



Universiteit
Leiden
The Netherlands

Massively collaborative machine learning

Rijn, J.N. van

Citation

Rijn, J. N. van. (2016, December 19). *Massively collaborative machine learning*. *IPA Dissertation Series*. Retrieved from <https://hdl.handle.net/1887/44814>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/44814>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/44814> holds various files of this Leiden University dissertation

Author: Rijn, Jan van

Title: Massively collaborative machine learning

Issue Date: 2016-12-19

Bibliography

- [1] S. M. Abdulrahman and P. Brazdil. Measures for Combining Accuracy and Time for Meta-learning. In *Meta-Learning and Algorithm Selection Workshop at ECAI 2014*, pages 49–50, 2014.
- [2] S. M. Abdulrahman, P. Brazdil, J. N. van Rijn, and J. Vanschoren. Algorithm selection via meta-learning and sample-based active testing. In J. Vanschoren, P. Brazdil, C. Giraud-Carrier, and L. Kotthoff, editors, *Proceedings of the 2015 International Workshop on Meta-Learning and Algorithm Selection (MetaSel)*, number 1455 in CEUR Workshop Proceedings, pages 55–66, Aachen, 2015.
- [3] S. M. Abdulrahman, P. Brazdil, J. N. van Rijn, and J. Vanschoren. Speeding up algorithm selection using average ranking and active testing by introducing runtime. *Machine Learning, Special Issue on Metalearning and Algorithm Selection*, forthcoming, 2016.
- [4] D. W. Aha. Generalizing from case studies: a case study. In *Proceedings of the ninth international workshop on Machine learning*, pages 1–10, San Francisco, CA, USA, 1992. Morgan Kaufmann Publishers Inc.
- [5] E. L. Allwein, R. E. Schapire, and Y. Singer. Reducing multiclass to binary: A unifying approach for margin classifiers. *Journal of machine learning research*, 1(Dec):113–141, 2000.
- [6] C. Apté and S. Weiss. Data mining with decision trees and decision rules. *Future generation computer systems*, 13(2):197–210, 1997.

- [7] M. Atzmueller. Subgroup discovery. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 5(1):35–49, 2015.
- [8] J. Bergstra and Y. Bengio. Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 13(Feb):281–305, 2012.
- [9] J. S. Bergstra, R. Bardenet, Y. Bengio, and B. Kégl. Algorithms for hyper-parameter optimization. In *Advances in Neural Information Processing Systems*, pages 2546–2554, 2011.
- [10] J. Beringer and E. Hüllermeier. Efficient instance-based learning on data streams. *Intelligent Data Analysis*, 11(6):627–650, 2007.
- [11] A. Bifet and R. Gavaldà. Learning from Time-Changing Data with Adaptive Windowing. In *SDM*, volume 7, pages 139–148. SIAM, 2007.
- [12] A. Bifet and R. Gavaldà. Adaptive learning from evolving data streams. In *Advances in Intelligent Data Analysis VIII*, pages 249–260. Springer, 2009.
- [13] A. Bifet, G. Holmes, R. Kirkby, and B. Pfahringer. MOA: Massive Online Analysis. *J. Mach. Learn. Res.*, 11:1601–1604, 2010.
- [14] A. Bifet, G. Holmes, and B. Pfahringer. Leveraging Bagging for Evolving Data Streams. In *Machine Learning and Knowledge Discovery in Databases*, volume 6321 of *Lecture Notes in Computer Science*, pages 135–150. Springer, 2010.
- [15] A. Bifet, E. Frank, G. Holmes, and B. Pfahringer. Ensembles of restricted hoeffding trees. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 3(2):30, 2012.
- [16] B. Bischl, M. Lang, L. Kotthoff, J. Schiffner, J. Richter, E. Studerus, G. Casalicchio, and Z. M. Jones. mlr: Machine Learning in R. *Journal of Machine Learning Research*, 17(170):1–5, 2016.
- [17] A. L. Blum and R. L. Rivest. Training a 3-node neural network is np-complete. *Neural Networks*, 5(1):117–127, 1992.
- [18] L. Bottou. Stochastic Learning. In *Advanced lectures on machine learning*, pages 146–168. Springer, 2004.
- [19] P. Brazdil and C. Soares. A Comparison of Ranking Methods for Classification Algorithm Selection. In *Machine Learning: ECML 2000*, pages 63–75. Springer, 2000.

