penyelidikan ini melibatkan beberapa fasa pelaksanaan iaitu pemilihan pembolehubah, implementasi rangkaian neural, pemilihan parameter rangkaian neural dan pembangunan prototaip sebuah enjin carian ringkas. Bagi tujuan penilaian, 53 dokumen telah dimuat naik ke dalam pangkalan dokumen. Dokumen tersebut diperolehi daripada halaman web Persidangan Antarabangsa Ketujuh World Wide Web. Dokumen tersebut seterusnya diuji dengan menggunakan dua kata kunci berbeza iaitu '*metadata*' dan '*multimedia*'. Ujian terhadap kata kunci '*metadata*' menunjukkan capaian semula adalah 100 peratus dan ketepatan capaian adalah 50 peratus. Manakala, ujian terhadap kata kunci '*multimedia*' menunjukkan capaian semula adalah 75 peratus dan ketepatan capaian adalah 60 peratus. Keputusan tersebut menunjukkan pendekatan Rangkaian Neural telah berjaya menghasilkan capaian semula yang tinggi. Hasil capaian tersebut juga diuji menggunakan pengukuran *fall-out* dan anggaran umum. *Fall-out* bagi kedua-dua kata kunci pula adalah masing-masing 6 dan 5.666 peratus manakala anggaran umum pula adalah masing-masing 4.08 dan 7.54 peratus.

ABSTRACT

Information retrieval (IR) is one of the Computer Science branches that deals with accessing relevant information from a database. Several search engines have been developed to assist users in retrieving the relevant information from the Internet. However, due to information overload, some search engines are still incapable of returning only the most relevant documents to the users. Hence, this research aims to explore the use of Artificial Intelligence (AI) technique, particularly neural network (NN) in measuring the relevancy of each document compared to the users requests. Backpropagation learning algorithm has been used as a basis for learning in this study. Several phases are involved, namely as the identification of the document's attributes, implementation of NN, identification of NN parameters and development of simple search engine prototype. 53 documents have been uploaded into the database for evaluation purpose. These documents have been downloaded from the Seventh International World Wide Web Conference. The documents are then used to test with two different queries; '*metadata*' and '*multimedia*'. A test for '*metadata*' query achieved 100 percent recall and 50 percent precision. Whereas, the test for '*multimedia*' query achieved 75 percent recall and 60 percent precision. The result shows that the usage of NN approaches has produced a high recall. The result is also tested using fallout and generality measurement. Fallout for both queries are 6 and 5.666 percent respectively. Whereas, the generality for both queries are 4.08 and 7.54 respectively.

PENGHARGAAN

Alhamdulillah, syukur ke hadrat Illahi kerana dengan limpah dan rahmatNya dapat saya menyiapkan tesis yang bertajuk "*Penentuan Kerelevanan Dokumen Menggunakan Rangkaian Rambatan Balik*" bagi memenuhi syarat pengijazahan Ijazah Sarjana Sains Teknologi Maklumat melalui penyelidikan.

Setinggi penghargaan ditujukan kepada penyelia saya iaitu Prof. Madya Fadzilah Siraj dan Prof. Madya Dr. Wan Rozaini Sheikh Osman, di atas tunjuk ajar serta masa yang diluahkan untuk sesi perbincangan.

Jutaan terima kasih kepada Puan Azizah Ahmad selaku Pengerusi Ijazah Lanjutan Sekolah Teknologi Maklumat dan kakitangan Sekolah Siswazah di atas kerjasama yang telah diberikan. Penghargaan ini ditujukan juga kepada Kementerian Sains, Teknologi dan Alam Sekitar (MOSTE) serta Universiti Utara Malaysia (UUM) di atas bantuan kewangan serta elaun yang diberikan sepanjang penyelidikan ini.

Penghargaan dan terima kasih ini juga ditujukan kepada suami, Wan Hussain Wan Ishak dan ibu bapa yang banyak memberikan dorongan dan semangat. Ribuan terima kasih juga diucapkan kepada rakan-rakan seperjuangan di atas sokongan dan buah fikiran yang diberikan.

KANDUNGAN

Abstrak	iii
Abstract	iv
Penghargaan	v
Daftar Jadual	viii
Daftar Rajah	ix
1.0 Pengenalan	1
1.1 Pernyataan Masalah	3
1.2 Objektif Penyelidikan	7
1.3 Skop Penyelidikan	7
1.4 Susun Atur Tesis	8
2.0 Ulasan Karya	9
2.1 Sistem Capaian Maklumat	9
2.2 Isu Kerelevanan Dokumen Dalam Sistem Capaian Maklumat	12
2.3 Rangkaian Neural Dalam Sistem Capaian Maklumat	17
2.3.1 Pembelajaran Rangkaian Neural	20
2.3.2 Aplikasi Rangkaian Neural Dalam Sistem Capaian Maklumat	22
2.4 Rumusan	28
3.0 Metodologi Penyelidikan	30
3.1 Pemilihan dan Perwakilan Pembolehubah	30
3.1.1 Penentuan Pembolehubah	31
3.1.2 Sumber Data	38
3.2 Rangkaian Rambatan Balik	38
3.2.1 Senibina Rangkaian	39
3.2.2 Nilai Awalan dan Pengemaskinian Pemberat	40
3.2.3 Algoritma Pembelajaran	42
3.2.4 Algoritma Aplikasi	47
3.3 Permodelan Rangkaian Rambatan Balik Dalam Penentuan Kerelevanan Dokumen	47
3.3.1 Penentuan Parameter bagi Rangkaian Neural	47
3.3.2 Implementasi Pendekatan	48
3.3.3 Pembangunan Prototaip	50
3.4 Rumusan	57
4.0 Hasil dan Penilaian Penyelidikan	58
4.1 Pengenalan	58
4.2 Parameter Rangkaian Neural	59
4.2.1 Penentuan Unit Tersembunyi	59

