参考文献

1. Li X , Fang M , Zhang J J , et al. Sample selection for visual domain adaptation via sparse coding[J]. Signal Processing: Image Communication, 2016, 44:92-100.
2. Y. Liao and X. Pan, "A New Method of Training Sample Selection in Text Classification," 2010 Second International Workshop on Education Technology and Computer Science, Wuhan, 2010, pp. 211-214.
3. Jiantao X , Mingyi H , Yuying W , et al. A fast training algorithm for support vector machine via boundary sample selection[C]// International Conference on Neural Networks & Signal Processing. IEEE, 2004.
4. X.T. Niu , Fcm-lssvm based on training sample selection, Metall. Mining Ind. 7 (9) (2015) 751–757 .
5. Fa Zhu, Ning Ye, Wei Yu, Sheng Xu, Guobao Li,Boundary detection and sample reduction for one-class Support Vector Machines,Neurocomputing,Volume 123,2014,Pages 166-173,ISSN 0925-2312,https://doi.org/10.1016/j.neucom.2013.07.002.
6. Zhai, Junhai, Wang, et al. Sample selection based on K-L divergence for effectively training SVM[C]// 2013.
7. 郝红卫, 蒋蓉蓉. 基于最近邻规则的神经网络训练样本选择方法[J]. 自动化学报, 2007(12).
8. J.M. Fernández Jaramillo, R. Mayerle,Sample selection via angular distance in the space of the arguments of an artificial neural network,Computers & Geosciences,Volume 114,2018, Pages 98-106,ISSN 0098-3004,https://doi.org/10.1016/j.cageo.2018.02.003.
9. Huiyong Li, Weiwei Bao, Jinfeng Hu, Julan Xie, Ruixin Liu,A training samples selection method based on system identification for STAP, Signal Processing,Volume 142,2018,Pages 119-124,ISSN 0165-1684, <https://doi.org/10.1016/j.sigpro.2017.07.008>.
10. DongDong Li, Zhe Wang, Chenjie Cao, Yu Liu,Information entropy based sample reduction for support vector data description,Applied Soft Computing,Volume 71,2018,Pages 1153-1160,ISSN 1568-4946,https://doi.org/10.1016/j.asoc.2018.02.053.
11. Jianpei Wang, Shihong Yue, Xiao Yu, Yaru Wang,An efficient data reduction method and its application to cluster analysis,Neurocomputing,Volume 238,2017,Pages 234-244,ISSN 0925-2312,https://doi.org/10.1016/j.neucom.2017.01.059.
12. Y. Yang, D. Chen and H. Wang, "Active Sample Selection Based Incremental Algorithm for Attribute Reduction With Rough Sets," in IEEE Transactions on Fuzzy Systems, vol. 25, no. 4, pp. 825-838, Aug. 2017.doi: 10.1109/TFUZZ.2016.2581186
13. Z. Jiang, X. Zhu, W. Tan and R. Liston, "Training sample selection for deep learning of distributed data," 2017 IEEE International Conference on Image Processing (ICIP), Beijing, 2017, pp. 2189-2193.doi: 10.1109/ICIP.2017.8296670
14. Z. Li, Y. Zhang, H. Liu and Y. Guo, "Training samples selection method for space-time adaptive processing based on clutter covariance matrix reconstruction," in Electronics Letters, vol. 53, no. 14, pp. 953-954, 6 7 2017.
15. <http://hadoop.apache.org/>.
16. <http://spark.apache.org/>.
17. <http://mahout.apache.org/>.
18. <http://spark.apache.org/mllib/>.
19. Chen X W, Lin X. Big Data Deep Learning: Challenges and Perspectives [J]. IEEE Access, 2014, 2:514-525
20. Bechini A, Marcelloni F, Segatori A. A MapReduce solution for associative classification of big data [J]. Information Sciences, 2016, 332:33-55.
21. Ludwig S A. MapReduce-based fuzzy c-means clustering algorithm: implementation and scalability [J]. International Journal of Machine Learning & Cybernetics, 2015, 6(6):923-934.
22. Panigrahi S, Lenka R K, Stitipragyan A. A Hybrid Distributed Collaborative Filtering Recommender Engine Using Apache Spark [J]. Procedia Computer Science, 2016, 83:1000-1006.
23. Maillo J, Ramírez S, Triguero I, et al. kNN-IS: An Iterative Spark-based design of the k-Nearest Neighbors Classifier for Big Data [J]. Knowledge-Based Systems, 2017, 117:3-15.
24. K. Nikolaidis, J.Y. Goulermas, Q.H. Wu, A class boundary preserving algorithm for data condensation, Pattern Recognit. 44 (3) (2011) 704–715.
25. H.A. Fayed, A.F. Atiya, A novel template reduction approach for the k-nearest neighbor method, IEEE Trans. Neural Netw. 20 (5) (2009) 890–896.
26. Y.H. Li, L. Maguire, Selecting critical patterns based on local geometrical and statistical information, IEEE Trans. Pattern Anal. Mach. Intell. 33 (6) (2011)1189–1201.
27. G.O. César, H.G.A. de, G.P. Nicolás, Democratic instance selection: a linear complexity instance selection algorithm based on classifier ensemble concepts, Artif. Intell. 174 (5–6) (2010) 410–441.
28. I. Triguero, D. Peralta, J. Bacardit, et al., MRPR: a MapReduce solution for prototype reduction in big data classification, Neurocomputing 150(Part A)(2015) 331–345.
29. E. Leyva, A. González, R. Pérez, A set of complexity measures designed for applying meta-learning to instance selection, IEEE Trans. Knowl. Data Eng.27 (2) (2015) 354–367.
30. W.C. Lin, C.F. Tsai, S.W. Ke, et al., Learning to detect representative data for large scale instance selection, J. Syst. Softw. 106 (2015) 1–8.
31. J. Hamidzadeh, R. Monsefi, H.S. Yazdi, IRAHC: instance reduction algorithm using hyperrectangle clustering, Pattern Recognit. 48 (5) (2015) 1878–1889.

R. M. Marone, F. Camara and S. Ndiaye, "LSIS: Large scale instance selection algorithm for big data," *2017 3rd IEEE International Conference on Computer and Communications (ICCC)*, Chengdu, 2017, pp. 2353-2356.

Malhat M , Menshawy M E , Mousa H , et al. Improving instance selection methods for big data classification[C]// International Computer Engineering Conference. IEEE, 2018.