- [20] P. Brazdil, J. Gama, and B. Henery. Characterizing the applicability of classification algorithms using meta-level learning. In *Machine Learning: ECML-94*, pages 83–102. Springer, 1994.
- [21] P. Brazdil, C. Soares, and J. P. Da Costa. Ranking learning algorithms: Using IBL and meta-learning on accuracy and time results. *Machine Learning*, 50(3): 251–277, 2003.
- [22] P. Brazdil, C. Giraud-Carrier, C. Soares, and R. Vilalta. *Metalearning: Applications to data mining*. Springer Science & Business Media, 2009.
- [23] L. Breiman. Bagging Predictors. *Machine learning*, 24(2):123–140, 1996.
- [24] L. Breiman. Random Forests. *Machine learning*, 45(1):5–32, 2001.
- [25] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen. *Classification and regression trees*. CRC press, 1984.
- [26] E. Brochu, V. M. Cora, and N. De Freitas. A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. *arXiv preprint arXiv:1012.2599*, 2010.
- [27] C. E. Brodley. Addressing the selective superiority problem: Automatic algorithm/model class selection. In *Proceedings of the Tenth International Conference on Machine Learning*, pages 17–24, 1993.
- [28] L. Buitinck, G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, R. Layton, J. VanderPlas, A. Joly, B. Holt, and G. Varoquaux. API design for machine learning software: experiences from the scikit-learn project. In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, pages 108–122, 2013.
- [29] J. Carpenter. May the best analyst win. *Science*, 331(6018):698–699, 2011.
- [30] R. Caruana, A. Niculescu-Mizil, G. Crew, and A. Ksikes. Ensemble selection from libraries of models. In *Proceedings of the twenty-first international conference on Machine learning*, page 18. ACM, 2004.
- [31] D. M. Chickering. Learning bayesian networks is np-complete. In *Learning from data*, pages 121–130. Springer, 1996.
- [32] W. W. Cohen. Fast effective rule induction. In *Proceedings of the twelfth international conference on machine learning*, pages 115–123, 1995.

- [33] P. Compton, G. Edwards, B. Kang, L. Lazarus, R. Malor, P. Preston, and A. Srinivasan. Ripple down rules: Turning knowledge acquisition into knowledge maintenance. *Artif. Intell. Med.*, 4(6):463–475, Dec. 1992.
- [34] G. F. Cooper. The computational complexity of probabilistic inference using bayesian belief networks. *Artificial intelligence*, 42(2-3):393–405, 1990.
- [35] T. Cover and P. Hart. Nearest neighbor pattern classification. *IEEE transactions on information theory*, 13(1):21–27, 1967.
- [36] J. Demšar. Statistical Comparisons of Classifiers over Multiple Data Sets. *The Journal of Machine Learning Research*, 7:1–30, 2006.
- [37] T. G. Dietterich. Ensemble methods in machine learning. In *International workshop on multiple classifier systems*, pages 1–15. Springer, 2000.
- [38] P. Domingos. A few useful things to know about machine learning. *Communications of the ACM*, 55(10):78–87, 2012.
- [39] P. Domingos and G. Hulten. Mining High-Speed Data Streams. In *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 71–80, 2000.
- [40] P. Domingos and G. Hulten. A general framework for mining massive data streams. *Journal of Computational and Graphical Statistics*, 12(4):945–949, 2003.
- [41] H. J. Escalante, M. Montes, and L. E. Sucar. Particle Swarm Model Selection. *The Journal of Machine Learning Research*, 10:405–440, 2009.
- [42] M. Fernández-Delgado, E. Cernadas, S. Barro, and D. Amorim. Do we need hundreds of classifiers to solve real world classification problems? *Journal of Machine Learning Research*, 15:3133–3181, 2014.
- [43] M. Feurer, J. T. Springenberg, and F. Hutter. Initializing bayesian hyperparameter optimization via meta-learning. In *AAAI*, pages 1128–1135, 2015.
- [44] R. A. Fisher. The use of multiple measurements in taxonomic problems. *Annals of eugenics*, 7(2):179–188, 1936.
- [45] Y. Freund and R. E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. In *European conference on computational learning theory*, pages 23–37. Springer, 1995.

