REFERENCES

1. Amy JC Trappey, Fu-Chiang Hsu, Charles V Trappey & ChiaI L in 2006, 'Development of a Patent Document Classification and Search Platform using a Back-Propagation Network', Expert Systems with Applications, vol. 31, pp. 755-765.

2. Antonie, ML & Zaiane, OR 2002, 'Text Document Categorization by Term Association', In Proceedings of IEEE International Conference on Data Mining,pp.19–26.

3. Aronson, AR & Lang, FM 2010, 'An Overview of Meta Map: Historical Perspective and Recent Advances', Historical Article AGROVOC from http://aims.fao.org/standards/agrovoc/about.

4. Aseervatham, S & Bennani, Y 2009, 'Semi-Structured Document Categorization with Semantic Kernel Pattern Recognition', vol.429, pp.2067-2076.

5. Astrakhantsev, NA & Turdakov, DY 2013, 'Automatic Construction and Enrichment of Informal Ontologies', A Survey, Programming and Computer Software, vol.391,pp.34-42.

6. Bai, R, Wang, X & Liao, J 2010, 'Using an Integrated Ontology Database to Categorize Web Pages', Paper Presented at the Proceedings of the International Conference on Advances in Computer Science and Information Technology, Miyazaki, Japan.

7. Bashar Tahayna, Ramesh Kumar Ayyasamy & Saadat Alhashmi 2010, 'A Novel Weighting Scheme for Efficient Document Indexing and Classification', Information Technology ITSim, vol. 2, pp. 783-788.

8. Blanco Fernández, Y, Pazos-Arias, JJ, Gil-Solla, A, Ramos-Cabrer, M, López- Nores, M, García-Duque, J, Caropreso, MF, Matwin, S & Sebastiani, F 2001, 'A learner-Independent Evaluation of the Usefulness of Statistical Phrases for Automated Text Categorization', In A. G. Chin Ed., Text Databases & Document Management pp. 78-102.

9. Blei, DM, Ng, AY & Jordan, MI 2003, 'Latent Dirichlet Allocation', Journal of Machine Learning Research, vol. 3, no. 4-5, pp.993–1022.

10. Bloehdorn, S & Hotho, A 2006, 'Boosting for Text Classification with Semantic Features', Paper Presented at the Proceedings of the 6th International Conference on Knowledge Discovery on the Web: Advances in Web Mining and Web Usage Analysis, Seattle, WA.

11. Bodenreider, O2004, 'The Unified Medical Language System UMLS: Integrating Biomedical Terminology', Nucleic Acids Research, vol. 32, No. 1, pp. D267-D270.

12. Bulskov, H, Knappe, R & Andreasen, T 2002, 'On Measuring Similarity for Conceptual Querying', Proceedings of the 5th International Conference on Flexible Query Answering Systems, Copenhagen, Denmark.

13. Carmen Banea, Di Chen, Radha Mihalcea, Claire Cardie & Janyce Wiebe 2014, 'Using Deep Learning Word Embeddings to Assess Cross-level Similarity', Proceedings of Eighth International Workshop on Semantic Evaluation, pp.560-565, Dublin Ireland.

14. Crossno, PJ, Wilson, AT, Shead, TM & Dunlavy, DM 2011, 'Visually Comparing Topic Models of Text Collections', Paper Presented at the Proceedings of the IEEE 23rd International Conference on Tools with Artificial Intelligence.

15. Dagan, I, Karov, Y & Roth, D 1997, 'Mistake-Driven Learning in Text Categorization', In Proceedings of CORR.

16. Damiano Spina, Julio Gonzalo & Enrique Amigo 2013, 'Discovering Filter Keywords for Company Name Disambiguation in Twitter'.

17. David, B, Bracewell, Jiajun Yan & Fuji Ren 2009,'Category Classification and Topic Discovery of Japanese and English News Articles', Electronic Notes in Theoretical Computer Science, vol. 225, pp. 51-65.

18. Deerwester, S, Dumais, ST, Furnas, GW, Landauer, TK & Harshman, R 1990, 'Indexing by Latent Semantic Analysis', Journal of the American Society for Information Science, vol. 41, no. 6, pp.391–407.