4.2.2	Benih Pemberat	60
4.2.3	Kadar Pembelajaran	61
4.2.4	Momentum	62
4.3	Penilaian Kecekapan Capaian Dokumen	62
4.3.1	Sampel Dokumen	64
4.3.2	Pengujian Kata Kunci	65
4.4	Rumusan	67
5.0	Kesimpulan dan Penyelidikan Lanjutan	69
5.1	Faedah dan Kebaikan	69
5.2	Kekangan Pendekatan	70
5.3	Penyelidikan Lanjutan	71
5.4	Sumbangan	73
Bibliografi		74
Lampiran		
Lampiran A:	Perwakilan Dokumen	86
Lampiran B:	Persidangan Antarabangsa ke 7 WWW	92
Lampiran C:	Sampel Dokumen	105
Lampiran D:	Hasil Carian Kata kunci “Metadata”	108
Lampiran E:	Hasil Carian Kata kunci “Multimedia”	109
Lampiran F:	Paparan Antaramuka InForest	110
Lampiran G:	Senarai Pembentangan dan Penerbitan	113

DAFTAR JADUAL

Jadual	Perkara	Halaman
Jadual 3.1	Penggunaan Pembolehubah dalam Penyelidikan yang Berkaitan	32
Jadual 3.2	Nilai setiap Pembolehubah Terpilih Mengikut Kewujudan	38
Jadual 4.1	Pemilihan bagi Unit Tersembunyi	60
Jadual 4.2	Pemilihan bagi Benih Pemberat	61
Jadual 4.3	Pemilihan bagi Kadar Pembelajaran	61
Jadual 4.4	Pemilihan bagi Momentum	62
Jadual 4.5	Penilaian Capaian Dokumen	64
Jadual 4.6	Ukuran Capaian	64
Jadual 4.7	Peratusan Panggil semula dan Ketepatan bagi Kata kunci “metadata”	65
Jadual 4.8	Peratusan Panggil semula dan Ketepatan bagi Kata kunci “multimedia”	66

DAFTAR RAJAH

Rajah	Perkara	Halaman
Rajah 1.1	Jaringan Maklumat	2
Rajah 1.2	Skop Penyelidikan	8
Rajah 2.1	Komponen Asas Sistem Capaian Maklumat	10
Rajah 2.2	Model Asas Neuron	17
Rajah 2.3	Neuron Biologi	18
Rajah 2.4	Fungsi Identiti	19
Rajah 2.5	Fungsi Binary Step	19
Rajah 2.6	Fungsi Binary Sigmoid	19
Rajah 2.7	Fungsi Bipolar Sigmoid	19
Rajah 2.8	Rangkaian Neural Satu Lapisan	21
Rajah 2.9	Rangkaian Neural Berbilang Lapisan	21
Rajah 2.10	Rangkaian Neural dengan Lapisan Kompetitif	22
Rajah 3.1	Pemilihan dan Perwakilan Pembolehubah	31
Rajah 3.2	Contoh URL	33
Rajah 3.3	Rangkaian Rambatan Balik Berbilang Lapisan	40
Rajah 3.4	Perwakilan Pembolehubah ke dalam Rangkaian Rambatan Balik	48
Rajah 3.5	Senibina Sistem	51
Rajah 3.6	Antaramuka Sistem InForest	52
Rajah 3.7	Simulator Rambatan Balik InForest	53
Rajah 3.8	Kod Pseudo bagi fasa Suap Hadapan	54
Rajah 3.9	Kod Pseudo bagi fasa Rambatan Balik Ralat	55
Rajah 3.10	Kod Pseudo bagi fasa Kemaskini Pemberat	56
Rajah 4.1	Ketepatan dan panggil semula bagi Kata kunci “metadata”	66
Rajah 4.2	Ketepatan dan panggil semula bagi Kata kunci “multimedia”	67

BAB 1

Pengenalan

Maklumat merupakan unsur terpenting bagi meningkatkan pengetahuan. Oleh itu, keperluan terhadap maklumat yang berkualiti amat kritikal terutama dalam bidang pengurusan dan akademik. Penggunaan kertas sebagai medium penyimpanan dan penyebaran maklumat menjadi begitu penting ekoran daripada kepesatan pembangunan ekonomi dan sosial masyarakat yang didorong oleh pelbagai faktor seperti perniagaan, pendidikan, penyelidikan dan komunikasi. Walau bagaimanapun, ia merupakan satu amalan tradisional dan melibatkan penggunaan kos dan ruang yang besar.

Perkembangan dan kemajuan dalam penciptaan teknologi seperti komputer telah menghasilkan satu kaedah penyimpanan maklumat secara elektronik. Melalui kaedah ini maklumat dapat disimpan dan dicapai dengan mudah. Selain itu, jumlah atau saiz medium penyimpanan maklumat seperti saiz storan sentiasa dapat disesuaikan dengan keperluan pengguna pada kos yang rendah. Perkembangan ini juga dapat membantu meningkatkan proses penyimpanan dan capaian maklumat serta secara tidak langsung membantu memudahkan penyebaran dan perkongsiannya.

