

References

- [1] W. Fan and A. Bifet, "Mining big data: current status, and forecast to the future," *ACM SIGKDD Explor. Newsl.*, vol. 14, no. 2, pp. 1–5, 2013.
- [2] "Big data analytics: What it is and why it matters | SAS." [Online]. Available: http://www.sas.com/en_us/insights/analytics/big-data-analytics.html. [Accessed: 14-Mar-2016].
- [3] J. Han and M. Kamber, *Data mining: concepts and techniques*, 3rd ed. Burlington, MA: Elsevier, 2011.
- [4] U. M. Fayyad, Ed., *Advances in knowledge discovery and data mining*. Menlo Park, Calif.: AAAI Press [u.a.], 1996.
- [5] S.-H. Liao, P.-H. Chu, and P.-Y. Hsiao, "Data mining techniques and applications – A decade review from 2000 to 2011," *Expert Syst. Appl.*, vol. 39, no. 12, pp. 11303–11311, Sep. 2012.
- [6] G. C. Nsofor, "A Comparative Analysis of Predictive Data-Mining Techniques," in *A Thesis Presented for the Master of Science Degree, The University of Tennessee, Knoxville*, 2006.
- [7] I.-C. Yeh and C. Lien, "The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients," *Expert Syst. Appl.*, vol. 36, no. 2, pp. 2473–2480, Mar. 2009.
- [8] C. Phua, V. Lee, K. Smith, and R. Gayler, "A comprehensive survey of data mining-based fraud detection research," *ArXiv Prepr. ArXiv10096119*, 2010.
- [9] A. Sharma and P. K. Panigrahi, "A review of financial accounting fraud detection based on data mining techniques," *ArXiv Prepr. ArXiv13093944*, 2013.
- [10] A. F. Atiya, "Bankruptcy prediction for credit risk using neural networks: A survey and new results," *Neural Netw. IEEE Trans. On*, vol. 12, no. 4, pp. 929–935, 2001.
- [11] V. Bapat and A. Nagale, "Comparison of Bankruptcy Prediction Models: Evidence from India," *Account. Finance Res.*, vol. 3, no. 4, Oct. 2014.
- [12] M. Alam, C. Hao, and K. Carling, "Review of the literature on credit risk modeling: development of the past 10 years," *Banks Bank Syst.*, vol. 5, no. 3, pp. 43–60, 2010.

- [13] L. K. Sheng and T. Y. Wah, "A comparative study of data mining techniques in predicting consumers' credit card risk in banks," *Afr. J. Bus. Manag.*, vol. 5, no. 20, p. 8307, 2011.
- [14] "Non-performing Loans Increase In Non-Banking, Leasing Sectors | The Sunday Leader." [Online]. Available: <http://www.thesundayleader.lk/2015/06/21/non-performing-loans-increase-in-non-banking-leasing-sectors/>. [Accessed: 04-Mar-2016].
- [15] "Non-Performing Loan Npl Definition from Financial Times Lexicon." [Online]. Available: http://lexicon.ft.com/Term?term=non_performing-loan--NPL. [Accessed: 04-Mar-2016].
- [16] "12.ar08_part_III_e.pdf." [Online]. Available: http://www.cbsl.gov.lk/pics_n_docs/10_pub/_docs/efr/annual_report/ar2008e/12.ar08_part_III_e.pdf. [Accessed: 04-Mar-2016].
- [17] W.-S. Chen and Y.-K. Du, "Using neural networks and data mining techniques for the financial distress prediction model," *Expert Syst. Appl.*, vol. 36, no. 2, pp. 4075–4086, Mar. 2009.
- [18] N. Uddin, "Consumer Credit Customers' Financial Distress Prediction by Using Two-Group Discriminant Analysis: A Case Study," *Int. J. Econ. Finance*, vol. 5, no. 6, May 2013.
- [19] S. DATA, "DATA MINING TECHNIQUES FOR STRUCTURED AND," stanford university, 2000.
- [20] H. A. Bekhet and S. F. K. Eletter, "Credit risk assessment model for Jordanian commercial banks: Neural scoring approach," *Rev. Dev. Finance*, vol. 4, no. 1, pp. 20–28, Jan. 2014.
- [21] E. Kirkos and Y. Manolopoulos, "Data mining in finance and accounting: a review of current research trends," in *Proceedings of the 1st international conference on enterprise systems and accounting (ICESAcc)*, 2004, pp. 63–78.
- [22] M. Ahmed and M. E. Mollah, "Data Mining Research Challenges in E-Commerce."
- [23] Two Crows Corporation, *Introduction to data mining and knowledge discovery*. Potomac, MD: Two Crows Corp., 1999.
- [24] "data mining | computer science | Britannica.com." [Online]. Available: <http://global.britannica.com/technology/data-mining>. [Accessed: 27-Feb-2016].