19. Drew, T & Kalaska, J 2008, 'Computational Neuro Science: Theoretical Insights into Brain Function', Elsevier.

20. Dumais, S & Chen 2000, 'Hierarchical Classification of Web Content', in Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 256– 263, New York: ACM.

21. Duoqian Miao, Qiguo Duan, Hongyun Zhang & Na Jiao 2009, 'Rough Set Based Hybrid Algorithm for Text Classification', Expert Systems with Applications, vol. 36, pp. 9168-9174.

22. Durmaz, O & Bilge, HS 2011, 'Effects of Dimensionality Reduction and Feature Selection in Text Classification', In Signal Processing and Communications Applications SIU, IEEE 19th Conference on pp. 21-24.

23. Eneko Agirre, Aitor Soroa & Mark Stevenson 2010, 'Graph-based Word Sense Disambiguation of Biomedical Documents', Journal of Bioinformatics, vol.26, no.22, pp. 2889-2896.

24. Erdem Alparslan, Adem Karahoca & Hayretdin Bahsi 2011 'Classification of Confidential Documents by Using Adaptive Neuro Fuzzy inference systems', Procedia Computer Science, vol.3, pp.1412-1417.

25. Fabian M Suchanek, Gjergji Kasneci & Gerhard Weikum2007, 'Yago - A Core of Semantic Knowledge', 16th International World Wide Web Conference WWW.

26. Faralli, S & Navigli, R 2013, 'A Java framework for Multilingual Definition and Hypernym Extraction', Proc. of ACL, Comp. Volume.

27. Feldman & Sanger 2007, 'The Text Mining Handbook-Advanced Approaches in Analyzing Unstructured Data', USA: New York.

28. Fernandez-Vilas, A, Diaz-Redondo, RP & Bermejo-Munoz, J 2008, 'A Flexible Semantic Inference Methodology to Reason about User Preferences in Knowledge-Based Recommender Systems', Knowledge Based System., vol. 21, no.4.

29. Galavotti, L, Sebastiani, F & Simi, M 2000, 'Experiments on the use of Feature Selection and Negative Evidence in Automated Text Categorization', In Proceedings of ECDL-00,4th European Conference on Research and Advanced Technology for Digital Libraries Lisbon, Portugal , pp.59–68.

30. George A Miller 1995, 'WordNet: a Lexical Database for English'. Communications of the Association for Computing Machinery, vol.3811, pp.39–41.

31. Glorot, A, Bordes & Bengio, Y 2011, 'Domain Adaptation for Large-Scale Sentiment Classification: A Deep Learning Approach', In Proceedings of the Twenty-Eight International Conference on Machine Learning, ICML.

32. Gruber, TR 1995, 'Toward Principles for the Design of Ontologies Used for Knowledge Sharing', International Journal of Human-Computer Studies, vol.435-6, pp.907-928.

33. Han, EHS, Karypis & Kumar, G2001,'Text Categorization Using Weight Adjusted k-Nearest Neighbor Classification', 5th pacific Asia Conference on Knowledge Discovery and Data Mining, PAKDD, pp. 53-65.

34. Hinton, GE & Salakhutdinov, RR 2006, 'Reducing the Dimensionality of Data with Neural Networks', Science, vol. 313, no. 5786, pp.504–507.

35. Hinton, GE 2002, 'Training Products of Experts by Minimizing Contrastive Divergence', Neural Computation, vol.14, pp.1771–1800.

36. Hinton, GE 2007, 'Boltzmann Machine', Scholarpedia, vol.25, pp.16-68.

37. Hinton, GE 2007, 'Learning Multiple Layers of Representation', Trends in Cognitive Sciences, vol. 1110, pp.428–434.

38. Hofmann, T 1999, 'Probabilistic Latent Semantic Indexing', In Proceedings of the 22nd Annual International ACM SI-GIR Conference on Research and Development in Information Retrieval, pp.50–57, ACM, Berkeley, California, USA.

39. Holland, JH 1992, 'Genetic Algorithms. Scientific American', vol. 4, pp. 44–50.

40. How Jing, Yu Tsao, Kuan-Yu Chen & Hsin-Min Wang 2013, 'Semantic Naive Bayes Classifier for Document Classification', International Joint Conference on Natural Language Processing, pp. 1117–1123.