The contents of
the thesis is for
internal user
only

BIBLIOGRAFI

- Abdul Bashah Mat Ali (1996). Internet: Kesan Sosial. *Seminar Internet Anjuran Perbadanan Perpustakaan Awam Kedah*, 19-25.
- Alan, J. (2001). Perspectives on Information Retrieval and Speech. *Dalam Coden, Brown and Srinivasan (Eds.), Information retrieval Techniques for Speech Applications* (ms: 1-10).
- Asnicar, F. A., dan Tasso, C. (1997). ifWeb: A Prototype of User Model-Based Intelligent Agent for Document Filtering and Navigation in the World Wide Web. *Workshop on "Adaptive Systems and User Modeling on the World Wide Web". Sixth International Conference on User Modeling 2-5 June 1997. Sardinia. Dimuat-turun dari <http://www.dimi.uniud.it/~ift/positionp.html>*
- Azman Yasin, Azizi Zakaria, Fadzilah Siraj, Mohamad Shamrie Sainin dan Muhammad Yusof. (2000). Capaian Maklumat Teks Menggunakan Teknik Genetik Algoritma. *Prosiding Persidangan Ulang Tahun ke-30 UKM*, (ms: 158-169) 5-6 Sept. 2000. Universiti Kebangsaan Malaysia.
- Azman Yasin, Mohamed Yusof, Tengku Mohamed Tengku Sembuk, dan Mohd Shahrul Azman Mohd Noah. (2001). An Information Filtering Agent to Learn User's Preferences. *Proceedings of Artificial Intelligence Seminar (AIS) 2001 (CD-ROM)*, Universiti Utara Malaysia, Sintok pada 1-3 Nov. 2001
- Baeza-Yates, R., dan Robeiro-Neto, B. (1999). *Modern Information Retrieval*. New York: ACM Press .
- Balabanovic, M., dan Shoham, Y. (1995). Learning Information Retrieval Agent: Experiments with Automated Web Browsing. *Proceedings of the {AAAI} Spring Symposium on Information Gathering from Heterogenous, Distributed Resources*, 13-18.
- Bhandarkar, A., Chandrasekar, R., Ramani, S. dan Bhatnagar, A. (1989). Intelligent Categorization, Archival and Retrieval of Information. Dalam Siekmann, J. (Ed). *Lecture Notes in Artificial Intelligence*, (ms:309-320), New York, London.
- Billsus, D., dan Pazzani, M. (1996). Revising User Profiles: The Search for Interesting Web Sites. *International Multi-strategy Learning Conference*. Harpers Ferry, VA. Dimuat-turun dari <http://www.ics.uci.edu/~pazzani/Publications/Publications.html>.

- Bloedorn, E., Mani, I., dan MacMillan, R. (1996). Representational Issues in Machine Learning of User Profiles. *Proceedings of the Thirteenth National Conference on Artificial Intelligence (AAAI-96)*, Portland, OR, August, 1996.
- Brake, D. (1997). Lost In Cyberspace. *New Scientist, IPC Megazine Limited*.
- Brin, S., dan Page, L. (1998). The Anatomy of a Large Scale Hypertextual Web Search Engine. *Computer Network and ISDN Systems*. 30 (1-7), 107-117. Dimuat-turun dari <http://www.citeseer.nj.nec.com/brin98anatomy.html>.
- Budzik, J., Hammond, K. J., dan Birnbaum, L. (2001). Information Access in Context. *Knowledge Based System*, 37-53.
- Chen, H. (1995). Machine Learning for Information Retrieval: Neural Networks, Symbolic Learning, and Genetic Algorithms, *JASIS* 46(3), 194-216.
- Cheung, D. W., Kao, B., dan Lee, J. (1998). Discovering user Access Patterns on the World Wide Web. *Knowledge Based System Journal*, 10(7), Elsevier Science, May 1998.
- Cho, J., Garcia-Molina, H., dan Page, L. (1998). Efficient Crawling Through URL Ordering. *WWW7/Computer Networks* 30 (1-7), 161-172.
- Chowdhury, G. G. (1999). *Introduction to Modern Information Retrieval*. Library Associated Publishing.
- Cooley, R., Mobasher, B., dan Srivastava, J. (1997). Web Mining: Information and Pattern Discovery on the World Wide Web. *Proceedings of the 9th International Conference on Tools with Artificial Intelligence (ICTAI'97)*. Dibawa turun dari <http://winter.cs.umn.edu/~cooley/>.
- Crestani, F. (2000). Neural Relevance Feedback for Information Retrieval. Dalam B. Bouchon-Meunier, R.R Yager, dan L. A Zadeh (eds). *Uncertainty in Intelligent and Information Systems*. World Scientific, Singapore.
- Crestani, F. (1993a). An Adaptive Information Retrieval System Based on Neural Networks. Dalam *New Trends in Neural Computation: International Workshop on Artificial Neural Networks*, Sitges, Spain, Junes 1993. Printed in *Lecture Notes in Computer Science*, 686, 732-737, Springer-Verlag.
- Crestani, F. (1993b). Learning Strategies for an Adaptive Information Retrieval System Using Neural Networks. *Proceedings of the 1993 IEEE International Conference on Neural Networks*, 244-249, Sun Francisco, California, USA.
- Crestani, F. (1994). Domain Knowledge Acquisition for Information Retrieval Using Neural Networks. *International Journal of Applied Expert System*. 2(2), 101-115.
- Crestani, F. (1995). Implementation and Evaluation of a Relevance Feedback Device Based on Neural Networks. Dalam *From Natural to Artificial Neural*

Computation: International Workshop on Artificial Neural Networks, Malaga, Spain, Jun 1995. Printed in Lecture Notes in Computer Science, 930, 597-604, Springer-Verlag.

