

Bibliography

- [1] Abe, N. & Mamitsuka, H. (1997). *Machine Learning*, **29** (2), 275–301.
- [2] Agrawal, R. & Srikant, R. (1994). In: *Proceedings of the 20th International Conference on Very Large Data Bases VLDB '94* pp. 487–499, San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.
- [3] Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). *Journal of Molecular Biology*, **215** (3), 403 – 410.
- [4] Altschul, S. F., Madden, T. L., Schffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). *Nucleic Acids Research*, **25** (17), 3389–3402.
- [5] Ana, L. N. F. & Jain, A. K. (2003). In: *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings.* volume 2 pp. II–128–II–133 vol.2, Washington, DC, USA: IEEE Computer Society.
- [6] Avis, D. & Fukuda, K. (1996). *Discrete Applied Mathematics*, **65** (1), 21 – 46.
- [7] Barabasi, A. L. & Oltvai, Z. N. (2004). *Nat. Rev. Genet.* **5** (2), 101–113.
- [8] Basu, S. (2005). *Semi-supervised Clustering: Probabilistic Models, Algorithms and Experiments*. PhD thesis University of Texas at Austin Austin, TX, USA.
- [9] Belkin, M. & Niyogi, P. (2003). *Neural Comput.* **15** (6), 1373–1396.
- [10] Bento, A. P., Gaulton, A., Hersey, A., Bellis, L. J., Chambers, J., Davies, M., Krüger, F. A., Light, Y., Mak, L., McGlinchey, S., Nowotka, M., Papadatos, G., Santos, R., & Overington, J. P. (2014). *Nucleic Acids Research*, **42** (D1), D1083–D1090.
- [11] Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). *J. Mach. Learn. Res.* **3**, 993–1022.
- [12] Borg, I. & Groenen, P. (1997). *Modern Multidimensional Scaling: Theory and Applications*. Springer series in statistics. Berlin, Heidelberg: Springer.
- [13] Borgwardt, K. M. & Kriegel, H.-P. (2005). In: *Proceedings of the Fifth IEEE International Conference on Data Mining ICDM '05* pp. 74–81, Washington, DC, USA: IEEE Computer Society.
- [14] Breiman, L. (2001). *Machine Learning*, **45** (1), 5–32.
- [15] Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1984). *Classification and Regression Trees*. Monterey, CA: Wadsworth and Brooks.

- [16] Burges, C., Shaked, T., Renshaw, E., Lazier, A., Deeds, M., Hamilton, N., & Hullender, G. (2005). In: *Proceedings of the 22Nd International Conference on Machine Learning ICML '05* pp. 89–96, New York, NY, USA: ACM.
- [17] Cadez, I., Heckerman, D., Meek, C., Smyth, P., & White, S. (2003). *Data Mining and Knowledge Discovery*, **7** (4), 399–424.
- [18] Cai, X., Bazerque, J. A., & Giannakis, G. B. (2013). *PLoS Computational Biology*, **9** (5), 1–13.
- [19] Chapelle, O., Metzler, D., Zhang, Y., & Grinspan, P. (2009). In: *Proceedings of the 18th ACM Conference on Information and Knowledge Management CIKM '09* pp. 621–630, New York, NY, USA: ACM.
- [20] Cheng, J. & Greiner, R. (1999). In: *Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence UAI'99* pp. 101–108, San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.
- [21] Choi, H. & Baraniuk, R. G. (2001). *IEEE Transactions on Image Processing*, **10** (9), 1309–1321.
- [22] Chomsky, N. (1956). *IRE Transactions on Information Theory*, **2**, 113–124.
- [23] Chung, F. R. K. (1997). *Spectral Graph Theory*. Providence, RI USA: American Mathematical Society.
- [24] Cover, T. M. & Thomas, J. A. (2006). *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. New York, NY, USA: Wiley-Interscience.
- [25] Dhillon, I. S., Guan, Y., & Kulis, B. (2004). In: *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining KDD '04* pp. 551–556, New York, NY, USA: ACM.
- [26] Diligenti, M., Frasconi, P., & Gori, M. (2003). *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **25** (4), 519–523.
- [27] Drozdetskiy, A., Cole, C., Procter, J., & Barton, G. J. (2015). *Nucleic Acids Res.* **43** (W1), W389–394.
- [28] duVerle, D. A. & Mamitsuka, H. (2012). *Brief. Bioinformatics*, **13** (3), 337–349.
- [29] Eddy, S. R. (1995). *Proc Int Conf Intell Syst Mol Biol*, **3**, 114–120.
- [30] Eisen, M. B., Spellman, P. T., Brown, P. O., & Botstein, D. (1998). *Proceedings of the National Academy of Sciences*, **95** (25), 14863–14868.
- [31] Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., Potter, S. C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G. A., Tate, J., & Bateman, A. (2016). *Nucleic Acids Res.* **44** (D1), D279–285.
- [32] Friedman, J., Hastie, T., & Tibshirani, R. (2000). *Ann. Statist.* **28** (2), 337–407.
- [33] Grauman, K. & Darrell, T. (2007). *J. Mach. Learn. Res.* **8**, 725–760.

