Tutorial Reference List for TC304 Student Contest

1   Probability Theory

[1]   Sivia, D.S., and Skilling, J. (2006). *Data Analysis: A Bayesian Tutorial.* Oxford University Press, New York. **(Google citations: 2535)**

[2]   Ang, A.H.-S., and Tang, W.H. (2007). *Probability concepts in engineering, Vol. I Emphasis on applications to civil and environmental engineering, 2nd Ed*. Wiley, New York. **(Google citations: 1186)**

2   Introductory Materials on Machine Learning

[3]   Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer-Verlag New York. **(Google citations: 29990)**

[4]   Barber, D. (2012). *Bayesian Reasoning and Machine Learning*. Cambridge University Press. **(Google citations: 875)**

3   Supervised learning

3.1   Decision tree learning

[5]   Breiman, Leo, Friedman, J. H., Olshen, R. A. and Stone, C. J. (1984). *Classification and regression trees.* Monterey, CA: Wadsworth and Brooks/Cole Advanced Books and Software**.** **(Google citations: 37373)**

[6]   James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). An Introduction to Statistical Learning (Vol. 112). New York: springer. **(Google citations: 2673)**

3.2   Artificial neural networks

[7]   McCulloch, W. S., and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4), 115-133. **(Google citations: 15445)**

[8]   Ripley, B. D. (1996). *Pattern recognition and neural networks*. Cambridge University Press. **(Google citations: 7483)**

3.3   Deep learning

[9]   Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, 61, 85-117. **(Google citations: 2870)**

[10] LeCun, Y., Bengio, Y., Hinton, G. (2015). Deep Learning. *Nature,* 521, 436-444. **(Google citations: 6899)**

3.4   Inductive logic programming

[11] Muggleton, S., and De Raedt, L. (1994). Inductive Logic Programming: Theory and methods. *The Journal of Logic Programming*, 19–20, 629–679. **(Google citations: 1708)**

3.5   Support vector machines

[12] Chang, C. C., and Lin, C. J. (2011). LIBSVM: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology. 2 (3). **(Google citations: 36612)**

3.6   Similarity and metric learning

[13] Davis, J. V., Kulis, B., Jain, P., Sra, S., and Dhillon, I. S. (2007). Information-theoretic metric learning. *International Conference in Machine Learning (ICML)*: 209–216. **(Google citations: 1580)**

3.7   Sparse dictionary learning

[14] Kreutz-Delgado, K., Murray, J. F., Rao, B. D., Engan, K., Lee, T. W., and Sejnowski, T. J. (2003). Dictionary learning algorithms for sparse representation**.** *Neural Computation*, 15(2), 349-396. **(Google citations: 714)**

[15] Aharon, M., Elad, M., and Bruckstein, A. (2006). rmK-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. IEEE Transactions on Signal Processing, 54(11), 4311-4322. **(Google citations: 6842)**

4   Semi-supervised learning

[16] Chapelle, O., Scholkopf, B., and Zien, A. (2006). *Semi-Supervised Learning*, MIT Press, Cambridge, MA. **(Google citations: 3929)**

5   Reinforcement learning

[17] Sutton, R. S., and Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge: MIT press. **(Google citations: 26015)**

6   Unsupervised learning

6.1   Clustering

[18] Han, J., Pei, J., and Kamber, M. (2011). Data mining: concepts and techniques. Elsevier. **(Google citations: 39607)**

6.2   Feature learning (Dimensionality reduction)

[19] Kohonen, T. (1998). The self-organizing map. *Neurocomputing*, 21(1), 1-6. **(Google citations: 23430)**

[20] Jolliffe, I. T. (1986). Principal Component Analysis and Factor Analysis. In *Principal Component Analysis* (pp. 115-128). Springer New York. **(Google citations: 33332)**

6.3   Outlier detection

[21] Breunig, M. M., Kriegel, H. P., Ng, R. T., and Sander, J. (2000). LOF: identifying density-based local outliers. In ACM sigmod record, 29(2), 93-104). ACM. **(Google citations: 3563)**

6.4   Generative adversarial networks

[22] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., & Bengio, Y. (2014). Generative adversarial nets. *Advances in Neural Information Processing Systems,* 27 (NIPS 2014): 2672-2680. **(Google citations: 2829)**

7   Bayesian machine learning

7.1   Bayesian network

[23] Heckerman, D., Geiger, D., and Chickering, D. M. (1995). Learning Bayesian networks: The combination of knowledge and statistical data. Machine Learning, 20(3), 197-243. **(Google citations: 4819)**

7.2   Bayesian neural network

[24] Neal, R. M. (2012). Bayesian learning for neural networks (Vol. 118). Springer Science and Business Media**. (Google citations: 2915)**

7.3   Gaussian processes

[25] Rasmussen, C. E., and Williams, C. K. (2006). Gaussian processes for machine learning (Vol. 1). Cambridge: MIT press. **(Google citations: 1530)**

7.4   Relevance vector machine (sparse Bayesian learning)

[26] Tipping, M. E. (2001). Sparse Bayesian learning and the relevance vector machine. *Journal of Machine Learning Research*, 1(Jun), 211-244. **(Google citations: 4949)**

7.5   Bayesian deep learning

[27] Gal, Y., and Ghahramani, Z. (2016). Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. In *International Conference on Machine Learning* (pp. 1050-1059). **(Google citations: 302)**

7.6   Bayesian model class selection and system identification

[28] Burnham, K. P., and Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods and Research*, 33(2), 261-304. **(Google citations: 4977)**

[29] Yuen, K.V. (2010). Bayesian Methods for Structural Dynamics and Civil Engineering. John Wiley & Sons, ISBN: 978-0-470-82454-2. **(Google citations: 262)**

7.7   Simulation-based methods for Bayesian inference

[30] MacKay, D.J.C. (1998). Introduction to Monte Carlo methods. In M. Jordan, editor, Learning in graphical models. MIT Press. **(Google citations: 589)**