- [25] W. A. Aldana, "Data mining industry: emerging trends and new opportunities," Massachusetts Institute of Technology, Dept. of Electrical Engineering and Computer Science, 2000.
- [26] C. Edeki and S. Pandya, "Comparative Study of Data Mining and Statistical Learning Techniques for Prediction of Cancer Survivability," *Mediterr. J. Soc. Sci.*, 2012.
- [27] G. C. Nsofor, "A Comparative Analysis of Predictive Data-Mining Techniques," in *A Thesis Presented for the Master of Science Degree, The University of Tennessee, Knoxville*, 2006.
- [28] L. Tapak, H. Mahjub, O. Hamidi, and J. Poorolajal, "Real-Data Comparison of Data Mining Methods in Prediction of Diabetes in Iran," *Healthc. Inform. Res.*, vol. 19, no. 3, p. 177, 2013.
- [29] "Data mining | Better Evaluation." [Online]. Available: http://betterevaluation.org/evaluation-options/data_mining. [Accessed: 27-Feb-2016].
- [30] O. Maimon and L. Rokach, "Introduction to knowledge discovery in databases," in *Data Mining and Knowledge Discovery Handbook*, Springer, 2005, pp. 1–17.
- [31] S. Srivastava and G. Anchal, "DATA MINING FOR CREDIT CARD RISK ANALYSIS: A REVIEW," *Int. J. Comput. Sci.*, 2013.
- [32] P. Richhariya, P. K. Singh, E. Duneja, B. BITS, and I. S. C. Softwares, "A Survey on Financial Fraud Detection Methodologies," *Int. J. Comput. Appl.*, vol. 45, no. 22, 2012.
- [33] R. Gupta and N. S. Gill, "A Data Mining Framework for Prevention and Detection of Financial Statement Fraud," *Int. J. Comput. Appl.*, vol. 50, no. 8, 2012.
- [34] M. F. Santos, P. Cortez, J. Pereira, and H. Quintela, "Corporate bankruptcy prediction using data mining techniques," 2006, vol. 1, pp. 349–357.
- [35] M. M. Shiri and M. Ahangary, "Corporate Bankruptcy Prediction using Data Mining Techniques: Evidence from Iran," *Afr. J. Sci. Res. Vol.*, vol. 8, no. 1, 2012.
- [36] Z. e i nab rashidi and S. ayed mohsen hosscini, "Bankruptcy Prediction of Listed Corporations in Tehran Stock Exchange Using Data Mining Techniques," Jun-2013.
- [37] M. Moorthi, "A Review on Data Mining and its Techniques in Finance Sector," 2015.

- [38] M. Siami, Z. Hajimohammadi, and others, "Credit scoring in banks and financial institutions via data mining techniques: A literature review," *J. AI Data Min.*, vol. 1, no. 2, pp. 119–129, 2013.
- [39] I. G. N. N. Mandala, C. B. Nawangpalupi, and F. R. Praktiko, "Assessing Credit Risk: An Application of Data Mining in a Rural Bank," *Procedia Econ. Finance*, vol. 4, pp. 406–412, 2012.
- [40] M.-C. Lee, "Enterprise Credit Risk Evaluation models: A Review of Current Research Trends," *Int. J. Comput. Appl.* 0975–8887, vol. 44, no. 11, pp. 37–44, 2012.
- [41] "What Is Data Mining?" [Online]. Available: https://docs.oracle.com/cd/E11882_01/datamine.112/e16808/process.htm#DMC ON123. [Accessed: 03-Mar-2016].
- [42] "What is Primary Data? definition and meaning." [Online]. Available: http://www.investorwords.com/19322/primary_data.html. [Accessed: 03-Mar-2016].
- [43] "Data_preprocessing.pdf." [Online]. Available: http://mercury.webster.edu/aleshunus/Support%20Materials/Data_preprocessing.pdf. [Accessed: 03-Mar-2016].
- [44] "What is the Statistical Package for the Social Sciences (SPSS)? - Definition from Techopedia." [Online]. Available: <https://www.techopedia.com/definition/12401/statistical-package-for-the-social-sciences-spss>. [Accessed: 03-Mar-2016].
- [45] B. Ratner, "Variable selection methods in regression: Ignorable problem, outing notable solution," *J. Target. Meas. Anal. Mark.*, vol. 18, no. 1, pp. 65–75, Mar. 2010.
- [46] K. Abhishek, "Stock Prediction using Artificial Neural Networks." Dept. of Computer Science and Engineering, IIT Kanpur.
- [47] "Neural Network Data Normalization and Encoding -- Visual Studio Magazine." [Online]. Available: <https://visualstudiomagazine.com/articles/2013/07/01/neural-network-data-normalization-and-encoding.aspx>. [Accessed: 07-Mar-2016].
- [48] "Data Mining for Visualization." [Online]. Available: <http://jcsites.juniata.edu/faculty/rhodes/ida/dm4viz.htm>. [Accessed: 07-Mar-2016].