41. Huang, EH, Socher, R, Manning, CD & Ng,AY 2012, 'Improving Word Representations via Global Context and Multiple Word Prototypes', In Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics ACL ' vol.12, pp. 873–882.

42. Hugo Larochelle & Yoshua Bengio 2008, 'Classification using Discriminative Restricted Boltzmann Machines', In William W. Cohen, Andrew McCallum, and Sam T. Roweis, Editors, Proceedings of the Twenty-Fifth International Conference on Machine Learning ICML'08, pp.536–543 ACM.

43. Hugo Larochelle, Michael Mandel, Razvan Pascanu & Yoshua Bengio, 2012, 'Learning Algorithms for the Classification Restricted Boltzmann Machine', Journal of Machine Learning Research, pp.643-669.

44. Jemma Wu 2012, 'A Framework for Learning Comprehensible Theories in XML Document Classification', IEEE transactions on knowledge and Data Engineering, vol. 24, no.1, pp. 1-14.

45. Jianping Zeng, Jiangjiao Duan, Wenjun Cao & Chengrong Wu 2012, 'Topics Modeling Based on Selective Zipf Distribution', Expert Systems with Applications, vol. 39, pp. 6541-6546.

46. Jiawei Han, Micheline Kamber & Jian Pei 2012, 'Data Mining: Concepts and Techniques', Third Edition, The Morgan Kaufmann Series in Data Management Systems.

47. Karki, M 1997, 'Patent Citation Analysis: A policy Analysis tool'. World Patent Information, vol.194, pp. 269–272.

48. Khan, A, Baharudin, B& Khan, K 2010, 'Semantic Based Features Selection and Weighting Method for Text Classification', In Information Technology ITSim, International Symposium in IEEE, vol. 2, pp. 850-855.

49. Ko, Y & Seo, J 2000, 'Automatic Text Categorization by Unsupervised Learning', In Proceedings of the 17th Conference on Computational Linguistics COLING'2000, Saarbrucken, Germany, pp. 453–459.

50. Leis, V, Kemper, A & Neumann, T 2013,'The Adaptive Radix Tree: ARTful indexing for main-memory databases', ICDE.

51. Li Dong, Furu Wei,Ming Zhou & Ke Xu 2014, 'Adaptive Multi-Compositionality for Recursive Neural Models with Applications to Sentiment Analysis', Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, pp. 1537 – 1543

52. Li, Z, Li, P, Wei, W, Liu, H, He, J, Liu, T & Du, X 2009, 'Auto PCS: A Phrase-Based Text Categorization System for Similar Texts. In Q'.

53. Li, L, Feng, J, Pei, S, Wang, X, Zhou & Zhu,QM 2005, 'Advances in Data and Web Management ,Springer Berlin / Heidelberg: vol. 5446, pp. 369-380.

54. Li, Z, Lee, KC, Zheng, B, Lee, WC, Lee, DL & Wang, X 2011, 'IR-tree: An Efficient Index for Geographic Document Search. Knowledge and Data Engineering', IEEE Transactions on, vol.234, pp.585-599.

55. Lili Yang, Chunping Li, Qiang Ding & Li Li 2013, 'Combining Lexical and Semantic Features for Short Text Classification', Procedia Computer Science, vol. 22, pp. 78-86.

56. Lin, D 1998, 'An Information-Theoretic Definition of Similarity', Proceedings of the 15th International Conference on Machine Learning, 24-27; Madison, Wisconsin, USA.

57. Lingling Meng, Runqing Huang & Junzhong Gu 2013, 'A Review of Semantic Similarity Measures in WordNet', International Journal of Hybrid Information Technology vol. 6, no. 1.

58. Liu, T, Chen, Z, Zhang, B, Ma, WY & Wu,G 2004, 'Improving Text Classification using Local Latent Semantic Indexing'. Paper Presented at the Proceedings of the Fourth IEEE International Conference on Data Mining.

59. Manning, CD, Raghavan, P & Schütze, H 2008, 'Introduction to Information Retrieval', New York, NY, USA: Cambridge University Press.

60. Massimiliano Ciaramita, Thomas Hofmann & Mark Johnson 2003, 'Hierarchical Semantic Classification: Word Sense Disambiguation with World Knowledge' in18th International Joint Conference on Artificial Intelligence.

