Knowledge Discovery in Databases with Exercises

Literature References

Dominik Probst

August 5, 2025

Contents

2 Introduction

[6] H. Du, *Data Mining Techniques and Applications: An Introduction*. Cengage Learning EMEA, May 2010, p. 336, ISBN: 978-1844808915.

[7] A. Geŕon, *Hands-on machine learning with Scikit-Learn and TensorFlow : concepts, tools, and techniques to build intelligent systems*, 2nd ed. O'Reilly Media, 2017, ISBN: 978-1491962299.

[9] T. Dasu and T. Johnson, *Exploratory Data Mining and Data Cleaning*. John Wiley, 2003, ISBN: 0-471-26851-8.

[10] U. M. Fayyad, G. G. Grinstein, and A. Wierse, Eds., *Information Visualization in Data Mining and Knowledge Discovery*. Morgan Kaufmann, 2001, ISBN: 1-55860-689-0. [Online]. Available: `http://www.mkp.com/books%5C_catalog/catalog.asp?ISBN=1-55860-689-0`.

[42] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification, 2nd Edition*. Wiley, 2001, ISBN: 9780471056690. [Online]. Available: `https://www.worldcat.org/oclc/41347061`.

[43] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Springer Series in Statistics). Springer New York, 2009, ISBN: 978-0-387-84857-0.

[45] T. M. Mitchell, *Machine learning* (McGraw Hill series in computer science). McGraw-Hill, 1997, ISBN: 978-0-07-042807-2. [Online]. Available: `http://www.worldcat.org/oclc/61321007`.

[48] P. Tan, M. S. Steinbach, A. Karpatne, and V. Kumar, *Introduction to Data Mining (Second Edition)*. Pearson, 2019. [Online]. Available: `https://www-users.cse.umn.edu/%5C%7Ekumar001/dmbook/index.php`.

[210] S. Chakrabarti, *Mining the Web: Analysis of Hypertext and Semi Structured Data*. Morgan Kaufmann, Aug. 2002, ISBN: 1558607544. [Online]. Available: `http://www.amazon.com/exec/obidos/redirect?tag=citeulike07-20%5C&path=ASIN/1558607544`.

[211] U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, Eds., *Advances in Knowledge Discovery and Data Mining*. USA: American Association for Artificial Intelligence, 1996, ISBN: 0262560976.

[212] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2011, ISBN: 0123814790.

[213] D. J. Hand, P. Smyth, and H. Mannila, *Principles of Data Mining*. Cambridge, MA, USA: MIT Press, 2001, ISBN: 026208290X.

[214] B. Liu, *Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data*. Berlin, Heidelberg: Springer-Verlag, 2009, ISBN: 3642072372.

[215] G. Piatetsky-Shapiro and W. J. Frawley, Eds., *Knowledge Discovery in Databases*. AAAI/MIT Press, 1991, ISBN: 0-262-62080-4.

[216] S. M. Weiss and N. Indurkhya, Eds., *Predictive data mining: a practical guide*. Morgan Kaufmann Publishers Inc., 1998, ISBN: 1558604030.

[217] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining, Fourth Edition: Practical Machine Learning Tools and Techniques*, 4th. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2016, ISBN: 0128042915.

[218] C. Shearer, "The crisp-dm model: The new blueprint for data mining," *Journal of Data Warehousing*, vol. 5, no. 4, 2000.

[219] T. Xie, S. Thummalapenta, D. Lo, and C. Liu, "Data mining for software engineering," *Computer*, vol. 42, no. 8, pp. 55–62, 2009. DOI: 10.1109/MC.2009.256.

[220] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, 2nd. OTexts, 2018.

3 Data

[4] S. S. Stevens, "On the theory of scales of measurement," *Science*, vol. 103, no. 2684, pp. 677–680, 1946, ISSN: 00368075, 10959203. Accessed: Apr. 9, 2025. [Online]. Available: http://www.jstor.org/stable/1671815.

[8] W. S. Cleveland, *Visualizing Data*. Hobart Press, Summit, New Jersey, 1993, ISBN: 978-0963488404.

[9] T. Dasu and T. Johnson, *Exploratory Data Mining and Data Cleaning*. John Wiley, 2003, ISBN: 0-471-26851-8.

[10] U. M. Fayyad, G. G. Grinstein, and A. Wierse, Eds., *Information Visualization in Data Mining and Knowledge Discovery*. Morgan Kaufmann, 2001, ISBN: 1-55860-689-0. [Online]. Available: http://www.mkp.com/books%5C_catalog/catalog.asp?ISBN=1-55860-689-0.

[11] L. Kaufman and P. J. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis*. John Wiley, 1990, ISBN: 978-0-47187876-6. DOI: 10.1002/9780470316801.

[12] D. Lomet and J. M. Hellerstein, Eds., *Special Issue on Data Reduction Techniques* vol. 20.4 1997, *Bulletin of the Technical Committee on Data Engineering*: *Bulletin of the Technical Committee on Data Engineering*. [Online]. Available: http://sites.computer.org/debull/97DEC-CD.pdf.

[13] D. A. Keim, "Information visualization and visual data mining," *IEEE Trans. Vis. Comput. Graph.*, vol. 8, no. 1, pp. 1–8, 2002. DOI: 10.1109/2945.981847. [Online]. Available: https://doi.org/10.1109/2945.981847.

[14] F. Naumann, "Data profiling revisited," *SIGMOD Rec.*, vol. 42, no. 4, pp. 40–49, 2013. DOI: 10.1145/2590989.2590995. [Online]. Available: https://doi.org/10.1145/2590989.2590995.

[15] D. Pyle, *Data Preparation for Data Mining*. Morgan Kaufmann, 1999, ISBN: 1-55860-529-0.

[16] S. Santini and R. C. Jain, "Similarity measures," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 9, pp. 871–883, 1999. DOI: 10.1109/34.790428. [Online]. Available: `https://doi.org/10.1109/34.790428`.

[17] E. R. Tufte, *The visual display of quantitative information*. Graphics Press, 2001, 2nd edition, ISBN: 978-0-9613921-4-7.

[18] C. Yu, Y. Zhong, T. Smith, I. Park, and W. Huang, "Visual data mining of multimedia data for social and behavioral studies," *Inf. Vis.*, vol. 8, no. 1, pp. 56–70, 2009. DOI: 10.1057/ivs.2008.32. [Online]. Available: `https://doi.org/10.1057/ivs.2008.32`.

[299] J. J. van Wijk and R. van Liere, "Hyperslice - visualization of scalar functions of many variables," in *4th IEEE Visualization Conference, IEEE Vis 1993, San Jose, CA, USA, October 25-29, 1993, Proceedings*, G. M. Nielson and R. D. Bergeron, Eds., IEEE Computer Society, 1993, pp. 119–125, ISBN: 0-8186-3940-7. DOI: 10.1109/VISUAL.1993.398859. [Online]. Available: `https://doi.org/10.1109/VISUAL.1993.398859`.

[301] G. W. Furnas and A. Buja, "Prosection views: Dimensional inference through sections and projections," *Journal of Computational and Graphical Statistics*, vol. 3, no. 4, p. 323, Dec. 1994, ISSN: 1061-8600. DOI: 10.2307/1390897. [Online]. Available: `http://dx.doi.org/10.2307/1390897`.

[302] J. Friedman and J. Tukey, "A projection pursuit algorithm for exploratory data analysis," *IEEE Transactions on Computers*, vol. C-23, no. 9, pp. 881–890, Sep. 1974, ISSN: 0018-9340. DOI: 10.1109/t-c.1974.224051. [Online]. Available: `http://dx.doi.org/10.1109/t-c.1974.224051`.

[303] G. G. Robertson, J. D. Mackinlay, and S. K. Card, "Cone trees: Animated 3d visualizations of hierarchical information," in *Conference on Human Factors in Computing Systems, CHI 1991, New Orleans, LA, USA, April 27 - May 2, 1991, Proceedings*, S. P. Robertson, G. M. Olson, and J. S. Olson, Eds., ACM, 1991, pp. 189–194, ISBN: 0-89791-383-3. DOI: 10.1145/108844.108883. [Online]. Available: `https://doi.org/10.1145/108844.108883`.

[305] Z. Abedjan, L. Golab, F. Naumann, and T. Papenbrock, *Data Profiling*. Morgan & Claypool Publishers LLC, Nov. 2018, vol. 10, pp. 1–154. DOI: 10.2200/s00878ed1v01y201810dtm052. [Online]. Available: `http://dx.doi.org/10.2200/s00878ed1v01y201810dtm052`.

4 Data Preprocessing

[10] U. M. Fayyad, G. G. Grinstein, and A. Wierse, Eds., *Information Visualization in Data Mining and Knowledge Discovery*. Morgan Kaufmann, 2001, ISBN: 1-55860-689-0. [Online]. Available: `http://www.mkp.com/books%5C_catalog/catalog.asp?ISBN=1-55860-689-0`.

[15] D. Pyle, *Data Preparation for Data Mining*. Morgan Kaufmann, 1999, ISBN: 1-55860-529-0.

[44] H. Liu and H. Motoda, *Feature Extraction, Construction and Selection: A Data Mining Perspective* (The Springer International Series in Engineering and Computer Science). Springer US, 1998, ISBN: 9780792381969.

[221] D. P. Ballou and G. K. Tayi, "Enhancing data quality in data warehouse environments," *Commun. ACM*, vol. 42, no. 1, pp. 73–78, Jan. 1999, ISSN: 0001-0782. DOI: 10.1145/291469.291471. [Online]. Available: https://doi.org/10.1145/291469.291471.

[222] A. Bruce, D. Donoho, and H.-Y. Gao, "Wavelet analysis [for signal processing]," *IEEE Spectrum*, vol. 33, no. 10, pp. 26–35, 1996. DOI: 10.1109/6.540087.

[223] R. Peck and J. Devore, *Statistics: The Exploration and Analysis of Data*. Cengage Learning, 2011, ISBN: 9781133171744.

[224] H. Galhardas, D. Florescu, D. Shasha, E. Simon, and C.-A. Saita, "Declarative data cleaning: Language, model, and algorithms," in *Proceedings of the 27th International Conference on Very Large Data Bases*, ser. VLDB '01, San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2001, pp. 371–380, ISBN: 1558608044.

[225] M. Hua and J. Pei, "Cleaning disguised missing data: A heuristic approach," Aug. 2007, pp. 950–958. DOI: 10.1145/1281192.1281294.

[226] D. Barbará et al., "The new jersey data reduction report," *IEEE Data Eng. Bull.*, vol. 20, no. 4, pp. 3–45, 1997.

[227] J. Olson, *Data Quality: The Accuracy Dimension* (The Morgan Kaufmann Series in Data Management Systems). Morgan Kaufmann Publishers Inc., Dec. 2002, p. 312, ISBN: 1-55860-891-5.

[228] V. Raman and J. M. Hellerstein, "Potter's wheel: An interactive data cleaning system," in *VLDB 2001, Proceedings of 27th International Conference on Very Large Data Bases, September 11-14, 2001, Roma, Italy*, Morgan Kaufmann, 2001, pp. 381–390.

[229] T. Redman, *Data Quality: The Field Guide* (Data management series). Elsevier Science, 2001, ISBN: 9781555582517.

[230] R. Y. Wang, V. C. Storey, and C. P. Firth, "A framework for analysis of data quality research," *IEEE Trans. Knowl. Data Eng.*, vol. 7, no. 4, pp. 623–640, 1995. DOI: 10.1109/69.404034.

[231] J. Shlens, "A tutorial on principal component analysis," *CoRR*, vol. abs/1404.1100, 2014. arXiv: 1404.1100. [Online]. Available: http://arxiv.org/abs/1404.1100.

5 Data Warehousing and Online Analytical Processing

[5] C. Sapia, M. Blaschka, G. Höfling, and B. Dinter, "Extending the E/R model for the multidimensional paradigm," in *Advances in Database Technologies, ER '98 Workshops on Data Warehousing and Data Mining, Mobile Data Access, and Collaborative Work Support and Spatio-Temporal Data Management, Singapore, November 19-20, 1998, Proceedings,*

Y. Kambayashi, D. L. Lee, E. Lim, M. K. Mohania, and Y. Masunaga, Eds., ser. Lecture Notes in Computer Science, vol. 1552, Springer, 1998, pp. 105–116. DOI: 10.1007/978-3-540-49121-7_9. [Online]. Available: https://doi.org/10.1007/978-3-540-49121-7%5C_9.

[19] S. Agarwal et al., "On the computation of multidimensional aggregates," in *VLDB'96, Proceedings of 22th International Conference on Very Large Data Bases, September 3-6, 1996, Mumbai (Bombay), India*, T. M. Vijayaraman, A. P. Buchmann, C. Mohan, and N. L. Sarda, Eds., Morgan Kaufmann, 1996, pp. 506–521, ISBN: 1-55860-382-4. [Online]. Available: http://www.vldb.org/conf/1996/P506.PDF.

[20] D. Agrawal, A. El Abbadi, A. Singh, and T. Yurek, "Efficient view maintenance at data warehouses," *Proceedings of the 1997 ACM SIGMOD international conference on Management of data - SIGMOD '97*, 1997. DOI: 10.1145/253260.253355. [Online]. Available: http://dx.doi.org/10.1145/253260.253355.

[21] R. Agrawal, A. Gupta, and S. Sarawagi, "Modeling multidimensional databases," in *Proceedings of the Thirteenth International Conference on Data Engineering, April 7-11, 1997, Birmingham, UK*, W. A. Gray and P. Larson, Eds., IEEE Computer Society, 1997, pp. 232–243, ISBN: 0-8186-7807-0. DOI: 10.1109/ICDE.1997.581777. [Online]. Available: https://doi.org/10.1109/ICDE.1997.581777.

[23] A. Gupta and I. S. Mumick, *Materialized Views: Techniques, Implementations, and Applications*. The MIT Press, May 1999, ISBN: 9780262287500. DOI: 10.7551/mitpress/4472.001.0001. [Online]. Available: https://doi.org/10.7551/mitpress/4472.001.0001.

[24] S. Chaudhuri and U. Dayal, "An overview of data warehousing and OLAP technology," *SIGMOD Rec.*, vol. 26, no. 1, pp. 65–74, 1997. DOI: 10.1145/248603.248616. [Online]. Available: https://doi.org/10.1145/248603.248616.