- [46] Y. Freund and R. E. Schapire. Experiments with a new boosting algorithm. In *Proceedings of the Thirteenth International Conference on Machine Learning (ICML 1996)*, pages 148–156. Morgan Kaufmann, 1996.
- [47] P. W. Frey and D. J. Slate. Letter recognition using holland-style adaptive classifiers. *Machine Learning*, 6:161–182, 1991.
- [48] J. Friedman, T. Hastie, and R. Tibshirani. Additive logistic regression: a statistical view of boosting. *Annals of Statistics*, 28:2000, 1998.
- [49] J. H. Friedman. Multivariate adaptive regression splines. *The annals of statistics*, pages 1–67, 1991.
- [50] J. H. Friedman. On Bias, Variance, 0/1-Loss, and the Curse-of-Dimensionality. *Data Mining and Knowledge Discovery*, 1(1):55–77, 1997.
- [51] J. H. Friedman. Greedy function approximation: a gradient boosting machine. *Annals of statistics*, pages 1189–1232, 2001.
- [52] J. Fürnkranz and J. Petrak. An Evaluation of Landmarking Variants. In *Working Notes of the ECML/PKDD 2000 Workshop on Integrating Aspects of Data Mining, Decision Support and Meta-Learning*, pages 57–68, 2001.
- [53] J. Gama and P. Brazdil. Cascade Generalization. *Machine Learning*, 41(3): 315–343, 2000.
- [54] J. Gama and P. Kosina. Recurrent concepts in data streams classification. *Knowledge and Information Systems*, 40(3):489–507, 2014.
- [55] J. Gama, P. Medas, G. Castillo, and P. Rodrigues. Learning with Drift Detection. In *SBIA Brazilian Symposium on Artificial Intelligence*, volume 3171 of *Lecture Notes in Computer Science*, pages 286–295. Springer, 2004.
- [56] J. Gama, P. Medas, and R. Rocha. Forest Trees for On-line Data. In *Proceedings of the 2004 ACM symposium on Applied computing*, pages 632–636. ACM, 2004.
- [57] J. Gama, R. Sebastião, and P. P. Rodrigues. Issues in Evaluation of Stream Learning Algorithms. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 329–338. ACM, 2009.
- [58] J. Gama, R. Sebastião, and P. P. Rodrigues. On evaluating stream learning algorithms. *Machine Learning*, 90(3):317–346, 2013.
- [59] C. Giraud-Carrier. Beyond predictive accuracy: What? Technical report, University of Bristol, Bristol, UK, UK, 1998.