- Croft, W. B. (1995). What Do People Want From Information Retrieval. *D-Lib Magazine*.
- Croft, W. B., Cronen-Twonsend, S., dan Lavrenko, V. (2001). Relevance Feedback and Personalization: A Language Modeling Perspective. *Proceedings of the DELOS-NSF Workshop on Personalization and Recommender Systems in Digital Libraries*, (ms. 49-54).
- Croft, W. B., dan Lewis, D. D. (1987). An Approach to Natural Language Processing for Document Retrieval. Dalam *Tenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. 26-32. New Orleans, LA.
- Cronen-Townsend, S. dan Croft, W. B. (2002). Quantifying Query Ambiguity. In *the Proceedings of HLT2002*, (ms. 94-98). Dimuat-turun dari <http://ciir.cs.umass.edu/pubfiles/ir-241.pdf>
- Dasigi, V., dan Mann, R.C. (1996). Neural Net Learning Issues in Classification of Free Text Documents. *American association for Artificial Intelligence (AAAI) Spring symposium on machine Learning in Information Access*, Stanford, CA.
- Sarjon Defit dan Mohd Nor Md Sap (2001). Organizing and Searching Collections of Textual Documents Using WEBSOM Method. *Artificial Intelligence Seminar (AIS) 2001 (CD ROM)*, Universiti Utara Malaysia, Sintok pada 1-3 Nov.
- Dominich, S. (2000a). Formal Foundation of Information Retrieval. *Technology Letters*, 4(2), 5-12.
- Dominich, S. (2000b). Application, Analysis, and Evaluation of neural network-based Interaction Information Retrieval. Center for Information Retrieval (<http://dcs.vein.hu/CIR/>). Dimuat-turun pada 04 Jun 2002 dari http://dcs.vein.hu/CIR/cikkek/IPM_Dominich.pdf
- Fadhilah Mat Yamin dan Fadzilah Siraj (2001). Searching For Information: Relevancy Measurement. *National Conference on Management Science: New Paradigms for the Knowledge Economy* (19-20 June), Universiti Putra Malaysia, Serdang, Selangor.
- Fadhilah Mat Yamin, Fadzilah Siraj dan Wan Rozaini Sheikh Osman (2001a). Pengukuran Kerelevanan Dokumen Menggunakan Rangkaian Neural. *Artificial Intelligence Seminar (AIS) 2001(CD ROM)*, Universiti Utara Malaysia, Sintok pada 1-3 Nov.

- Fadhilah Mat Yamin, Fadzilah Siraj, Wan Rozaini Sheikh Osman (2001b). Determining The Relevancy Of The Document Using Backpropagation. Dibentangkan di *Seminar Bulanan Kepintaran Buatan 2001 Siri 3/2001*, Sekolah Teknologi Maklumat, Universiti Utara Malaysia, Sintok.
- Faloutsos, C., dan Oard, D. (1996). A Survey of Information Retrieval and Filtering Methods. *Technical Report CS-TR-3514*, Computer Science Department, University of Maryland. Dimuat-turun dari <http://www.cs.umd.edu/library/TRS>.
- Fausett, L. (1994). *Fundamentals of Neural Networks: Architectures, Algorithms and Applications*. Prentice Hall, England Cliffs.
- Feinstein, Y. Z., Goldman, C. V., Mor, Y., dan Rosenschein, J. S. (1997). Relevancy Ranking of Web Pages Using Shallow Parsing. *Proceedings of the Practical Application of Knowledge Discovery and Data Mining (PADD97)*, (ms: 125-136).
- France, R. K., dan Fox, E. A. (1988). An Artificial Intelligence Environment for Information Retrieval Research. Virginia Polytechnic Ins and State University. Dimuat-turun dari http://cs-tr.cs.cornell.edu:80/Dienst/UI...rize/ncstrl.vatech_cs/TR-88-10?abstract. Pada 15 February 2001.
- Giles, L., Bollacker, K. dan Lawrence, S. (1998). CiteSeer: An Autonomous Web Agent for Automatic Retrieval and Identification of Interesting Publications. Dalam *proceeding of the 2nd International ACM Conference on Autonomous Agents*, 116.
- Glover, E. J., Flake, G. W., Lawrence, S., Birmingham, W. P., Kruger, A., Giles, C. L., dan Pennock, D. M. (2001). Improving Category Specific Web Search by Learning Query Modifications. *Symposium on Applications and the Internet, SAINT 2001*, January 8-12, San Diego, California.
- Gravano, L., Chang, C. C. K., Morlina, H. G., Paepcke, A. (1997). STARTS: Stanford Protocol Proposal for Internet Retrieval and Search. Working Papers. SIDL-WP-1996-0043. Stanford Digital Library Technologies. Dimuat-turun dari <http://www.diglib.stanford.edu/WP/WWW/WPTitles.html>.
- Harter, S., P. (1986). *Online Information Retrieval: Concepts, Principles and Techniques*. Library and Information Science Series, ms: 3.
- Hayes, D. (2000). Business Accelerator Launches Research Service on Internet. *TheStar* Dimuat-turun dari <http://www.kcstar.com/item/pages/business.pat,business/37745a3d.330> (13 Januari 2002).
- Haynes, T. (1998). A Comparison of Random Search Versus Genetic Programming as Engine for Collective Adaptation. Dalam V. William Porto(Ed). *Proceedings of the Seventh International Conference on Evolutionary Programming*.