[25] E. F. Codd, S. B. Codd, and C. T. Salley, "Beyond decision support," *Computer World*, vol. 27, no. 30, pp. 87–89, 1993.

[26] J. Gray et al., "Data cube: A relational aggregation operator generalizing group-by, crosstab, and sub totals," *Data Min. Knowl. Discov.*, vol. 1, no. 1, pp. 29–53, 1997. DOI: 10.1023/A:1009726021843. [Online]. Available: https://doi.org/10.1023/A:1009726021843.

[27] J. Han, "Towards on-line analytical mining in large databases," *SIGMOD Rec.*, vol. 27, no. 1, pp. 97–107, 1998. DOI: 10.1145/273244.273273. [Online]. Available: https://doi.org/10.1145/273244.273273.

[28] V. Harinarayan, A. Rajaraman, and J. D. Ullman, "Implementing data cubes efficiently," in *Proceedings of the 1996 ACM SIGMOD International Conference on Management of Data, Montreal, Quebec, Canada, June 4-6, 1996*, H. V. Jagadish and I. S. Mumick, Eds., ACM Press, 1996, pp. 205–216, ISBN: 978-0-89791-794-0. DOI: 10.1145/233269.233333. [Online]. Available: https://doi.org/10.1145/233269.233333.

[30] J. M. Hellerstein, P. J. Haas, and H. J. Wang, "Online aggregation," in *SIGMOD 1997, Proceedings ACM SIGMOD International Conference on Management of Data, May 13-15, 1997, Tucson, Arizona, USA*, J. Peckham, Ed., ACM Press, 1997, pp. 171–182, ISBN: 978-0-89791-911-1. DOI: 10.1145/253260.253291. [Online]. Available: https://doi.org/10.1145/253260.253291.

[31] C. Imhoff, N. Galemmo, and J. Geiger, *Mastering Data Warehouse Design: Relational and Dimensional Techniques*. Wiley, 2003, ISBN: 978-0-47132-421-8.

[32] W. H. Inmon, *Building the Data Warehouse*. Wiley, 2005, 4th edition, ISBN: 978-076459-944-6.

[33] R. Kimball and M. Ross, *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling*. Wiley, 2013, 3rd edition, ISBN: 978-1-118-53080-1.

[34] R. Kimball and J. Caserta, *The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data*. Wiley, 2004, ISBN: 978-0764567575.

[35] P. E. O'Neil and G. Graefe, "Multi-table joins through bitmapped join indices," *SIGMOD Rec.*, vol. 24, no. 3, pp. 8–11, 1995. DOI: 10.1145/211990.212001. [Online]. Available: https://doi.org/10.1145/211990.212001.

[36] P. E. O'Neil and D. Quass, "Improved query performance with variant indexes," in *SIGMOD 1997, Proceedings ACM SIGMOD International Conference on Management of Data, May 13-15, 1997, Tucson, Arizona, USA*, J. Peckham, Ed., ACM Press, 1997, pp. 38–49, ISBN: 978-0-89791-911-1. DOI: 10.1145/253260.253268. [Online]. Available: https://doi.org/10.1145/253260.253268.

[38] A. Bauer and H. Günzel, *Data Warehouse Systeme – Architektur, Entwicklung, Anwendung*. dpunkt.verlag GmbH, 2004, 4th edition, ISBN: 978-3-89864-785-4.

6 Mining Frequent Patterns, Associations, and Correlations

Basic Concepts of Frequent-pattern Mining

[232] R. Agrawal, T. Imieliński, and A. Swami, "Mining association rules between sets of items in large databases," *SIGMOD Rec.*, vol. 22, no. 2, pp. 207–216, Jun. 1993, ISSN: 0163-5808. DOI: 10.1145/170036.170072. [Online]. Available: https://doi.org/10.1145/170036.170072.

[233] R. J. Bayardo, "Efficiently mining long patterns from databases," *SIGMOD Rec.*, vol. 27, no. 2, pp. 85–93, Jun. 1998, ISSN: 0163-5808. DOI: 10.1145/276305.276313. [Online]. Available: https://doi.org/10.1145/276305.276313.

[234] N. Pasquier, Y. Bastide, R. Taouil, and L. Lakhal, "Discovering frequent closed itemsets for association rules," in *Proceedings of the 7th International Conference on Database Theory*, ser. ICDT '99, Berlin, Heidelberg: Springer-Verlag, 1999, pp. 398–416, ISBN: 3540654526.

[235] R. Agrawal and R. Srikant, "Mining sequential patterns," in *Proceedings of the Eleventh International Conference on Data Engineering*, 1995, pp. 3–14. DOI: `10.1109/ICDE.1995.380415`.

A Priori and its Improvements

[236] R. Agrawal and R. Srikant, "Fast algorithms for mining association rules in large databases," in *Proceedings of the 20th International Conference on Very Large Data Bases*, ser. VLDB '94, San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1994, pp. 487–499, ISBN: 1558601538.

[238] A. Savasere, E. Omiecinski, and S. B. Navathe, "An efficient algorithm for mining association rules in large databases," in *Proceedings of the 21th International Conference on Very Large Data Bases*, ser. VLDB '95, San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1995, pp. 432–444, ISBN: 1558603794.

[239] J. S. Park, M.-S. Chen, and P. S. Yu, "An effective hash-based algorithm for mining association rules," *SIGMOD Rec.*, vol. 24, no. 2, pp. 175–186, May 1995, ISSN: 0163-5808. DOI: `10.1145/568271.223813`. [Online]. Available: `https://doi.org/10.1145/568271.223813`.

[240] H. Toivonen, "Sampling large databases for association rules," in *Proceedings of the 22th International Conference on Very Large Data Bases*, ser. VLDB '96, San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1996, pp. 134–145, ISBN: 1558603824.

[241] S. Brin, R. Motwani, J. D. Ullman, and S. Tsur, "Dynamic itemset counting and implication rules for market basket data," *SIGMOD Rec.*, vol. 26, no. 2, pp. 255–264, Jun. 1997, ISSN: 0163-5808. DOI: `10.1145/253262.253325`. [Online]. Available: `https://doi.org/10.1145/253262.253325`.

[242] S. Sarawagi, S. Thomas, and R. Agrawal, "Integrating association rule mining with relational database systems: Alternatives and implications," in *Proceedings of the 1998 ACM SIGMOD International Conference on Management of Data*, ser. SIGMOD '98, Seattle, Washington, USA: Association for Computing Machinery, 1998, pp. 343–354, ISBN: 0897919955. DOI: `10.1145/276304.276335`. [Online]. Available: `https://doi.org/10.1145/276304.276335`.

Depth-first, Projection-based FP Mining

[243] R. C. Agarwal, C. C. Aggarwal, and V. Prasad, "A tree projection algorithm for generation of frequent item sets," *J. Parallel Distrib. Comput.*, vol. 61, no. 3, pp. 350–371, Mar. 2001, ISSN: 0743-7315. DOI: `10.1006/jpdc.2000.1693`. [Online]. Available: `https://doi.org/10.1006/jpdc.2000.1693`.

[244] G. Grahne and J. Zhu, "Efficiently using prefix-trees in mining frequent itemsets," in *FIMI '03, Frequent Itemset Mining Implementations, Proceedings of the ICDM 2003 Workshop on Frequent Itemset Mining Implementations, 19 December 2003, Melbourne, Florida, USA*, B. Goethals and M. J. Zaki, Eds., ser. CEUR Workshop Proceedings, vol. 90, CEUR-WS.org, 2003. [Online]. Available: `http://ceur-ws.org/Vol-90/grahne.pdf`.

[245] G. Grahne and J. Zhu, "Reducing the main memory consumptions of fpmax*and fpclose," in *Proc. Workshop Frequent Item Set Mining Implementations (FIMI 2004, Brighton, UK), Aachen, Germany*, 2004, p. 75.

[246] B. Goethals and M. J. Zaki, "Advances in frequent itemset mining implementations: Introduction to FIMI03," in *FIMI '03, Frequent Itemset Mining Implementations, Proceedings of the ICDM 2003 Workshop on Frequent Itemset Mining Implementations, 19 December 2003, Melbourne, Florida, USA*, B. Goethals and M. J. Zaki, Eds., ser. CEUR Workshop Proceedings, vol. 90, CEUR-WS.org, 2003. [Online]. Available: `http://ceur-ws.org/Vol-90/intro.pdf`.

[247] J. Han, J. Pei, and Y. Yin, "Mining frequent patterns without candidate generation," *SIGMOD Rec.*, vol. 29, no. 2, pp. 1–12, May 2000, ISSN: 0163-5808. DOI: `10.1145/335191.335372`. [Online]. Available: `https://doi.org/10.1145/335191.335372`.

[248] J. Han et al., "Prefixspan: Mining sequential patterns efficiently by prefix-projected pattern growth," in *proceedings of the 17th international conference on data engineering*, IEEE Piscataway, NJ, USA, 2001, pp. 215–224.

[250] J. Han, J. Wang, Y. Lu, and P. Tzvetkov, "Mining top-k frequent closed patterns without minimum support," in *2002 IEEE International Conference on Data Mining, 2002. Proceedings.*, 2002, pp. 211–218. DOI: `10.1109/ICDM.2002.1183905`.

[251] J. Wang, J. Han, and J. Pei, "CLOSET+: searching for the best strategies for mining frequent closed itemsets," in *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, August 24 - 27, 2003*, L. Getoor, T. E. Senator, P. M. Domingos, and C. Faloutsos, Eds., ACM, 2003, pp. 236–245. DOI: `10.1145/956750.956779`. [Online]. Available: `https://doi.org/10.1145/956750.956779`.

Vertical Format and Row Enumeration Methods

[252] M. J. Zaki, S. Parthasarathy, M. Ogihara, and W. Li, "Parallel algorithms for discovery of association rules," *Data Min. Knowl. Discov.*, vol. 1, no. 4, pp. 343–373, 1997. DOI: `10.1023/A:1009773317876`. [Online]. Available: `https://doi.org/10.1023/A:1009773317876`.

[253] M. J. Zaki and C.-J. Hsiao, "Charm: An efficient algorithm for closed itemset mining," in *Proceedings of the 2002 SIAM International Conference on Data Mining (SDM)*, 2002, pp. 457–473. DOI: `10.1137/1.9781611972726.27`.

[254] C. Bucila, J. Gehrke, D. Kifer, and W. M. White, "Dualminer: A dual-pruning algorithm for itemsets with constraints," in *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, July 23-26, 2002, Edmonton, Alberta, Canada*, ACM, 2002, pp. 42–51. DOI: `10.1145/775047.775054`. [Online]. Available: `https://doi.org/10.1145/775047.775054`.

[255] H. Liu, J. Han, D. Xin, and Z. Shao, "Mining interesting patterns from very high dimensional data: A top-down row enumeration approach," in *Proceedings of the Sixth SIAM International Conference on Data Mining, April 20-22, 2006, Bethesda, MD, USA*, J. Ghosh, D. Lambert, D. B. Skillicorn, and J. Srivastava, Eds., SIAM, 2006, pp. 282–293. DOI: `10.1137/1.9781611972764.25`. [Online]. Available: `https://doi.org/10.1137/1.9781611972764.25`.

Mining Correlations and Interesting Rules

[1] G. Liu, H. Lu, J. X. Yu, W. Wang, and X. Xiao, "Afopt: An efficient implementation of pattern growth approach.," in *FIMI*, 2003, pp. 1–10.

[2] F. Pan, G. Cong, A. K. Tung, J. Yang, and M. J. Zaki, "Carpenter: Finding closed patterns in long biological datasets," in *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, 2003, pp. 637–642.

[3] H. Jagadish, J. Madar, and R. T. Ng, "Semantic compression and pattern extraction with fascicles," in *VLDB*, vol. 99, 1999, pp. 186–97.

[256] S. Brin, R. Motwani, and C. Silverstein, "Beyond market baskets: Generalizing association rules to correlations," in *SIGMOD 1997, Proceedings ACM SIGMOD International Conference on Management of Data, May 13-15, 1997, Tucson, Arizona, USA*, J. Peckham, Ed., ACM Press, 1997, pp. 265–276. DOI: `10.1145/253260.253327`. [Online]. Available: `https://doi.org/10.1145/253260.253327`.

[257] M. Klemettinen, H. Mannila, P. Ronkainen, H. Toivonen, and A. I. Verkamo, "Finding interesting rules from large sets of discovered association rules," in *Proceedings of the Third International Conference on Information and Knowledge Management (CIKM'94), Gaithersburg, Maryland, USA, November 29 - December 2, 1994*, ACM, 1994, pp. 401–407. DOI: `10.1145/191246.191314`. [Online]. Available: `https://doi.org/10.1145/191246.191314`.

[258] R. J. Hilderman and H. J. Hamilton, *Knowledge Discovery and Measures of Interest*. USA: Kluwer Academic Publishers, 2001, ISBN: 0792375076.

[259] X. Yan, J. Han, and R. Afshar, "Clospan: Mining: Closed sequential patterns in large datasets," in *Proceedings of the 2003 SIAM international conference on data mining*, SIAM, 2003, pp. 166–177.

[260] X. Yan and J. Han, "Gspan: Graph-based substructure pattern mining," in *2002 IEEE International Conference on Data Mining, 2002. Proceedings.*, IEEE, 2002, pp. 721–724.

[261] F. Zhu, X. Yan, J. Han, and P. S. Yu, "Gprune: A constraint pushing framework for graph pattern mining," in *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, Springer, 2007, pp. 388–400.

[262] D. Xin, J. Han, X. Li, and B. W. Wah, "Star-cubing: Computing iceberg cubes by top-down and bottom-up integration," in *Proceedings 2003 VLDB Conference*, Elsevier, 2003, pp. 476–487.

[263] J. Pei, X. Zhang, M. Cho, H. Wang, and P. S. Yu, "Maple: A fast algorithm for maximal pattern-based clustering," in *Third ieee international conference on data mining*, IEEE, 2003, pp. 259–266.

[264] H. Cheng, X. Yan, J. Han, and C.-W. Hsu, "Discriminative frequent pattern analysis for effective classification," in *2007 IEEE 23rd international conference on data engineering*, IEEE, 2006, pp. 716–725.