- [60] C. Giraud-Carrier and F. Provost. Toward a justification of meta-learning: Is the no free lunch theorem a show-stopper. In *Proceedings of the ICML-2005 Workshop on Meta-learning*, pages 12–19, 2005.
- [61] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten. The WEKA Data Mining Software: An Update. *ACM SIGKDD explorations newsletter*, 11(1):10–18, 2009.
- [62] M. A. Hall. *Correlation-based Feature Subset Selection for Machine Learning*. PhD thesis, University of Waikato, Hamilton, New Zealand, 1998.
- [63] D. Hand. Classifier technology and the illusion of progress. *Statistical Science*, 21(1):1–14, 2006.
- [64] L. Hansen and P. Salamon. Neural Network Ensembles. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 12(10):993–1001, 1990.
- [65] T. Hastie, R. Tibshirani, and J. Friedman. *The elements of statistical learning, 2nd edition*. Springer New York, 2009.
- [66] J. L. Hintze and R. D. Nelson. Violin plots: a box plot-density trace synergism. *The American Statistician*, 52(2):181–184, 1998.
- [67] H. Hirsh. Data mining research: Current status and future opportunities. *Statistical Analysis and Data Mining*, 1(2):104–107, 2008.
- [68] M. W. Hoffman, B. Shahriari, and N. de Freitas. Exploiting correlation and budget constraints in bayesian multi-armed bandit optimization. *arXiv preprint arXiv:1303.6746*, 2013.
- [69] R. C. Holte. Very simple classification rules perform well on most commonly used datasets. *Machine learning*, 11(1):63–90, 1993.
- [70] H. J. Hoogeboom, W. A. Kusters, J. N. van Rijn, and J. K. Vis. Acyclic Constraint Logic and Games. *ICGA Journal*, 37(1):3–16, 2014.
- [71] V. Hoste and W. Daelemans. Comparing learning approaches to coreference resolution. There is more to it than bias. In *Proceedings of the ICML'05 Workshop on Meta-learning*, pages 20–27, 2005.
- [72] J. Hühn and E. Hüllermeier. Furia: an algorithm for unordered fuzzy rule induction. *Data Mining and Knowledge Discovery*, 19(3):293–319, 2009.
- [73] G. Hulten, L. Spencer, and P. Domingos. Mining Time-changing Data Streams. In *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 97–106, 2001.

- [74] F. Hutter, H. H. Hoos, K. Leyton-Brown, and K. Murphy. Time-Bounded Sequential Parameter Optimization. In *Learning and intelligent optimization*, pages 281–298. Springer, 2010.
- [75] F. Hutter, H. H. Hoos, and K. Leyton-Brown. Sequential model-based optimization for general algorithm configuration. In *International Conference on Learning and Intelligent Optimization*, pages 507–523. Springer, 2011.
- [76] L. Hyafil and R. L. Rivest. Constructing optimal binary decision trees is np-complete. *Information Processing Letters*, 5(1):15–17, 1976.
- [77] A. Jain and D. Zongker. Feature Selection: Evaluation, Application, and Small Sample Performance. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 19(2):153–158, 1997.
- [78] G. H. John and P. Langley. Estimating continuous distributions in bayesian classifiers. In *Proceedings of the Eleventh conference on Uncertainty in artificial intelligence*, pages 338–345. Morgan Kaufmann Publishers Inc., 1995.
- [79] T. Kealey. *Sex, science and profits*. Random House, 2010.
- [80] E. Keogh and S. Kasetty. On the need for time series data mining benchmarks: A survey and empirical demonstration. *Data Mining and Knowledge Discovery*, 7(4):349–371, 2003.
- [81] R. D. King, S. H. Muggleton, A. Srinivasan, and M. Sternberg. Structure-activity relationships derived by machine learning: The use of atoms and their bond connectivities to predict mutagenicity by inductive logic programming. *Proceedings of the National Academy of Sciences*, 93(1):438–442, 1996.
- [82] R. Kohavi, D. H. Wolpert, et al. Bias plus variance decomposition for zero-one loss functions. In *ICML*, volume 96, pages 275–83, 1996.
- [83] K. K. Ladha. Condorcet’s jury theorem in light of de finetti’s theorem. *Social Choice and Welfare*, 10(1):69–85, 1993.
- [84] N. Landwehr, M. Hall, and E. Frank. Logistic model trees. *Machine Learning*, 59(1-2):161–205, 2005.
- [85] J. W. Lee and C. Giraud-Carrier. A metric for unsupervised metalearning. *Intelligent Data Analysis*, 15(6):827–841, 2011.
- [86] R. Leite and P. Brazdil. Predicting Relative Performance of Classifiers from Samples. In *Proceedings of the 22nd international conference on Machine learning*, pages 497–503. ACM, 2005.

