REFERENCES

1. Witten IH, Frank E, Hall MA, Pal CJ, 2016, Data Mining: Practical machine learning tools and techniques, Morgan Kaufmann.

2. Han, Jiawei, Jian Pei, Micheline Kamber, 2011, Data mining: concepts and techniques, Elsevier.

3. Kantardzic, M., 2003, Data mining - Concepts, models, methods and algorithms, Wiley –IEEE Press.

4. Kesavaraj, G. and Sukumaran, S., 2013, A study on classification techniques in data mining. In Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), IEEE, pp. 1-7.

5. Fu, Tak-chung, 2011, A review on time series data mining, Engineering Applications of Artificial Intelligence, Vol.24, pp. 164-181.

6. Gorunescu, F., 2011. Data Mining: Concepts, models and techniques, Springer Science & Business Media, Vol. 12.

7. Xu, Rui & Wunsch, Donald. 2005, Survey of Clustering Algorithms, IEEE Transactions on Neural Networks, Vol.16, pp. 645 - 678.

8. Saxena, Amit, et al., 2017, A review of clustering techniques and developments, Neurocomputing 267, pp. 664-681.

9. Jain, A. K., 2010, Data clustering: 50 years beyond K-means. Pattern recognition letters, 31(8), pp. 651-666.

10. Zhang, C. and Zhang, S., 2002, Association rule mining: models and algorithms, Springer-Verlag.

11. Cooley, R., Mobasher, B. and Srivastava, J., 1997, Web Mining: Information and Pattern Discovery on the World Wide Web. In ICTAI, Vol. 97, pp. 558-567.

12. Mennis, J. and Guo, D., 2009, Spatial data mining and geographic knowledge discovery - An introduction, Computers, Environment and Urban Systems, 33(6), pp.403-408.

13. Gupta, V. and Lehal, G.S., 2009, A survey of text mining techniques and applications, Journal of emerging technologies in web intelligence, 1(1), pp.60-76.

14. Pujari, A.K., 2001, Data mining techniques, Universities press.

15. Cook, D.J. and Holder, L.B. eds., 2006, Mining graph data. John Wiley & Sons.

16. Michie, D., Spiegelhalter, D.J. and Taylor, C.C., 1994, Machine learning. Neural and Statistical Classification, 13.

17. Bhavsar, H., & Ganatra, A. 2012, A comparative study of training algorithms for supervised machine learning. International Journal of Soft Computing and Engineering (IJSCE), Vol. 2(4), pp. 2231-2307.

18. Kotsiantis, Sotiris B., I. Zaharakis, and P. Pintelas, 2007, Supervised machine learning: A review of classification techniques, Emerging artificial intelligence applications in computer engineering, Vol. 160, pp. 3-24.

19. Rokach, L. and Maimon, O.Z., 2008, Data mining with decision trees: theory and applications, Vol. 69, World scientific.

20. Scholkopf, B. and Smola, A.J., 2001, Learning with kernels: Support vector machines, regularization, optimization, and beyond, MIT press.

21. Xuegong, Z., 2000, Introduction to statistical learning theory and support vector machines, Acta Automatica Sinica, Vol. 26(1), pp.32-42.

22. Patil, T.R. and Sherekar, S.S., 2013. Performance analysis of Naive Bayes and J48 classification algorithm for data classification. International journal of computer science and applications, Vol. 6(2), pp.256-261.

23. Zurada, J.M., 1992, Introduction to artificial neural systems, St. Paul: West publishing company, Vol. 8.

24. Smith, M., 1993, Neural networks for statistical modeling, New York: Van Nostrand Reinhold, Vol. 115.

25. Goodfellow, I., Bengio, Y. and Courville, A., 2016, Deep learning, MIT press.

26. LeCun, Y., Bengio, Y. and Hinton, G., 2015, Deep learning. Nature, Vol. 521(7553), pp.436.

27. Redei, G.P., 2008, Encyclopedia of genetics, genomics, proteomics, and informatics. Springer Science & Business Media.

28. Xia, X., 2018, Bioinformatics and the cell: modern computational approaches in genomics, proteomics and transcriptomics, Springer.

29. Geschwind, D.H. and Levitt, P., 2007, Autism spectrum disorders: developmental disconnection syndromes. Current opinion in neurobiology, Vol. 17(1), pp.103-111.

30. Johnson, C.P. and Myers, S.M., 2007, Identification and evaluation of children with autism spectrum disorders. Pediatrics, Vol. 120(5), pp.1183-1215.

31. Matson, J.L., Kozlowski, A.M., Hattier, M.A., Horovitz, M. and Sipes, M., 2012, DSM-IV vs DSM-5 diagnostic criteria for toddlers with autism. Developmental Neurorehabilitation, Vol. 15(3), pp.185-190.

