References

[1] M. Singla and K. K. Shukla, “Robust statistics-based support vector machine and its variants: a survey,” *Neural Computing and Applications*, 2019, pp. 1–22.

[2] D. S. Fong, L. Aiello, T. W. Gardner, G. L. King, G. Blankenship, J. D. Cavallerano, F. L. Ferris, and R. Klein, “Retinopathy in diabetes,” *Diabetes Care*, 2004, vol. 27, no. suppl 1, pp. s84–s87.

[3] T. Kauppi, V. Kalesnykiene, J. K. Kamarainen, L. Lensu, I. Sorri, J. Pietila, H. Kalviainen, and H. Uusitalo, “DIARETDB1-Standard Diabetic Retinopathy Database,” *Last Modified*, 2007, vol. 6, p. 19.

[4] K. Kourou, T. P. Exarchos, K. P. Exarchos, M. V. Karamouzis, and D. I. Fotiadis, “Machine learning applications in cancer prognosis and prediction,” *Computational and Structural Biotechnology Journal*, 2015, vol. 13, pp. 8–17.

[5] J. A. Nichols, H. W. H. Chan, and M. A. Baker, “Machine learning: applications of artificial intelligence to imaging and diagnosis,” *Biophysical Reviews*, 2019, vol. 11, no. 1, pp. 111–118.

[6] S. Kolouri, S. R. Park, M. Thorpe, D. Slepcev, and G. K. Rohde, “Optimal mass transport: Signal processing and machine-learning applications,” *IEEE Signal Processing Magazine*, 2017, vol. 34, no. 4, pp. 43–59.

[7] J. Vamathevan, D. Clark, P. Czodrowski, I. Dunham, E. Ferran, G. Lee, B. Li, A. Madabhushi, P. Shah, and M. Spitzer, “Applications of machine learning in drug discovery and development,” *Nature Reviews Drug Discovery*, 2019, vol. 18, no. 6, pp. 463–477.

[8] Y. Lei, B. Yang, X. Jiang, F. Jia, N. Li, and A. K. Nandi, “Applications of machine learning to machine fault diagnosis: A review and roadmap,” *Mechanical Systems and Signal Processing*, 2020, vol. 138, p. 106587.

[9] C. Cortes and V. Vapnik, “Support-vector networks,” *Machine Learning*, 1995, vol. 20, no. 3, pp. 273–297.

[10] H. Drucker, C. J. Burges, L. Kaufman, A. Smola, and V. Vapnik, “Support vector regression machines,” *Advances in Neural Information Processing Systems*, 1997, vol. 9, pp. 155–161.

[11] S. Huang, N. Cai, P. P. Pacheco, S. Narrandes, Y. Wang, and W. Xu, “Applications of support vector machine (SVM) learning in cancer genomics,” *Cancer Genomics-Proteomics*, 2018, vol. 15, no. 1, pp. 41–51.

[12] E. Byvatov and G. Schneider, “Support vector machine applications in bioinformatics.” *Applied Bioinformatics*, 2003, vol. 2, no. 2, pp. 67–77.

[13] B. B. Sahoo, R. Jha, A. Singh, and D. Kumar, “Application of support vector regression for modeling low flow time series,” *Journal of Civil Engineering*, 2019, vol. 23, no. 2, pp. 923–934.

[14] C. Venkatesan, P. Karthigaikumar, A. Paul, S. Satheeskumaran, and R. Kumar, “ECG signal preprocessing and SVM classifier-based abnormality detection in remote healthcare applications,” *IEEE Access*, 2018, vol. 6, pp. 9767–9773.

[15] A. Rojas-Domínguez, L. C. Padierna, J. M. C. Valadez, H. J. Puga-Soberanes, and H. J. Fraire, “Optimal hyper-parameter tuning of svm classifiers with application to medical diagnosis,” *IEEE Access*, 2017, vol. 6, pp. 7164–7176.

[16] D. Conforti and R. Guido, “Kernel based support vector machine via semidefinite programming: Application to medical diagnosis,” *Computers & Operations Research*, 2010, vol. 37, no. 8, pp. 1389–1394.

[17] M. Wang and H. Chen, “Chaotic multi-swarm whale optimizer boosted support vector machine for medical diagnosis,” *Applied Soft Computing*, 2020, vol. 88, p. 105946.

[18] J. A. K. Suykens and J. Vandewalle, “Least squares support vector machine classifiers,” *Neural Processing Letters*, 1999, vol. 9, no. 3, pp. 293–300.

[19] C. F. Lin and S. D. Wang, “Fuzzy support vector machines,” *IEEE Transactions on Neural Networks*, 2002, vol. 13, no. 2, pp. 464–471.

[20] C. F. Lin, “Training algorithms for fuzzy support vector machines with noisy data,” *Pattern Recognition Letters*, 2004, vol. 25, no. 14, pp. 1647–1656.

[21] C. T. Lin, S. F. Liang, C. M. Yeh, and K. W. Fan, "Fuzzy neural network design using support vector regression for function approximation with outliers," in *IEEE International Conference on Systems, Man and Cybernetics*, vol. 3. IEEE, 2005, pp. 2763–2768.

[22] Q. Tao and J. Wang, "A new fuzzy support vector machine based on the weighted margin," *Neural Processing Letters*, 2004, vol. 20, no. 3, pp. 139–150.

[23] Y. Wang, S. Wang, and K. K. Lai, "A new fuzzy support vector machine to evaluate credit risk," *IEEE Transactions on Fuzzy Systems*, 2005, vol. 13, no. 6, pp. 820–831.

[24] T.-Y. Wang and H.-M. Chiang, "Fuzzy support vector machine for multi-class text categorization," *Information Processing & Management*, 2007, vol. 43, no. 4, pp. 914–929.

[25] X. Yang, Q. Song, and Y. Wang, "A weighted support vector machine for data classification," *International Journal of Pattern Recognition and Artificial Intelligence*, 2007, vol. 21, no. 05, pp. 961–976.

