

Prerequisites

- Klein, Philip N. (2017). *Coding The Matrix: Linear Algebra Through Computer Science Applications*. URL: <http://codingthematrix.com/>.
- Lehman, Eric, F. Thomson Leighton, and Albert R. Meyer (2015). *Mathematics for Computer Science*. MIT OpenCourseWare. URL: <https://courses.csail.mit.edu/6.042/spring15/mcs.pdf>.
- Prasad, Devi (2012). *Elementary linear algebra*. Second edition. Alpha Science International Ltd.

Books, textbooks and courses material

- Aggarwal, Charu C (2015). *Data mining: the textbook*. Springer.
- (2017). *Outlier analysis*. Springer.
- Ellenberg, Jordan (2015). *How Not to Be Wrong: The Power of Mathematical Thinking*. (Quiz Q0a.10). Penguin Press. URL: <https://medium.com/@penguinpress/an-excerpt-from-how-not-to-be-wrong-by-jordan-ellenberg-664e708cfc3d>.
- Friedman, Jerome, Trevor Hastie, and Robert Tibshirani (2001). *The elements of statistical learning*. Springer.
- Goldberg, Sharon (2014). *CAS CS 237: Probability in Computing*. (Exercise 3.3). Boston University.
- Leskovec, Jure, Anand Rajaraman, and Jeff Ullman (2014). *Mining of Massive Datasets, Chapter 4: Mining Data Streams*. (Exercises 3.2 and 3.4). URL: <http://www.mmids.org/>.
- Stamp, Mark (2018). *A Revealing Introduction to Hidden Markov Models*. URL: <https://www.cs.sjsu.edu/~stamp/RUA/HMM.pdf>.

Tutorials

- Batal, Iyad (n.d.). *Introduction to Temporal Data Mining*. URL: <https://people.cs.pitt.edu/~iyad/temp.pdf>.
- Berwick, R. (n.d.). *An Idiot's guide to Support vector machines (SVMs)*. URL: <http://web.mit.edu/6.034/wwwbob/svm-notes-long-08.pdf>.
- Elements of AI, online open course* (n.d.). URL: <https://course.elementsofai.com/>.
- Kriegel, Hans-Peter, Peer Kröger, and Arthur Zimek (2010). *Outlier Detection Techniques*. URL: <https://archive.siam.org/meetings/sdm10/tutorial3.pdf>.
- NIST/SEMATECH (1998). *Handbook of Statistical Methods, Chapter 6.2: Introduction to Time Series Analysis*. URL: <https://itl.nist.gov/div898/handbook/pmc/section4/pmc4.htm>.

Research articles

- Datar, M. et al. (2002). "Maintaining Stream Statistics over Sliding Windows." In: *SIAM Journal on Computing* 31.6, pp. 1794–1813. doi: [10.1137/S0097539701398363](https://doi.org/10.1137/S0097539701398363).
- Domingos, Pedro and Geoff Hulten (2000). "Mining High-speed Data Streams." In: *Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '00. ACM, pp. 71–80. doi: [10.1145/347090.347107](https://doi.org/10.1145/347090.347107).
- Gupta, M. et al. (2014). "Outlier Detection for Temporal Data: A Survey." In: *IEEE Transactions on Knowledge and Data Engineering* 26.9, pp. 2250–2267. doi: [10.1109/TKDE.2013.184](https://doi.org/10.1109/TKDE.2013.184).
- Hulten, Geoff, Laurie Spencer, and Pedro Domingos (2001). "Mining Time-changing Data Streams." In: *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '01. ACM, pp. 97–106. doi: [10.1145/502512.502529](https://doi.org/10.1145/502512.502529).
- Keogh, E., J. Lin, and A. Fu (2005). "HOT SAX: efficiently finding the most unusual time series subsequence." In: *Proceedings of the Fifth IEEE International Conference on Data Mining*. doi: [10.1109/ICDM.2005.79](https://doi.org/10.1109/ICDM.2005.79).
- Liu, Fei Tony, Kai Ming Ting, and Zhi-Hua Zhou (2008). "Isolation Forest." In: *Proceedings of the 2008 Eighth IEEE International Conference on Data Mining*. ICDM '08. IEEE Computer Society, pp. 413–422. doi: [10.1109/ICDM.2008.17](https://doi.org/10.1109/ICDM.2008.17).