[266] C. Silverstein, S. Brin, R. Motwani, and J. D. Ullman, "Scalable techniques for mining causal structures," in *VLDB'98, Proceedings of 24rd International Conference on Very Large Data Bases, August 24-27, 1998, New York City, New York, USA*, A. Gupta, O. Shmueli, and J. Widom, Eds., Morgan Kaufmann, 1998, pp. 594–605. [Online]. Available: `http://www.vldb.org/conf/1998/p594.pdf`.

[267] J. Wang and J. Han, "Bide: Efficient mining of frequent closed sequences," in *Proceedings. 20th international conference on data engineering*, IEEE, 2004, pp. 79–90.

[268] P. Tan, V. Kumar, and J. Srivastava, "Selecting the right interestingness measure for association patterns," in *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, July 23-26, 2002, Edmonton, Alberta, Canada*, ACM, 2002, pp. 32–41. DOI: 10.1145/775047.775053. [Online]. Available: `https://doi.org/10.1145/775047.775053`.

[269] E. Omiecinski, "Alternative interest measures for mining associations in databases," *IEEE Trans. Knowl. Data Eng.*, vol. 15, no. 1, pp. 57–69, 2003. DOI: 10.1109/TKDE.2003.1161582. [Online]. Available: `https://doi.org/10.1109/TKDE.2003.1161582`.

[270] T. Wu, Y. Chen, and J. Han, "Association mining in large databases: A re-examination of its measures," in *Knowledge Discovery in Databases: PKDD 2007, 11th European Conference on Principles and Practice of Knowledge Discovery in Databases, Warsaw, Poland, September 17-21, 2007, Proceedings*, J. N. Kok, J. Koronacki, R. L. de Mántaras, S. Matwin, D. Mladenic, and A. Skowron, Eds., ser. Lecture Notes in Computer Science, vol. 4702, Springer, 2007, pp. 621–628. DOI: 10.1007/978-3-540-74976-9_66. [Online]. Available: `https://doi.org/10.1007/978-3-540-74976-9%5C_66`.

[271] T. Wu, Y. Chen, and J. Han, "Re-examination of interestingness measures in pattern mining: A unified framework," *Data Min. Knowl. Discov.*, vol. 21, no. 3, pp. 371–397, 2010. DOI: 10.1007/s10618-009-0161-2. [Online]. Available: `https://doi.org/10.1007/s10618-009-0161-2`.

7 Classification

Books on Classification

[39] E. Alpaydin, *Introduction to Machine Learning*. MIT Press, 2004, ISBN: 9780262012119.

[40] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*. Wadsworth, 1984, ISBN: 0-534-98053-8.

[41] C. M. Bishop, *Pattern recognition and machine learning, 5th Edition* (Information science and statistics). Springer, 2007, ISBN: 9780387310732. [Online]. Available: `https://www.worldcat.org/oclc/71008143`.

[42] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification, 2nd Edition*. Wiley, 2001, ISBN: 9780471056690. [Online]. Available: `https://www.worldcat.org/oclc/41347061`.

[43] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Springer Series in Statistics). Springer New York, 2009, ISBN: 978-0-387-84857-0.

[44] H. Liu and H. Motoda, *Feature Extraction, Construction and Selection: A Data Mining Perspective* (The Springer International Series in Engineering and Computer Science). Springer US, 1998, ISBN: 9780792381969.

[45] T. M. Mitchell, *Machine learning* (McGraw Hill series in computer science). McGraw-Hill, 1997, ISBN: 978-0-07-042807-2. [Online]. Available: `http://www.worldcat.org/oclc/61321007`.

[46] S. Marsland, *Machine Learning - An Algorithmic Perspective* (Chapman and Hall / CRC machine learning and pattern recognition series). CRC Press, 2009, ISBN: 978-1-4200-6718-7. [Online]. Available: `http://www.crcpress.com/product/isbn/9781420067187`.

[47] J. R. Quinlan, *C4.5: Programs for Machine Learning*. Morgan Kaufmann, 1993, ISBN: 1-55860-238-0.

[48] P. Tan, M. S. Steinbach, A. Karpatne, and V. Kumar, *Introduction to Data Mining (Second Edition)*. Pearson, 2019. [Online]. Available: `https://www-users.cse.umn.edu/%5C%7Ekumar001/dmbook/index.php`.

[49] S. M. Weiss and C. A. Kulikowski, *Computer Systems That Learn: Classification and Prediction Methods from Statistics, Neural Nets, Machine Learning and Expert Systems*. Morgan Kaufmann, 1990, ISBN: 1-55860-065-5.

[50] S. M. Weiss and N. Indurkhya, *Predictive data mining - a practical guide*. Morgan Kaufmann, 1997, ISBN: 978-1-55860-403-2.

[51] I. H. Witten and E. Frank, *Data mining - practical machine learning tools and techniques, Second Edition* (The Morgan Kaufmann series in data management systems). Morgan Kaufmann, 2005, ISBN: 978-0-12-088407-0.

[52] M. Ankerst, C. Elsen, M. Ester, and H. Kriegel, "Visual classification: An interactive approach to decision tree construction," in *Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Diego, CA, USA, August 15-18, 1999*, U. M. Fayyad, S. Chaudhuri, and D. Madigan, Eds., ACM, 1999, pp. 392–396, ISBN: 1-58113-143-7. DOI: `10.1145/312129.312298`. [Online]. Available: `https://doi.org/10.1145/312129.312298`.

[54] C. Apté and S. M. Weiss, "Data mining with decision trees and decision rules," *Future Gener. Comput. Syst.*, vol. 13, no. 2-3, pp. 197–210, 1997. DOI: `10.1016/S0167-739X(97)00021-6`. [Online]. Available: `https://doi.org/10.1016/S0167-739X(97)00021-6`.

[55] C. E. Brodley and P. E. Utgoff, "Multivariate decision trees," *Mach. Learn.*, vol. 19, no. 1, pp. 45–77, 1995. DOI: `10.1007/BF00994660`. [Online]. Available: `https://doi.org/10.1007/BF00994660`.

[56] P. K. Chan and S. J. Stolfo, "Learning arbiter and combiner trees from partitioned data for scaling machine learning," in *Proceedings of the First International Conference on Knowledge Discovery and Data Mining (KDD-95), Montreal, Canada, August 20-21, 1995*, U. M. Fayyad and R. Uthurusamy, Eds., AAAI Press, 1995, pp. 39–44, ISBN: 0-929280-82-2. [Online]. Available: `http://www.aaai.org/Library/KDD/1995/kdd95-047.php`.

[58] U. M. Fayyad, "Branching on attribute values in decision tree generation," in *Proceedings of the 12th National Conference on Artificial Intelligence, Seattle, WA, USA, July 31 - August 4, 1994, Volume 1*, B. Hayes-Roth and R. E. Korf, Eds., AAAI Press / The MIT Press, 1994, pp. 601–606, ISBN: 0-262-61102-3. [Online]. Available: `http://www.aaai.org/Library/AAAI/1994/aaai94-091.php`.

[60] M. Mehta, R. Agrawal, and J. Rissanen, "SLIQ: A fast scalable classifier for data mining," in *Advances in Database Technology - EDBT'96, 5th International Conference on Extending Database Technology, Avignon, France, March 25-29, 1996, Proceedings*, P. M. G. Apers, M. Bouzeghoub, and G. Gardarin, Eds., ser. Lecture Notes in Computer Science, vol. 1057, Springer, 1996, pp. 18–32, ISBN: 3-540-61057-X. DOI: `10.1007/BFb0014141`. [Online]. Available: `https://doi.org/10.1007/BFb0014141`.

[62] J. Gehrke, R. Ramakrishnan, and V. Ganti, "Rainforest - A framework for fast decision tree construction of large datasets," in *VLDB'98, Proceedings of 24rd International Conference on Very Large Data Bases, August 24-27, 1998, New York City, New York, USA*, A. Gupta, O. Shmueli, and J. Widom, Eds., Morgan Kaufmann, 1998, pp. 416–427, ISBN: 1-55860-566-5. [Online]. Available: `http://www.vldb.org/conf/1998/p416.pdf`.

[63] J. Gehrke, V. Ganti, R. Ramakrishnan, and W. Loh, "Boat-optimistic decision tree construction," in *SIGMOD 1999, Proceedings ACM SIGMOD International Conference on Management of Data, June 1-3, 1999, Philadelphia, Pennsylvania, USA*, A. Delis, C. Faloutsos, and S. Ghandeharizadeh, Eds., ACM Press, 1999, pp. 169–180, ISBN: 1-58113-084-8. DOI: `10.1145/304182.304197`. [Online]. Available: `https://doi.org/10.1145/304182.304197`.

[65] S. K. Murthy, "Automatic construction of decision trees from data: A multi-disciplinary survey," *Data Min. Knowl. Discov.*, vol. 2, no. 4, pp. 345–389, 1998. DOI: `10.1023/A:1009744630224`. [Online]. Available: `https://doi.org/10.1023/A:1009744630224`.

[66] J. R. Quinlan, "Induction of decision trees," *Mach. Learn.*, vol. 1, no. 1, pp. 81–106, 1986. DOI: `10.1023/A:1022643204877`. [Online]. Available: `https://doi.org/10.1023/A:1022643204877`.

[67] J. R. Quinlan and R. L. Rivest, "Inferring decision trees using the minimum description length principle," *Inf. Comput.*, vol. 80, no. 3, pp. 227–248, 1989. DOI: `10.1016/0890-5401(89)90010-2`. [Online]. Available: `https://doi.org/10.1016/0890-5401(89)90010-2`.

[68] J. R. Quinlan, "Learning logical definitions from relations," *Mach. Learn.*, vol. 5, pp. 239–266, 1990. DOI: `10.1007/BF00117105`. [Online]. Available: `https://doi.org/10.1007/BF00117105`.

[69] R. S. Michalski, I. Mozetic, J. Hong, and N. Lavrac, "The multi-purpose incremental learning system aq15 and its testing application to three medical domains," in *Proc. AAAI*, vol. 1986, 1986, pp. 1–041.

[70] P. Clark and T. Niblett, "The cn2 induction algorithm," *Machine learning*, vol. 3, no. 4, pp. 261–283, 1989.

[71] W. W. Cohen et al., "Fast effective rule induction," in *Proceedings of the twelfth international conference on machine learning*, 1995, pp. 115–123.

[72] R. Rastogi and K. Shim, "PUBLIC: A decision tree classifier that integrates building and pruning," in *VLDB'98, Proceedings of 24rd International Conference on Very Large Data Bases, August 24-27, 1998, New York City, New York, USA*, A. Gupta, O. Shmueli, and J. Widom, Eds., Morgan Kaufmann, 1998, pp. 404–415, ISBN: 1-55860-566-5. [Online]. Available: `http://www.vldb.org/conf/1998/p404.pdf`.

[74] J. C. Shafer, R. Agrawal, and M. Mehta, "SPRINT: A scalable parallel classifier for data mining," in *VLDB'96, Proceedings of 22th International Conference on Very Large Data Bases, September 3-6, 1996, Mumbai (Bombay), India*, T. M. Vijayaraman, A. P. Buchmann, C. Mohan, and N. L. Sarda, Eds., Morgan Kaufmann, 1996, pp. 544–555, ISBN: 1-55860-382-4. [Online]. Available: `http://www.vldb.org/conf/1996/P544.PDF`.

[75] Y. Shih, "Families of splitting criteria for classification trees," *Stat. Comput.*, vol. 9, no. 4, pp. 309–315, 1999. DOI: `10.1023/A:1008920224518`. [Online]. Available: `https://doi.org/10.1023/A:1008920224518`.

[76] C. M. Bishop, *Neural Networks for Pattern Recognition*. USA: Oxford University Press, Inc., 1995, ISBN: 0198538642.

[77] Y. Chauvin and D. E. Rumelhart, *Backpropagation: Theory, Architectures, and Applications*. Lawrence Erlbaum Associates, 1995, ISBN: 0-8058-1259-8. DOI: `10.4324/9780203763247`.

[78] J. W. Shavlik, R. J. Mooney, and G. G. Towell, "Symbolic and neural learning algorithms: An experimental comparison," *Mach. Learn.*, vol. 6, pp. 111–143, 1991. DOI: 10.1007/BF00114160. [Online]. Available: https://doi.org/10.1007/BF00114160.

[79] S. Haykin, *Neural Networks and Learning Machines*. Prentice Hall, 2008, 3rd edition, ISBN: 978-0-13-147139-9. DOI: 10.4324/9780203763247.

[80] J. A. Hertz, A. Krogh, and R. G. Palmer, *Introduction to the theory of neural computation* (The advanced book program). Addison-Wesley, 1991, vol. 1, ISBN: 0201515601. [Online]. Available: https://www.worldcat.org/oclc/21522159.

[81] R. Hecht-Nielsen, *Neurocomputing*. Addison-Wesley, 1990, ISBN: 978-0-201-09355-1.

[82] B. D. Ripley, *Pattern Recognition and Neural Networks*. Cambridge University Press, 1996, ISBN: 978-0-521-46086-6. DOI: 10.1017/CBO9780511812651. [Online]. Available: https://doi.org/10.1017/CBO9780511812651.

[84] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Min. Knowl. Discov.*, vol. 2, no. 2, pp. 121–167, 1998. DOI: 10.1023/A:1009715923555. [Online]. Available: https://doi.org/10.1023/A:1009715923555.

[85] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press, 2010, ISBN: 978-0-521-78019-3. DOI: 10.1017/CBO9780511801389.

[86] H. Drucker, C. J. C. Burges, L. Kaufman, A. J. Smola, and V. Vapnik, "Support vector regression machines," in *Advances in Neural Information Processing Systems 9, NIPS, Denver, CO, USA, December 2-5, 1996*, M. Mozer, M. I. Jordan, and T. Petsche, Eds., MIT Press, 1996, pp. 155–161. [Online]. Available: http://papers.nips.cc/paper/1238-support-vector-regression-machines.

[88] J. C. Platt, "Fast training of support vector machines using sequential minimal optimization," in *Advances in Kernel Methods: Support Vector Learning*. Cambridge, MA, USA: MIT Press, 1999, pp. 185–208, ISBN: 0262194163.

[89] B. Schölkopf, P. L. Bartlett, A. J. Smola, and R. C. Williamson, "Shrinking the tube: A new support vector regression algorithm," in *Advances in Neural Information Processing Systems 11, [NIPS Conference, Denver, Colorado, USA, November 30 - December 5, 1998]*, M. J. Kearns, S. A. Solla, and D. A. Cohn, Eds., The MIT Press, 1998, pp. 330–336, ISBN: 0-262-11245-0. [Online]. Available: http://papers.nips.cc/paper/1563-shrinking-the-tube-a-new-support-vector-regression-algorithm.