[26] K. Bennett and A. Demiriz, "Semi-supervised support vector machines," *Advances in Neural Information Processing Systems*, 1999, pp. 368–374.

[27] J. Wang, X. Shen, and W. Pan, "On transductive support vector machines," *Contemporary Mathematics*, 2007, vol. 443, pp. 7–20.

[28] O. L. Mangasarian and E. W. Wild, "Proximal support vector machine classifiers," in *Proceedings in Knowledge Discovery and Data Mining.* Citeseer, 2001.

[29] ——, "Multisurface proximal support vector machine classification via generalized eigenvalues," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, vol. 28, no. 1, pp. 69–74.

[30] M. R. Guarracino, C. Cifarelli, O. Seref, and P. M. Pardalos, "A classification method based on generalized *E*igenvalue problems," *Optimisation Methods and Software*, 2007, vol. 22, no. 1, pp. 73–81.

[31] Y. H. Shao, N. Y. Deng, W. J. Chen, and Z. Wang, "Improved generalized *E*igenvalue proximal support vector machine," *IEEE Signal Processing Letters*, 2012, vol. 20, no. 3, pp. 213–216.

[32] R. Khemchandani, S. Chandra *et al.*, "Twin support vector machines for pattern classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, vol. 29, no. 5, pp. 905–910.

[33] X. Peng, "$TSVR$: an efficient twin support vector machine for regression," *Neural Networks*, 2010, vol. 23, no. 3, pp. 365–372.

[34] D. M. J. Tax, "One-class classification: Concept learning in the absence of counter-examples." 2002.

[35] M. Singla, D. Ghosh, and K. K. Shukla, "A survey of robust optimization based machine learning with special reference to support vector machines," *International Journal of Machine Learning and Cybernetics*, 2019, pp. 1–27.

[36] L. Wang, H. Jia, and J. Li, "Training robust support vector machine with smooth ramp loss in the primal space," *Neurocomputing*, 2008, vol. 71, no. 13-15, pp. 3020–3025.

[37] X. Huang, L. Shi, and J. A. Suykens, "Support vector machine classifier with pinball loss," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, vol. 36, no. 5, pp. 984–997.

[38] A. Le Thi Hoai and P. D. Tao, "Solving a class of linearly constrained indefinite quadratic problems by *DC* algorithms," *Journal of Global Optimization*, 1997, vol. 11, no. 3, pp. 253–285.

[39] Q. Song, W. Hu, and W. Xie, "Robust support vector machine with bullet hole image classification," *IEEE Transactions on Systems, Man, and Cybernetics*, 2002, vol. 32, no. 4, pp. 440–448.

[40] E. M. Jordaan and G. F. Smits, "Robust outlier detection using svm regression," in *IEEE International Joint Conference on Neural Networks*, vol. 3, 2004, pp. 2017–2022.

[41] Y. Zhao and J. Sun, "Robust support vector regression in the primal," *Neural Networks*, 2008, vol. 21, no. 10, pp. 1548–1555.

[42] C. Angulo, X. Parra, and A. Catala, "K-svcr. a support vector machine for multi-class classification," *Neurocomputing*, 2003, vol. 55, no. 1-2, pp. 57–77.

[43] S. M. H. Bamakan, H. Wang, and Y. Shi, “Ramp loss k-support vector classification-regression; a robust and sparse multi-class approach to the intrusion detection problem,” *Knowledge-Based Systems*, 2017, vol. 126, pp. 113–126.

[44] W. Zhu, Y. Song, and Y. Xiao, “A new support vector machine plus with pinball loss,” *Journal of Classification*, 2018, vol. 35, no. 1, pp. 52–70.

[45] J. A. K. Suykens, J. De Brabanter, L. Lukas, and J. Vandewalle, “Weighted least squares support vector machines: robustness and sparse approximation,” *Neurocomputing*, 2002, vol. 48, no. 1-4, pp. 85–105.

[46] C. C. Chuang, S. F. Su, J. T. Jeng, and C. C. Hsiao, “Robust support vector regression networks for function approximation with outliers,” *IEEE Transactions on Neural Networks*, 2002, vol. 13, no. 6, pp. 1322–1330.

[47] D. Liu, Y. Shi, Y. Tian, and X. Huang, “Ramp loss least squares support vector machine,” *Journal of Computational Science*, 2016, vol. 14, pp. 61–68.

[48] C. C. Chuang, “Fuzzy weighted support vector regression with a fuzzy partition,” *IEEE Transactions on Systems, Man, and Cybernetics*, 2007, vol. 37, no. 3, pp. 630–640.

[49] G. H. Lee, J. S. Taur, and C. W. Tao, “A robust fuzzy support vector machine for two-class pattern classification,” *International Journal of Fuzzy Systems*, 2006, vol. 8, no. 2, pp. 76–86.

[50] Q. Wu and R. Law, “Fuzzy support vector regression machine with penalizing gaussian noises on triangular fuzzy number space,” *Expert Systems with Applications*, 2010, vol. 37, no. 12, pp. 7788–7795.

[51] Y. Zhang, F. Xie, D. Huang, and M. Ji, “Support vector classifier based on fuzzy C-means and mahalanobis distance,” *Journal of Intelligent Information Systems*, 2010, vol. 35, no. 2, pp. 333–345.

[52] C. Y. Liu, L. Sun, and Z. J. Zhou, “Weighted support vector machine based on association rules,” in *International Conference on Machine Learning and Cybernetics*, vol. 1. IEEE, 2013, pp. 381–386.

[53] Y. Wu and Y. Liu, “Robust truncated hinge loss support vector machines,” *Journal of the American Statistical Association*, 2007, vol. 102, no. 479, pp. 974–983.

[54] O. Chapelle, "Training a support vector machine in the primal," *Neural Computation*, 2007, vol. 19, no. 5, pp. 1155–1178.