32. Lauritsen, M.B., 2013, Autism spectrum disorders. European child & adolescent psychiatry, Vol. 22(1), pp.37-42.

33. Geschwind, D.H. and Levitt, P., 2007, Autism spectrum disorders: developmental disconnection syndromes. Current opinion in neurobiology, Vol. 17(1), pp.103-111.

34. Sztainberg, Y. and Zoghbi, H.Y., 2016, Lessons learned from studying syndromic autism spectrum disorders. Nature neuroscience, Vol. 19(11), p.1408.

35. Jones, R.M. and Lord, C., 2013, Diagnosing autism in neurobiological research studies. Behavioural brain research, Vol. 251, pp.113-124.

36. Hormozdiari, F., Penn, O., Borenstein, E. and Eichler, E.E., 2015, The discovery of integrated gene networks for autism and related disorders. Genome research, Vol. 25(1), pp.142-154.

37. Jamal, W., Das, S., Oprescu, I.A., Maharatna, K., Apicella, F. and Sicca, F., 2014, Classification of autism spectrum disorder using supervised learning of brain connectivity measures extracted from synchrostates, Journal of neural engineering, Vol. 11(4), pp.046019.

38. Latkowski, T. and Osowski, S., 2015, Computerized system for recognition of autism on the basis of gene expression microarray data. Computers in biology and medicine, Vol. 56, pp.82-88

39. Liu, L., Lei, J., Sanders, S.J., Willsey, A.J., Kou, Y., Cicek, A.E., Klei, L., Lu, C., He, X., Li, M. and Muhle, R.A., 2014, DAWN: a framework to identify autism genes and subnetworks using gene expression and genetics. Molecular autism, Vol. 5(1), p.22.

40. Krishnan, A., Zhang, R., Yao, V., Theesfeld, C.L., Wong, A.K., Tadych, A., Volfovsky, N., Packer, A., Lash, A. and Troyanskaya, O.G., 2016, Genome-wide prediction and functional characterization of the genetic basis of autism spectrum disorder. Nature neuroscience, Vol. 19(11), p.1454.

41. Kou, Y., Betancur, C., Xu, H., Buxbaum, J.D. and Ma'Ayan, A., 2012, May. Network-and attribute-based classifiers can prioritize genes and pathways for autism spectrum disorders and

intellectual disability, American Journal of Medical Genetics Part C: Seminars in Medical Genetics, Vol. 160(2), pp. 130-142.

42. Jiao, Y., Chen, R., Ke, X., Chu, K., Lu, Z. and Herskovits, E.H., 2010. Predictive models of autism spectrum disorder based on brain regional cortical thickness. Neuroimage, Vol. 50(2), pp.589-599.

43. Brunak S, Engelbrecht J, Knudsen S. 1991, Prediction of human mRNA donor and acceptor sites from the DNA sequence, J. Mol. Biol, Vol. 220(1), pp:49–65.

44. Hebsgaard SM, Korning PG, Tolstrup N, Engelbrecht J, Rouz´e P, Brunak S, 1996, Splice site prediction in Arabidopsis thaliana pre-mRNA by combining local and global sequence information, Nucleic Acids Res., Vol. 24(17), pp.3439–52.

45. Fakoor, R., Ladhak, F., Nazi, A. and Huber, M., 2013, June. Using deep learning to enhance cancer diagnosis and classification. In Proceedings of the international conference on machine learning, New York, USA: ACM, Vol. 28.

46. Tan, J., Ung, M., Cheng, C. and Greene, C.S., 2014, Unsupervised feature construction and knowledge extraction from genome-wide assays of breast cancer with denoising autoencoders, In Pacific Symposium on Biocomputing Co-Chairs, pp. 132-143.

47. Danaee P, Ghaeini R, Hendrix DA, 2016, A deep learning approach for cancer detection and relevant gene identification, Pac Symp Biocomput, Vol. 22, pp. 219–229.

48. Singh, V., Baranwal, N., Sevakula, R.K., Verma, N.K. and Cui, Y., 2016, Layerwise feature selection in Stacked Sparse Auto-Encoder for tumor type prediction, IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 1542-1548.

49. Zhang, J.M. and Kamath, G.M., 2016, Learning the Language of the Genome using RNNs.

50. Quang, D. and Xie, X., 2016, DanQ: a hybrid convolutional and recurrent deep neural network for quantifying the function of DNA sequences, Nucleic acids research, Vol. 44(11), pp. e100-e107.

51. Yang, B., Liu, F., Ren, C., Ouyang, Z., Xie, Z., Bo, X. and Shu, W., 2017, BiRen: predicting enhancers with a deep-learning-based model using the DNA sequence alone, Bioinformatics, Vol. 33(13), pp.1930-1936.

