$\alpha_i > 0$ are called support vectors. These are the only examples that contribute to the classifier.

4. Conclusions

Data Mining has two major goals as sate below

- To generate descriptive models to solve problems.
- To generate predictive model to solve problems.

Descriptive models have already been developed in statistical methodology. It is predictive model building that caught attention because this is the activity that involves both supervised and unsupervised learning. Supervised learning is a generalization of many statistical methods. However, classical statistical methods do not have a separate test data set. The concept of dividing the available data into training, validation and test sets came about only due to machines learning and Data Mining. As consequences, predictive model is are being used more and more in marketing, insurance, banking, manufacturing, supply chain management, customer relation management, and so on. This paper has avoided a discussion. On linear and logistic regressions because they are classical statistical models. SVMs are discussed in the detail because they are the gift of Data Mining to statistics.

References

[1] Abe, N., Pednault, E.,Wang, H., Zadrozny, B., Fan,W., and Apte, C. (2002). Empirical comparison of various reinforcement learning strategies for sequential targeted marketing. In Proceedings of the IEEE International Conference on Data Mining, pages 3–10. IEEE.
[2] Ernst, D., Geurts, P., and Wehenkel, L. (2005). Tree-based batch mode reinforcement learning. Journal of Machine Learning Research, 6(1):503–556.
[3] Huang, J., Smola, A., Gretton, A., Borgwardt, K. M., and Sch¨olkopf, B. (2006). Correcting sample selection bias by unlabeled data. In Proceedings of the Neural Information Processing Systems Conference (NIPS 2006).
[4] Jonas, J. and Harper, J. (2006). Effective counterterrorism and the limited role of predictive data mining. Technical report, Cato Institute.
[5] Michie, D., Spiegelhalter, D. J., and Taylor, C. C. (1994). Machine Learning, Neural and Statistical Classification. Ellis Horwood.
[6] Murphy, S. A. (2005). A generalization error for Q-learning. Journal of Machine Learning Research, 6:1073–1097.
[7] Neumann, G. (2008). Batch-mode reinforcement learning for continuous state spaces: A survey. O¨GAI Journal, 27(1):15–23.
[8] Powell, W. B. (2007). Approximate Dynamic Programming. John Wiley & Sons, Inc.
[9] Riedmiller, M. (2005). Neural fitted Q iteration–first experiences with a data efficient neural reinforcement learning method. In Proceedings of the 16th European Conference on Machine Learning (ECML), pages 317–328.
[10] Simester, D. I., Sun, P., and Tsitsiklis, J. N. (2006). Dynamic catalog mailing policies. Management Science, 52(5):683–696.
[11] Tang, L. and Liu, H. (2009a). Relational learning via latent social dimensions. In Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 817–826. ACM.
[12] Tang, L. and Liu, H. (2009b). Scalable learning of collective behavior based on sparse social dimensions. In Proceeding of the 18th ACM Conference on Information and Knowledge Management (CIKM), pages 1107–1116. ACM.
[13] Vert, J.-P. and Jacob, L. (2008). Machine learning for in silico virtual screening and chemical genomics: New strategies. Combinatorial Chemistry & High Throughput Screening, 11(8):677–685(9).
[14] Viviani, P. and Flash, T. (1995). Minimum-jerk, two thirds power law, and isochrony: converging approaches to movement planning. Journal of Experimental Psychology, 21:32–53.
[15] Yu, V. (2007). Approximate dynamic programming for blood inventory management. Honors thesis, Princeton University.
[16] Zhu, X. and Goldberg, A. B. (2009). Introduction to semi-supervised learning. Synthesis Lectures on Artificial Intelligence and Machine Learning, 3(1):1–130