

REFERENCES

- [Agarwal et al.1998] Agarwal, R., Gehrke, J., Gunopulos, D., Raghavan, P., “Automatic subspace clustering of high dimensional data for data mining applications”, Proceedings of ACM SIGMOD conference, pp. 94 – 105, 1998
- [Al-Razgan et. al., 2007] Al-Razgan M., Domeniconi C., and Barbara D. "Clustering Ensembles for Categorical Data", in Proceedings of the *Workshop on Supervised and Unsupervised Ensemble Methods and their Applications*, held in conjunction with the *Third Iberian Conference on Pattern Recognition and Image Analysis*, 2007
- [Al-Shaqsi and Wang 2010] Al-Shaqsi, J. Wenjia Wang, "A clustering ensemble method for clustering mixed data", in proceedings of the international Joint Conference on Neural Networks (IJCNN), Barcelona, July 2010.
- [Amir and Lipika 2007] Amir Ahmad, Lipika Dey, “A k-mean clustering algorithm for mixed numeric and categorical data”, Elsevier Journal of Data & Knowledge Engineering, (2007), Vol. 63, pp. 503–527.
- [Anderberg, 1973] Anderberg, M. R., “Cluster Analysis for Applications”, 1973, Academic Press, New York.
- [Aranganayagi and Thangavel 2010] Aranganayagi S. and Thangavel K., “Incremental Algorithm to Cluster the Categorical Data with Frequency Based Similarity Measure”, International Journal of Information and Mathematical Sciences, Vol 6, Issue:1, 2010.
- [Arun.2001] Arun.K.Pujari, “Data Mining Techniques”, University Press, 2001.

[Aynur and Craig 2004] Aynur Dayanik, Craig G. Nevill-Manning, "Clustering in Relational Biological Data", ICML-2004 Workshop on Statistical Relational Learning and Connections to Other Fields, pp: 42-47, 2004.

[Ben and Elouedi 2008] Ben Hariz,A., Elouedi, Z., "An Incremental Clustering Approach within Belief Function Framework", In Proceedings of the Artificial Intelligence and Soft Computing, 2008.

[Bo Liu *et al.* 2006] Bo Liu, Jiuhui Pan and Bob McKay, "Incremental Clustering Based on Swarm Intelligence", Simulated Evolution and Learning, Lecture Notes in Computer Science, 2006, Vol: 4247/2006, pp.189-196

[Can 1993] Fazil Can, "Incremental Clustering for Dynamic Information Processing", ACM Transactions on Information Systems, April 1993, Vol. 11, No. 2, pp. 143-164.

[Carpenter and Grossberg 1990] Carpenter G. and Grossberg S. 1990. ART3: Hierarchical search using chemical transmitters in self-organizing pattern recognition architectures. *Neural Networks* 3, 129–152.

[Charikar et al.1997] M. Charikar, C. Chekuri, T. Feder, and R. Motwani, "Incremental clustering and dynamic information retrieval," 29th Symposium on Theory of Computing, 1997, pp. 626—635.

[Chen et. al., 1996] Chen M.S., Han J., Yu P. S., " Data Mining: An Overview from Database Perspective" *IEEE Trans. On Knowledge and Data Engineering*, 5(1):866-883, December 1996.

[Chen *et al.* 2002] Chien-Yu Chen, Shien-Ching Hwang, and Yen-Jen Oyang, "An Incremental Hierarchical Data Clustering Algorithm Based on Gravity Theory",

Proceedings of the 6th Pacific- Asia Conference on Advances in Knowledge Discovery and Data Mining, 2002, pp: 237 – 250.

[Chen H.L et. al., 2005] Chen H.L., Chuang K.T., Chen M.S., “Labeling Unclustered Categorical Data into Clusters Based on the Important Attribute Values,” ICDM '05 Proceedings of the Fifth IEEE International Conference on Data Mining, 2005, pp. 106 - 113.

[Chen *et al.* 2005] Chien-Yu Chen, Shien-Ching Hwang, Yen-Jen Oyang, “A statistics-based approach to control the quality of subclusters in incremental gravitational clustering”, The Journal of Pattern Recognition Society, vol: 38, pp. 2256 – 2269, 2005

[Chen *et al.* 2007] Chen Zhuo, Liu Xiang-shuang, Zhuang Xiao-dong, "A Fast Incremental Clustering Algorithm Based on Grid and Density", Third International Conference on Natural Computation, Vol: 5, pp: 207 - 211, 2007.