[90] M. J. Kearns, S. A. Solla, and D. A. Cohn, Eds., *Advances in Neural Information Processing Systems 11, [NIPS Conference, Denver, Colorado, USA, November 30 - December 5, 1998]*, The MIT Press, 1999, ISBN: 0-262-11245-0. [Online]. Available: https://proceedings.neurips.cc/paper/1998.

[91] H. Yu, J. Yang, and J. Han, "Classifying large data sets using svms with hierarchical clusters," in *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, August 24 - 27, 2003*, L. Getoor, T. E. Senator, P. M. Domingos, and C. Faloutsos, Eds., ACM, 2003, pp. 306–315, ISBN: 1-58113-737-0. DOI: 10.1145/956750.956786. [Online]. Available: https://doi.org/10.1145/956750.956786.

[92] L. Getoor, T. E. Senator, P. M. Domingos, and C. Faloutsos, Eds., *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, August 24 - 27, 2003*, ACM, 2003, ISBN: 1-58113-737-0.

[93] H. Cheng, X. Yan, J. Han, and C. Hsu, "Discriminative frequent pattern analysis for effective classification," in *Proceedings of the 23rd International Conference on Data Engineering, ICDE 2007, The Marmara Hotel, Istanbul, Turkey, April 15-20, 2007*, R. Chirkova, A. Dogac, M. T. Özsu, and T. K. Sellis, Eds., IEEE Computer Society, 2007, pp. 716–725, ISBN: 1-4244-0802-4. DOI: 10.1109/ICDE.2007.367917. [Online]. Available: https://doi.org/10.1109/ICDE.2007.367917.

[95] H. Cheng, X. Yan, J. Han, and P. S. Yu, "Direct discriminative pattern mining for effective classification," in *Proceedings of the 24th International Conference on Data Engineering, ICDE 2008, April 7-12, 2008, Cancún, Mexico*, G. Alonso, J. A. Blakeley, and A. L. P. Chen, Eds., IEEE Computer Society, 2008, pp. 169–178, ISBN: 978-1-4244-1836-7. DOI: 10.1109/ICDE.2008.4497425. [Online]. Available: https://doi.org/10.1109/ICDE.2008.4497425.

[97] G. Cong, K. Tan, A. K. H. Tung, and X. Xu, "Mining top-k covering rule groups for gene expression data," in *Proceedings of the ACM SIGMOD International Conference on Management of Data, Baltimore, Maryland, USA, June 14-16, 2005*, F. Özcan, Ed., ACM, 2005, pp. 670–681, ISBN: 1-59593-060-4. DOI: 10.1145/1066157.1066234. [Online]. Available: https://doi.org/10.1145/1066157.1066234.

[99] G. Dong and J. Li, "Efficient mining of emerging patterns: Discovering trends and differences," in *Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Diego, CA, USA, August 15-18, 1999*, U. M. Fayyad, S. Chaudhuri, and D. Madigan, Eds., ACM, 1999, pp. 43–52, ISBN: 1-58113-143-7. DOI: 10.1145/312129.312191. [Online]. Available: https://doi.org/10.1145/312129.312191.

[100] H. Kim, S. Kim, T. Weninger, J. Han, and T. F. Abdelzaher, "Ndpmine: Efficiently mining discriminative numerical features for pattern-based classification," in *Machine Learning and Knowledge Discovery in Databases, European Conference, ECML PKDD 2010, Barcelona, Spain, September 20-24, 2010, Proceedings, Part II*, J. L. Balcázar, F. Bonchi, A. Gionis, and M. Sebag, Eds., ser. Lecture Notes in Computer Science, vol. 6322, Springer, 2010, pp. 35–50, ISBN: 978-3-642-15882-7. DOI: 10.1007/978-3-642-15883-4_3. [Online]. Available: https://doi.org/10.1007/978-3-642-15883-4%5C_3.

[102] W. Li, J. Han, and J. Pei, "CMAR: accurate and efficient classification based on multiple class-association rules," in *Proceedings of the 2001 IEEE International Conference on Data Mining, 29 November - 2 December 2001, San Jose, California, USA*, N. Cercone, T. Y. Lin, and X. Wu, Eds., IEEE Computer Society, 2001, pp. 369–376, ISBN: 0-7695-1119-8. DOI: `10.1109/ICDM.2001.989541`. [Online]. Available: `https://doi.org/10.1109/ICDM.2001.989541`.

[104] B. Liu, W. Hsu, and Y. Ma, "Integrating classification and association rule mining," in *Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining (KDD-98), New York City, New York, USA, August 27-31, 1998*, R. Agrawal, P. E. Stolorz, and G. Piatetsky-Shapiro, Eds., AAAI Press, 1998, pp. 80–86, ISBN: 1-57735-070-7. [Online]. Available: `http://www.aaai.org/Library/KDD/1998/kdd98-012.php`.

[106] J. Wang and G. Karypis, "HARMONY: efficiently mining the best rules for classification," in *Proceedings of the 2005 SIAM International Conference on Data Mining, SDM 2005, Newport Beach, CA, USA, April 21-23, 2005*, H. Kargupta, J. Srivastava, C. Kamath, and A. Goodman, Eds., SIAM, 2005, pp. 205–216, ISBN: 978-0-89871-593-4. DOI: `10.1137/1.9781611972757.19`. [Online]. Available: `https://doi.org/10.1137/1.9781611972757.19`.

[108] S. L. Crawford, "Extensions to the CART algorithm," *Int. J. Man Mach. Stud.*, vol. 31, no. 2, pp. 197–217, 1989. DOI: `10.1016/0020-7373(89)90027-8`. [Online]. Available: `https://doi.org/10.1016/0020-7373(89)90027-8`.

[110] J. R. Quinlan and R. M. Cameron-Jones, "FOIL: A midterm report," in *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5-7, 1993, Proceedings*, P. Brazdil, Ed., ser. Lecture Notes in Computer Science, vol. 667, Springer, 1993, pp. 3–20, ISBN: 3-540-56602-3. DOI: `10.1007/3-540-56602-3\_124`. [Online]. Available: `https://doi.org/10.1007/3-540-56602-3%5C_124`.

[113] X. Yin and J. Han, "CPAR: classification based on predictive association rules," in *Proceedings of the Third SIAM International Conference on Data Mining, San Francisco, CA, USA, May 1-3, 2003*, D. Barbará and C. Kamath, Eds., SIAM, 2003, pp. 331–335, ISBN: 978-0-89871-545-3. DOI: `10.1137/1.9781611972733.40`. [Online]. Available: `https://doi.org/10.1137/1.9781611972733.40`.

[115] A. Aamodt and E. Plaza, "Case-based reasoning: Foundational issues, methodological variations, and system approaches," *AI Commun.*, vol. 7, no. 1, pp. 39–59, 1994. DOI: `10.3233/AIC-1994-7104`. [Online]. Available: `https://doi.org/10.3233/AIC-1994-7104`.

[116] T. M. Cover and P. E. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. Theory*, vol. 13, no. 1, pp. 21–27, 1967. DOI: `10.1109/TIT.1967.1053964`. [Online]. Available: `https://doi.org/10.1109/TIT.1967.1053964`.

[117] B. Dasarathy, *Nearest Neighbor (NN) Norms: NN Pattern Classification Techniques* (IEEE Computer Society Press tutorial). IEEE Computer Society Press, 1991, ISBN: 9780818689307.

[118] J. L. Kolodner, *Case-Based Reasoning.* Morgan Kaufmann, 1993, ISBN: 9781558602373. DOI: `10.1016/c2009-0-27670-7`. [Online]. Available: `https://doi.org/10.1016/c2009-0-27670-7`.

[119] A. Veloso, W. M. Jr., and M. J. Zaki, "Lazy associative classification," in *Proceedings of the 6th IEEE International Conference on Data Mining (ICDM 2006), 18-22 December 2006, Hong Kong, China*, IEEE Computer Society, 2006, pp. 645–654, ISBN: 0-7695-2701-9. DOI: `10.1109/ICDM.2006.96`. [Online]. Available: `https://doi.org/10.1109/ICDM.2006.96`.

[121] A. J. Dobson and A. G. Barnett, *An Introduction to Generalized Linear Models.* Chapmann and Hall/CRC, 2018, ISBN: 9781138741515.

[122] D. Heckerman, D. Geiger, and D. M. Chickering, "Learning bayesian networks: The combination of knowledge and statistical data," *Mach. Learn.*, vol. 20, no. 3, pp. 197–243, 1995. DOI: `10.1007/BF00994016`. [Online]. Available: `https://doi.org/10.1007/BF00994016`.

[123] G. F. Cooper and E. Herskovits, "A bayesian method for the induction of probabilistic networks from data," *Mach. Learn.*, vol. 9, pp. 309–347, 1992. DOI: `10.1007/BF00994110`. [Online]. Available: `https://doi.org/10.1007/BF00994110`.

[124] A. Darwiche, "Bayesian networks," *Commun. ACM*, vol. 53, no. 12, pp. 80–90, 2010. DOI: `10.1145/1859204.1859227`. [Online]. Available: `https://doi.org/10.1145/1859204.1859227`.

[125] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 39, no. 1, pp. 1–22, 1977. DOI: `https://doi.org/10.1111/j.2517-6161.1977.tb01600.x`. [Online]. Available: `https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.2517-6161.1977.tb01600.x`.

[126] F. V. Jensen, *Introduction to Bayesian Networks*, 1st. Berlin, Heidelberg: Springer-Verlag, 1996, ISBN: 0387915028.

[127] D. Koller and N. Friedman, *Probabilistic Graphical Models - Principles and Techniques.* MIT Press, 2009, ISBN: 978-0-262-01319-2. [Online]. Available: `http://mitpress.mit.edu/catalog/item/default.asp?ttype=2%5C&tid=11886`.

[128] J. Pearl, *Probabilistic reasoning in intelligent systems - networks of plausible inference* (Morgan Kaufmann series in representation and reasoning). Morgan Kaufmann, 1989.

[129] S. J. Russell, J. Binder, D. Koller, and K. Kanazawa, "Local learning in probabilistic networks with hidden variables," in *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence, IJCAI 95, Montréal Québec, Canada, August 20-25 1995, 2 Volumes*, Morgan Kaufmann, 1995, pp. 1146–1152. [Online]. Available: `http://ijcai.org/Proceedings/95-2/Papers/017.pdf`.

[131] V. Vapnik, *Statistical learning theory.* Wiley, 1998, ISBN: 978-0-471-03003-4.

[132] D. E. Goldberg, *Genetic Algorithms in Search Optimization and Machine Learning.* Addison-Wesley, 1989, ISBN: 0-201-15767-5.

[133] S. A. Harp, T. Samad, and A. Guha, "Designing application-specific neural networks using the genetic algorithm," in *Advances in Neural Information Processing Systems 2, [NIPS Conference, Denver, Colorado, USA, November 27-30, 1989]*, D. S. Touretzky, Ed., Morgan Kaufmann, 1989, pp. 447–454, ISBN: 1-55860-100-7. [Online]. Available: `http://papers.nips.cc/paper/263-designing-application-specific-neural-networks-using-the-genetic-algorithm`.

[135] Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs, Third Revised and Extended Edition.* Springer, 1996, ISBN: 978-3-540-60676-5. DOI: `10.1007/978-3-662-03315-9`. [Online]. Available: `https://doi.org/10.1007/978-3-662-03315-9`.

[136] M. Mitchell, *An introduction to genetic algorithms.* MIT Press, 1998, ISBN: 978-0-262-63185-3.

[137] Z. Pawlak, *Rough sets - theoretical aspects of reasoning about data* (Theory and decision library : series D). Kluwer, 1991, vol. 9, ISBN: 978-0-7923-1472-1. DOI: `10.1007/978-94-011-3534-4`. [Online]. Available: `https://doi.org/10.1007/978-94-011-3534-4`.

[138] R. R. Yager and L. A. Zadeh, *Fuzzy sets, neural networks, and soft computing.* Van Nostrand Reinhold, 1994, ISBN: 978-0-442-01621-0.

[139] O. Chapelle, B. Schölkopf, and A. Zien, Eds., *Semi-Supervised Learning.* The MIT Press, 2006, ISBN: 9780262033589. DOI: `10.7551/mitpress/9780262033589.001.0001`. [Online]. Available: `https://doi.org/10.7551/mitpress/9780262033589.001.0001`.

[140] T. G. Dietterich and G. Bakiri, "Solving multiclass learning problems via error-correcting output codes," *J. Artif. Intell. Res.*, vol. 2, pp. 263–286, 1995. DOI: `10.1613/jair.105`. [Online]. Available: `https://doi.org/10.1613/jair.105`.

[141] W. Dai, Q. Yang, G. Xue, and Y. Yu, "Boosting for transfer learning," in *Machine Learning, Proceedings of the Twenty-Fourth International Conference (ICML 2007), Corvallis, Oregon, USA, June 20-24, 2007*, Z. Ghahramani, Ed., ser. ACM International Conference Proceeding Series, vol. 227, ACM, 2007, pp. 193–200, ISBN: 978-1-59593-793-3. DOI: `10.1145/1273496.1273521`. [Online]. Available: `https://doi.org/10.1145/1273496.1273521`.

[142] Z. Ghahramani, Ed., *Machine Learning, Proceedings of the Twenty-Fourth International Conference (ICML 2007), Corvallis, Oregon, USA, June 20-24, 2007*, vol. 227, ACM International Conference Proceeding Series, ACM, 2007, ISBN: 978-1-59593-793-3.

[143] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, 2010. DOI: `10.1109/TKDE.2009.191`. [Online]. Available: `https://doi.org/10.1109/TKDE.2009.191`.

[144] Q. Yang, Y. Zhang, W. Dai, and S. J. Pan, *Transfer Learning.* Cambridge University Press, 2020, ISBN: 978-1107016903. DOI: `10.1017/9781139061773`.

[145]   B. Settles, "Active learning literature survey," University of Wisconsin-Madison Department of Computer Sciences, Tech. Rep. TR1648, Jan. 2009, p. 47. [Online]. Available: http://digital.library.wisc.edu/1793/60660.