- Hoenkamp, E., Schomaker, L., van Bommel, P., Koster, C. H. A., dan Weide, T.P. (1996). Profile-A Proactive Information Filter. *Technical Note CSI-N9602*, Computing Science Institute, University of Nijmegen, Nijmegen, The Netherlands.
- Hui, B. (1998). Applying NLP to IR: Why and How. *Lecture Notes on "Intelligence Interface"*.
- Huibers, T.W.C., Lalmas, M., dan van-Rijbergen, C.J. (1996). Information Retrieval and Situation Theory. *ACM Press*, 30, 11-25.
- IEI (1999). Neural Network Web Site Placement on the Altavista Search Engine. Imagination Engines Inc. Dimuat-turun dari URL: <http://www.imagination-engines.com/altavista/pparse.htm>. pada 04 September 2001
- Izhar Che Zainol Rashid, Fadzilah Siraj dan Nur Azzah Abu Bakar (2001). Personalized Web Agent Approach for Web Content Mining. *Artificial Intelligence Seminar (AIS) 2001 (CD-ROM)*, Universiti Utara Malaysia, Sintok pada 1-3 Nov.
- Jansen, J. (1996). Using an Intelligent Agent to Enhance Search Engine Performance. *FirstMonday-Peer-Reviewed Journal on the Internet*. Dimuat-turun dari http://www.firstmonday.dk/issues2_3/jansen/ pada 05 Sept. 2001.
- Kaski, S. (1999). Fast Winner Search for SOM-Based Monitoring and Retrieval of High-Dimensional Data. *Proceedings of ICANN99, Ninth International Conference on Artificial Neural Networks*, 2, 940-945, IEE, London.
- Kaski, S., Honkela, T., Lagus, K., dan Kohonen, T. (1996). Creating an Order in Digital Libraries with Self-Organizing Maps. Dalam *Proceedings of WCNN'96, World Congress on Neural Networks*, September 15-18, San Diego, California, 814-817. Lawrence Erlbaum and INNS Press, Mahwah, NJ.
- Keim, M., Lewis, D. D., dan Madigan, D. (1997). Bayesian Information Retrieval: Preliminary Evaluation. Dalam D. Madigan and P. Smyth, (Eds.) *Preliminary Papers of the Sixth international Workshop Artificial Intelligence and Statistic*, 303-310. Ft. Lauderdale, Florida
- Kinsella, P. (1985). *The Techniques of Writing (4th)*. San Diego: Harcourt Brace Jovanovich.
- Kohonen, T. (1997). Exploration of Very Large Databases by Self-Organizing Maps. Dalam *Proceedings of ICNN'97, International Conference on Neural Networks*, pages PL1-PL6. IEEE Service Center, Piscataway, NJ
- Korfhage, R., R. (1997). *Information Storage and Retrieval*. Library of Congress Cataloging in Publication Data. Wiley Computer Publishing.

- Koskela, M., Laaksonen, J., Laakso, S., dan Oja, E. (2000). The PicSOM Retrieval System: Description and Evaluations. Dalam *Proceedings of Challenge of Image Retrieval (CIR 2000)*. Brighton, UK. May 2000. Dimuat-turun dari <http://www.cis.hut.fi/picsom/publications.html>
- Kruschwitz, U. (2000). UKSearch-Web Search with Knowledge Rich Indices. *Proceedings of the AAAI-2000 Workshop on AI for Web Search*, (ms. 41-45), Austin; AAAI Press.
- Laakso, S., Laaksonen, J., Koskela, M., dan Oja, E. (2001). Self-Organizing Maps of Web Link Information. Dalam *Advances in Self-Organising Maps* (eds. N. Allinson, H. Yin, L. Allinson, and J. Slack). June 2001. Dimuat-turun dari <http://www.cis.hut.fi/picsom/publications.html>
- Laaksonen, J., Koskela, M., dan Oja, E. (1999a). PicSOM - A Framework for Content-Based Image Database Retrieval using Self-Organizing Maps. *Proceedings of 11th Scandinavian Conference on Image Analysis (SCIA'99)*. Kangerlussuaq, Greenland. June 1999. Dimuat-turun dari <http://www.cis.hut.fi/picsom/publications.html>
- Laaksonen, J., Koskela, M., dan Oja, E. (1999b). PicSOM: Self-Organizing Maps for Content-Based Image Retrieval. Dalam *Proceedings of IEEE International Joint Conference on Neural Networks (IJCNN'99)*. Washington, DC. July 1999. Dimuat-turun dari <http://www.cis.hut.fi/picsom/publications.html>
- Laaksonen, J., Koskela, M., dan Oja, E. (1999c). Application of Tree Structured Self-Organizing Maps in Content-Based Image Retrieval. *Proceedings of 9th International Conference on Artificial Neural Networks (ICANN99)*. Edinburgh, UK. September 1999 Dimuat-turun dari <http://www.cis.hut.fi/picsom/publications.html>
- Lagus, K. (1997). Map of WSOM'97 Abstracts--Alternative Index. *Proceedings of WSOM'97, Workshop on Self-Organizing Maps*, Espoo, Finland, June 4-6, pages 368-372. Helsinki University of Technology, Neural Networks Research Centre, Espoo, Finland.
- Lagus, K. (2000). Text Retrieval Using Self-Organized Document Maps. *Technical Report A61*, Helsinki University of Technology, Laboratory of Computer and Information Science. ISBN 951-22-5145-0.
- Lagus, K., Honkela, T., Kaski, S., dan Kohonen, T. (1996a). Self-organizing Maps of Document Collections: A New Approach to Interactive Exploration. Dalam Simoudis, E., Han, J., and Fayyad, U., editors, *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, 238-243. AAAI Press, Menlo Park, California.
- Lagus, K., Kaski, S., Honkela, T., and Kohonen, T. (1996b). Browsing Digital Libraries with the Aid of Self-Organizing Maps. *Proceedings of the Fifth International World Wide Web Conference WWW5*, May 6-10, Paris, France, volume Poster Proceedings, pages 71-79. EPGL.