[55] X. Huang, L. Shi, and J. A. K. Suykens, "Ramp loss linear programming support vector machine," *The Journal of Machine Learning Research*, 2014, vol. 15, no. 1, pp. 2185–2211.

[56] R. Collobert, F. Sinz, J. Weston, and L. Bottou, "Trading convexity for scalability," in *Proceedings of the 23rd International Conference on Machine Learning*, 2006, pp. 201–208.

[57] X. Shen, L. Niu, Z. Qi, and Y. Tian, "Support vector machine classifier with truncated pinball loss," *Pattern Recognition*, 2017, vol. 68, pp. 199–210.

[58] K. Wang and P. Zhong, "Robust non-convex least squares loss function for regression with outliers," *Knowledge-Based Systems*, 2014, vol. 71, pp. 290–302.

[59] S. Y. Park and Y. Liu, "Robust penalized logistic regression with truncated loss functions," *Canadian Journal of Statistics*, 2011, vol. 39, no. 2, pp. 300–323.

[60] L. Yang, Z. Ren, Y. Wang, and H. Dong, "A robust regression framework with laplace kernel-induced loss," *Neural Computation*, 2017, vol. 29, no. 11, pp. 3014–3039.

[61] Y. F. Ye, Y. H. Shao, N. Y. Deng, C. N. Li, and X. Y. Hua, "Robust L_p-norm least squares support vector regression with feature selection," *Applied Mathematics and Computation*, 2017, vol. 305, pp. 32–52.

[62] C. Chen, Y. Li, C. Yan, J. Guo, and G. Liu, "Least absolute deviation-based robust support vector regression," *Knowledge-Based Systems*, 2017, vol. 131, pp. 183–194.

[63] H. Cevikalp and V. Franc, "Large-scale robust transductive support vector machines," *Neurocomputing*, 2017, vol. 235, pp. 199–209.

[64] S. Mehrkanoon, X. Huang, and J. A. K. Suykens, "Non-parallel support vector classifiers with different loss functions," *Neurocomputing*, 2014, vol. 143, pp. 294–301.

[65] Y. Xu, Z. Yang, and X. Pan, "A novel twin support-vector machine with pinball loss," *IEEE Transactions on Neural Networks and Learning Systems*, 2016, vol. 28, no. 2, pp. 359–370.

[66] L. Tang, Y. Tian, C. Yang, and P. M. Pardalos, "Ramp-loss non-parallel support vector regression: robust, sparse and scalable approximation," *Knowledge-Based Systems*, 2018, vol. 147, pp. 55–67.

[67] Y. Xiao, H. Wang, and W. Xu, "Ramp loss based robust one-class SVM," *Pattern Recognition Letters*, 2017, vol. 85, pp. 15–20.

[68] H. J. Xing and M. Ji, "Robust one-class support vector machine with rescaled hinge loss function," *Pattern Recognition*, 2018, vol. 84, pp. 152–164.

[69] T. Liu and D. Tao, "Classification with noisy labels by importance reweighting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, vol. 38, no. 3, pp. 447–461.

[70] J. Saeedi, S. M. Ahadi, and K. Faez, "Robust voice activity detection directed by noise classification," *Signal, Image and Video Processing*, 2015, vol. 9, no. 3, pp. 561–572.

[71] A. Roy, J. Singha, S. S. Devi, and R. H. Laskar, "Impulse noise removal using svm classification based fuzzy filter from gray scale images," *Signal Processing*, 2016, vol. 128, pp. 262–273.

[72] V. Vijayalakshmi, M. S. Babu, and R. P. Lakshmi, "Kfcm algorithm for effective brain stroke detection through svm classifier," in *IEEE International Conference on System, Computation, Automation and Networking*. IEEE, 2018, pp. 1–6.

[73] S. Mohdiwale, T. P. Sahu, R. K. Chaurasia, N. K. Nagwani, and S. Verma, "Detection and classification of noise using bark domain features," in *Proceedings of the 6th International Conference on Communications and Broadband Networking*, 2018, pp. 18–21.

[74] S. Gurumurthy, C. Sushama, M. Ramu, and K. S. Nikhitha, "Design and implementation of intelligent system to detect malicious facebook posts using support vector machine (SVM)," in *Soft Computing and Medical Bioinformatics*. Springer, 2019, pp. 17–24.

[75] V. Goel, H. Raj, K. Muthigi, S. S. Kumar, D. Prasad, and V. Nath, "Development of human detection system for security and military applications," in *Proceedings of the Third International Conference on Microelectronics, Computing and Communication Systems*. Springer, 2019, pp. 195–200.

[76] Z. Rustam, J. Pandelaki, and A. Siahaan, "Kernel spherical k-means and support vector machine for acute sinusitis classification," in *IOP Conference Series: Materials Science and Engineering*, vol. 546, no. 5. IOP Publishing, 2019, p. 052011.

[77] R. Wang, W. Li, R. Li, and L. Zhang, "Automatic blur type classification via ensemble svm," *Signal processing: Image Communication*, 2019, vol. 71, pp. 24–35.

[78] G. Xu, Z. Cao, B.-G. Hu, and J. C. Principe, "Robust support vector machines based on the rescaled hinge loss function," *Pattern Recognition*, 2017, vol. 63, pp. 139–148.

[79] T. Yang, M. Mahdavi, R. Jin, and S. Zhu, "An efficient primal dual prox method for non-smooth optimization," *Machine Learning*, 2015, vol. 98, no. 3, pp. 369–406.

[80] B. Van Rooyen, A. K. Menon, and R. C. Williamson, "Learning with symmetric label noise: The importance of being unhinged," *arXiv preprint arXiv:1505.07634*, 2015.

[81] A. Singh, R. Pokharel, and J. Principe, "The C-loss function for pattern classification," *Pattern Recognition*, 2014, vol. 47, no. 1, pp. 441–453.