- [49] “Multicollinearity statistics with SPSS : AnnMaria’s Blog.” [Online]. Available: <http://www.thejuliagroup.com/blog/?p=1405>. [Accessed: 07-Mar-2016].
- [50] C. Bounsaythip and E. Rinta-Runsala, “Overview of Data Mining for Customer Behavior Modeling,” VVT Information Technology, 29-Jun-2001.
- [51] I. H. Witten and E. Frank, *Data mining: practical machine learning tools and techniques*, 2nd ed. Amsterdam ; Boston, MA: Morgan Kaufman, 2005.
- [52] E. Osmanbegović and M. Suljić, “Data mining approach for predicting student performance,” *Econ. Rev.*, vol. 10, no. 1, 2012.
- [53] “CiteSeerX — A comparison of event models for Naive Bayes text classification.” [Online]. Available: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.65.9324>. [Accessed: 14-Mar-2016].
- [54] S.-B. Kim, K.-S. Han, H.-C. Rim, and S. H. Myaeng, “Some effective techniques for naive bayes text classification,” *Knowl. Data Eng. IEEE Trans. On*, vol. 18, no. 11, pp. 1457–1466, 2006.
- [55] S. Zhang, C. Tjortjis, X. Zeng, H. Qiao, I. Buchan, and J. Keane, “Comparing Data Mining Methods with Logistic Regression.”
- [56] S. Zhang, C. Tjortjis, X. Zeng, H. Qiao, I. Buchan, and J. Keane, “Comparing data mining methods with logistic regression in childhood obesity prediction,” *Inf. Syst. Front.*, vol. 11, no. 4, pp. 449–460, 2009.
- [57] L. Rokach and O. Maimon, “Decision trees,” in *Data mining and knowledge discovery handbook*, Springer, 2005, pp. 165–192.
- [58] W. Chen, G. Xiang, Y. Liu, and K. Wang, “Credit risk Evaluation by hybrid data mining technique,” *Syst. Eng. Procedia*, vol. 3, pp. 194–200, 2012.
- [59] E. Turban, R. Sharda, Delen Dursun, and David King, *BUSINESS INTELLIGENCE : A Managerial Approach*, Second. .
- [60] V. Pacelli and M. Azzollini, “An Artificial Neural Network Approach for Credit Risk Management,” *J. Intell. Learn. Syst. Appl.*, vol. 3, no. 2, pp. 103–112, 2011.
- [61] G. M. M. Sujeewa, “Artificial neural network model for business failure prediction of distressed firms in Colombo Stock Exchange,” *Acad. Int. Multidiscip. Res. J.*, vol. 4, no. 9, pp. 140–156, 2014.
- [62] M. Adya and F. Collopy, “How effective are neural networks at forecasting and prediction? A review and evaluation,” *J Forecast.*, vol. 17, pp. 481–495, 1998.

- [63] "Interrater reliability: the kappa statistic." [Online]. Available: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3900052/>. [Accessed: 05-Mar-2016].
- [64] J. R. Landis and G. G. Koch, "The Measurement of Observer Agreement for Categorical Data," *Biometrics*, vol. 33, no. 1, p. 159, Mar. 1977.
- [65] D. Rodríguez, "Making predictions on new data using Weka," *Univ. Alcala*.
- [66] K. Hajian-Tilaki, "Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation," Department of Social Medicine and Health, Babol University of Medical Sciences, Babol, Iran., 2013.
- [67] "The Area Under an ROC Curve." [Online]. Available: <http://gim.unmc.edu/dxtests/roc3.htm>. [Accessed: 05-Mar-2016].
- [68] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature," *Geosci. Model Dev.*, vol. 7, no. 3, pp. 1247–1250, Jun. 2014.
- [69] "Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE)." [Online]. Available: http://www.eumetcal.org/resources/ukmeteocal/verificationSAV/www/english/m_sg/ver_cont_var/uos3/uos3_ko1.htm. [Accessed: 10-Mar-2016].
- [70] "Data Mining Systems." [Online]. Available: http://www.tutorialspoint.com/data_mining/dm_systems.htm. [Accessed: 03-Mar-2016].