[Chen *et al.* 2009] Chen J., Yang Z., Yin J., Yang X., Huang L., "An Incremental Clustering with Attribute Unbalance Considered for Categorical Data", Computational Intelligence and Intelligent Systems, Communications in Computer and Information Science, Volume 51. Springer-Verlag, 2009, p. 433

[Chen H.L et al. 2009] Chen H.L, Chen M.S and Chen L.S , "Catching the Trend: A Framework for Clustering Concept-Drifting Categorical Data", IEEE Transactions on Knowledge and Data Engineering, Vol 21, No.5, May 2009, pp. 652 - 665

[Chen and Meng 2004] Zhuo Chen, Qing-Chun Meng, “An incremental clustering algorithm based on swarm intelligence theory”, Proc. of the 3rd Int. Conf. on Machine Learning and Cybernetics, 2004, Vol : 3.

[Dan and Namita 2004] Dan Simovici, Namita Singla, "Metric Incremental Clustering of Nominal Data", ICDM, pp: 523-526, 2004.

[Dunham 2003] Dunham, M.H., "Data Mining – Introductory and Advanced Topics", Pearson Education, 2003.

[Edwin 2008] Edwin Lughofer, "Extensions of vector quantization for incremental clustering", Pattern Recognition, Vol: 41, No: 3, pp: 995-1011, 2008.

[Ester et al.1996] Ester, M., Kriegel, H., Sander, J., Xu X., "A density-based algorithm for discovering clustering in large spatial databases with noise", Proceedings of ACM SIGMOD conference, pp. 226 – 231, 1996.

[Ester et al.1998] Ester, M., Kriegel, H., Sander, J., Xu X., Wimmer, M., "Incremental Clustering for Mining in a Data Warehousing Environment", Proceedings of the 24th International Conference on Very Large Databases (VLDB'98), New York, USA, 1998, pp. 323-333.

[Everitt 1980] Everitt, B., "Cluster analysis", Halsted Press, 1980.

[Fahim *et al.* 2006] Fahim A.M., Salem A.M., Torkey F.A., Ramadan M.A., "An efficient enhanced k-means clustering algorithm", Journal of Zhejiang University science, vol. 7, no.10, pp.1626-1633, 2006.

[Fayyad et. al., 1996] Fayyad, U.M., Piatsky-shapiro, G., Smyth, P. "From Data Mining to knowledge discovery in databases ", AI magazine, Vol. 7, pp. 37 – 54, 1996.

[Fisher 1987] Fisher D., "Knowledge acquisition via incremental conceptual clustering," Machine Learning, vol. 2, 1987, pp.139-172.

[Fisher et al. 1993] Fisher , D., Xu, L., Carnes, R., Rich, Y., Fenves, S.J., Chen, J., Shiavi, R., Biswas, G., and Weinberg, J. 1993. Applying AI clustering to engineering tasks. *IEEE Expert* 8, 51–60.

[Fei R. et al. 2008] Fei Ren, Liang Hu, Hao Liang, Xiaobo Liu, Weiwu Ren, "Using Density-Based Incremental Clustering for Anomaly Detection", International Conference on Computer Science and Software Engineering, Vol: 3, pp: 986 - 989, 2008.

[Fotakis 2006] Fotakis D., "Incremental algorithms for Facility Location and k-Median", Theoretical Computer Science, Vol: 361, No: 2-3, pp: 275-313, 2006.

[Ganti et al. 1996] Ganti V., Gehrke J., Ramakrishna R., "CACTUS: Clustering categorical data using summaries", Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '99), 1999.pp. 73-83

[Gennary et al. 1989] J. Gennary, P. Langley, and D. Fisher, "Models of Incremental Concept Formation," Artificial Intelligence Journal, vol. 40, 1989, pp. 11-61.