52. Shen, Z., Bao, W. and Huang, D.S., 2018, Recurrent Neural Network for Predicting Transcription Factor Binding Sites. Scientific reports, Vol. 8(1), pp.15270.

53. Goodfellow, I., Lee, H., Le, Q.V., Saxe, A. and Ng, A.Y., 2009, Measuring invariances in deep networks, In Advances in neural information processing systems, pp. 646-654.

54. Agostinelli, F., Hoffman, M., Sadowski, P. and Baldi, P., 2014, Learning activation functions to improve deep neural networks, arXiv preprint arXiv: 1412.6830.

55. Lauzon, F.Q., 2012, An introduction to deep learning, In 11th International Conference on Information Science, Signal Processing and their Applications (ISSPA), pp. 1438-1439.

56. Buduma, N. and Locascio, N., 2017, Fundamentals of deep learning: Designing next-generation machine intelligence algorithms, O'Reilly Media, Inc.

57. Canziani, A., Paszke, A. and Culurciello, E., 2016, An analysis of deep neural network models for practical applications, arXiv preprint arXiv: 1605.07678.

58. Bengio, Y., 2012, June. Deep learning of representations for unsupervised and transfer learning. In Proceedings of ICML workshop on unsupervised and transfer learning, pp. 17-36.

59. Shin, H.C., Roth, H.R., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D. and Summers, R.M., 2016, Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE transactions on medical imaging, Vol. 35(5), pp.1285-1298.

60. Oquab, M., Bottou, L., Laptev, I. and Sivic, J., 2014, Learning and transferring mid-level image representations using convolutional neural networks. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1717-1724.

61. Larochelle, H. and Bengio, Y., 2008, Classification using discriminative restricted Boltzmann machines. In Proceedings of the 25th international conference on Machine learning ACM, pp. 536-543.

62. Hinton, G.E., 2012, A practical guide to training restricted Boltzmann machines. In Neural networks: Tricks of the trade, Springer, pp. 599-619.

63. Fischer, A. and Igel, C., 2012, An introduction to restricted Boltzmann machines. In Iberoamerican Congress on Pattern Recognition, Springer, pp. 14-36.

64. Lee, H., Grosse, R., Ranganath, R. and Ng, A.Y., 2009, Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations. In Proceedings of the 26th annual international conference on machine learning, ACM, pp. 609-616.

65. Le Roux, N. and Bengio, Y., 2008, Representational power of restricted Boltzmann machines and deep belief networks. Neural computation, Vol. 20(6), pp.1631-1649.

66. Lipton, Z.C., Berkowitz, J. and Elkan, C., 2015, A critical review of recurrent neural networks for sequence learning, arXiv preprint arXiv:1506.00019

67. Zaremba, W., Sutskever, I. and Vinyals, O., 2014, Recurrent neural network regularization, arXiv preprint arXiv: 1409.2329.

68. Boden, M., 2002, A guide to recurrent neural networks and backpropagation. The Dallas project.

69. Ranzato, M.A., Chopra, S., Auli, M. and Zaremba, W., 2015, Sequence level training with recurrent neural networks, arXiv preprint arXiv:1511.06732.

70. Sutskever, I., Martens, J. and Hinton, G.E., 2011. Generating text with recurrent neural networks. In Proceedings of the 28th International Conference on Machine Learning (ICML-11), pp. 1017-1024.

71. Hochreiter, S., Bengio, Y., Frasconi, P. and Schmidhuber, J., 2001, Gradient flow in recurrent nets: the difficulty of learning long-term dependencies.

72. Jozefowicz, R., Zaremba, W. and Sutskever, I., 2015, An empirical exploration of recurrent network architectures. In International Conference on Machine Learning, pp. 2342-2350.

73. Hochreiter, S. and Schmidhuber, J., 1997, LSTM can solve hard long time lag problems. In Advances in neural information processing systems, pp. 473-479

74. Sundermeyer, M., Schlüter, R. and Ney, H., 2012, LSTM neural networks for language modeling. In Thirteenth annual conference of the international speech communication association.

75. Lipton, Z.C., Kale, D.C., Elkan, C. and Wetzel, R., 2015, Learning to diagnose with LSTM recurrent neural networks. arXiv preprint arXiv:1511.03677.

76. Chung, J., Gulcehre, C., Cho, K. and Bengio, Y., 2014, Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv preprint arXiv:1412.3555.

77. Irie, K., Tüske, Z., Alkhouli, T., Schlüter, R. and Ney, H., 2016, September. LSTM, GRU, Highway and a Bit of Attention: An Empirical Overview for Language Modeling in Speech Recognition. In Interspeech, pp. 3519-3523.