61. McCallum, A & Nigam, K 1998, 'A Comparison of Event Models for Naive Bayes Text Classification', In Proceeding of the AAAI'98 Workshop on Learning for Text Categorization, Madison, pp. 41–48.

62. Miguel E Ruiz & Padmini Srinivasan 1998, 'Automatic Text Categorization Using Neural Network', In Proceedings of the 8th ASIS SIG/CR Workshop on Classification Research, pp. 59-72.

63. Miroslava Drazic, Dragan Kukolj & Milana Vitas 2013, 'Technology Matching of the Patent Documents using Clustering Algorithms', In IEEE International Symposium on Computational Intelligence and Informatics, pp. 405-409

64. Mitchell, T 1997, 'Machine Learning', McGraw Hill, ISBN 0-07-042807-7.

65. Mohamed, AR, Dahl, G & Hinton, GE 2009, 'Deep Belief Networks for Phone Recognition', In NIPS 22 Workshop on Deep Learning for Speech Recognition.

66. Mohammad Khabbaz, Keivan Kianmehr & Reda Alhajj 2012, 'Employing Structural and Textual Feature Extraction for Semi Structured Document Classification', IEEE Transaction, vol. 42, no. 6, pp. 1566-1578.

67. Nikos Tsimboukakis & George Tambouratzis 2011, 'Word-Map Systems for Content-Based Document Classification', IEEE transactionson Systems, Man, and Cybernetics, vol. 41, no. 5, pp. 662-673.

68. Patil, L & Atique, M 2013, 'A Novel Approach for Feature Selection Method TF-IDF in Document Clustering'. In Advance Computing Conference IACC, IEEE 3rd International pp. 858-862.

69. Pei Yi Hao, Jung-Hsien Chiang & Yi-Kun Tu 2007, 'Hierarchically SVM classification Based on Support Vector Clustering Method and its Application to Document Categorization', Expert Systems with Applications, vol. 33, pp. 627-635.

70. Porter, MF 1980, 'An Algorithm for Suffix Stripping Program', vol.143, pp.130-137.

71. Rada, R, Mili, H, Bicknell, E & Blettner, M 1989, 'Development and Application of a Metric in Semantic Nets', IEEE Transactions on Systems, Man and Cybernetics, vol. 19, issue 1, pp. 17 - 30.

72. Raghavan, P, Amer-Yahia, S & Gravano, L 2004, 'Structure in Text: Extraction and Exploitation', In. Proceeding of the 7th International Workshop on the Web and Databases WebDB, ACM SIGMOD/PODS, ACM Press, vol.67.

73. Resnik, P1995, 'Using Information Content to Evaluate Semantic Similarity', Proceedings of the 14th International Joint Conference on Artificial Intelligence, pp.20-25; Montréal Quebec, Canada.

74. Ricardo Baeza-Yates & Berthier Ribeiro-Neto 1999, 'Modern Information Retrieval', Addison-Wesley,England, pp.513.

75. Roberto Navigli, Stefano Faralli, Aitor Soroa, Oier de Lacalle & Eneko Agirre 2008, 'Two Birds with One Stone: Learning Semantic Models for Text Categorization and Word'.

76. Ronald L Rivest 1987, 'Learning Decision Lists', Machine Learning2, Kluwer Academic Publishers, Boston,, Netherland, pp.229-246.

77. Ronen Feldman & James Sanger 2007, 'The Text Mining Handbook : Advanced Approaches in Analyzing Unstructured Data', Cambridge University Press.

78. Salakhutdinov, RR & Hinton, GE 2007, 'Semantic Hashing'. In Proceedings of the Workshop on Information Retrieval and Applications of Graphical Models SIGIR 2007 Amsterdam. Elsevier.

79. Salakhutdinov, RR & Hinton, GE 2009, 'Deep Boltzmann Machines'. In Proceedings of the International Conference on Artificial Intelligence and Statistics.

80. Salakhutdinov, RR & Hinton, GE 2009,'Replicated Softmax: An Undirected Topic Model', In Advances in Neural Information Processing Systems, vol.22.

81. Salton, G, Wong, A & Yang, CS 1975,'A Vector Space Model for Information Retrieval', Communications of the ACM, vol.1811, pp. 613–620.