- Larkey, L. S. dan Connell, M. E. (2001). Arabic Information Retrieval at UMass in TREC-10. *Proceedings of the Tenth Text REtrieval Conference (TREC-2001)* Gaithersburg, Maryland, November 13-16, 2001, Dimuat-turun dari <http://trec.nist.gov/pubs.html> pada 25 April 2002
- Lawrence, S. (2000). Context in Web Search. *IEEE Data Engineering Bulletin*, 23(3), 25-32.
- Lawrence, S., dan Giles, C. L. (1998a). Context and Page Analysis for Improved Web Search. *IEEE Internet Computing*, 2(4), 38-46.
- Lawrence, S., dan Giles, C. L. (1998b). Inquirus, the NECI Meta Search Engine. *Proceedings of the 7th International World Wide Web Conference*, Brisbane, Australia, Elsevier Science, ms: 95-105.
- Lesteven, S., Poincot, P., dan Murtagh, F. (1996). Neural Networks and Information Extraction in Astronomical Information Retrieval. *Vistas in Astronomy*, 40, ms: 395-400.
- Leuski, A. dan Allan, J. (2000). Details of Lighthouse. *CIIR Tech. Report IR-212*.
- Liddy, E. D., Paik, W., Yu, E., S., dan McKenna, M. (1995). Document Retrieval Using Linguistic Knowledge. *RAIO '94 Proceedings*. Rockefeller University, New York.
- Lin, C., dan Chen, H. (1994). An Automatic Indexing and Neural Network Approach to Concept Retrieval and Classification of Multilingual (Chinese-English) Documents. Dimuat-turun dari <http://ai.bpa.arizona.edu/papers/chinese93/chinese93.html>
- Loke, S. W., Davison, A., dan Sterling, L. (1996). CiFi: An Intelligent Agent for Citation Finding on the World Wide Web. *Technical Report 96/4*. The University of Melbourne, Australia.
- Mandl, T. (2000). Tolerant and Adaptive Information Retrieval with Neural Networks. Dalam *Global Dialogue. Science and Technology Thinking the Future at Expo 2000 Hannover, 11-13 Julai 2000*, Dimuat-turun dari http://www.shaping-the-future.de/Pages/abstracts/abstract_190.htm pada 19 Mei 2002.
- Manmatha, R., dan Sever, H. (2002). A Formal Approach to Score Normalization for Meta-Search. *Proceedings of HLT 2002*, 88-93. Dimuat-turun dari <http://ciir.cs.umass.edu/> pada: 25 April 2002.
- Marzuki Khalid (1994). Rangkaian Neural Dan Penggunaannya. *Seminar Sains Matematik Dalam Industri MDI'94; Simbiosis Antara Matematik Dengan Industri*, 18-32, Universiti Teknologi Malaysia, Johor.

- McCallum, A. K., Nigam, K., Rennie, J., dan Seymore, K. (2000). Automating the Construction of Internet Portals with Machine Learning. *Information Retrieval Journal*, 3, 127-163. Kluwer Academic Pub.
- McCallum, A. K., Nigam, K., Rennie, J., dan Seymore, K. (1999a). A Machine Learning Approach to Building Domain-Specific Search Engines. *The Sixteenth International Joint Conference of Artificial Intelligence*. Dimuat-turun dari <http://www-2.cs.cmu.edu/~mccallum/papers/cora-ijcai99.ps.gz> pada 04 Jun 2002.
- McCallum, A. K., Nigam, K., Rennie, J., dan Seymore, K. (1999b). Building Domain-Specific Search Engines with Machine Learning Techniques. *AAAI-99 Spring Symposium*. Dimuat-turun dari <http://www-2.cs.cmu.edu/~mccallum/papers/cora-aaaiss99.ps.gz>. pada 04 Jun 2002
- Meadow, C. T., Boyce, B. R., dan Kraft, D. H. (2000). *Text Information Retrieval System (2nd Edition)*. Academy Press.
- Mendelson, A. O., Mihaila, G. A., dan Milo, T. (1997). Querying the WWW. *International Journal on Digital Libraries*. 1(1), 54-67.
- Mizzaro, S. (1996a). How Many Relevances in Information Retrieval? In C. W. Johnson and M. Dunlop (eds), *Proceedings of the Workshop 'Information Retrieval and Human Computer Interaction'*, GIST Technical Report GR96-2, Glasgow University, The British Computer Society, Glasgow, UK, 57-60.
- Mizzaro, S. (1996b). On the Foundations of Information Retrieval. *Proceedings of the Annual Conference AICA'96 , 24-27 September 1996*, (ms. 363-386), Rome, Italy.
- Moukas, A. (1996). Amalthea: Information Discovery and Filtering using a Multiagent Evolving Ecosystem. *Proceedings of the Conference on Practical Applications of Agents and Multiagent Technology*, London.
- Mukherjea, S., Hirata, K., dan Hara, Y. (1997). Towards a Multimedia World Wide Web Information Retrieval Engine. *Sixth International World Wide Web Conference, April 1997*, Santa Clara, CA.
- Muniyandi, R., C. (2000). Neural Networks: An Exploration in Document Retrieval System. *TENCON Proceedings: Intelligent Systems and Technologies for the New Millennium 24-27 Sept. Vol. I*, 156-161.
- Nasariah Mansor, Nafisah Hj. Mahmud, Nuraini Yusoff, Rohizah Halim, Abdull Syukur Shaari dan Mohd. Isha Awang. (2001). *Pengantar Penulisan. Sekolah Sains Kognitif dan Pemikiran*. Universiti Utara Malaysia.
- Noriha Muhammad. (2000). Perubahan Teknik Pencarian Maklumat: Perbandingan Sistem Manual dan Digital di Perpustakaan UKM. *Tesis Ijazah Sarjana Teknologi Maklumat*, UKM, Bangi.