[82] J. T. Barron, "A general and adaptive robust loss function," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 4331–4339.

[83] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, 2009, vol. 2, no. 1, pp. 183–202.

[84] I. Chivers and J. Sleightholme, "An introduction to algorithms and the big o notation," in *Introduction to Programming with Fortran*. Springer, 2015, pp. 359–364.

[85] A. Das, R. Panda, and A. K. Roy-Chowdhury, "Continuous adaptation of multi-camera person identification models through sparse non-redundant representative selection," *Computer Vision and Image Understanding*, 2017, vol. 156, pp. 66–78.

[86] X. Peng, "Building sparse twin support vector machine classifiers in primal space," *Information Sciences*, 2011, vol. 181, no. 18, pp. 3967–3980.

[87] X. Peng and D. Xu, "Twin mahalanobis distance-based support vector machines for pattern recognition," *Information Sciences*, 2012, vol. 200, pp. 22–37.

[88] Y.-H. Shao, W.-J. Chen, and N.-Y. Deng, "Non-parallel hyperplane support vector machine for binary classification problems," *Information Sciences*, 2014, vol. 263, pp. 22–35.

[89] Q. Ye, C. Zhao, S. Gao, and H. Zheng, "Weighted twin support vector machines with local information and its application," *Neural Networks*, 2012, vol. 35, pp. 31–39.

[90] Z. Qi, Y. Tian, and Y. Shi, "Robust twin support vector machine for pattern classification," *Pattern Recognition*, 2013, vol. 46, no. 1, pp. 305–316.

[91] L.-W. Huang, Y.-H. Shao, J. Zhang, Y.-T. Zhao, and J.-Y. Teng, "Robust rescaled hinge loss twin support vector machine for imbalanced noisy classification," *IEEE Access*, 2019, vol. 7, pp. 65 390–65 404.

[92] M. Tanveer, A. Sharma, and P. N. Suganthan, "General twin support vector machine with pinball loss function," *Information Sciences*, 2019, vol. 494, pp. 311–327.

[93] Z. Wang, Y.-H. Shao, L. Bai, C.-N. Li, L.-M. Liu, and N.-Y. Deng, "Insensitive stochastic gradient twin support vector machines for large scale problems," *Information Sciences*, 2018, vol. 462, pp. 114–131.

[94] H. Pant, M. Sharma, S. Soman *et al.*, "Twin neural networks for the classification of large unbalanced datasets," *Neurocomputing*, 2019, vol. 343, pp. 34–49.

[95] Y. Xu and L. Wang, "A weighted twin support vector regression," *Knowledge-Based Systems*, 2012, vol. 33, pp. 92–101.

[96] Y.-P. Zhao, J. Zhao, and M. Zhao, "Twin least squares support vector regression," *Neurocomputing*, 2013, vol. 118, pp. 225–236.

[97] X. Chen, J. Yang, and L. Chen, "An improved robust and sparse twin support vector regression via linear programming," *Soft Computing*, 2014, vol. 18, no. 12, pp. 2335–2348.

[98] Y.-F. Ye, L. Bai, X.-Y. Hua, Y.-H. Shao, Z. Wang, and N.-Y. Deng, "Weighted lagrange ε-twin support vector regression," *Neurocomputing*, 2016, vol. 197, pp. 53–68.

[99] S. Balasundaram and D. Gupta, "Training lagrangian twin support vector regression via unconstrained convex minimization," *Knowledge-Based Systems*, 2014, vol. 59, pp. 85–96.

[100] P. Anagha, S. Balasundaram, and Y. Meena, "On robust twin support vector regression in primal using squared pinball loss," *Journal of Intelligent & Fuzzy Systems*, 2018, vol. 35, no. 5, pp. 5231–5239.

[101] L. Chen, H. Qu, J. Zhao, B. Chen, and J. C. Principe, "Efficient and robust deep learning with correntropy-induced loss function," *Neural Computing and Applications*, 2016, vol. 27, no. 4, pp. 1019–1031.

[102] D. Solow, "Linear and nonlinear programming," *Wiley Encyclopedia of Computer Science and Engineering*, 2007.

[103] S. Boyd, N. Parikh, and E. Chu, *Distributed optimization and statistical learning via the alternating direction method of multipliers.* Now Publishers Inc, 2011.

[104] "MS Windows 10 kaggle datasets," https://www.kaggle.com/datasets, accessed on: 2019-05-27.

[105] M. Lichman, "UCI machine learning repository," https://archive.ics.uci.edu/ml/index.php, 2013.

[106] "MS Windows 10 libstat dataset," http://lib.stat.cmu.edu/datasets, accessed on: 2019-05-27.

[107] W. R. Rice, "Analyzing tables of statistical tests," *Evolution*, 1989, vol. 43, no. 1, pp. 223–225.

[108] Y. Zheng, S. H. Ley, and F. B. Hu, "Global aetiology and epidemiology of type 2 diabetes mellitus and its complications," *Nature Reviews Endocrinology*, 2018, vol. 14, no. 2, p. 88.

[109] S. Roychowdhury, D. D. Koozekanani, and K. K. Parhi, "*DREAM*: diabetic retinopathy analysis using machine learning," *IEEE Journal of Biomedical and Health Informatics*, 2013, vol. 18, no. 5, pp. 1717–1728.

[110] R. Gargeya and T. Leng, "Automated identification of diabetic retinopathy using deep learning," *Ophthalmology*, 2017, vol. 124, no. 7, pp. 962–969.

[111] A. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piece-wise threshold probing of a matched filter response." in *Proceedings of the AMIA Symposium.* American Medical Informatics Association, 1998, p. 931.

[112] G. Azzopardi, N. Strisciuglio, M. Vento, and N. Petkov, "Vessels delineation in retinal images using *COSFIRE* filters," *Reconstruction*, 2006, vol. 25, no. 9, pp. 1200–1213.