- Liu, Fei Tony, Kai Ming Ting, and Zhi-Hua Zhou (2012). “Isolation-Based Anomaly Detection.” In: *ACM Trans. Knowl. Discov. Data* 6.1, 3:1–3:39. DOI: [10.1145/2133360.2133363](https://doi.org/10.1145/2133360.2133363).
- Lumbroso, Jérémie O. (2018). “How Flajolet Processed Streams with Coin Flips.” In: *CoRR* abs/1805.00612. URL: <http://arxiv.org/abs/1805.00612>.
- Patel, P. et al. (2002). “Mining motifs in massive time series databases.” In: *Proceedings of the 2002 IEEE International Conference on Data Mining*, pp. 370–377. DOI: [10.1109/ICDM.2002.1183925](https://doi.org/10.1109/ICDM.2002.1183925).
- Pei, Jian et al. (2001). “PrefixSpan: mining sequential patterns efficiently by prefix-projected pattern growth.” In: *Proceedings of the 17th International Conference on Data Engineering*, pp. 215–224. DOI: [10.1109/ICDE.2001.914830](https://doi.org/10.1109/ICDE.2001.914830).
- Vlachos, Michail et al. (2003). “Indexing Multi-dimensional Time-series with Support for Multiple Distance Measures.” In: *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '03. ACM, pp. 216–225. DOI: [10.1145/956750.956777](https://doi.org/10.1145/956750.956777).
- Vries, T. de, S. Chawla, and M. E. Houle (2010). “Finding Local Anomalies in Very High Dimensional Space.” In: *Proceedings of the 2010 IEEE International Conference on Data Mining*, pp. 128–137. DOI: [10.1109/ICDM.2010.151](https://doi.org/10.1109/ICDM.2010.151).

Online resources

- An Interactive Guide To The Fourier Transform* (n.d.). URL: <https://betterexplained.com/articles/an-interactive-guide-to-the-fourier-transform/>.
- Basemap* (n.d.). URL: <https://matplotlib.org/basemap/>.
- Bourret Sicotte, Xavier (n.d.). *Support Vector Machine: Python implementation using CVXOPT*. URL: https://xavierbourretsicotte.github.io/SVM_implementation.html.
- Brownlee, Jason (n.d.). *Machine Learning Mastery: Machine Learning Algorithms From Scratch*. URL: <https://machinelearningmastery.com/category/algorithms-from-scratch/>.
- Fletcher, Tristan (n.d.). *Support Vector Machines Explained*. URL: <https://static1.squarespace.com/static/58851af9ebbd1a30e98fb283/t/58902fbae4fcb5398aeb7505/1485844411772/SVM+Explained.pdf>.
- Fortmann-Roe, Scott (n.d.). *Understanding the Bias-Variance Tradeoff*. URL: <http://scott.fortmann-roe.com/docs/BiasVariance.html>.
- Jordan, Jeremy (n.d.). *Evaluating a machine learning model*. URL: <https://www.jeremyjordan.me/evaluating-a-machine-learning-model/>.
- Matplotlib* (n.d.). URL: <https://matplotlib.org/>.
- Scikit-learn* (n.d.). URL: <https://scikit-learn.org>.
- Tulloch, Andrew (n.d.). *Support Vector Machine python implementation*. URL: <https://github.com/ajtulloch/svmpy>.
- Voronoi diagram python implementation* (n.d.). URL: https://github.com/nlpenguin/py_geo_voronoi.

Datasets

- Department of Statistical Science, Duke University (n.d.). *Time-series datasets*. URL: http://www2.stat.duke.edu/~mw/ts_data_sets.html.
- Finnish Meteorological Institute (n.d.). *Weather observations*. URL: <https://en.ilmatieteenlaitos.fi/download-observations#!/>.
- GPSseuranta (n.d.). *GPS-tracking for outdoor events*. URL: <https://www.tulospalvelu.fi/gps/>.
- Kaggle (n.d.). URL: <https://www.kaggle.com/datasets>.
- Keogh, Eamonn (n.d.). *Time-series datasets*. URL: <http://www.cs.ucr.edu/~eamonn/discords/>.
- Stony Brook University (n.d.). *Outlier Detection DataSets (ODDS)*. URL: <http://odds.cs.stonybrook.edu/>.
- UCI ML Repository* (n.d.). URL: <https://archive.ics.uci.edu>.