[146]   X. Zhu, "Semi-supervised learning literature survey," University of Wisconsin-Madison Department of Computer Sciences, Tech. Rep. TR1530, Sep. 2005, p. 39. [Online]. Available: http://digital.library.wisc.edu/1793/60444.

[147]   L. Breiman, "Bagging predictors," *Mach. Learn.*, vol. 24, no. 2, pp. 123–140, 1996. DOI: 10.1007/BF00058655. [Online]. Available: https://doi.org/10.1007/BF00058655.

[148]   L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001. DOI: 10.1023/A:1010933404324. [Online]. Available: https://doi.org/10.1023/A:1010933404324.

[149]   C. Elkan, "The foundations of cost-sensitive learning," in *Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence, IJCAI 2001, Seattle, Washington, USA, August 4-10, 2001*, B. Nebel, Ed., Morgan Kaufmann, 2001, pp. 973–978, ISBN: 1-55860-777-3. [Online]. Available: http://ijcai.org/proceedings/2001-1.

[151]   B. Efron and R. Tibshirani, *An Introduction to the Bootstrap*. Springer, 1993, ISBN: 978-1-4899-4541-9. DOI: 10.1007/978-1-4899-4541-9. [Online]. Available: https://doi.org/10.1007/978-1-4899-4541-9.

[152]   J. H. Friedman and B. E. Popescu, "Predictive learning via rule ensembles," *The Annals of Applied Statistics*, vol. 2, no. 3, pp. 916–954, 2008. DOI: 10.1214/07-AOAS148. [Online]. Available: https://doi.org/10.1214/07-AOAS148.

[153]   T. Lim, W. Loh, and Y. Shih, "A comparison of prediction accuracy, complexity, and training time of thirty-three old and new classification algorithms," *Mach. Learn.*, vol. 40, no. 3, pp. 203–228, 2000. DOI: 10.1023/A:1007608224229. [Online]. Available: https://doi.org/10.1023/A:1007608224229.

[154]   J. Magidson, "The chaid approach to segmentation modeling: Chi-squared automatic interaction detection," *Advanced methods of marketing research*, pp. 118–159, 1994.

[155]   J. R. Quinlan, "Bagging, boosting, and C4.5," in *Proceedings of the Thirteenth National Conference on Artificial Intelligence and Eighth Innovative Applications of Artificial Intelligence Conference, AAAI 96, IAAI 96, Portland, Oregon, USA, August 4-8, 1996, Volume 1*, W. J. Clancey and D. S. Weld, Eds., AAAI Press / The MIT Press, 1996, pp. 725–730, ISBN: 0-262-51091-X. [Online]. Available: http://www.aaai.org/Library/AAAI/1996/aaai96-108.php.

[157]   G. Seni and J. F. E. IV, *Ensemble Methods in Data Mining: Improving Accuracy Through Combining Predictions* (Synthesis Lectures on Data Mining and Knowledge Discovery). Morgan & Claypool Publishers, 2010. DOI: 10.2200/S00240ED1V01Y200912DMK002. [Online]. Available: https://doi.org/10.2200/S00240ED1V01Y200912DMK002.

[158] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *J. Comput. Syst. Sci.*, vol. 55, no. 1, pp. 119–139, 1997. DOI: `10.1006/jcss.1997.1504`. [Online]. Available: `https://doi.org/10.1006/jcss.1997.1504`.

[159] Y. Freund and R. E. Schapire, "Experiments with a new boosting algorithm," in *Machine Learning, Proceedings of the Thirteenth International Conference (ICML '96), Bari, Italy, July 3-6, 1996*, L. Saitta, Ed., Morgan Kaufmann, 1996, pp. 148–156, ISBN: 1-55860-419-7.

[264] H. Cheng, X. Yan, J. Han, and C.-W. Hsu, "Discriminative frequent pattern analysis for effective classification," in *2007 IEEE 23rd international conference on data engineering*, IEEE, 2006, pp. 716–725.

Decision Trees

[52] M. Ankerst, C. Elsen, M. Ester, and H. Kriegel, "Visual classification: An interactive approach to decision tree construction," in *Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Diego, CA, USA, August 15-18, 1999*, U. M. Fayyad, S. Chaudhuri, and D. Madigan, Eds., ACM, 1999, pp. 392–396, ISBN: 1-58113-143-7. DOI: `10.1145/312129.312298`. [Online]. Available: `https://doi.org/10.1145/312129.312298`.

[54] C. Apté and S. M. Weiss, "Data mining with decision trees and decision rules," *Future Gener. Comput. Syst.*, vol. 13, no. 2-3, pp. 197–210, 1997. DOI: `10.1016/S0167-739X(97)00021-6`. [Online]. Available: `https://doi.org/10.1016/S0167-739X(97)00021-6`.

[55] C. E. Brodley and P. E. Utgoff, "Multivariate decision trees," *Mach. Learn.*, vol. 19, no. 1, pp. 45–77, 1995. DOI: `10.1007/BF00994660`. [Online]. Available: `https://doi.org/10.1007/BF00994660`.

[56] P. K. Chan and S. J. Stolfo, "Learning arbiter and combiner trees from partitioned data for scaling machine learning," in *Proceedings of the First International Conference on Knowledge Discovery and Data Mining (KDD-95), Montreal, Canada, August 20-21, 1995*, U. M. Fayyad and R. Uthurusamy, Eds., AAAI Press, 1995, pp. 39–44, ISBN: 0-929280-82-2. [Online]. Available: `http://www.aaai.org/Library/KDD/1995/kdd95-047.php`.

[58] U. M. Fayyad, "Branching on attribute values in decision tree generation," in *Proceedings of the 12th National Conference on Artificial Intelligence, Seattle, WA, USA, July 31 - August 4, 1994, Volume 1*, B. Hayes-Roth and R. E. Korf, Eds., AAAI Press / The MIT Press, 1994, pp. 601–606, ISBN: 0-262-61102-3. [Online]. Available: `http://www.aaai.org/Library/AAAI/1994/aaai94-091.php`.

[60] M. Mehta, R. Agrawal, and J. Rissanen, "SLIQ: A fast scalable classifier for data mining," in *Advances in Database Technology - EDBT'96, 5th International Conference on Extending Database Technology, Avignon, France, March 25-29, 1996, Proceedings*, P. M. G. Apers, M. Bouzeghoub, and G. Gardarin, Eds., ser. Lecture Notes in Computer Science, vol. 1057, Springer, 1996, pp. 18–32, ISBN: 3-540-61057-X. DOI: 10.1007/BFb0014141. [Online]. Available: https://doi.org/10.1007/BFb0014141.

[62] J. Gehrke, R. Ramakrishnan, and V. Ganti, "Rainforest - A framework for fast decision tree construction of large datasets," in *VLDB'98, Proceedings of 24rd International Conference on Very Large Data Bases, August 24-27, 1998, New York City, New York, USA*, A. Gupta, O. Shmueli, and J. Widom, Eds., Morgan Kaufmann, 1998, pp. 416–427, ISBN: 1-55860-566-5. [Online]. Available: http://www.vldb.org/conf/1998/p416.pdf.

[63] J. Gehrke, V. Ganti, R. Ramakrishnan, and W. Loh, "Boat-optimistic decision tree construction," in *SIGMOD 1999, Proceedings ACM SIGMOD International Conference on Management of Data, June 1-3, 1999, Philadelphia, Pennsylvania, USA*, A. Delis, C. Faloutsos, and S. Ghandeharizadeh, Eds., ACM Press, 1999, pp. 169–180, ISBN: 1-58113-084-8. DOI: 10.1145/304182.304197. [Online]. Available: https://doi.org/10.1145/304182.304197.

[65] S. K. Murthy, "Automatic construction of decision trees from data: A multi-disciplinary survey," *Data Min. Knowl. Discov.*, vol. 2, no. 4, pp. 345–389, 1998. DOI: 10.1023/A:1009744630224. [Online]. Available: https://doi.org/10.1023/A:1009744630224.

[66] J. R. Quinlan, "Induction of decision trees," *Mach. Learn.*, vol. 1, no. 1, pp. 81–106, 1986. DOI: 10.1023/A:1022643204877. [Online]. Available: https://doi.org/10.1023/A:1022643204877.

[67] J. R. Quinlan and R. L. Rivest, "Inferring decision trees using the minimum description length principle," *Inf. Comput.*, vol. 80, no. 3, pp. 227–248, 1989. DOI: 10.1016/0890-5401(89)90010-2. [Online]. Available: https://doi.org/10.1016/0890-5401(89)90010-2.

[72] R. Rastogi and K. Shim, "PUBLIC: A decision tree classifier that integrates building and pruning," in *VLDB'98, Proceedings of 24rd International Conference on Very Large Data Bases, August 24-27, 1998, New York City, New York, USA*, A. Gupta, O. Shmueli, and J. Widom, Eds., Morgan Kaufmann, 1998, pp. 404–415, ISBN: 1-55860-566-5. [Online]. Available: http://www.vldb.org/conf/1998/p404.pdf.

[74] J. C. Shafer, R. Agrawal, and M. Mehta, "SPRINT: A scalable parallel classifier for data mining," in *VLDB'96, Proceedings of 22th International Conference on Very Large Data Bases, September 3-6, 1996, Mumbai (Bombay), India*, T. M. Vijayaraman, A. P. Buchmann, C. Mohan, and N. L. Sarda, Eds., Morgan Kaufmann, 1996, pp. 544–555, ISBN: 1-55860-382-4. [Online]. Available: http://www.vldb.org/conf/1996/P544.PDF.

[75] Y. Shih, "Families of splitting criteria for classification trees," *Stat. Comput.*, vol. 9, no. 4, pp. 309–315, 1999. DOI: `10.1023/A:1008920224518`. [Online]. Available: `https://doi.org/10.1023/A:1008920224518`.

Neural Networks

[76] C. M. Bishop, *Neural Networks for Pattern Recognition*. USA: Oxford University Press, Inc., 1995, ISBN: 0198538642.

[77] Y. Chauvin and D. E. Rumelhart, *Backpropagation: Theory, Architectures, and Applications*. Lawrence Erlbaum Associates, 1995, ISBN: 0-8058-1259-8. DOI: `10.4324/9780203763247`.

[78] J. W. Shavlik, R. J. Mooney, and G. G. Towell, "Symbolic and neural learning algorithms: An experimental comparison," *Mach. Learn.*, vol. 6, pp. 111–143, 1991. DOI: `10.1007/BF00114160`. [Online]. Available: `https://doi.org/10.1007/BF00114160`.

[79] S. Haykin, *Neural Networks and Learning Machines*. Prentice Hall, 2008, 3rd edition, ISBN: 978-0-13-147139-9. DOI: `10.4324/9780203763247`.

[80] J. A. Hertz, A. Krogh, and R. G. Palmer, *Introduction to the theory of neural computation* (The advanced book program). Addison-Wesley, 1991, vol. 1, ISBN: 0201515601. [Online]. Available: `https://www.worldcat.org/oclc/21522159`.

[81] R. Hecht-Nielsen, *Neurocomputing*. Addison-Wesley, 1990, ISBN: 978-0-201-09355-1.

[82] B. D. Ripley, *Pattern Recognition and Neural Networks*. Cambridge University Press, 1996, ISBN: 978-0-521-46086-6. DOI: `10.1017/CBO9780511812651`. [Online]. Available: `https://doi.org/10.1017/CBO9780511812651`.

Support Vector Machines

[84] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Min. Knowl. Discov.*, vol. 2, no. 2, pp. 121–167, 1998. DOI: `10.1023/A:1009715923555`. [Online]. Available: `https://doi.org/10.1023/A:1009715923555`.

[85] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press, 2010, ISBN: 978-0-521-78019-3. DOI: `10.1017/CBO9780511801389`.

[86] H. Drucker, C. J. C. Burges, L. Kaufman, A. J. Smola, and V. Vapnik, "Support vector regression machines," in *Advances in Neural Information Processing Systems 9, NIPS, Denver, CO, USA, December 2-5, 1996*, M. Mozer, M. I. Jordan, and T. Petsche, Eds., MIT Press, 1996, pp. 155–161. [Online]. Available: `http://papers.nips.cc/paper/1238-support-vector-regression-machines`.

[88] J. C. Platt, "Fast training of support vector machines using sequential minimal optimization," in *Advances in Kernel Methods: Support Vector Learning*. Cambridge, MA, USA: MIT Press, 1999, pp. 185–208, ISBN: 0262194163.

[89] B. Schölkopf, P. L. Bartlett, A. J. Smola, and R. C. Williamson, "Shrinking the tube: A new support vector regression algorithm," in *Advances in Neural Information Processing Systems 11, [NIPS Conference, Denver, Colorado, USA, November 30 - December 5, 1998]*, M. J. Kearns, S. A. Solla, and D. A. Cohn, Eds., The MIT Press, 1998, pp. 330–336, ISBN: 0-262-11245-0. [Online]. Available: `http://papers.nips.cc/paper/1563-shrinking-the-tube-a-new-support-vector-regression-algorithm`.

[90] M. J. Kearns, S. A. Solla, and D. A. Cohn, Eds., *Advances in Neural Information Processing Systems 11, [NIPS Conference, Denver, Colorado, USA, November 30 - December 5, 1998]*, The MIT Press, 1999, ISBN: 0-262-11245-0. [Online]. Available: `https://proceedings.neurips.cc/paper/1998`.

[91] H. Yu, J. Yang, and J. Han, "Classifying large data sets using svms with hierarchical clusters," in *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, August 24 - 27, 2003*, L. Getoor, T. E. Senator, P. M. Domingos, and C. Faloutsos, Eds., ACM, 2003, pp. 306–315, ISBN: 1-58113-737-0. DOI: `10.1145/956750.956786`. [Online]. Available: `https://doi.org/10.1145/956750.956786`.

[92] L. Getoor, T. E. Senator, P. M. Domingos, and C. Faloutsos, Eds., *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, August 24 - 27, 2003*, ACM, 2003, ISBN: 1-58113-737-0.

Pattern-based Classification

[93] H. Cheng, X. Yan, J. Han, and C. Hsu, "Discriminative frequent pattern analysis for effective classification," in *Proceedings of the 23rd International Conference on Data Engineering, ICDE 2007, The Marmara Hotel, Istanbul, Turkey, April 15-20, 2007*, R. Chirkova, A. Dogac, M. T. Özsu, and T. K. Sellis, Eds., IEEE Computer Society, 2007, pp. 716–725, ISBN: 1-4244-0802-4. DOI: `10.1109/ICDE.2007.367917`. [Online]. Available: `https://doi.org/10.1109/ICDE.2007.367917`.