[113] J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. Van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE Transactions on Medical Imaging*, 2004, vol. 23, no. 4, pp. 501–509.

[114] S. Roychowdhury, D. D. Koozekanani, and K. K. Parhi, "Iterative vessel segmentation of fundus images," *IEEE Transactions on Biomedical Engineering*, 2015, vol. 62, no. 7, pp. 1738–1749.

[115] J. Lachure, A. Deorankar, S. Lachure, S. Gupta, and R. Jadhav, "Diabetic retinopathy using morphological operations and machine learning," in *IEEE International Advance Computing Conference.* IEEE, 2015, pp. 617–622.

[116] J. Amin, M. Sharif, and M. Yasmin, "A review on recent developments for detection of diabetic retinopathy," *Scientifica*, 2016, vol. 2016.

[117] D. S. Sisodia, S. Nair, and P. Khobragade, "Diabetic retinal fundus images: preprocessing and feature extraction for early detection of diabetic retinopathy," *Biomedical and Pharmacology Journal*, 2017, vol. 10, no. 2, pp. 615–626.

[118] S. Kumar and B. Kumar, "Diabetic retinopathy detection by extracting area and number of microaneurysm from colour fundus image," in *5th International Conference on Signal Processing and Integrated Networks.* IEEE, 2018, pp. 359–364.

[119] P. Powar and C. Jadhav, "Retinal disease identification by segmentation techniques in diabetic retinopathy," in *Information and Communication Technology for Sustainable Development.* Springer, 2018, pp. 255–265.

[120] R. Raju, "Optimisation of image processing networks for neuronal membrane detection," Ph.D. dissertation, University of Nottingham, 2016.

[121] F. Y. Shih and O. R. Mitchell, "Threshold decomposition of gray-scale morphology into binary morphology," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1989, vol. 11, no. 1, pp. 31–42.

[122] J. B. Saaddine, A. A. Honeycutt, K. M. V. Narayan, X. Zhang, R. Klein, and J. P. Boyle, "Projection of diabetic retinopathy and other major eye diseases among people with diabetes mellitus: United states, 2005-2050," *Archives of Ophthalmology*, 2008, vol. 126, no. 12, pp. 1740–1747.

[123] P. J. Burt, "Fast filter transform for image processing," *Computer Graphics and Image Processing*, 1981, vol. 16, no. 1, pp. 20–51.

[124] V. Vapnik, "24 transductive inference and semi-supervised learning," 2006.

[125] V. Vapnik and A. Sterin, "On structural risk minimization or overall risk in a problem of pattern recognition," *Automation and Remote Control*, 1977, vol. 10, no. 3, pp. 1495–1503.

[126] O. Chapelle, V. Sindhwani, and S. S. Keerthi, "Optimization techniques for semi-supervised support vector machines," *Journal of Machine Learning Research*, 2008, vol. 9, pp. 203–233.

[127] O. Chapelle and A. Zien, "Semi-supervised classification by low density separation." in *AISTATS*, vol. 2005. Citeseer, 2005, pp. 57–64.

[128] X. Lu and D. Chen, "Cancer classification through filtering progressive transductive support vector machine based on gene expression data," in *AIP Conference Proceedings*, vol. 1864, no. 1. AIP Publishing LLC, 2017, p. 020101.

[129] N. Zemmal, N. Azizi, M. Sellami, and N. Dey, "Automated classification of mammographic abnormalities using transductive semi supervised learning algorithm," in *Proceedings of the Mediterranean Conference on Information & Communication Technologies*, vol. 381. Springer, 2016, pp. 657–662.

[130] N. Zemmal, N. Azizi, M. Sellami, D. Zenakhra, S. Cheriguene, N. Dey, and A. S. Ashour, "Robust feature selection algorithm based on transductive SVM wrapper and genetic algorithm: application on computer-aided glaucoma classification,"

International Journal of Intelligent Systems Technologies and Applications, 2018, vol. 17, no. 3, pp. 310–346.

[131] H. Cevikalp, M. Elmas, and S. Ozkan, "Large-scale image retrieval using transductive support vector machines," *Computer Vision and Image Understanding*, 2018, vol. 173, pp. 2–12.

[132] Q. Oliveau and H. Sahbi, "Transductive attributes for ship category recognition," in *2017 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2017, pp. 894–897.

[133] Y. Li, Y. Wang, J. Zhou, and X. Jiang, "Robust transductive support vector machine for multi-view classification," *Journal of Circuits, Systems and Computers*, 2018, vol. 27, no. 12, p. 1850185.

[134] Y. Xu, J. Hua, H. Zhang, R. Hu, X. Huang, J. Liu, and F. Guo, "Improved transductive support vector machine for a small labelled set in motor imagery-based brain-computer interface," *Computational Intelligence and Neuroscience*, 2019, vol. 2019, p. 2087132.

[135] Y. Lin, "A note on margin-based loss functions in classification," *Statistics & Probability Letters*, 2004, vol. 68, no. 1, pp. 73–82.

[136] T. Joachims, "Transductive inference for text classification using support vector machines," in *International Conference on Machine Learning*, vol. 99, 1999, pp. 200–209.

[137] R. Collobert, F. Sinz, J. Weston, and L. Bottou, "Large scale transductive SVMs," *Journal of Machine Learning Research*, 2006, vol. 7, no. Aug, pp. 1687–1712.

[138] Y. Xiao, J. Feng, and B. Liu, "A new transductive learning method with universum data," *Applied Intelligence*, 2021, pp. 1–13.

[139] M. Belkin, P. Niyogi, and V. Sindhwani, "Manifold regularization: A geometric framework for learning from labeled and unlabeled examples." *Journal of Machine Learning Research*, 2006, vol. 7, no. 11.

[140] X. Zheng, L. Zhang, and Z. Xu, "L_1-norm laplacian support vector machine for data reduction in semi-supervised learning," *Neural Computing and Applications*, 2021, pp. 1–18.