- Nur Izura Hj. Udzir, Md. Nasir Sulaiman, Ali Mamat, Ramlan Mahmud dan Fatimah Ahmad. (1997). Rangkaian Neural Dalam Dapatan Pangkalan Data. *National Conference on Research and Development in Computer Science and its Applications*, 93-97.
- Pannu, A. dan Sycara, K. (1996). Learning Text Filtering Preferences. *1996 AAAI Symposium on Machine Learning and Information Access*, Dibawa turun dari <http://citeseer.nj.nec.com/49897.html>.
- Pao, M., L. (1989). *Concepts of Information Retrieval*. Library of Congress Cataloging –in-Publication Data.
- Pathak, P. (2000). Use of Genetic Algorithms in Information Retrieval: Adapting Matching Function. *Ph.D. Dissertation*. The University of Michigan.
- Pazzani, M., dan Billsus, D. (1997). Learning and Revising User Profiles: The Identification of Interesting Web Sites. *Machine Learning 27*, Kluwer Academic Publisher, 313-331.
- Pazzani, M., Muramatsu, J., dan Billsus, D. (1996). Syskill & Webert: Identifying Interesting Web Sites. *Proceedings of the National Conference on Artificial Intelligent*, Portland, OR.
- Pazzani, M., Nguyen, L., dan Mantik, S. (1995). Learning from Hotlists and Coldlists: towards a WWW Information Filtering and Seeking Agent. *In Proceedings of the Seventh International Conference on Tools with Artificial Intelligence*, Dibawa turun dari <http://www.ics.uci.edu/~pazzani/Coldlist.pdf>.
- Pickens, J. (2001). A Survey of Feature Selection Techniques for Music Information Retrieval. *CIIR Technical Report*.
- Ra, D-Y, Park, E-K dan Jang, J-S. (2001). Yonse/ETRI at Trec 10: Utilizing Web Document Properties. *Proceeding of the Tenth Text REtrieval Conference (TREC 2001)*, NIST.
- Rennie, J., dan McCallum, A. K. (1999). Using Reinforcement Learning to Spider the Web Efficiently. *Proceeding of the Sixteenth International Conference on Machine Learning*. Dimuat-turun dari <http://www-2.cs.cmu.edu/~mccallum/papers/rlspider-icml99.ps.gz> pada 04 Jun 2002.
- Richardson, M., dan Domingos, P. (2002). The Intelligent Surfer: Probabilistic Combination of Link and Content Information in PageRank. Dalam T.G Dietterich, S. Becker dan Z. Ghahramani (eds). *Advances in Neural Information Processing System 14*, Chambridge, MA; MIT Press.
- Rijsbergen, C, J, V. (1979). *Information Retrieval (2nd)*. London, Butterworths.
- Ruthven, I. (1999). Abduction, Explanation and Relevance Feedback. *Searching for Information; Artificial Intelligence and Information Retrieval Approaches*.

- Glasgow, 1999, poster paper. Dimuat-turun dari <http://www.dcs.gla.ac.uk/~igr/Papers/> pada 04 Jun 2002.
- Ruthven, I., dan Lalmas, M (1999). Integrating Information Use Into Relevance Feedback. Departememt of Computer Science, Queen Mary & Westfield College, March 1999. Dimuat-turun dari <http://www.dcs.gla.ac.uk/~igr/Papers/> pada 04 Jun 2002.
- Ruthven, I., dan van-Rijsbergen, C. J. (1996). Context Generation in Information Retrieval. *Florida Artificial Intelligence Research Symposium, Flairs '96*. Key West, April 20-22. Dimuat-turun dari <http://www.dcs.gla.ac.uk/~igr/Papers/> pada 04 Jun 2002.
- Ruthven, I., Lalmas, M., dan van-Rijsbergen, C. J. (1999). Retrieval Through Explanation: An Abductive Inference Approach to Relevance Feedback. *10th Irish Conference on Artificial Intelligence & Cognitive Science. Cork*. Dimuat-turun dari <http://www.dcs.gla.ac.uk/~igr/Papers/> pada 04 Jun 2002.
- Sallehudin Mustaffa (1996). Pengenalan Kepada WWW. *Seminar Internet Anjuran Perbadanan Perpustakaan Awam Kedah, Kedah, 29-38*.
- Saracevic, T. (1996). Modeling Interaction in Information Retrieval (IR): A Review and Proposal. *Proceedings of the American Society for Information Science*, 33, 3-9.
- Sarle, W. S. (1994). Neural Network Implementation in SAS Software. *Proceedings of the nineteenth Annual SAS Users Group International Conference*, (ms: 551-573), Cary, NC: SAS Institutes.
- Shavlik, J., Calcari, S., Rad, T. E., dan Solock, J. (1999). An Instructable, Adaptive, Interface for Discovering and Monitoring Information on the WWW. *Proceedings of the 1999 International Conference on Intelligent User Interfaces*, Redondo Beach (Los Angeles, CA, USA), January 5-8, 1999.
- Shavlik, J., dan Rad, T. E. (1998a). Building Intelligent Agents for Web-Based Tasks: A Theory-Refinement Approach. *Proceedings of the CONALD Workshop on Learning from Text and the Web*, June 1998.
- Shavlik, J., dan Rad, T. E. (1998b). Intelligent Agents for Web-Based Tasks: An Advice-Taking Approach. *Working notes of the AAAI/ICML Workshop on Learning for Text Catagorization*, July 1998.
- SoloSearch (2000a). Mid-West Start-Up Introduces An "Intelligent Search" Manager To Eliminate Internet Information Overload. *Press Release March 28, 2000*. Dimuat-turun dari <http://www.solosearch.com/press.asp#mid>.
- SoloSearch (2000b). Searchhound.Com Merges With Pan International And Solosearch.Com To Create Powerful New Public B2b2c Web Company. *Press release May 15, 2000*. <http://www.solosearch.com/pressSearchHound.asp#mid>