[Gibson et al. 2000] Gibson D., Kleinberg J., Raghavan P., "Clustering categorical data: An approach based on dynamical systems", The VLDB Journal — The International Journal on Very Large Data Bases archive, Volume 8 Issue 3-4, February 2000,Pages 222-236

[Gluck and Corter 1985] Gluck, M. A., and Corter, J. E., "Information, uncertainty, and the utility of categories", Proceedings of the Seventh Annual Conference of the Cognitive Science Society, pp. 283 – 287, 1985.

[Gorunescu and Dumitrescu 2003] Gorunescu R. and Dumitrescu D. “Evolutionary Clustering Using An Incremental Technique” Studia University. Babes-Bolyai, INFORMATICA, Volume XLVIII, Number 2, 2003.

[Guha et al.1998] Guha, S., Rastogi, R., Shim, K., “CURE: An efficient clustering algorithm for large databases”, Proceedings of ACM SIGMOD conference, pp. 73 – 84,1998.

[Guha et al.1999] Guha, S., Rastogi, R., Shim, K., “ROCK: A robust clustering algorithm for categorical attributes”, Proceedings of the 15th International conference on Data Engineering, pp. 512 – 521,1999.

[Hastie et al. 2009] Hastie, T., Tibshirani, R.; Friedman, J. (2009). "[Hierarchical clustering](#)" (PDF). *The Elements of Statistical Learning* (2nd ed.). New York: Springer. pp. 520–528. ISBN 0-387-84857-6. <http://www-stat.stanford.edu/~tibs/ElemStatLearn/>.

[Han and Kamber 2001] Han, J., Kamber, M., “ Data Mining: Concepts and Techniques”, Morgan Kaufmann Publishers, 2001.

[Hariz and Elouedi 2008] Hariz A.B. and Elouedi Z., "An Incremental Clustering Approach within Belief Function Framework", In Proceedings of the Artificial Intelligence and Soft Computing, 2008.

[Hartigan 1975] Hartigan, J.A. 1975. *Clustering Algorithms*. John Wiley and Sons, Inc., New York, NY.

[Hawwash and Nasraoui 2010] Hawwash B. and Nasraoui O., “Robust Clustering of Data Streams using Incremental Optimization”, Proceedings of the 4th Alberto Mendelzon International Workshop on Foundations of Data Management, 2010.

[Hsu and Huang 2008] Hsu C.C and Huang Y.P : "Incremental clustering of mixed data based on distance hierarchy", *Expert Systems with Applications*, Vol: 35, No: 3, pp: 1177-1185, 2008.

[Hsu and Wang 2005] Hsu, C. C. and Wang, S.-H., "An integrated framework for visualized and exploratory pattern discovery in mixed data", *IEEE Transactions on Knowledge and Data Engineering*, Vol: 18, No: 2, pp: 161–173, 2005.

[Huang 1997] Huang Z, "Clustering Large Data Sets With Mixed Numeric And Categorical Values", *The First Pacific-Asia Conference on Knowledge Discovery and Data Mining (1997)*, pp. 21-34.

[Huang 1998] Huang Z, "Extensions to the k -Means Algorithm for Clustering Large Data Sets With Categorical Values", *Data Mining and Knowledge Discovery*, 1998, Vol. 2, pp. 283–304.

Iris dataset : <http://archive.ics.uci.edu/ml/datasets/Iris>

[Jize et. al., 2007] Jize Chen, Zhimin Yang, Jian Yin, Xiaobo Yang, Li Huang, "An Incremental Clustering with Attribute Unbalance Considered for Categorical Data", *Computational Intelligence and Intelligent Systems, Communications in Computer and Information Science*, Volume 51. 2009, p. 433

[Judd et al. 1996] Judd, D., Mckinley, P., Jain, A. K. 1996. "Large-scale parallel data clustering", *Proceedings of the International Conference on Pattern Recognition (Vienna, Austria)*, pp. 488–493.

[Jain *et al.* 1999] Jain A.K., Murthy M.N., Flynn P.J., "Data Clustering: A Review", *ACM Computing Surveys*, September 1999, Vol. 31, No.3, pp. 264 – 323.

[Jain and Dubes 1988] Jain, A.K., Dubes R.C., "Algorithms for clustering data", Prentice Hall, 1988.

[Jing *et al.* 2005] Jing Gao, Jianzhong Li, Zhaogong Zhang, Pang-Ning Tan, “An Incremental Data Stream Clustering Algorithm Based on Dense Units Detection”, Proceedings of the 9th Pacific-Asia conference on Advances in Knowledge Discovery and Data Mining (PAKDD'05), 2005, pp. 420-425.