[141] X. Li, H. Meng, and X. Peng, "Semi-supervised learning for fault identification in electricity distribution networks," in *Twelfth International Conference on Signal Processing Systems*, vol. 11719. International Society for Optics and Photonics, 2021, p. 117190B.

[142] E. M. Roopa Devi and R. C. Suganthe, "Enhanced transductive support vector machine classification with grey wolf optimizer cuckoo search optimization for intrusion detection system," *Concurrency and Computation: Practice and Experience*, 2020, vol. 32, no. 4, p. e4999.

[143] M. Thiyagaraj and G. Suseendran, "Enhanced prediction of heart disease using particle swarm optimization and rough sets with transductive support vector machines classifier," in *Data Management, Analytics and Innovation*. Springer, 2020, pp. 141–152.

[144] A. L. Yuille and A. Rangarajan, "The concave-convex procedure," *Neural Computation*, 2003, vol. 15, no. 4, pp. 915–936.

[145] L. Bottou, "Large-scale machine learning with stochastic gradient descent," in *Proceedings of COMPSTAT*. Springer, 2010, pp. 177–186.

[146] C.-C. Chang and C.-J. Lin, "*LIBSVM*: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, 2011, vol. 2, no. 3, p. 27.

[147] S. Zeng, Y. Y. Tang, Y. Wei, and Y. Wang, "Algorithm of ε-SVR based on a large-scale sample set: Step-by-step search," *International Journal of Wavelets, Multiresolution and Information Processing*, 2011, vol. 9, no. 02, pp. 197–210.

[148] A. Asuncion and D. Newman, "UCI machine learning repository," 2007.

[149] R. A. Rossi and N. K. Ahmed, "The network data repository with interactive graph analytics and visualization," in *Association for the Advancement of Artificial Intelligence*, 2015.

[150] S. A. Alvarez, "An exact analytical relation among recall, precision, and classification accuracy in information retrieval," *Boston College, Boston, Technical Report*, 2002, pp. 1–22.

[151] G. Rätsch, "Benchmark repository," *Intelligent Data Analysis Group, Fraunhofer-FIRST, Technical Repository*, 2005.

[152] A. Gokaslan, V. Ramanujan, D. Ritchie, K. I. Kim, and J. Tompkin, "Improving shape deformation in unsupervised image-to-image translation," in *Proceedings of the European Conference on Computer Vision*, 2018, pp. 649–665.

[153] H. Li, H. Liu, X. Ji, G. Li, and L. Shi, "Cifar10-dvs: an event-stream dataset for object classification," *Frontiers in Neuroscience*, 2017, vol. 11, p. 309.

[154] A. Ravi, "Pre-trained convolutional neural network features for facial expression recognition," *arXiv preprint arXiv:1812.06387*, 2018.

[155] J. Zhou, X. Yang, L. Zhang, S. Shao, and G. Bian, "Multisignal VGG19 network with transposed convolution for rotating machinery fault diagnosis based on deep transfer learning," *Shock and Vibration*, 2020, vol. 2020.

[156] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley Interdisciplinary Reviews: Computational Statistics*, 2010, vol. 2, no. 4, pp. 433–459.

Appendices

Appendix A

LS-SVM

The optimization problem of LS-SVM solves linear equations using equality constraints. This variant was proposed to ease the implementation of conventional SVM. This extension made use of the square loss function. The optimization problem for LS-SVM is given by [18]

$$\begin{aligned} \min_{w,b,\xi} F(w,b,\xi) \quad & \frac{1}{2}\left\|w\right\|^2 + \frac{\lambda}{2}\sum_{i=1}^{m}\xi_i^2 \\ \text{s.t.} \quad & y_i(w^T x_i + b) = 1 - \xi_i, \;\; i = 1,\dots,m. \end{aligned} \tag{A.1}$$

Langrangian function for this equation is

$$L(w,b,\xi;\alpha) = F(w,b,\xi) - \sum_{i=1}^{m}\alpha_i\{y_i(w^T x_i + b) - 1 + \xi_i\}, \tag{A.2}$$

where α_i is the Lagrangian multiplier.

The KKT optimality conditions for $L(w,b,\xi;\alpha)$ are

$$\begin{aligned} &\frac{\partial L}{\partial w} = 0 \rightarrow w = \sum_{i=1}^{m}\alpha_i y_i x_i \\ &\frac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^{m}\alpha_i y_i = 0 \\ &\frac{\partial L}{\partial \xi_i} = 0 \rightarrow \alpha_i = \lambda\xi_i, \qquad\qquad i = 1,\dots,m \\ &\frac{\partial L}{\partial \alpha_i} = 0 \rightarrow y_i(w^T x_i + b) - 1 + \xi_i = 0, \qquad i = 1,\dots,m. \end{aligned} \tag{A.3}$$

These can be written as the solution to the following linear equations [18]

$$\left[\begin{array}{ccc|c} I & 0 & 0 & -Z^T \\ 0 & 0 & 0 & -Y^T \\ 0 & 0 & \lambda I & -I \\ \hline Z & Y & I & 0 \end{array}\right] \quad \left[\begin{array}{c} w \\ b \\ \xi \\ \hline \alpha \end{array}\right] = \left[\begin{array}{c} 0 \\ 0 \\ 0 \\ \hline \vec{1} \end{array}\right] \tag{A.4}$$

where $Z = [x_1^T y_1, x_2^T y_2, \ldots, x_m^T y_m]$, $Y = [y_1, y_2, \ldots, y_m]$, $\vec{1} = [1, 1, \ldots, 1]$, $\xi = [\xi_1, \xi_2, \ldots, \xi_m]$ and $\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_m]$. The solution can be written as

$$\left[\begin{array}{c|c} 0 & -Y^T \\ \hline Y & ZZ^T + \lambda^{-1} I \end{array}\right] \quad \left[\begin{array}{c} b \\ \hline \alpha \end{array}\right] = \left[\begin{array}{c} 0 \\ \hline \vec{1} \end{array}\right] \tag{A.5}$$

Therefore, the solution of LS-SVM is found by solving the linear set of equation (A.5) instead of quadratic programming [18].