[95] H. Cheng, X. Yan, J. Han, and P. S. Yu, "Direct discriminative pattern mining for effective classification," in *Proceedings of the 24th International Conference on Data Engineering, ICDE 2008, April 7-12, 2008, Cancún, Mexico*, G. Alonso, J. A. Blakeley, and A. L. P. Chen, Eds., IEEE Computer Society, 2008, pp. 169–178, ISBN: 978-1-4244-1836-7. DOI: `10.1109/ICDE.2008.4497425`. [Online]. Available: `https://doi.org/10.1109/ICDE.2008.4497425`.

[97] G. Cong, K. Tan, A. K. H. Tung, and X. Xu, "Mining top-k covering rule groups for gene expression data," in *Proceedings of the ACM SIGMOD International Conference on Management of Data, Baltimore, Maryland, USA, June 14-16, 2005*, F. Özcan, Ed., ACM, 2005, pp. 670–681, ISBN: 1-59593-060-4. DOI: `10.1145/1066157.1066234`. [Online]. Available: `https://doi.org/10.1145/1066157.1066234`.

[99] G. Dong and J. Li, "Efficient mining of emerging patterns: Discovering trends and differences," in *Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Diego, CA, USA, August 15-18, 1999*, U. M. Fayyad, S. Chaudhuri, and D. Madigan, Eds., ACM, 1999, pp. 43–52, ISBN: 1-58113-143-7. DOI: `10.1145/312129.312191`. [Online]. Available: `https://doi.org/10.1145/312129.312191`.

[100] H. Kim, S. Kim, T. Weninger, J. Han, and T. F. Abdelzaher, "Ndpmine: Efficiently mining discriminative numerical features for pattern-based classification," in *Machine Learning and Knowledge Discovery in Databases, European Conference, ECML PKDD 2010, Barcelona, Spain, September 20-24, 2010, Proceedings, Part II*, J. L. Balcázar, F. Bonchi, A. Gionis, and M. Sebag, Eds., ser. Lecture Notes in Computer Science, vol. 6322, Springer, 2010, pp. 35–50, ISBN: 978-3-642-15882-7. DOI: `10.1007/978-3-642-15883-4\_3`. [Online]. Available: `https://doi.org/10.1007/978-3-642-15883-4%5C_3`.

[102] W. Li, J. Han, and J. Pei, "CMAR: accurate and efficient classification based on multiple class-association rules," in *Proceedings of the 2001 IEEE International Conference on Data Mining, 29 November - 2 December 2001, San Jose, California, USA*, N. Cercone, T. Y. Lin, and X. Wu, Eds., IEEE Computer Society, 2001, pp. 369–376, ISBN: 0-7695-1119-8. DOI: `10.1109/ICDM.2001.989541`. [Online]. Available: `https://doi.org/10.1109/ICDM.2001.989541`.

[104] B. Liu, W. Hsu, and Y. Ma, "Integrating classification and association rule mining," in *Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining (KDD-98), New York City, New York, USA, August 27-31, 1998*, R. Agrawal, P. E. Stolorz, and G. Piatetsky-Shapiro, Eds., AAAI Press, 1998, pp. 80–86, ISBN: 1-57735-070-7. [Online]. Available: `http://www.aaai.org/Library/KDD/1998/kdd98-012.php`.

[106] J. Wang and G. Karypis, "HARMONY: efficiently mining the best rules for classification," in *Proceedings of the 2005 SIAM International Conference on Data Mining, SDM 2005, Newport Beach, CA, USA, April 21-23, 2005*, H. Kargupta, J. Srivastava, C. Kamath, and A. Goodman, Eds., SIAM, 2005, pp. 205–216, ISBN: 978-0-89871-593-4. DOI: `10.1137/1.9781611972757.19`. [Online]. Available: `https://doi.org/10.1137/1.9781611972757.19`.

Rule Induction

[68] J. R. Quinlan, "Learning logical definitions from relations," *Mach. Learn.*, vol. 5, pp. 239–266, 1990. DOI: `10.1007/BF00117105`. [Online]. Available: `https://doi.org/10.1007/BF00117105`.

[69] R. S. Michalski, I. Mozetic, J. Hong, and N. Lavrac, "The multi-purpose incremental learning system aq15 and its testing application to three medical domains," in *Proc. AAAI*, vol. 1986, 1986, pp. 1–041.

[70]   P. Clark and T. Niblett, "The cn2 induction algorithm," *Machine learning*, vol. 3, no. 4, pp. 261–283, 1989.

[71]   W. W. Cohen et al., "Fast effective rule induction," in *Proceedings of the twelfth international conference on machine learning*, 1995, pp. 115–123.

[108] S. L. Crawford, "Extensions to the CART algorithm," *Int. J. Man Mach. Stud.*, vol. 31, no. 2, pp. 197–217, 1989. DOI: `10.1016/0020-7373(89)90027-8`. [Online]. Available: `https://doi.org/10.1016/0020-7373(89)90027-8`.

[110] J. R. Quinlan and R. M. Cameron-Jones, "FOIL: A midterm report," in *Machine Learning: ECML-93, European Conference on Machine Learning, Vienna, Austria, April 5-7, 1993, Proceedings*, P. Brazdil, Ed., ser. Lecture Notes in Computer Science, vol. 667, Springer, 1993, pp. 3–20, ISBN: 3-540-56602-3. DOI: `10.1007/3-540-56602-3\_124`. [Online]. Available: `https://doi.org/10.1007/3-540-56602-3%5C_124`.

[113] X. Yin and J. Han, "CPAR: classification based on predictive association rules," in *Proceedings of the Third SIAM International Conference on Data Mining, San Francisco, CA, USA, May 1-3, 2003*, D. Barbará and C. Kamath, Eds., SIAM, 2003, pp. 331–335, ISBN: 978-0-89871-545-3. DOI: `10.1137/1.9781611972733.40`. [Online]. Available: `https://doi.org/10.1137/1.9781611972733.40`.

k-NN & Case-based Reasoning

[115] A. Aamodt and E. Plaza, "Case-based reasoning: Foundational issues, methodological variations, and system approaches," *AI Commun.*, vol. 7, no. 1, pp. 39–59, 1994. DOI: `10.3233/AIC-1994-7104`. [Online]. Available: `https://doi.org/10.3233/AIC-1994-7104`.

[116] T. M. Cover and P. E. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. Theory*, vol. 13, no. 1, pp. 21–27, 1967. DOI: `10.1109/TIT.1967.1053964`. [Online]. Available: `https://doi.org/10.1109/TIT.1967.1053964`.

[117] B. Dasarathy, *Nearest Neighbor (NN) Norms: NN Pattern Classification Techniques* (IEEE Computer Society Press tutorial). IEEE Computer Society Press, 1991, ISBN: 9780818689307.

[118] J. L. Kolodner, *Case-Based Reasoning*. Morgan Kaufmann, 1993, ISBN: 9781558602373. DOI: `10.1016/c2009-0-27670-7`. [Online]. Available: `https://doi.org/10.1016/c2009-0-27670-7`.

[119] A. Veloso, W. M. Jr., and M. J. Zaki, "Lazy associative classification," in *Proceedings of the 6th IEEE International Conference on Data Mining (ICDM 2006), 18-22 December 2006, Hong Kong, China*, IEEE Computer Society, 2006, pp. 645–654, ISBN: 0-7695-2701-9. DOI: `10.1109/ICDM.2006.96`. [Online]. Available: `https://doi.org/10.1109/ICDM.2006.96`.

Bayesian Method & Statistical Models

[121] A. J. Dobson and A. G. Barnett, *An Introduction to Generalized Linear Models.* Chapmann and Hall/CRC, 2018, ISBN: 9781138741515.

[122] D. Heckerman, D. Geiger, and D. M. Chickering, "Learning bayesian networks: The combination of knowledge and statistical data," *Mach. Learn.*, vol. 20, no. 3, pp. 197–243, 1995. DOI: 10.1007/BF00994016. [Online]. Available: https://doi.org/10.1007/BF00994016.

[123] G. F. Cooper and E. Herskovits, "A bayesian method for the induction of probabilistic networks from data," *Mach. Learn.*, vol. 9, pp. 309–347, 1992. DOI: 10.1007/BF00994110. [Online]. Available: https://doi.org/10.1007/BF00994110.

[124] A. Darwiche, "Bayesian networks," *Commun. ACM*, vol. 53, no. 12, pp. 80–90, 2010. DOI: 10.1145/1859204.1859227. [Online]. Available: https://doi.org/10.1145/1859204.1859227.

[125] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 39, no. 1, pp. 1–22, 1977. DOI: https://doi.org/10.1111/j.2517-6161.1977.tb01600.x. [Online]. Available: https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.2517-6161.1977.tb01600.x.

[126] F. V. Jensen, *Introduction to Bayesian Networks*, 1st. Berlin, Heidelberg: Springer-Verlag, 1996, ISBN: 0387915028.

[127] D. Koller and N. Friedman, *Probabilistic Graphical Models - Principles and Techniques.* MIT Press, 2009, ISBN: 978-0-262-01319-2. [Online]. Available: http://mitpress.mit.edu/catalog/item/default.asp?ttype=2%5C&tid=11886.

[128] J. Pearl, *Probabilistic reasoning in intelligent systems - networks of plausible inference* (Morgan Kaufmann series in representation and reasoning). Morgan Kaufmann, 1989.

[129] S. J. Russell, J. Binder, D. Koller, and K. Kanazawa, "Local learning in probabilistic networks with hidden variables," in *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence, IJCAI 95, Montréal Québec, Canada, August 20-25 1995, 2 Volumes*, Morgan Kaufmann, 1995, pp. 1146–1152. [Online]. Available: http://ijcai.org/Proceedings/95-2/Papers/017.pdf.

[131] V. Vapnik, *Statistical learning theory.* Wiley, 1998, ISBN: 978-0-471-03003-4.

Semi-Supervised & Multi-Class Learning

[139] O. Chapelle, B. Schölkopf, and A. Zien, Eds., *Semi-Supervised Learning.* The MIT Press, 2006, ISBN: 9780262033589. DOI: 10.7551/mitpress/9780262033589.001.0001. [Online]. Available: https://doi.org/10.7551/mitpress/9780262033589.001.0001.

[140] T. G. Dietterich and G. Bakiri, "Solving multiclass learning problems via error-correcting output codes," *J. Artif. Intell. Res.*, vol. 2, pp. 263–286, 1995. DOI: 10.1613/jair.105. [Online]. Available: https://doi.org/10.1613/jair.105.

[141] W. Dai, Q. Yang, G. Xue, and Y. Yu, "Boosting for transfer learning," in *Machine Learning, Proceedings of the Twenty-Fourth International Conference (ICML 2007), Corvallis, Oregon, USA, June 20-24, 2007*, Z. Ghahramani, Ed., ser. ACM International Conference Proceeding Series, vol. 227, ACM, 2007, pp. 193–200, ISBN: 978-1-59593-793-3. DOI: `10.1145/1273496.1273521`. [Online]. Available: `https://doi.org/10.1145/1273496.1273521`.

[142] Z. Ghahramani, Ed., *Machine Learning, Proceedings of the Twenty-Fourth International Conference (ICML 2007), Corvallis, Oregon, USA, June 20-24, 2007*, vol. 227, ACM International Conference Proceeding Series, ACM, 2007, ISBN: 978-1-59593-793-3.

[143] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, 2010. DOI: `10.1109/TKDE.2009.191`. [Online]. Available: `https://doi.org/10.1109/TKDE.2009.191`.

[144] Q. Yang, Y. Zhang, W. Dai, and S. J. Pan, *Transfer Learning*. Cambridge University Press, 2020, ISBN: 978-1107016903. DOI: `10.1017/9781139061773`.

[145] B. Settles, "Active learning literature survey," University of Wisconsin-Madison Department of Computer Sciences, Tech. Rep. TR1648, Jan. 2009, p. 47. [Online]. Available: `http://digital.library.wisc.edu/1793/60660`.

[146] X. Zhu, "Semi-supervised learning literature survey," University of Wisconsin-Madison Department of Computer Sciences, Tech. Rep. TR1530, Sep. 2005, p. 39. [Online]. Available: `http://digital.library.wisc.edu/1793/60444`.

Genetic Algorithms & Rough/Fuzzy Sets

[132] D. E. Goldberg, *Genetic Algorithms in Search Optimization and Machine Learning*. Addison-Wesley, 1989, ISBN: 0-201-15767-5.

[133] S. A. Harp, T. Samad, and A. Guha, "Designing application-specific neural networks using the genetic algorithm," in *Advances in Neural Information Processing Systems 2, [NIPS Conference, Denver, Colorado, USA, November 27-30, 1989]*, D. S. Touretzky, Ed., Morgan Kaufmann, 1989, pp. 447–454, ISBN: 1-55860-100-7. [Online]. Available: `http://papers.nips.cc/paper/263-designing-application-specific-neural-networks-using-the-genetic-algorithm`.

[135] Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs, Third Revised and Extended Edition*. Springer, 1996, ISBN: 978-3-540-60676-5. DOI: `10.1007/978-3-662-03315-9`. [Online]. Available: `https://doi.org/10.1007/978-3-662-03315-9`.

[136] M. Mitchell, *An introduction to genetic algorithms*. MIT Press, 1998, ISBN: 978-0-262-63185-3.

[137] Z. Pawlak, *Rough sets - theoretical aspects of reasoning about data* (Theory and decision library : series D). Kluwer, 1991, vol. 9, ISBN: 978-0-7923-1472-1. DOI: `10.1007/978-94-011-3534-4`. [Online]. Available: `https://doi.org/10.1007/978-94-011-3534-4`.

[138] R. R. Yager and L. A. Zadeh, *Fuzzy sets, neural networks, and soft computing*. Van Nostrand Reinhold, 1994, ISBN: 978-0-442-01621-0.

Model Evaluation & Ensemble Methods

[147] L. Breiman, "Bagging predictors," *Mach. Learn.*, vol. 24, no. 2, pp. 123–140, 1996. DOI: `10.1007/BF00058655`. [Online]. Available: `https://doi.org/10.1007/BF00058655`.

[148] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001. DOI: `10.1023/A:1010933404324`. [Online]. Available: `https://doi.org/10.1023/A:1010933404324`.