82. Schutze, D, Hull, A &Pedersen, JO 1995, 'A Comparison of Classifiers and Document Representations for the Routing Problem', In Proceedings of SIGIR-95, 18th ACM International Conference on Research and Development in Information Retrieval, pp. 229–237, Seattle, US.

83. Seaghdha, DO 2009, 'Semantic Classification with WordNet kernels', Paper Presented at the Proceedings of Human Language Technologies: The Annual Conference of the North American Chapter of the Association for Computational Linguistics, Companion Volume: Short Papers, Boulder, Colorado.

84. Sebastiani,F 2002, 'Machine Learning in Automated Text Categorization', in ACM Computing Surveys Archive vol. 34 , issue 1, pp. 1 – 47.

85. Sebastiani, F 2005, 'Text Categorization, in A. Zanasi ed., Text Mining and its Applications to Intelligence', CRM and Knowledge Management, pp. 109-129, WIT Press, Southampton, UK.

86. Selamat, A& Omatu, S 2004,'Web Page Feature Selection and Classification using Neural Networks', Information Sciences, vol.158, pp.69–88.

87. Sergio Greco, Francesco Gullo, Giovanni Ponti & Andrea Tagarelli 2011, 'Collaborative Clustering of XML Documents', Journal of Computer and System Sciences, vol. 77,pp.988–1008.

88. Sofia & Bulgaria Nyberg, K 2011, 'Document Classification Using Machine Learning and Ontologies'.

89. Tao Liu, Shengping Liu, Zheng Chen & Wei-Ying Ma2003, 'An Evaluation on Feature Selection for Text Clustering', Proceedings of the Twentieth International Conference on Machine Learning ICML, Washington DC.

90. Tikk, D, Biró & Törcsvári 2005, 'Experiment with a Hierarchical Text Categorization Method on WIPO Patent Collections', Applied Research in Uncertainty Modeling and Analysis, International Series in Intelligent Technologies, vol.20, pp.283-302.

91. Turban, E & Aronson, JE 2001, 'Decision Support Systems and Intelligent Systems',6th edition. Upper Saddle River, NJ: Prentice-Hall International.

92. Varelas, G, Voutsakis, E, Raftopoulou, P, Petrakis, GM & Milios, EE 2005, 'Semantic Similarity Methods in WordNet and their Application to Information Retrieval on the Web', Proceedings of the 7th Annual ACM International Workshop on Web Information and Data Management, Bremen, Germany.

93. Versky, T 1977, 'Features of Similarity', Psycological Review, vol. 84, no. 4.

94. Wang, Y & Wang, XJ 2005, 'A New Approach to Feature Selection in Text Classification', Proceedings of 4thInternational Conference on Machine Learning and Cybernetics, IEEE, vol.6, pp. 3814-3819.

95. Wu, Z & Palmer, M1994, 'Verb Semantics and Lexical Selection', Proceedings of 32nd annual Meeting of the Association for Computational Linguistics, Las Cruces, New Mexico.

96. Xavier Glorot, Antoine Bordes & Yoshua Bengio 2011, 'Domain Adaptation for Large-Scale Sentiment Classification', A Deep Learning Approach, Proceedings of the 28th International Conference on Machine Learning, Bellevue, WA, USA.

97. Xinhui Tu, Tingting He, Long Chen, Jing Luo & Maoyuan Zhang 2010,'Wikipedia-Based Semantic Smoothing for the Language Modeling Approach to Information Retrieval', Lecture Notes in Computer Science, vol. 5993, pp. 370-381.

98. YAGO2s: 'A High-Quality Knowledge Base', Last Access 2013, from http://www.mpiinf.mpg.de/yago-naga/yago

99. Yan Liu, Sheng-hua Zhong & Wen-jie Li 2012, 'Query-Oriented Unsupervised Multi-document Summarization via Deep Learning', In Proceedings of 26th Conference on Artificial Intelligence AAAI'12.

100. Yang Yimin & Pedersen, JO 1997, 'A Comparative Study on Feature Selection in Text Categorization'. Proceedings of the 14th International Conference on Machine Learning ICML-97, pp.412-420.

101. Yang, Y& Pedersen, J 1997, 'A Comparative Study on Feature Selection in Text Categorization', In Proceeding of the 14th International Conference on Machine Learning, ICML, pp.412-420, Nashville, TN.