78. Golmohammadi, M., Ziyabari, S., Shah, V., Von Weltin, E., Campbell, C., Obeid, I. and Picone, J., 2017, Gated recurrent networks for seizure detection, In IEEE Signal Processing in Medicine and Biology Symposium (SPMB), pp. 1-5.

79. Stenson et al, 2003, The Human Gene Mutation Database (HGMD): 2003 Update. Hum Mutat., Vol. 21, pp. 577-581.

80. Banerjee-Basu, Sharmila, and Alan Packer, 2010, SFARI Gene: an evolving database for the autism research community. Disease Models & Mechanisms Vol. 3, pp. 133-135, Web. https://gene.sfari.org/

81. Geer LY, Marchler-Bauer A, Geer RC, Han L, He J, He S, Liu C, Shi W, Bryant SH, 2010, The NCBI BioSystems database, Nucleic Acids Res, Vol. 38 (Database issue): D492-6.

82. Online Mendelian Inheritance in Man, OMIM, McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University, Web URL: https://omim.org/

83. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J, 1990, Basic local alignment search tool, J. Mol. Biol., Vol. 215, pp. 403-410.

84. Altschul, S.F., 1991, Amino acid substitution matrices from an information theoretic perspective, J. Mol. Biol, Vol. 219, pp. 555-565.

85. Henikoff, S. and Henikoff, J.G., 1992, Amino acid substitution matrices from protein blocks, Proceedings of the National Academy of Sciences, Vol. 89(22), pp.10915-10919.

86. G Ojeme, Blessing and Audrey Mbogho, 2016, Selecting Learning Algorithms for Simultaneous Identification of Depression and Comorbid Disorders, KES.

87. Johan Brodin, 2016, Working with emotions: Recommending subjective labels to music tracks using machine learning

88. Julio H. Zaragoza, L. Enrique Sucar Eduardo F. Morales, Concha Bielza Pedro Larranaga, 2011, Bayesian Chain Classifiers for Multidimensional Classification Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence, pp. 2192-2197.

89. Peter R. de Waal and Linda C. van der Gaag, 2007, Inference and learning in multi-dimensional bayesian network classifiers, Lecture Notes in Artificial Intelligence, Vol. 4724, pp. 501–511.

90. C. Bielza, G. Li, and P. Larranaga, 2011, Multi-dimensional classification with Bayesian networks, Int. J. Approx. Reason., Vol. 52(6), pp.705–727.

91. J.D. Rodríguez, J.A. Lozano, 2008, Multi-objective learning of multi-dimensional Bayesian classifiers, Proceedings of the Eighth International Conference on Hybrid Intelligent Systems, pp. 501–506.

92. L. van der Gaag, P.R. de Waal, 2006, Multi-dimensional Bayesian network classifiers, in: Third European Conference on Probabilistic Graphical Models, pp. 107–114.

93. L. Tenenboim, L. Rokach, B. Shapira, 2010, Identification of label dependencies for multi-label classification, Proceedings of the 2nd International Workshop on Learning from Multi-Label Data (MLD'10), pp. 53–60

94. G. Tsoumakas, I. Katakis, I. Vlahavas, 2008, Effective and efficient multilabel classification in domains with large number of labels, Proceedings of ECML/PKDD 2008 Workshop on Mining Multidimensional Data (MMD'08), pp. 30–44.

95. J. Read, B. Pfahringer. G. Holmes, 2008, Multi-label classification using ensembles of pruned sets, Proceedings of the Eighth IEEE International Conference on Data Mining (ICDM'08), IEEE Computer Society, pp. 995-1000

96. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P, 2015, The Molecular Signatures Database (MSigDB) hallmark gene set collection, Cell Syst, Vol.1(6), pp. 417-425, url:http://www.broadinstitute.org/gsea/msigdb/index.jsp

97. Srivastava, R.K., Greff, K. and Schmidhuber, J., 2015. Training very deep networks. In Advances in neural information processing systems, pp. 2377-2385.

98. Schuster, M. and Paliwal, K.K., 1997, Bidirectional recurrent neural networks, IEEE Transactions on Signal Processing, Vol. 45(11), pp.2673-2681

99. Berglund, M., Raiko, T., Honkala, M., Karkkainen, L., Vetek, A. and Karhunen, J.T., 2015, Bidirectional recurrent neural networks as generative models. In Advances in Neural Information Processing Systems, pp. 856-864.

100. Sak,H., Senior, Beaufays, 2014, Long short-term memory based recurrent neural network architectures for large vocabulary speech recognition. arXiv preprint arXiv:1402.1128

101. Dey, R. and Salemt, F.M., 2017, Gate-variants of gated recurrent unit (GRU) neural networks. In IEEE 60th international midwest symposium on circuits and systems (MWSCAS), pp. 1597-1600.