[149] C. Elkan, "The foundations of cost-sensitive learning," in *Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence, IJCAI 2001, Seattle, Washington, USA, August 4-10, 2001*, B. Nebel, Ed., Morgan Kaufmann, 2001, pp. 973–978, ISBN: 1-55860-777-3. [Online]. Available: `http://ijcai.org/proceedings/2001-1`.

[151] B. Efron and R. Tibshirani, *An Introduction to the Bootstrap*. Springer, 1993, ISBN: 978-1-4899-4541-9. DOI: `10.1007/978-1-4899-4541-9`. [Online]. Available: `https://doi.org/10.1007/978-1-4899-4541-9`.

[152] J. H. Friedman and B. E. Popescu, "Predictive learning via rule ensembles," *The Annals of Applied Statistics*, vol. 2, no. 3, pp. 916–954, 2008. DOI: `10.1214/07-AOAS148`. [Online]. Available: `https://doi.org/10.1214/07-AOAS148`.

[153] T. Lim, W. Loh, and Y. Shih, "A comparison of prediction accuracy, complexity, and training time of thirty-three old and new classification algorithms," *Mach. Learn.*, vol. 40, no. 3, pp. 203–228, 2000. DOI: `10.1023/A:1007608224229`. [Online]. Available: `https://doi.org/10.1023/A:1007608224229`.

[154] J. Magidson, "The chaid approach to segmentation modeling: Chi-squared automatic interaction detection," *Advanced methods of marketing research*, pp. 118–159, 1994.

[155] J. R. Quinlan, "Bagging, boosting, and C4.5," in *Proceedings of the Thirteenth National Conference on Artificial Intelligence and Eighth Innovative Applications of Artificial Intelligence Conference, AAAI 96, IAAI 96, Portland, Oregon, USA, August 4-8, 1996, Volume 1*, W. J. Clancey and D. S. Weld, Eds., AAAI Press / The MIT Press, 1996, pp. 725–730, ISBN: 0-262-51091-X. [Online]. Available: `http://www.aaai.org/Library/AAAI/1996/aaai96-108.php`.

[157] G. Seni and J. F. E. IV, *Ensemble Methods in Data Mining: Improving Accuracy Through Combining Predictions* (Synthesis Lectures on Data Mining and Knowledge Discovery). Morgan & Claypool Publishers, 2010. DOI: `10.2200/S00240ED1V01Y200912DMK002`. [Online]. Available: `https://doi.org/10.2200/S00240ED1V01Y200912DMK002`.

[158] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *J. Comput. Syst. Sci.*, vol. 55, no. 1, pp. 119–139, 1997. DOI: `10.1006/jcss.1997.1504`. [Online]. Available: `https://doi.org/10.1006/jcss.1997.1504`.

[159]   Y. Freund and R. E. Schapire, "Experiments with a new boosting algorithm," in *Machine Learning, Proceedings of the Thirteenth International Conference (ICML '96), Bari, Italy, July 3-6, 1996*, L. Saitta, Ed., Morgan Kaufmann, 1996, pp. 148–156, ISBN: 1-55860-419-7.

8 Cluster Analysis

[11]   L. Kaufman and P. J. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis*. John Wiley, 1990, ISBN: 978-0-47187876-6. DOI: 10.1002/9780470316801.

[171]   M. M. Breunig, H. Kriegel, R. T. Ng, and J. Sander, "LOF: identifying density-based local outliers," in *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data, May 16-18, 2000, Dallas, Texas, USA*, W. Chen, J. F. Naughton, and P. A. Bernstein, Eds., ACM, 2000, pp. 93–104, ISBN: 1-58113-217-4. DOI: 10.1145/342009.335388. [Online]. Available: https://doi.org/10.1145/342009.335388.

[194]   E. M. Knorr and R. T. Ng, "Algorithms for mining distance-based outliers in large datasets," in *VLDB'98, Proceedings of 24rd International Conference on Very Large Data Bases, August 24-27, 1998, New York City, New York, USA*, A. Gupta, O. Shmueli, and J. Widom, Eds., Morgan Kaufmann, 1998, pp. 392–403, ISBN: 1-55860-566-5. [Online]. Available: http://www.vldb.org/conf/1998/p392.pdf.

[272]   R. Agrawal, J. Gehrke, D. Gunopulos, and P. Raghavan, "Automatic subspace clustering of high dimensional data for data mining applications," in *SIGMOD 1998, Proceedings ACM SIGMOD International Conference on Management of Data, June 2-4, 1998, Seattle, Washington, USA*, L. M. Haas and A. Tiwary, Eds., ACM Press, 1998, pp. 94–105.

[273]   M. Anderberg and O. O. T. A. F. S. S. K. A. N. MEX., *Cluster Analysis for Applications* (Probability and Mathematical Statistics : a series of monographs and textbooks). Academic Press, 1973, ISBN: 9780120576500.

[274]   M. Ankerst, M. M. Breunig, H. Kriegel, and J. Sander, "OPTICS: ordering points to identify the clustering structure," in *SIGMOD 1999, Proceedings ACM SIGMOD International Conference on Management of Data, June 1-3, 1999, Philadelphia, Pennsylvania, USA*, A. Delis, C. Faloutsos, and S. Ghandeharizadeh, Eds., ACM Press, 1999, pp. 49–60.

[275]   A. Banerjee and R. N. Davé, "Validating clusters using the hopkins statistic," in *IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2004, Budapest, Hungary, July 25-29, 2004*, IEEE, 2004, pp. 149–153. DOI: 10.1109/FUZZY.2004.1375706. [Online]. Available: https://doi.org/10.1109/FUZZY.2004.1375706.

[276]   F. Beil, M. Ester, and X. Xu, "Frequent term-based text clustering," in *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, July 23-26, 2002, Edmonton, Alberta, Canada*, ACM, 2002, pp. 436–442. DOI: 10.1145/775047.775110. [Online]. Available: https://doi.org/10.1145/775047.775110.

[277] M. Ester, H. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96), Portland, Oregon, USA*, E. Simoudis, J. Han, and U. M. Fayyad, Eds., AAAI Press, 1996, pp. 226–231.

[278] M. Ester, H. Kriegel, and X. Xu, "Knowledge discovery in large spatial databases: Focusing techniques for efficient class identification," in *Advances in Spatial Databases, 4th International Symposium, SSD'95, Portland, Maine, USA, August 6-9, 1995, Proceedings*, M. J. Egenhofer and J. R. Herring, Eds., ser. Lecture Notes in Computer Science, vol. 951, Springer, 1995, pp. 67–82.

[279] D. H. Fisher, "Knowledge acquisition via incremental conceptual clustering," *Mach. Learn.*, vol. 2, no. 2, pp. 139–172, 1987. DOI: `10.1007/BF00114265`. [Online]. Available: `https://doi.org/10.1007/BF00114265`.

[280] L. Fu and E. Medico, "Flame, a novel fuzzy clustering method for the analysis of DNA microarray data," *BMC Bioinform.*, vol. 8, 2007. DOI: `10.1186/1471-2105-8-3`. [Online]. Available: `https://doi.org/10.1186/1471-2105-8-3`.

[281] D. Gibson, J. M. Kleinberg, and P. Raghavan, "Clustering categorical data: An approach based on dynamical systems," in *VLDB'98, Proceedings of 24rd International Conference on Very Large Data Bases, August 24-27, 1998, New York City, New York, USA*, A. Gupta, O. Shmueli, and J. Widom, Eds., Morgan Kaufmann, 1998, pp. 311–322. [Online]. Available: `http://www.vldb.org/conf/1998/p311.pdf`.

[282] V. Ganti, J. Gehrke, and R. Ramakrishnan, "CACTUS - clustering categorical data using summaries," in *Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Diego, CA, USA, August 15-18, 1999*, U. M. Fayyad, S. Chaudhuri, and D. Madigan, Eds., ACM, 1999, pp. 73–83. DOI: `10.1145/312129.312201`. [Online]. Available: `https://doi.org/10.1145/312129.312201`.

[283] S. Guha, R. Rastogi, and K. Shim, "CURE: an efficient clustering algorithm for large databases," in *SIGMOD 1998, Proceedings ACM SIGMOD International Conference on Management of Data, June 2-4, 1998, Seattle, Washington, USA*, L. M. Haas and A. Tiwary, Eds., ACM Press, 1998, pp. 73–84. DOI: `10.1145/276304.276312`. [Online]. Available: `https://doi.org/10.1145/276304.276312`.

[284] S. Guha, R. Rastogi, and K. Shim, "ROCK: A robust clustering algorithm for categorical attributes," in *Proceedings of the 15th International Conference on Data Engineering, Sydney, Australia, March 23-26, 1999*, M. Kitsuregawa, M. P. Papazoglou, and C. Pu, Eds., IEEE Computer Society, 1999, pp. 512–521. DOI: `10.1109/ICDE.1999.754967`. [Online]. Available: `https://doi.org/10.1109/ICDE.1999.754967`.

[285] A. Hinneburg and D. A. Keim, "An efficient approach to clustering in large multimedia databases with noise," in *Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining (KDD-98), New York City, New York, USA, August 27-31,*

1998, R. Agrawal, P. E. Stolorz, and G. Piatetsky-Shapiro, Eds., AAAI Press, 1998, pp. 58–65.

[286] A. K. Jain and R. C. Dubes, *Algorithms for Clustering Data*. Prentice-Hall, 1988.

[287] G. Karypis, E. Han, and V. Kumar, "Chameleon: Hierarchical clustering using dynamic modeling," *IEEE Computer*, vol. 32, no. 8, pp. 68–75, 1999. DOI: `10.1109/2.781637`. [Online]. Available: `https://doi.org/10.1109/2.781637`.

[288] G. J. McLachlan and K. E. Basford, *Mixture models: Inference and applications to clustering*. M. Dekker New York, 1988, vol. 38.

[289] R. T. Ng and J. Han, "Efficient and effective clustering methods for spatial data mining," in *VLDB'94, Proceedings of 20th International Conference on Very Large Data Bases, September 12-15, 1994, Santiago de Chile, Chile*, J. B. Bocca, M. Jarke, and C. Zaniolo, Eds., Morgan Kaufmann, 1994, pp. 144–155. [Online]. Available: `http://www.vldb.org/conf/1994/P144.PDF`.

[290] L. Parsons, E. Haque, and H. Liu, "Subspace clustering for high dimensional data: A review," *SIGKDD Explor.*, vol. 6, no. 1, pp. 90–105, 2004. DOI: `10.1145/1007730.1007731`. [Online]. Available: `https://doi.org/10.1145/1007730.1007731`.

[291] E. Schikuta, "Grid-clustering: An efficient hierarchical clustering method for very large data sets," in *13th International Conference on Pattern Recognition, ICPR 1996, Vienna, Austria, 25-19 August, 1996*, IEEE Computer Society, 1996, pp. 101–105. DOI: `10.1109/ICPR.1996.546732`. [Online]. Available: `https://doi.org/10.1109/ICPR.1996.546732`.

[292] G. Sheikholeslami, S. Chatterjee, and A. Zhang, "Wavecluster: A multi-resolution clustering approach for very large spatial databases," in *VLDB'98, Proceedings of 24rd International Conference on Very Large Data Bases, August 24-27, 1998, New York City, New York, USA*, A. Gupta, O. Shmueli, and J. Widom, Eds., Morgan Kaufmann, 1998, pp. 428–439. [Online]. Available: `http://www.vldb.org/conf/1998/p428.pdf`.

[293] A. K. H. Tung, R. T. Ng, L. V. S. Lakshmanan, and J. Han, "Constraint-based clustering in large databases," in *Database Theory - ICDT 2001, 8th International Conference, London, UK, January 4-6, 2001, Proceedings*, J. V. den Bussche and V. Vianu, Eds., ser. Lecture Notes in Computer Science, vol. 1973, Springer, 2001, pp. 405–419. DOI: `10.1007/3-540-44503-X\_26`. [Online]. Available: `https://doi.org/10.1007/3-540-44503-X%5C_26`.

[294] A. K. H. Tung, J. Hou, and J. Han, "Spatial clustering in the presence of obstacles," in *Proceedings of the 17th International Conference on Data Engineering, April 2-6, 2001, Heidelberg, Germany*, D. Georgakopoulos and A. Buchmann, Eds., IEEE Computer Society, 2001, pp. 359–367. DOI: `10.1109/ICDE.2001.914848`. [Online]. Available: `https://doi.org/10.1109/ICDE.2001.914848`.

[295] H. Wang, W. Wang, J. Yang, and P. S. Yu, "Clustering by pattern similarity in large data sets," in *Proceedings of the 2002 ACM SIGMOD International Conference on Management of Data, Madison, Wisconsin, USA, June 3-6, 2002*, M. J. Franklin, B. Moon, and A. Ailamaki, Eds., ACM, 2002, pp. 394–405. DOI: `10.1145/564691.564737`. [Online]. Available: `https://doi.org/10.1145/564691.564737`.

[296] W. Wang, J. Yang, and R. R. Muntz, "STING: A statistical information grid approach to spatial data mining," in *VLDB'97, Proceedings of 23rd International Conference on Very Large Data Bases, August 25-29, 1997, Athens, Greece*, M. Jarke, M. J. Carey, K. R. Dittrich, F. H. Lochovsky, P. Loucopoulos, and M. A. Jeusfeld, Eds., Morgan Kaufmann, 1997, pp. 186–195. [Online]. Available: `http://www.vldb.org/conf/1997/P186.PDF`.

[297] T. Zhang, R. Ramakrishnan, and M. Livny, "BIRCH: an efficient data clustering method for very large databases," in *Proceedings of the 1996 ACM SIGMOD International Conference on Management of Data, Montreal, Quebec, Canada, June 4-6, 1996*, H. V. Jagadish and I. S. Mumick, Eds., ACM Press, 1996, pp. 103–114.

[298] X. Yin, J. Han, and P. S. Yu, "Linkclus: Efficient clustering via heterogeneous semantic links," in *Proceedings of the 32nd International Conference on Very Large Data Bases, Seoul, Korea, September 12-15, 2006*, U. Dayal et al., Eds., ACM, 2006, pp. 427–438. [Online]. Available: `http://dl.acm.org/citation.cfm?id=1164165`.

[306] J. Macqueen, "Some methods for classification and analysis of multivariate observations," in *In 5-th Berkeley Symposium on Mathematical Statistics and Probability*, 1967, pp. 281–297.