[Kaufman and Rousseeuw, 1990] Kaufman, L. and Rousseeuw, P.J. 1990. Finding Groups in Data—An Introduction to Cluster Analysis. Wiley.

[Khalilian and Mustapha] Khalilian M. and Mustapha N., “ Data Stream Clustering: Challenges and Issues”, Proceedings of the International MultiConference of Engineers and Computer Scientists (IMECS 2010), Vol 1.

[Li Jie *et al.* 2003] Li Jie, Gao Xinbo and Jiao Li-cheng, "A GA-based clustering algorithm for large data sets with mixed and categorical values", in Proceedings of the Fifth International Conference on Computational Intelligence and Multimedia Applications, pp, 102, September 2003.

[Li Taoying *et. al.*, 2010] Li Taoying, Chne Yan, Qu Lili, Mu Xiangwei “Incremental clustering for categorical data using clustering ensemble”, 29th Chinese Control Conference (CCC), IEEE Conference Publications, 2010, pp : 2519 – 2524.

[Lin *et. al.*, 2004] Lin J., Vlachos M., Keogh E., Gunopulos D., "Iterative Incremental Clustering of Time Series", In Proc. IX Conf. on Extending Database Technology, Crete, Greece, March 14--18, 2004.

[Liu *et. al.*, 2006] Liu B., Shi Y., Wang Z., Wang W., Shi B., "Dynamic Incremental Data Summarization for Hierarchical Clustering", WAIM 2006, pp: 410-421, 2006.

[MacQueen 1967] MacQueen, J. B., "Some Methods for classification and Analysis of Multivariate Observations", in Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability. pp. 281–297, 1967.

[Mao and Jain 1996] MAO, J. AND JAIN, A. K., "A self-organizing network for hyperellipsoidal clustering (HEC)", *IEEE Transactions on Neural Networks*, 1996, Vol 7, pp. 16–29.

[Ming et al. 2010] Ming-Yi Shih, Jar-Wen Jheng and Lien-Fu Lai, "A Two-Step Method for Clustering Mixed Categorical and Numeric Data", *Tamkang Journal of Science and Engineering*, Vol. 13, No. 1, pp. 11-19, 2010.

[Mirkin 1996] Mirkin, B., "Mathematic classification and clustering", Kluwer Academic. Mining.

[Pham and Afify 2007] Pham, D.T. and Afify, A.A. "Clustering techniques and their applications in engineering", *Proceedings- Institution of Mechanical Engineers Part C Journal of Mechanical Engineering Science*, Vol: 221; No: 11, pp: 1445-1460, 2007.

[Pham et. al., 2007] Pham, D. T. , Otri, S., Afify, A. A. , Mahmuddin, M. , Al-Jabbouli, H. , "Data clustering using the Bees Algorithm", *Proc 40th CIRP Int. Manufacturing Systems Seminar*, Liverpool, 2007.

[Ravichandra 2003] Ravichandra Rao, I.K., "Data Mining and Clustering Techniques", *DRTC Workshop on Semantic Web*, Bangalore, 8th – 10th December, 2003.

[Samuel and Nick 2006] Samuel Sambasivam, Nick Theodosopoulos, "Advanced Data Clustering Methods of Mining Web Documents", *The Journal of Issues in Informing Science and Information Technology*, Vol: 3, 2006.

[Sebastian and Mihai 2009] Sebastian Lühr and Mihai Lazarescu, "Incremental clustering of dynamic data streams using connectivity based representative points", *Journal of Data & Knowledge Engineering*, Vol: 68 Issue 1, pp: 1-27, 2009.

[Serban and Campan 2005] Serban G., Campan A., "Incremental Clustering Using a Core-Based Approach", *Lecture Notes in Computer Science*, Springer Berlin, Vol: 3733, pp: 854-863, 2005.

[Serhat and Gungel 2009] Serhat S.B., Gungel B., "Incremental Clustering via Nonnegative Matrix Factorization", *Pattern Recognition*, vol: 42, No: 5, pp: 788-797, 2009.