- [34] Griffiths, T. L. & Steyvers, M. (2004). *Proceedings of the National Academy of Sciences*, **101** (suppl 1), 5228–5235.
- [35] Hagen, L. & Kahng, A. B. (2006). *Trans. Comp.-Aided Des. Integ. Cir. Sys.* **11** (9), 1074–1085.
- [36] Han, J., Pei, J., & Yin, Y. (2000). *SIGMOD Rec.* **29** (2), 1–12.
- [37] Hand, D. J. & Till, R. J. (2001). *Machine Learning*, **45** (2), 171–186.
- [38] Hanley, J. A. & McNeil, B. J. (1982). *Radiology*, **143** (1), 29–36.
- [39] Hashimoto, K., Aoki-Kinoshita, K. F., Ueda, N., Kanehisa, M., & Mamitsuka, H. (2006). In: *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining KDD '06* pp. 177–186, New York, NY, USA: ACM.
- [40] Hashimoto, K., Aoki-Kinoshita, K. F., Ueda, N., Kanehisa, M., & Mamitsuka, H. (2008). *ACM Trans. Knowl. Discov. Data*, **2** (1), 6:1–6:30.
- [41] Hauser, A. S., Chavali, S., Masuho, I., Jahn, L. J., Martemyanov, K. A., Gloriam, D. E., & Babu, M. M. (2018). *Cell*, **172** (1), 41 – 54.e19.
- [42] Hoerl, A. E. & Kennard, R. W. (1970). *Technometrics*, **12** (1), 55–67.
- [43] Hofmann, T. (1999). In: *Proceedings of the 22Nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval SIGIR '99* pp. 50–57, New York, NY, USA: ACM.
- [44] Hopcroft, J. E., Motwani, R., & Ullman, J. D. (2006). *Introduction to Automata Theory, Languages, and Computation (3rd Edition)*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc.
- [45] Horton, P., Park, K. J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C. J., & Nakai, K. (2007). *Nucleic Acids Res.* **35** (Web Server issue), W585–587.
- [46] Hubert, L. & Arabie, P. (1985). *Journal of Classification*, **2** (1), 193–218.
- [47] Jaccard, P. (1912). *New Phytologist*, **11** (2), 37–50.
- [48] Jebara, T., Kondor, R., & Howard, A. (2004). *J. Mach. Learn. Res.* **5**, 819–844.
- [49] Karasuyama, M. & Mamitsuka, H. (2013a). *IEEE Transactions on Neural Networks and Learning Systems*, **24** (12), 1999–2012.
- [50] Karasuyama, M. & Mamitsuka, H. (2013b). In: *Advances in Neural Information Processing Systems 26*, (Burges, C. J. C., Bottou, L., Welling, M., Ghahramani, Z., & Weinberger, K. Q., eds) pp. 1547–1555. Curran Associates, Inc.
- [51] Karasuyama, M. & Mamitsuka, H. (2017). *Machine Learning*, **106** (2), 307–335.
- [52] Katoh, K. & Standley, D. M. (2013). *Molecular Biology and Evolution*, **30** (4), 772–780.

