**参考文献：**

[1]Ngufor C,Wojtusiak J.Extreme logistic regression[J]. Advances in Data Analysis and Classification, 2016, 10(1):27-52.

[2]Mackenzie C A,Trafalis TB,Barker K. A Bayesian beta kernel model for binary classification and online learning problems[J].Statistical Analysis and Data Mining, 2014, 7(6):434-449.

[3]Yang L,Qian Y. A sparse logistic regression framework by difference of convex functions programming[J]. Applied Intelligence, 2016, 45(2):1-14.

[4]Bao J, Chen Y, Yu L, et al. A multi-scale kernel learning method and its application in image classification[J]. Neurocomputing, 2017, 257: 16-23.

[5]Avron H, Clarkson K L, Woodruff D P. Faster kernel ridge regression using sketching and preconditioning[J]. SIAM Journal on Matrix Analysis and Applications, 2017, 38(4): 1116-1138.

[6]Koppel A, Warnell G, Stump E, et al. Parsimonious online learning with kernels via sparse projections in function space[C]. Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on. IEEE, 2017: 4671-4675.

[7]Kang Z, Peng C, Cheng Q. Kernel-driven similarity learning[J]. Neurocomputing, 2017, 267: 210-219.

[8]Vapnik V. The nature of statistical learning theory[M]. Springer science & business media, 2013.

[9]Zhou J, Duan B, Huang J, et al. Incorporating prior knowledge and multi-kernel into linear programming support vector regression[J]. Soft Computing, 2015, 19(7): 2047-2061.

[10]Shaobo L, Jinshan Z. Fast Learning With Polynomial Kernels[J]. IEEE Transactions on Cybernetics, 2018:1-13.

[11] Boyd S, Vandenberghe L. Convex optimization[M]. Cambridge university press, 2004.

[12]Maalouf M, Trafalis T B, Adrianto I. Kernel logistic regression using truncated Newton method[J]. Computational management science, 2011, 8(4): 415-428.

[13]Platt J C. 12 fast training of support vector machines using sequential minimal optimization[J]. Advances in kernel methods, 1999: 185-208.

[14] Keerthi S S, Duan K B, Shevade S K, et al. A fast dual algorithm for kernel logistic regression[J]. Machine learning, 2005, 61(1-3): 151-165.

[15] Rossmann T. The average size of the kernel of a matrix and orbits of linear groups[J]. Proceedings of the London Mathematical Society, 2018, 117(3): 574-616.

[16]Ubaru S, Saad Y, Seghouane A K. Fast estimation of approximate matrix ranks using spectral densities[J]. Neural computation, 2017, 29(5): 1317-1351.

[17]Stražar M, Curk T. Learning the kernel matrix via predictive low-rank approximations[OL]. arXiv:1601.04366, 2016.

[18]Phillips J M, Tai W M. Improved coresets for kernel density estimates[C]. Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms. Society for Industrial and Applied Mathematics, 2018: 2718-2727.

[19]Needell D, Srebro N, Ward R. Stochastic gradient descent, weighted sampling, and the randomized Kaczmarz algorithm[J]. Mathematical Programming, 2016, 155(1-2):549-573.

[20]Pourkamali-Anaraki F, Becker S. Randomized Clustered Nystrom for Large-Scale Kernel Machines[OL]. arXiv:1612.06470, 2016.

[21]Zhao W, Du S. Spectral–spatial feature extraction for hyperspectral image classification: A dimension reduction and deep learning approach[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(8): 4544-4554.

[22] Gower R M, Richtárik P. Randomized iterative methods for linear systems[J]. SIAM Journal on Matrix Analysis and Applications, 2015, 36(4): 1660-1690.

[23]Xia J, Falco N, Benediktsson J A, et al. Hyperspectral image classification with rotation random forest via KPCA[J]. IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens, 2017, 10(4): 1601-1609.

[24] Zhu Z B, Song Z H. A novel fault diagnosis system using pattern classification on kernel FDA subspace[J]. Expert Systems with Applications, 2011, 38(6): 6895-6905.

[25]Gönen M, Alpaydın E. Multiple kernel learning algorithms[J]. Journal of machine learning research, 2011, 12(Jul): 2211-2268.

[26]Jain A, Vishwanathan S V N, Varma M. SPF-GMKL: generalized multiple kernel learning with a million kernels[C]. Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2012: 750-758.

[27]Rakotomamonjy A, Bach F R, Canu S, et al. SimpleMKL[J]. Journal of Machine Learning Research, 2008, 9(Nov): 2491-2521.

[28]Cortes C, Mohri M, Rostamizadeh A. Algorithms for learning kernels based on centered alignment[J]. Journal of Machine Learning Research, 2012, 13(Mar): 795-828.

[29]Thiagarajan J J, Ramamurthy K N, Spanias A. Multiple kernel sparse representations for supervised and unsupervised learning[J]. IEEE transactions on Image Processing, 2014, 23(7): 2905-2915.

[30]Lee J H, Lee G M. On optimality conditions and duality theorems for robust semi-infinite multiobjective optimization problems[J]. Annals of Operations Research, 2018, 269(1-2): 419-438.

[31]He L, Li Y, Zhang X, et al. Incremental spectral clustering via fastfood features and its application to stream image segmentation[J]. Symmetry, 2018, 10(7): 272-289.

[32]Xia J, Chanussot J, Du P, et al. Rotation-based support vector machine ensemble in classification of hyperspectral data with limited training samples[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(3): 1519-1531.

[33]Li J, Marpu P R, Plaza A, et al. Generalized composite kernel framework for hyperspectral image classification[J]. IEEE transactions on geoscience and remote sensing, 2013, 51(9): 4816-4829.

[34]Fang L, Li S, Duan W, et al. Classification of hyperspectral images by exploiting spectral–spatial information of superpixel via multiple kernels[J]. IEEE transactions on geoscience and remote sensing, 2015, 53(12): 6663-6674.