- [87] R. Leite and P. Brazdil. Active Testing Strategy to Predict the Best Classification Algorithm via Sampling and Metalearning. In *ECAI*, pages 309–314, 2010.
- [88] R. Leite, P. Brazdil, and J. Vanschoren. Selecting Classification Algorithms with Active Testing. In *Machine Learning and Data Mining in Pattern Recognition*, pages 117–131. Springer, 2012.
- [89] L. Li, K. Jamieson, G. DeSalvo, A. Rostamizadeh, and A. Talwalkar. Efficient hyperparameter optimization and infinitely many armed bandits. *arXiv preprint arXiv:1603.06560*, 2016.
- [90] M. Lichman. UCI machine learning repository, 2013. URL <http://archive.ics.uci.edu/ml>.
- [91] I. Manolescu and S. Manegold. Performance evaluation in database research: principles and experience. *Proceedings of the International Conference on Extending Database Technology (EDBT)*, page 1156, 2008.
- [92] M. Meeng and A. Knobbe. Flexible enrichment with cortana–software demo. In *Proceedings of BeneLearn*, pages 117–119, 2011.
- [93] D. Michie, D. Spiegelhalter, and C. Taylor. *Machine learning, neural and statistical classification*. Ellis Horwood, Upper Saddle River, NJ, USA, 1994. ISBN 0-13-106360-X.
- [94] H.-L. Nguyen, Y.-K. Woon, W.-K. Ng, and L. Wan. Heterogeneous Ensemble for Feature Drifts in Data Streams. In *Advances in Knowledge Discovery and Data Mining*, pages 1–12. Springer, 2012.
- [95] P. Nguyen, M. Hilario, and A. Kalousis. Using Meta-mining to Support Data Mining Workflow Planning and Optimization. *Journal of Artificial Intelligence Research*, 51:605–644, 2014.
- [96] M. Nielsen. The future of science: Building a better collective memory. *APS Physics*, 17(10), 2008.
- [97] M. Nielsen. *Reinventing discovery: the new era of networked science*. Princeton University Press, 2012.
- [98] W. S. Noble. What is a support vector machine? *Nature biotechnology*, 24(12): 1565–1567, 2006.
- [99] I. Olier, C. Grosan, N. Sadawi, L. Soldatova, and R. D. King. Meta-qsar: Learning how to learn qsars. In J. Vanschoren, P. Brazdil, C. Giraud-Carrier, and

- L. Kotthoff, editors, *Proceedings of the 2015 International Workshop on Meta-Learning and Algorithm Selection (MetaSel)*, number 1455 in CEUR Workshop Proceedings, pages 104–105, Aachen, 2015.
- [100] E. Ostrom. Collective action and the evolution of social norms. *The Journal of Economic Perspectives*, pages 137–158, 2000.
- [101] N. C. Oza. Online Bagging and Boosting. In *Systems, man and cybernetics, 2005 IEEE international conference on*, volume 3, pages 2340–2345. IEEE, 2005.
- [102] T. Pedersen. Empiricism is not a matter of faith. *Computational Linguistics*, 34: 465–470, 2008.
- [103] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
- [104] Y. Peng, P. Flach, C. Soares, and P. Brazdil. Improved dataset characterisation for meta-learning. *Lecture Notes in Computer Science*, 2534:141–152, Jan 2002.
- [105] C. Perlich, F. Provost, and J. Simonoff. Tree induction vs. logistic regression: A learning-curve analysis. *Journal of Machine Learning Research*, 4:211–255, 2003.
- [106] A. H. Peterson and T. Martinez. Estimating The Potential for Combining Learning Models. In *In Proc. of the ICML Workshop on Meta-Learning*, pages 68–75, 2005.
- [107] J. Petrak. Fast Subsampling Performance Estimates for Classification Algorithm Selection. In *Proceedings of the ECML-00 Workshop on Meta-Learning: Building Automatic Advice Strategies for Model Selection and Method Combination*, pages 3–14, 2000.
- [108] B. Pfahringer, H. Bensusan, and C. Giraud-Carrier. Tell me who can learn you and I can tell you who you are: Landmarking Various Learning Algorithms. In *Proceedings of the 17th international conference on machine learning*, pages 743–750, 2000.
- [109] B. Pfahringer, G. Holmes, and R. Kirkby. New Options for Hoeffding Trees. In *AI 2007: Advances in Artificial Intelligence*, pages 90–99. Springer, 2007.