[Seokkyung and Dennis 2005] Seokkyung Chung and Dennis McLeod, "Dynamic Pattern Mining: An Incremental Data Clustering Approach", *Journal on Data Semantics*, Vol. 2, pp. 85-112, 2005.

[Selim and Ismail 1984] Selim, S.Z., and Ismail, M.A., "K-means type algorithms: A generalized convergence theorem and characterization of local optimality", *IEEE Transactions on PAMI-6*, Vol 1, 1984

[Shaw and Xu 2009] Shaw G. and Xu Y., "Enhancing an Incremental Clustering Algorithm for Web Page Collections", *Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*, Vol: 3, pp: 81-84, 2009.

[Sowjanya and Shashi 2010] Sowjanya A.M. and Shashi M., "Cluster Feature-Based Incremental Clustering Approach (CFICA) For Numerical Data", *IJCSNS International Journal of Computer Science and Network Security*, Vol.10, No.9, September 2010.

[Sotirios 2011] Sotirios P. Chatzis, "A fuzzy c-means-type algorithm for clustering of data with mixed numeric and categorical attributes employing a probabilistic

dissimilarity functional", Expert Systems with Applications Vol. 38, no. 7, pp. 8684-8689, July 2011.

[Tao and Sheng, 2004] Tao Li and Sheng Ma, "IFD: Iterative feature and data clustering, Proceedings of the SIAM International Conference on Data Mining, 2004.

Wine dataset : <http://archive.ics.uci.edu/ml/datasets/Wine>

[Witten 2000] Witten, I. H., "Data mining: practical machine learning tools and techniques with Java implementations", San Francisco, California: Morgan Kaufmann, 2000.

[Xiaoke *et al.* 2009] Xiaoke Su, Yang Lan, Renxia Wan, and Yuming Qin, "A Fast Incremental Clustering Algorithm", Proceedings of the 2009 International Symposium on Information Processing (ISIP'09), 2009, pp. 175-178

Yeast dataset : <http://archive.ics.uci.edu/ml/datasets/Yeast>

[Zadeh 1965]. Zadeh, L. A. 1965. Fuzzy sets. Inf. Control, Vol 8, pp. 338–353.

[Zengyou *et al.* 2005] Zengyou He, Xiaofei Xu, Shengchun Deng, "Clustering mixed numeric and categorical data: A cluster ensemble approach", Computing Research Repository (CoRR), abs/cs/0509011, 2005.

[Zhang T et al. 1996] Tian Zhang , Raghu Ramakrishnan , Miron Livny, "BIRCH: an efficient data clustering method for very large databases", Proceedings of the ACM SIGMOD International. Conference on Management of Data, pp 103-114, 1996.

[Zhang T et al. 1996] Tian Zhang , Raghu Ramakrishnan , Miron Livny, "BIRCH : A new data clustering algorithm and its applications", International conference on Data Mining and Knowledge Discovery, Vol. 1, No. 2., pp.141-182, 1997.

PUBLICATIONS

- 1) Cluster Feature-Based Incremental Clustering Approach (CFICA) for Numerical Data, **Sowjanya, A. M.**, Shashi, M., International Journal of Computer Science and Network Security (IJCSNS), Vol. 10, No.9, pp. 73-79, Sep 2010.
- 2) A Cluster Feature-Based Incremental Clustering Approach to Mixed Data, **Sowjanya, A. M.**, Shashi, M., Journal of Computer Science, Vol. 7(12), pp. 1875-1880, Science Publications 2011.
- 3) Performance Comparison of CFICA with Incremental DBSCAN and BIRCH, **Sowjanya, A. M.**, Shashi, M., International Journal of Systems, Algorithms & Applications (IJSAA) Volume 2, Issue ICRAET12, May 2012.
- 4) A New Distance Metric for Formation of Non-Uniformly Distributed Incremental Clusters, **Sowjanya, A. M.**, Shashi, M., ACM International Conference Proceedings, 8th International Conference on Information Processing, Management and Intelligent Information Technology, ICIPT 2013, Seoul, Korea, April 1-3, 2013, ISBN: 978-1-4503-1783-2.
- 5) Incremental Cluster Formation with emphasis on Non-uniformly distributed Entities Involving Mixed Attributes, **Sowjanya, A. M.**, Shashi, M. (communicated)