- [53] Kim, S., Thiessen, P. A., Bolton, E. E., Chen, J., Fu, G., Gindulyte, A., Han, L., He, J., He, S., Shoemaker, B. A., Wang, J., Yu, B., Zhang, J., & Bryant, S. H. (2016). *Nucleic Acids Res.* **44** (D1), D1202–1213.
- [54] Koller, D. & Friedman, N. (2009). *Probabilistic Graphical Models: Principles and Techniques - Adaptive Computation and Machine Learning*. Cambridge, MA, USA: The MIT Press.
- [55] Kvalseth, T. O. (1987). *IEEE Transactions on Systems, Man, and Cybernetics*, **17** (3), 517–519.
- [56] Lafferty, J. D., McCallum, A., & Pereira, F. C. N. (2001). In: *Proceedings of the Eighteenth International Conference on Machine Learning ICML '01* pp. 282–289, San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.
- [57] LeCun, Y. & an Geoffrey Hinton, Y. B. (2015). *Nature*, **521**, 436–444.
- [58] Leemis, L. M. & McQueston, J. T. (2008). *The American Statistician*, **62** (1), 45–53.
- [59] Leslie, C., Eskin, E., & Noble, W. S. (2002). *Pac. Symp. Biocomput.* , 564–575.
- [60] Lodhi, H., Saunders, C., Shawe-Taylor, J., Cristianini, N., & Watkins, C. (2002). *J. Mach. Learn. Res.* **2**, 419–444.
- [61] Luxburg, U. (2007). *Statistics and Computing*, **17** (4), 395–416.
- [62] Mamitsuka, H. (2006). *Pattern Recognition*, **39** (12), 2393–2404.
- [63] Mamitsuka, H., Okuno, Y., & Yamaguchi, A. (2003). *SIGKDD Explor. Newsl.* **5** (2), 113–121.
- [64] Mitchell, T. M. (2017). <http://www.cs.cmu.edu/~tom/NewChapters.html>.
- [65] Neuhaus, M. & Bunke, H. (2006). *Pattern Recognition*, **39** (10), 1852 – 1863.
- [66] Ng, A. Y., Jordan, M. I., & Weiss, Y. (2001). In: *Proceedings of the 14th International Conference on Neural Information Processing Systems: Natural and Synthetic NIPS'01* pp. 849–856, Cambridge, MA, USA: MIT Press.
- [67] Nguyen, C. H. & Mamitsuka, H. (2011). *IEEE Transactions on Neural Networks*, **22** (9), 1395–1405.
- [68] Pei, J., Han, J., Mortazavi-Asl, B., Pinto, H., Chen, Q., Dayal, U., & Hsu, M. (2001). In: *Proceedings of the 17th International Conference on Data Engineering* pp. 215–224, Washington, DC, USA: IEEE Computer Society.
- [69] Ramon, J. & Gärtner, T. (2003). In: *Proceedings of the First International Workshop on Mining Graphs, Trees and Sequences* pp. 65–74 : No Publisher.
- [70] Rand, W. M. (1971). *Journal of the American Statistical Association*, **66** (336), 846–850.
- [71] Rigoutsos, I. & Floratos, A. (1998). *Bioinformatics*, **14** (1), 55–67.
- [72] Rissanen, J. (1978). *Automatica*, **14** (5), 465–471.

- [73] Rogers, D. J. & Tanimoto, T. T. (1960). *Science*, **132** (3434), 1115–1118.
- [74] Roweis, S. T. & Saul, L. K. (2000). *SCIENCE*, **290**, 2323–2326.
- [75] Schapire, R. E. (1990). *Machine Learning*, **5** (2), 197–227.
- [76] Schapire, R. E. & Freund, Y. (2012). *Boosting: Foundations and Algorithms*. Cambridge, MA, USA: The MIT Press.
- [77] Schapire, R. E. & Singer, Y. (1999). *Machine Learning*, **37** (3), 297–336.
- [78] Shawe-Taylor, J. & Cristianini, N. (2004). *Kernel Methods for Pattern Analysis*. New York, NY, USA: Cambridge University Press.
- [79] Shervashidze, N., Vishwanathan, S. V. N., Petri, T., Mehlhorn, K., & Borgwardt, K. M. (2009). *Journal of Machine Learning Research - Proceedings Track*, **5**, 488–495.
- [80] Shi, J. & Malik, J. (2000). *IEEE Trans. Pattern Anal. Mach. Intell.* **22** (8), 888–905.
- [81] Shiga, M. & Mamitsuka, H. (2011). *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, **1** (6), 496–511.
- [82] Shiga, M. & Mamitsuka, H. (2012). *Pattern Recognition*, **45** (3), 1035–1049.
- [83] Shiga, M., Takigawa, I., & Mamitsuka, H. (2007). In: *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining KDD* pp. 647–656, New York, NY, USA: ACM.
- [84] Srikant, R. & Agrawal, R. (1996). In: *Proceedings of the 5th International Conference on Extending Database Technology: Advances in Database Technology EDBT '96* pp. 3–17, London, UK, UK: Springer-Verlag.
- [85] Strehl, A. & Ghosh, J. (2003). *J. Mach. Learn. Res.* **3**, 583–617.
- [86] Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., & Mesirov, J. P. (2005). *Proceedings of the National Academy of Sciences*, **102** (43), 15545–15550.
- [87] Takahashi, K., Takigawa, I., & Mamitsuka, H. (2013). *PLoS ONE*, **8** (12), e82890.
- [88] Takigawa, I. & Mamitsuka, H. (2011). *Machine Learning*, **82** (2), 95–121.
- [89] Takigawa, I., Tsuda, K., & Mamitsuka, H. (2011). *PLoS ONE*, **6** (2), e16999.
- [90] TheGeneOntologyConsortium (2017). *Nucleic Acids Research*, **45** (D1), D331–D338.
- [91] Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). *Nucleic Acids Research*, **22** (22), 4673–4680.
- [92] Vinh, N. X., Epps, J., & Bailey, J. (2010). *J. Mach. Learn. Res.* **11**, 2837–2854.
- [93] Vishwanathan, S. V. N., Borgwardt, K. M., & Schraudolph, N. N. (2006). In: *Proceedings of the 19th International Conference on Neural Information Processing Systems NIPS'06* pp. 1449–1456, Cambridge, MA, USA: MIT Press.