Appendix B

Behaviour of RSVM-PDProx on Synthetic Data Set

In this part of the thesis, the results shown in Subsection 3.5.1 are mentioned in tabular form.

Table B.1: Results of RSVM-PDProx and the Existing Methods on Synthetic Data Sets

Instances	Results	SVM	RSVM-RHHQ (η-val)	RSVM-PDProx					
				η=0.2	η=0.5	η=1	η=2	η=3	Parameter
100		92.00±4.01	93.00±2.76	95.33±3.39	96.63±3.34	**96.66±1.63**	94.66±1.63	96.00±2.49	10^{-2}
500	Accuracy	96.53±1.29	95.66±0.68	96.93±1.55	96.26±0.90	97.00±1.23	**97.28±1.22**	96.80±0.88	10^{-2}
1000	(0%	94.26±1.21	95.33±0.93	95.53±0.45	95.60±0.70	95.39±1.16	95.26±0.77	**96.06±0.67**	10^{-3}
5000	Noise)	96.61±0.28	96.70±0.54	**97.85±0.46**	97.26±1.26	97.00±0.42	97.33±0.16	96.66±0.22	10^{-2}
10000		96.04±0.30	96.15±0.35	96.18±0.17	**97.26±0.31**	96.13±0.36	96.66±1.22	96.43±0.72	10^{-2}
100		91.33±8.58	91.33±5.99	93.33±2.98	96.00±1.33	**96.10±4.80**	94.06±3.39	94.16±3.88	10^{-4}
500	Accuracy	95.60±1.08	96.00±1.42	96.73±0.90	**97.06±1.43**	96.96±0.59	96.80±1.48	96.53±1.65	10^{-2}
1000	(15%	94.16±0.53	95.10±1.16	**95.53±1.34**	95.46±0.80	95.00±0.96	95.16±0.82	95.06±1.54	10^{-4}
5000	Noise)	96.56±0.43	96.65±0.35	96.77±0.31	96.66±0.56	96.73±0.81	**97.23±0.24**	96.20±0.31	10^{-4}
10000		96.04±0.10	96.17±0.32	96.20±0.23	95.90±0.25	**96.40±0.37**	95.56±1.34	96.33±1.56	10^{-2}
100		89.33±4.42	92.00±3.95	92.00±3.39	**95.66±2.66**	95.34±1.63	94.00±2.49	94.00±1.33	10^{-4}
500	Accuracy	94.93±1.32	95.00±1.68	95.20±1.14	96.40±0.90	**96.93±0.90**	96.40±1.61	95.20±2.77	10^{-4}
1000	(30%	94.73±1.62	94.56±0.59	**95.60±0.57**	95.33±0.61	95.00±1.01	95.46±0.54	95.26±0.94	10^{-4}
5000	Noise)	96.36±0.33	96.48±0.28	**96.69±0.51**	96.40±0.16	95.93±0.86	96.31±0.32	95.96±0.85	10^{-4}
10000		96.16±0.31	95.24±0.30	96.19±0.48	95.76±0.32	96.13±0.37	**96.33±0.42**	96.00±0.82	10^{-2}

From Table B.1, it is observed that for majority of the cases, accuracy increases with the increase in number of instances. As training data increases, model gets adequate amount of instances to train itself, ultimately leads to the increase in accuracy.

However, when the number of features increase, accuracies almost remains same. For the synthetic data set with 1000 instances, accuracies with features 2, 5, 10, 100 and 100 are computed. Accuracies for all these cases are nearly same, i.e. 96%.

Appendix C

Analytical Proof showing Rescaled α-hinge loss is Superior than Rescaled Hinge Loss Function

Rescaled hinge loss function is given by

$$l_C(z) = \beta\left[1 - e^{\left(-\frac{l_{\text{hinge}(z)}}{2\sigma^2}\right)}\right]. \tag{C.1}$$

and rescaled α-hinge loss function is given by

$$l_{\alpha\text{-rhinge}}(z) = \beta\left[1 - e^{\left(-\eta l_{\alpha\text{-hinge}}(z)\right)}\right], \tag{C.2}$$

where $\eta = \frac{1}{2\sigma^2} \geq 0$ is a scaling constant and $\beta = \left(1 - e^{(-\eta)}\right)^{-1}$.

Note that

$$\begin{aligned} l_{\text{hinge}}(z) &= \max\{0, 1 - y(w^T x + b)\} \qquad \text{(C.3)} \\ &\geq \alpha(1 - y(w^T x + b)) \qquad \text{for any} \quad \alpha \in [0,1] \\ \text{Therefore,} \qquad \max\{0, 1 - y(w^T x + b)\} &\geq \max_{\alpha\in[0,1]} \alpha(1 - y(w^T x + b)) \\ \text{i.e.} \quad l_{\text{hinge}}(z) &\geq l_{\alpha\text{-hinge}}(z) \\ \implies -\eta l_{\alpha\text{-hinge}}(z) &\geq -\eta l_{\text{hinge}}(z) \qquad (\because \eta > 0) \\ \implies e^{-\eta l_{\alpha\text{-hinge}}(z)} &\geq e^{-\eta l_{\text{hinge}}(z)} \\ \implies \beta[1 - e^{-\eta l_{\alpha\text{-hinge}}(z)}] &\leq \beta[1 - e^{-\eta l_{\text{hinge}}(z)}] \\ \implies l_{\alpha\text{-rhinge}}(z) &\leq l_C(z) \end{aligned}$$

Appendix D

Results of Different Variants of SVM on DIARETDB1 Data Set

In Table D.1, different SVM variants discussed in Chapter 5 are compared on the basis of accuracy, sensitivity and specificity. From the table, it is observed that TWSVM with pinball loss is performing better than the rest of the techniques. The trade-off between sensitivity and specificity is also less in TWSVM with pinball loss function.

