Appendix of Table 1

　Table 1 には，代表的な機械学習手法の原理・アルゴリズムが詳しく解説されている総説論文や原著論文を列挙した。その他，俯瞰的に様々な機械学習法を解説した教科書として Bishop (2006)，Hastie, Tibshirani, & Friedman (2009)，Murphy (2012)なども挙げられる。以下に，それらの書誌情報を記載する。

Bank, D., Koenigstein, N., & Giryes, R. (2020). Autoencoders. *ArXiv, 2003.05991*. Retrieved from http://arxiv.org/abs/2003.05991

Bell, A. J., & Sejnowski, T. J. (1995). An Information-Maximization Approach to Blind Separation and Blind Deconvolution. *Neural Computation, 7*(6), 1129–1159. https://doi.org/10.1162/neco.1995.7.6.1129

Biau, G. (2012). Analysis of a Random Forests Model. *Journal of Machine Learning Research, 13*(1), 1063–1095.

Bishop, C. (2006). *Pattern Recognition and Machine Learning*. Berlin, Heidelberg: Springer-Verlag New York.

Blömer, J., Lammersen, C., Schmidt, M., & Sohler, C. (2016). Theoretical Analysis of the k-Means Algorithm – A Survey. In L. Kliemann & P. Sanders (Eds.), *Algorithm Engineering. Lecture Notes in Computer Science* (Vol. 9220 LNCS, pp. 81–116). Springer. https://doi.org/10.1007/978-3-319-49487-6_3

Breiman, L. (2001). Random forests. *Machine Learning, 45*, 5–32. https://doi.org/10.1023/A:1010933404324

Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L., & Lopez, A. (2020). A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing, 408*, 189–215. https://doi.org/10.1016/j.neucom.2019.10.118

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning, 20*(3), 273–297. https://doi.org/10.1007/BF00994018

Cox, D. R. (1958). The Regression Analysis of Binary Sequences Published by : Wiley for the Royal Statistical Society Stable URL : http://www.jstor.org/stable/2983890. *Journal of the Royal Statistical Society, 20*(2), 215–242.

Drucker, H., Surges, C. J. C., Kaufman, L., Smola, A., & Vapnik, V. (1997). Support vector regression machines. *Advances in Neural Information Processing Systems*, *1*, 155–161.

Fomby, T. B., Johnson, S. R., & Hill, R. C. (1984). Review of Ordinary Least Squares and Generalized Least Squares. In T. B. Fomby, S. R. Johnson, & R. C. Hill (Eds.), *Advanced Econometric Methods* (pp. 7–25). New York, NY: Springer New York. https://doi.org/10.1007/978-1-4419-8746-4_2

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, *29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning*. New York, NY: Springer New York. https://doi.org/10.1007/978-0-387-84858-7

Hinton, G. E. (2006). Reducing the Dimensionality of Data with Neural Networks. *Science*, *313*(5786), 504–507. https://doi.org/10.1126/science.1127647

Hyvarinen, A. (1999). Survey on independent component analysis. *Neural Computing Surveys*, *2*, 94–128.

Jiang, L., Wang, D., Cai, Z., & Yan, X. (2007). Survey of Improving Naive Bayes for Classification. In R. Alhajj, H. Gao, J. Li, X. Li, & O. R. Zaïane (Eds.), *Advanced Data Mining and Applications* (pp. 134–145). Berlin, Heidelberg: Springer Berlin Heidelberg.

Lachenbruch, P. A., & Goldstein, M. (1979). Discriminant Analysis. *Biometrics*, *35*(1), 69–85. https://doi.org/10.2307/2529937

Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature, 401*(6755), 788–791. https://doi.org/10.1038/44565

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, *1*, 281–297.

Melnykov, V., & Maitra, R. (2010). Finite mixture models and model-based clustering. *Statistics Surveys*, *4*, 80–116. https://doi.org/10.1214/09-SS053

Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective*. The MIT Press.

Murtagh, F., & Contreras, P. (2012). Algorithms for hierarchical clustering: an

overview. *WIREs Data Mining and Knowledge Discovery*, *2*(1), 86–97. https://doi.org/10.1002/widm.53

Natekin, A., & Knoll, A. (2013). Gradient boosting machines, a tutorial. *Frontiers in Neurorobotics*, *7*, 21. https://doi.org/10.3389/fnbot.2013.00021

Peterson, L. (2009). K-nearest neighbor. *Scholarpedia*, *4*(2), 1883. https://doi.org/10.4249/scholarpedia.1883

Rosipal, R., & Krämer, N. (2006). Overview and Recent Advances in Partial Least Squares. In C. Saunders, M. Grobelnik, S. Gunn, & J. Shawe-Taylor (Eds.), *Subspace, Latent Structure and Feature Selection* (pp. 34–51). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/11752790_2

Saeed, N., Nam, H., Haq, M. I. U., & Muhammad Saqib, D. B. (2018). A Survey on Multidimensional Scaling. *ACM Computing Surveys*, *51*(3), 1–25. https://doi.org/10.1145/3178155

Safavian, S. R., & Landgrebe, D. (1991). A survey of decision tree classifier methodology. *IEEE Transactions on Systems, Man, and Cybernetics*, *21*(3), 660–674. https://doi.org/10.1109/21.97458

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, *61*, 85–117. https://doi.org/10.1016/j.neunet.2014.09.003

Shlens, J. (2014). A Tutorial on Principal Component Analysis. *ArXiv*, 1404.1100. Retrieved from http://arxiv.org/abs/1404.1100

Wang, Y.-X., & Zhang, Y.-J. (2013). Nonnegative Matrix Factorization: A Comprehensive Review. *IEEE Transactions on Knowledge and Data Engineering*, *25*(6), 1336–1353. https://doi.org/10.1109/TKDE.2012.51