

References

- 1) P. H. A. Sneath and R. R. Sokal. "Numerical Taxonomy: The Principles and Practice of Numerical Classification." W. H. Freeman and Company, San Francisco, 1973.
- 2) Shyam Boriah, Varun Chandola and Vipin Kumar, "Similarity Measures for Categorical Data: A Comparative Evaluation", SIAM Data Mining Conference, April 2008, Atlanta, GA.
- 3) Matthew Michelson and Craig A. Knoblock, "Mining Heterogeneous Transformations for Record Linkage", (*IIWEB-07*), Vancouver, Canada, 2007.
- 4) Naresh Kumar Nagwani and Ashok Bhansali, "An Object Oriented Email Clustering Model Using Weighted Similarities between Emails Attributes", International Journal of Research and Reviews in Computer Science (IJRRCS) Vol. 1, No. 2, June 2010.
- 5) Naresh Kumar Nagwani, "OSIM: An Open Source Framework For Measuring Weighted Similarities Between Objects", International Journal of Emerging Technologies And Applications in Engineering, Technology and Sciences (IJ-ETA-ETS), ISSN: 0974-3588, Jan 2010.
- 6) Naresh Kumar Nagwani and Pradeep Singh, "Weight Similarity Measurement Model Based, Object Oriented Approach for Bug Databases Mining to Detect Similar and Duplicate Bugs", International Conference on Advances in Computing, Communication and Control (ICAC3'09).
- 7) Carina F. Dorneles, Carlos A. Heuser, Andrei E. N. Lima, Altigran da Silva and Edleno de Moura, "Measuring Similarity Between Collection of Values", ACM New York, NY, USA ©2004.

- 8) Melanie Weis and Felix Naumann , "DogmatiX Tracks down Duplicates in XML" , SIGMOD 2005 June 1416, 2005, Baltimore, Maryland, USA.
- 9) Mark A. Greenwood and Mark Stevenson, "Improving Semi-Supervised Acquisition of Relation Extraction Patterns", Workshop on Information Extraction Beyond The Document, Sydney, July 2006.
- 10) Jun Yao , "searching multiple databases for interesting complexes national university of Singapore", Department of information system and computer science National University of Singapore
- 11) Silvana Aciar , "Information source selection methodology for recommender systems " , Department of Electronic , Computer Science and Automated System University of Girona - Spain 2007.
- 12) Koby Carmmer ,Michael kearns and Jennifer wortman , "learning from multiple sources" , Journal of Machine Learning Research 2008.
- 13) Patrick Pantel , Andrew Philpot and Eduard Houy, "Matching and integration across heterogeneous data sources" , international conference on Digital government research ACM New York, NY, USA 2006.
- 14) Shela Tejada , Knoblok Steven Minton , "Learning Model for Multi Source Integration".
- 15) Hye-Chung Kum, Joong Hyuk Chang, Wei Wang, "Sequential Pattern Mining in Multi-Databases via Multiple Alignment". Data Min. Knowl. Discov. 12(2-3): 151-180 (2006).
- 16) Buttler, David, Matthew Coleman, Terence Critchlow, Renato Fileto, Wei Han, Calton Pu, Daniel Rocco, and Li Xiong, "Querying Multiple Bioinformatics Information Sources" Can Semantic Web Research Help? ", SIGMOD Record, Vol 31, No. 4 (December 2002).

- 17) Tae-Wan Ryu and Christoph F. EICK, "Similarity Measures for Multi-Valued Attributes for Database Clustering".
- 18) Gower, J.C, "A general Coefficient of Similarity and Some of Its Properties", *Biometrics* 27, 857-872.
- 19) M.lenzerin C Batini , S.B.navathe , "A comparative Analysis of Methodologies for database schema integration".
- 20) Shammugasun Dram Harsharan and Rengarmanujam Srinivasan "A comparison of Similarity measure for text documents", *journal information and knowledge management* (2008).
- 21) Michael W. Berry, Murray Brown, " lecture notes in data mining" World Scientific pub 2006.
- 22) Kardi Teknomo's Homepage, <http://people.revoledu.com>.
- 23) Hume 1748,"An Enquiry Concerning Human Understanding".
- 24) J.P Yoon and Kerschberb, "Framework for Knowledge Discovery and Evaluation in database", *IEEE transaction on knowledge and data engineering* 1993.
- 25) Tversky, A. (1977)," Features of similarity, *Psychological Review*", 84, 327-352.
- 26) Vasconcelos, N. "On the Efficient Evaluation of Probabilistic Similarity Functions For Image Retrieval" (2004). *IEEE Transactions on Information Theory*.
- 27) Tabata, H.," *History and Philosophy of Logic*", 21 (4): 265–295.
- 28) Michael W. Berry, Murray Brown "lecture Notes in Data mining" World Scientific pub 2006.
- 29) Cantor, G. Uber "A property of the Totality of all Real Algebraic Numbers", *Journal for the pure and applied mathematics* 77, 258{262}.

- 30) William W. Cohen, Natalie Glance, Charles Schafer, Roy Tromble, Yuk Wah Wong, "Context-Sensitive Similarity Metrics", siteseerx 2011.
- 31) Kamran Zamanifar and FarinazAlamiyan, "A New Similarity Measure for Instance Data Matching", International Conference on Computer Communication and Management Proc .of CSIT vol.5 IACSIT Press, Singapore (2011).
- 32) Rishi Sayal and V. Vijay Kumar ,"A Novel Similarity Measure for Clustering Categorical Data Sets",International Journal of Computer Applications (0975 – 8887)Volume 17– No.1, March 2011.