- Spink, A. dan Saracevic, T. (1998). Human-computer Interaction in Information Retrieval: Nature and Manifestations of Feedback. *Interacting with Computers - The Interdisciplinary Journal of Human-Computer Interaction*, 10 (3), 241-267.
- Spink, A., Jansen, B. J., Wolfram, D., dan Saracevic, T. (2002). From E-Sex to E-Commerce: Web Search Changes. *IEEE Computer*, 35(3), 107-109.
- Spink, A., Wolfram, D., Jansen, B. J., dan Saracevic, T. (2001). Searching the Web: The Public and Their Queries. *Journal of the American Society for Information Science and Technology*, 52 (3), 226-234.
- Strzalkowski, T., Guthrie, L., Karlgren, J., Leistensnider, J., Lin, F., Perez-Carballo, J., Straszheim, T., Wang, J., dan Wilding, J. (1996). Natural Language Information Retrieval: TREC-5 Report. Dalam Voorhees, E. M., and Harman, D. K. (Eds.), *Proceedings of the Fifth text retrieval Conference (TREC-5)* 20-22 November, (ms: 291-314).
- Theilmann, W., dan Rothermel, K. (1998). Domain Experts for Information Retrieval in the World Wide Web. *Proc. 2nd Int. Workshop on Cooperative Information Agents (CIA'98)*, Paris, July 4-7, 1998, M. Klusch, G. Weiß (Eds.), *Lecture Notes in Artificial Intelligence 1435*, (ms:216-227), Berlin, Heidelberg, New York: Springer-Verlag.
- Troina, G., dan Walker, N. (1996). Document Classification and Searching: A Neural Networks Approach. *ESA Bulletin Nr. 87*. Dimuat-turun dari <http://esapub.esa.int/bulletin/bullet87/troina87.htm>. Pada 3 Nov 2001.
- Tveter, D. R. (1997). The size of the Network: Backpropagator's Review. Dimuat-turun dari <http://www.dontrever.com/bpr/netsize.html>.
- Vrajitorn, D. (1997). Genetic Algorithms in Information Retrieval. AIDRI97: Learning: From Natural Principles to Artificial Methods., Geneve, June 1997.
- Vrajitorn, D. (2000). Large Population or Many Generations for Genetic Algorithms? Implications in Information Retrieval. Dalam F. Crestani, G. Pasi (Eds.): *Soft Computing in Information Retrieval. Techniques and Applications*, Physica-Verlag, Heidelberg, 199-222.
- Watters, C. R., Shepherd, M. A., dan Robertson, W. (1987). Towards An Experts System for Bibliographic Retrieval: A PROLOG Prototype. *Annual ACM Conference on R & D in IR: Proceedings of the Tenth Annual International ACM SIGIR Conference on R & D in IR*. 272-281. ACM Press.
- Welcher, G. (1998). Optimizing Documents for Search Retrieval : An Introduction to META Information. Dimuat-turun dari http://www.opc.on.ca/infosys/optim_doc.html pada 30 Desember 1999.

- Westerveld, T., Kraaij, W., dan Hiemstra, D. (2001). Retrieving Web Pages using Content, Links, URLs, and Anchors. *Proceedings of The Tenth Text Retrieval Conference (TREC 2001)*, NIST.
- WWW7 Consortium dan IW3C2 (1998). *7th International World Wide Web Conference*. Dimuat-turun dari <http://www7.scu.edu.au/programme/fullprog.html> Pada: 09/3/02
- Yang, J., Pai, P., Honavar, V., dan Miller, L (1998). Mobile Intelligent Agent for Document Classification and Retrieval: A Machine Learning Approach. *Proceedings of the European Symposium on Cybernetics and Systems Research*. Berlin:Springer-Verlag, 707-712.
- Yao, Y. Y., Hamilton, H. J., dan Wang, X. (2000). PagePrompter: An Intelligent Agent for Web Navigation Created Using Data Mining Techniques. *Technical Report TR 2000-08, ISBN 0-7731-0415-1*, Department of Computer Science, University of Regina.
- Zorman, M., Podgorelec, V., dan Kokol, P. (2001). Using Intelligent Search for Finding Medical Sites. *Health Informatics Europe Journal [on-line]*. Dimuat-turun dari http://www.hi-europe.info/files/1998_9/int_med_search.htm pada 3 Mac 2002.