- [110] F. Pinto, C. Soares, and J. Mendes-Moreira. Towards automatic generation of metafeatures. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pages 215–226. Springer, 2016.
- [111] M. J. Post, P. van der Putten, and J. N. van Rijn. Does Feature Selection Improve Classification? A Large Scale Experiment in OpenML. In *Advances in Intelligent Data Analysis XV*, pages 158–170. Springer, 2016.
- [112] J. Priem, P. Groth, and D. Taraborelli. The Altmetrics Collection. *PLoS ONE*, 11(7):e48753, 2012.
- [113] F. Provost, D. Jensen, and T. Oates. Efficient Progressive Sampling. In *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 23–32. ACM, 1999.
- [114] P. van der Putten and M. van Someren. A bias-variance analysis of a real world learning problem: The coil challenge 2000. *Machine Learning*, 57(1):177–195, 2004.
- [115] J. R. Quinlan. Induction of decision trees. *Machine learning*, 1(1):81–106, 1986.
- [116] J. R. Quinlan. *C4.5: programs for machine learning*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1993.
- [117] M. Radovanović, A. Nanopoulos, and M. Ivanović. Hubs in space: Popular nearest neighbors in high-dimensional data. *JMLR*, 11:2487–2531, 2010.
- [118] J. Read, A. Bifet, B. Pfahringer, and G. Holmes. Batch-Incremental versus Instance-Incremental Learning in Dynamic and Evolving Data. In *Advances in Intelligent Data Analysis XI*, pages 313–323. Springer, 2012.
- [119] J. R. Rice. The Algorithm Selection Problem. *Advances in Computers*, 15:65118, 1976.
- [120] J. N. van Rijn. Playing Games: The complexity of Klondike, Mahjong, Nonograms and Animal Chess. Master’s thesis, Leiden University, 2012.
- [121] J. N. van Rijn and J. Vanschoren. Sharing RapidMiner Workflows and Experiments with OpenML. In J. Vanschoren, P. Brazdil, C. Giraud-Carrier, and L. Kotthoff, editors, *Proceedings of the 2015 International Workshop on Meta-Learning and Algorithm Selection (MetaSel)*, number 1455 in CEUR Workshop Proceedings, pages 93–103, Aachen, 2015.