- [94] Vishwanathan, S. V. N. & Smola, A. J. (2002). In: *Proceedings of the 15th International Conference on Neural Information Processing Systems NIPS'02* pp. 585–592, Cambridge, MA, USA: MIT Press.
- [95] Vishwanathan, S. V. N. & Smola, A. J. (2004). In: *Kernel Methods in Computational Biology* pp. 113–130. MIT Press Cambridge, MA, USA.
- [96] Wagner, S. & Wagner, D. (2007). Technical Report 2006-04 Universität Karlsruhe (TH).
- [97] Wishart, D. S., Feunang, Y. D., Guo, A. C., Lo, E. J., Marcu, A., Grant, J. R., Sajed, T., Johnson, D., Li, C., Sayeeda, Z., Assempour, N., Iynkkaran, I., Liu, Y., Maciejewski, A., Gale, N., Wilson, A., Chin, L., Cummings, R., Le, D., Pon, A., Knox, C., & Wilson, M. (2018). *Nucleic Acids Research*, **46** (D1), D1074–D1082.
- [98] Xing, E. P., Jordan, M. I., & Karp, R. M. (2001). In: *Proceedings of the Eighteenth International Conference on Machine Learning ICML '01* pp. 601–608, San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.
- [99] Yamaguchi, A., Aoki, K. F., & Mamitsuka, H. (2004). *Information Processing Letters*, **92** (2), 57 – 63.
- [100] Yan, X. & Han, J. (2002). In: *Proceedings of the 2002 IEEE International Conference on Data Mining ICDM '02* pp. 721–724, Washington, DC, USA: IEEE Computer Society.
- [101] Yao, Y. Y. (2003). *Information-Theoretic Measures for Knowledge Discovery and Data Mining* pp. 115–136. Berlin, Heidelberg: Springer Berlin Heidelberg.
- [102] Yeung, K. Y. & Ruzzo, W. L. (2001). *Bioinformatics*, **17** (9), 763–774.
- [103] Yotsukura, S., Karasuyama, M., Takigawa, I., & Mamitsuka, H. (2017). *Briefings in Bioinformatics*, **18** (4), 619–633.
- [104] Yu, S., Tranchevent, L.-C., Moor, B., & Moreau, Y. (2013). *Kernel-based Data Fusion for Machine Learning: Methods and Applications in Bioinformatics and Text Mining*. Berlin, Heidelberg: Springer Publishing Company, Incorporated.
- [105] Zhang, L., Udaka, K., Mamitsuka, H., & Zhu, S. (2012). *Brief. Bioinformatics*, **13** (3), 350–364.
- [106] Zhou, D., Bousquet, O., Lal, T. N., Weston, J., & Schölkopf, B. (2003). In: *Proceedings of the 16th International Conference on Neural Information Processing Systems NIPS'03* pp. 321–328, Cambridge, MA, USA: MIT Press.