Table D.1: Comparison of Different SVM Variants On DIARETDB1 Data Set

	Results	**Linear SVM**	**TWSVM with Hinge Loss**	**TWSVM with Pinball Loss**
DIARETDB1 with noise	Accuracy (in %)	92.57±2.52	77.37±6.42	**94.72±4.61**
	Sensitivity	83.96±0.05	65.29±7.11	99.99±0.01
	Specificity	88.23±0.02	75.98±0.02	96.62±0.17
	Time (in seconds)	6.48	1.98	**0.78**
DIARETDB1 without noise	Accuracy (in %)	97.25±2.42	94.27±2.21	98.32±1.29
	Sensitivity	89.92±0.01	95.78±0.15	99.99±0.01
	Specificity	88.23±0.02	88.47±0.02	84.31±0.32
	Time (in seconds)	4.46	1.73	**1.09**

These methods were also compared on Messidor data set. Results are shown in Table D.2.

Table D.2: Comparison of Different SVM Variants On Messidor Data Set

	Results	**Linear SVM**	**TWSVM with Hinge Loss**	**TWSVM with Pinball Loss**
Messidor with noise	Accuracy (in %)	68.78±1.91	56.86±0.07	**69.47±1.82**
	Sensitivity	77.40±0.03	55.04±0.02	82.47±1.21
	Specificity	60.78±0.03	68.71±0.11	72.41±0.03
	Time (in seconds)	0.76	**0.20**	0.22
Messidor without noise	Accuracy (in %)	69.97±1.89	67.31±1.24	**71.56±2.61**
	Sensitivity	75.63±0.03	98.16±1.26	90.42±2.31
	Specificity	64.78±0.03	42.62±2.67	69.63±0.01
	Time (in seconds)	0.72	0.19	**0.14**

From Table D.2, it can be observed that TWSVM with pinball loss yields better accuracy, sensitivity and specificity as compared to the rest of the techniques. Please note that the training time corressponding to TWSVM with pinball loss function is also less when the data set was noise-free. However, on addition of noise, the training time of TWSVM with pinball loss function is comparable with the conventional TWSVM.

Appendix E

Comparison of the proposed Method and the Existing methods on COVID-19 data set

Table E.1 shows that the proposed approach, $\overline{\text{pin}}$-TSVM has performed better than SVM, TSVM and ramp-TSVM in terms of accuracy, sensitivity and specificity except two cases where specificity of TSVM is more than the rest of the techniques.

Table E.1: Comparison of $\overline{\text{pin}}$-TSVM and the Existing Methods over COVID-19 Data Set

Dats Sets	Noise Level	Results	SVM	TSVM	ramp-TSVM	$\overline{\text{pin}}$-TSVM	C
COVID-19 Data Set with PCA	0% Noise	Accuracy	93.78	94.69	93.12	**96.57**	(2,3,1,4)
		Sensitivity	96.84	96.89	96.08	**98.82**	
		Specificity	97.12	97.11	95.52	**97.80**	
	10% Noise	Accuracy	92.15	92.97	91.64	**94.75**	(2,3,1,1)
		Sensitivity	95.21	95.62	94.04	**96.91**	
		Specificity	89.12	92.10	88.23	**89.96**	
	15% Noise	Accuracy	88.47	92.97	91.79	**94.28**	(2,3,1,1)
		Sensitivity	92.62	95.61	94.08	**96.78**	
		Specificity	90.01	91.82	87.82	**94.23**	
	30% Noise	Accuracy	68.74	90.72	92.81	**93.51**	(2,3,1,3)
		Sensitivity	80.54	93.01	95.55	**96.18**	
		Specificity	82.43	**90.89**	86.86	90.14	
	40% Noise	Accuracy	68.18	90.32	90.07	**92.05**	(2,2,1,4)
		Sensitivity	80.01	92.01	89.99	**95.00**	
		Specificity	80.47	**90.04**	84.72	89.09	

LIST OF PUBLICATIONS

Refereed Journal Papers

1. Singla M, Shukla KK. Robust statistics-based support vector machine and its variants: a survey. Neural Computing and Applications. 2019 Dec 2:1-22.

2. Singla M, Ghosh D, Shukla KK. A survey of robust optimization based machine learning with special reference to support vector machines. International Journal of Machine Learning and Cybernetics. 2019 Dec 23:1-27.

3. Singla M, Ghosh D, Shukla KK, Pedrycz W. Robust twin support vector regression based on rescaled Hinge loss. Pattern Recognition. 2020 Sep 1;105:107395.

4. Singla M, Ghosh D, Shukla K K. Improved Sparsity of Support Vector Machine with Robustness Towards Label Noise Based on Rescaled α-hinge Loss with Non-smooth Regularizer. Neural Processing Letters. 2020 Dec;52(3):2211-39.

5. Singla M, Ghosh D, Shukla KK. $\overline{\text{pin}}$-TSVM: A Robust Transductive Support Vector Machine and its Application to the Detection of COVID-19 Infected Patients. Neural Processing Letters. 2021 Dec;53(6):3981-4010.

Refereed Conference Papers

1. Singla M, Soni S, Saini P, Chaudhary A, Shukla KK. Diabetic retinopathy detection using twin support vector machines. InAdvances in Bioinformatics, Multimedia, and Electronics Circuits and Signals 2020 (pp. 91-104). Springer, Singapore.

2. Singla M, Shukla KK. Experimental Evaluation of Nature-Inspired Algorithms on High Dimensions. InProceedings of 2nd International Conference on Communication, Computing and Networking 2019 (pp. 615-625). Springer, Singapore.