9 Outlier Analysis

[161] B. Abraham and G. E. P. Box, "Bayesian analysis of some outlier problems in time series," *Biometrika*, vol. 66, no. 2, pp. 229–236, 1979, ISSN: 1464-3510. DOI: `10.1093/biomet/66.2.229`. [Online]. Available: `http://dx.doi.org/10.1093/biomet/66.2.229`.

[162] M. Agyemang, K. Barker, and R. Alhajj, "A comprehensive survey of numeric and symbolic outlier mining techniques," *Intell. Data Anal.*, vol. 10, no. 6, pp. 521–538, 2006. [Online]. Available: `http://content.iospress.com/articles/intelligent-data-analysis/ida00266`.

[163] F. J. Anscombe, "Rejection of outliers," *Technometrics*, vol. 2, no. 2, pp. 123–146, May 1960, ISSN: 1537-2723. DOI: `10.1080/00401706.1960.10489888`. [Online]. Available: `http://dx.doi.org/10.1080/00401706.1960.10489888`.

[164] D. Agarwal, "Detecting anomalies in cross-classified streams: A bayesian approach," *Knowl. Inf. Syst.*, vol. 11, no. 1, pp. 29–44, 2007. DOI: `10.1007/s10115-006-0036-4`. [Online]. Available: `https://doi.org/10.1007/s10115-006-0036-4`.

[165] F. Angiulli and C. Pizzuti, "Outlier mining in large high-dimensional data sets," *IEEE Trans. Knowl. Data Eng.*, vol. 17, no. 2, pp. 203–215, 2005. DOI: `10.1109/TKDE.2005.31`. [Online]. Available: `https://doi.org/10.1109/TKDE.2005.31`.

[166] C. C. Aggarwal and P. S. Yu, "Outlier detection for high dimensional data," in *Proceedings of the 2001 ACM SIGMOD international conference on Management of data, Santa Barbara, CA, USA, May 21-24, 2001*, S. Mehrotra and T. K. Sellis, Eds., ACM, 2001, pp. 37–46, ISBN: 1-58113-332-4. DOI: `10.1145/375663.375668`. [Online]. Available: `https://doi.org/10.1145/375663.375668`.

[168] R. J. Beckman and R. D. Cook, "Outlier..........s," *Technometrics*, vol. 25, no. 2, p. 119, May 1983, ISSN: 0040-1706. DOI: `10.2307/1268541`. [Online]. Available: `http://dx.doi.org/10.2307/1268541`.

[169] I. Ben-Gal, "Outlier detection," in *Data Mining and Knowledge Discovery Handbook, 2nd ed*, O. Maimon and L. Rokach, Eds., Springer, 2010, pp. 117–130, ISBN: 978-0-387-09822-7. DOI: `10.1007/978-0-387-09823-4\_7`. [Online]. Available: `https://doi.org/10.1007/978-0-387-09823-4%5C_7`.

[171] M. M. Breunig, H. Kriegel, R. T. Ng, and J. Sander, "LOF: identifying density-based local outliers," in *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data, May 16-18, 2000, Dallas, Texas, USA*, W. Chen, J. F. Naughton, and P. A. Bernstein, Eds., ACM, 2000, pp. 93–104, ISBN: 1-58113-217-4. DOI: `10.1145/342009.335388`. [Online]. Available: `https://doi.org/10.1145/342009.335388`.

[173] D. Barbará, Y. Li, J. Lin, S. Jajodia, and J. Couto, "Bootstrapping a data mining intrusion detection system," in *Proceedings of the 2003 ACM Symposium on Applied Computing (SAC), March 9-12, 2003, Melbourne, FL, USA*, G. B. Lamont, H. Haddad, G. A. Papadopoulos, and B. Panda, Eds., ACM, 2003, pp. 421–425, ISBN: 1-58113-624-2. DOI: `10.1145/952532.952616`. [Online]. Available: `https://doi.org/10.1145/952532.952616`.

[175] Z. Bakar, R. Mohemad, A. Ahmad, and M. Deris, "A comparative study for outlier detection techniques in data mining," *2006 IEEE Conference on Cybernetics and Intelligent Systems*, Jun. 2006. DOI: `10.1109/iccis.2006.252287`. [Online]. Available: `http://dx.doi.org/10.1109/iccis.2006.252287`.

[176] S. D. Bay and M. Schwabacher, "Mining distance-based outliers in near linear time with randomization and a simple pruning rule," in *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, August 24 - 27, 2003*, L. Getoor, T. E. Senator, P. M. Domingos, and C. Faloutsos, Eds., ACM, 2003, pp. 29–38, ISBN: 1-58113-737-0. DOI: `10.1145/956750.956758`. [Online]. Available: `https://doi.org/10.1145/956750.956758`.

[177] D. Barbará, N. Wu, and S. Jajodia, "Detecting novel network intrusions using bayes estimators," *Proceedings of the 2001 SIAM International Conference on Data Mining*, Apr. 2001. DOI: `10.1137/1.9781611972719.28`. [Online]. Available: `http://dx.doi.org/10.1137/1.9781611972719.28`.

[178] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Comput. Surv.*, vol. 41, no. 3, 15:1–15:58, 2009. DOI: 10.1145/1541880.1541882. [Online]. Available: https://doi.org/10.1145/1541880.1541882.

[179] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection for discrete sequences: A survey," *IEEE Trans. Knowl. Data Eng.*, vol. 24, no. 5, pp. 823–839, 2012. DOI: 10.1109/TKDE.2010.235. [Online]. Available: https://doi.org/10.1109/TKDE.2010.235.

[180] D. Dasgupta and N. Majumdar, "Anomaly detection in multidimensional data using negative selection algorithm," *Proceedings of the 2002 Congress on Evolutionary Computation. CEC'02 (Cat. No.02TH8600)*, 2002. DOI: 10.1109/cec.2002.1004386. [Online]. Available: http://dx.doi.org/10.1109/cec.2002.1004386.

[181] E. Eskin, A. Arnold, M. Prerau, L. Portnoy, and S. Stolfo, "A geometric framework for unsupervised anomaly detection," in *Applications of Data Mining in Computer Security.* Boston, MA: Springer US, 2002, ch. 4, pp. 77–101, ISBN: 978-1-4615-0953-0. DOI: 10.1007/978-1-4615-0953-0_4. [Online]. Available: https://doi.org/10.1007/978-1-4615-0953-0%5C_4.

[182] E. Eskin, "Anomaly detection over noisy data using learned probability distributions," in *Proceedings of the Seventeenth International Conference on Machine Learning (ICML 2000), Stanford University, Stanford, CA, USA, June 29 - July 2, 2000*, P. Langley, Ed., Morgan Kaufmann, 2000, pp. 255–262, ISBN: 1-55860-707-2.

[184] T. Fawcett and F. J. Provost, "Adaptive fraud detection," *Data Min. Knowl. Discov.*, vol. 1, no. 3, pp. 291–316, 1997. DOI: 10.1023/A:1009700419189. [Online]. Available: https://doi.org/10.1023/A:1009700419189.

[185] V. J. Hodge and J. Austin, "A survey of outlier detection methodologies," *Artif. Intell. Rev.*, vol. 22, no. 2, pp. 85–126, 2004. DOI: 10.1023/B:AIRE.0000045502.10941.a9. [Online]. Available: https://doi.org/10.1023/B:AIRE.0000045502.10941.a9.

[186] D. M. Hawkins, "Identification of outliers," 1980. DOI: 10.1007/978-94-015-3994-4. [Online]. Available: http://dx.doi.org/10.1007/978-94-015-3994-4.

[187] Z. He, X. Xu, and S. Deng, "Discovering cluster-based local outliers," *Pattern Recognit. Lett.*, vol. 24, no. 9-10, pp. 1641–1650, 2003. DOI: 10.1016/S0167-8655(03)00003-5. [Online]. Available: https://doi.org/10.1016/S0167-8655(03)00003-5.

[188] W. Jin, A. K. H. Tung, and J. Han, "Mining top-n local outliers in large databases," in *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining, San Francisco, CA, USA, August 26-29, 2001*, D. Lee, M. Schkolnick, F. J. Provost, and R. Srikant, Eds., ACM, 2001, pp. 293–298, ISBN: 1-58113-391-X. DOI: 10.1145/502512.502554. [Online]. Available: https://doi.org/10.1145/502512.502554.

[190] W. Jin, A. K. H. Tung, J. Han, and W. Wang, "Ranking outliers using symmetric neighborhood relationship," in *Advances in Knowledge Discovery and Data Mining, 10th Pacific-Asia Conference, PAKDD 2006, Singapore, April 9-12, 2006, Proceedings*, W. K. Ng, M. Kitsuregawa, J. Li, and K. Chang, Eds., ser. Lecture Notes in Computer Science, vol. 3918, Springer, 2006, pp. 577–593, ISBN: 3-540-33206-5. DOI: 10.1007/11731139_68. [Online]. Available: https://doi.org/10.1007/11731139%5C_68.

[192] E. M. Knorr and R. T. Ng, "A unified notion of outliers: Properties and computation," in *Proceedings of the Third International Conference on Knowledge Discovery and Data Mining (KDD-97), Newport Beach, California, USA, August 14-17, 1997*, D. Heckerman, H. Mannila, and D. Pregibon, Eds., AAAI Press, 1997, pp. 219–222, ISBN: 1-57735-027-8. [Online]. Available: http://www.aaai.org/Library/KDD/1997/kdd97-044.php.

[194] E. M. Knorr and R. T. Ng, "Algorithms for mining distance-based outliers in large datasets," in *VLDB'98, Proceedings of 24rd International Conference on Very Large Data Bases, August 24-27, 1998, New York City, New York, USA*, A. Gupta, O. Shmueli, and J. Widom, Eds., Morgan Kaufmann, 1998, pp. 392–403, ISBN: 1-55860-566-5. [Online]. Available: http://www.vldb.org/conf/1998/p392.pdf.

[195] E. M. Knorr, R. T. Ng, and V. Tucakov, "Distance-based outliers: Algorithms and applications," *VLDB J.*, vol. 8, no. 3-4, pp. 237–253, 2000. DOI: 10.1007/s007780050006. [Online]. Available: https://doi.org/10.1007/s007780050006.

[196] H. Kriegel, M. Schubert, and A. Zimek, "Angle-based outlier detection in high-dimensional data," in *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Las Vegas, Nevada, USA, August 24-27, 2008*, Y. Li, B. Liu, and S. Sarawagi, Eds., ACM, 2008, pp. 444–452, ISBN: 978-1-60558-193-4. DOI: 10.1145/1401890.1401946. [Online]. Available: https://doi.org/10.1145/1401890.1401946.

[198] M. Markou and S. Singh, "Novelty detection: A review - part 1: Statistical approaches," *Signal Process.*, vol. 83, no. 12, pp. 2481–2497, 2003. DOI: 10.1016/j.sigpro.2003.07.018. [Online]. Available: https://doi.org/10.1016/j.sigpro.2003.07.018.

[199] M. Markou and S. Singh, "Novelty detection: A review - part 2: : Neural network based approaches," *Signal Process.*, vol. 83, no. 12, pp. 2499–2521, 2003. DOI: 10.1016/j.sigpro.2003.07.019. [Online]. Available: https://doi.org/10.1016/j.sigpro.2003.07.019.

[200] C. C. Noble and D. J. Cook, "Graph-based anomaly detection," in *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, August 24 - 27, 2003*, L. Getoor, T. E. Senator, P. M. Domingos, and C. Faloutsos, Eds., ACM, 2003, pp. 631–636, ISBN: 1-58113-737-0. DOI: 10.1145/956750.956831. [Online]. Available: https://doi.org/10.1145/956750.956831.

[201] S. Papadimitriou, H. Kitagawa, P. B. Gibbons, and C. Faloutsos, "LOCI: fast outlier detection using the local correlation integral," in *Proceedings of the 19th International Conference on Data Engineering, March 5-8, 2003, Bangalore, India*, U. Dayal, K. Ramamritham, and T. M. Vijayaraman, Eds., IEEE Computer Society, 2003, pp. 315–326, ISBN: 0-7803-7665-X. DOI: 10.1109/ICDE.2003.1260802. [Online]. Available: https://doi.org/10.1109/ICDE.2003.1260802.

[203] A. Patcha and J. Park, "An overview of anomaly detection techniques: Existing solutions and latest technological trends," *Comput. Networks*, vol. 51, no. 12, pp. 3448–3470, 2007. DOI: 10.1016/j.comnet.2007.02.001. [Online]. Available: https://doi.org/10.1016/j.comnet.2007.02.001.

[204] X. Song, M. Wu, C. M. Jermaine, and S. Ranka, "Conditional anomaly detection," *IEEE Trans. Knowl. Data Eng.*, vol. 19, no. 5, pp. 631–645, 2007. DOI: 10.1109/TKDE.2007.1009. [Online]. Available: https://doi.org/10.1109/TKDE.2007.1009.

[205] Y. Tao, X. Xiao, and S. Zhou, "Mining distance-based outliers from large databases in any metric space," in *Proceedings of the Twelfth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Philadelphia, PA, USA, August 20-23, 2006*, T. Eliassi-Rad, L. H. Ungar, M. Craven, and D. Gunopulos, Eds., ACM, 2006, pp. 394–403, ISBN: 1-59593-339-5. DOI: 10.1145/1150402.1150447. [Online]. Available: https://doi.org/10.1145/1150402.1150447.

[207] N. Ye and Q. Chen, "An anomaly detection technique based on a chi-square statistic for detecting intrusions into information systems," *Quality and Reliability Engineering International*, vol. 17, no. 2, pp. 105–112, 2001, ISSN: 1099-1638. DOI: 10.1002/qre.392. [Online]. Available: http://dx.doi.org/10.1002/qre.392.

[208] B. Yi, N. D. Sidiropoulos, T. Johnson, H. V. Jagadish, C. Faloutsos, and A. Biliris, "Online data mining for co-evolving time sequences," in *Proceedings of the 16th International Conference on Data Engineering, San Diego, California, USA, February 28 - March 3, 2000*, D. B. Lomet and G. Weikum, Eds., IEEE Computer Society, 2000, pp. 13–22, ISBN: 0-7695-0506-6. DOI: 10.1109/ICDE.2000.839383. [Online]. Available: https://doi.org/10.1109/ICDE.2000.839383.