- [122] J. N. van Rijn and J. K. Vis. Complexity and retrograde analysis of the game dou shou qi. In *BNAIC 2013: Proceedings of the 25th Benelux Conference on Artificial Intelligence*, 8 pages, 2013.
- [123] J. N. van Rijn and J. K. Vis. Endgame Analysis of Dou Shou Qi. *ICGA Journal*, 37(2):120–124, 2014.
- [124] J. N. van Rijn, B. Bischl, L. Torgo, B. Gao, V. Umaashankar, S. Fischer, P. Winter, B. Wiswedel, M. R. Berthold, and J. Vanschoren. OpenML: A Collaborative Science Platform. In *Machine Learning and Knowledge Discovery in Databases*, pages 645–649. Springer, 2013.
- [125] J. N. van Rijn, G. Holmes, B. Pfahringer, and J. Vanschoren. Algorithm Selection on Data Streams. In *Discovery Science*, volume 8777 of *Lecture Notes in Computer Science*, pages 325–336. Springer, 2014.
- [126] J. N. van Rijn, G. Holmes, B. Pfahringer, and J. Vanschoren. Towards Meta-learning over Data Streams. In J. Vanschoren, P. Brazdil, C. Soares, and L. Kotthoff, editors, *Proceedings of the 2014 International Workshop on Meta-learning and Algorithm Selection (MetaSel)*, number 1201 in *CEUR Workshop Proceedings*, pages 37–38, Aachen, 2014.
- [127] J. N. van Rijn, S. M. Abdulrahman, P. Brazdil, and J. Vanschoren. Fast Algorithm Selection using Learning Curves. In *Advances in Intelligent Data Analysis XIV*, pages 298–309. Springer, 2015.
- [128] J. N. van Rijn, G. Holmes, B. Pfahringer, and J. Vanschoren. Having a Blast: Meta-Learning and Heterogeneous Ensembles for Data Streams. In *Data Mining (ICDM), 2015 IEEE International Conference on*, pages 1003–1008. IEEE, 2015.
- [129] J. N. van Rijn, G. Holmes, B. Pfahringer, and J. Vanschoren. Case Study on Bagging Stable Classifiers for Data Streams. In *Proceedings of the 24th Belgian-Dutch Conference on Machine Learning (BeNeLearn 2015)*, 6 pages, 2015.
- [130] J. N. van Rijn, F. W. Takes, and J. K. Vis. The complexity of rummikub problems. In *Proceedings of the 27th Benelux Conference on Artificial Intelligence (BNAIC 2015)*, 8 pages, 2015.
- [131] J. N. van Rijn, S. M. Abdulrahman, P. Brazdil, and J. Vanschoren. On the Evaluation of Algorithm Selection Problems. In *Proceedings of the 25th Belgian-Dutch Conference on Machine Learning (BeNeLearn 2016)*, 2 pages, 2016.

- [132] L. Rokach and O. Maimon. Clustering methods. In *Data mining and knowledge discovery handbook*, pages 321–352. Springer, 2005.
- [133] A. L. D. Rossi, A. C. P. de Leon Ferreira, C. Soares, and B. F. De Souza. MetaStream: A meta-learning based method for periodic algorithm selection in time-changing data. *Neurocomputing*, 127:52–64, 2014.
- [134] C. Schaffer. A conservation law for generalization performance. In *Proceedings of the 11th international conference on machine learning*, pages 259–265, 1994.
- [135] R. E. Schapire. The Strength of Weak Learnability. *Machine learning*, 5(2): 197–227, 1990.
- [136] J. C. Schlimmer. *Concept Acquisition Through Representational Adjustment*. PhD thesis, University of California, Irvine, 1987.
- [137] J. C. Schlimmer and R. H. Granger Jr. Incremental learning from noisy data. *Machine learning*, 1(3):317–354, 1986.
- [138] A. Shaker and E. Hüllermeier. Recovery analysis for adaptive learning from non-stationary data streams: Experimental design and case study. *Neurocomputing*, 150:250–264, 2015.
- [139] S. Shalev-Shwartz, Y. Singer, N. Srebro, and A. Cotter. Pegasos: primal estimated sub-gradient solver for SVM. *Mathematical Programming*, 127(1):3–30, 2011.
- [140] M. R. Smith, T. Martinez, and C. Giraud-Carrier. An instance level analysis of data complexity. *Machine learning*, 95(2):225–256, 2014.
- [141] K. A. Smith-Miles. Cross-disciplinary Perspectives on Meta-Learning for Algorithm Selection. *ACM Computing Surveys (CSUR)*, 41(1):6, 2008.
- [142] S. Sonnenburg, M. Braun, C. Ong, S. Bengio, L. Bottou, G. Holmes, Y. LeCun, K. Muller, F. Pereira, C. Rasmussen, G. Ratsch, B. Scholkopf, A. Smola, P. Vincent, J. Weston, and R. Williamson. The need for open source software in machine learning. *Journal of Machine Learning Research*, 8:2443–2466, 2007.
- [143] W. N. Street and Y. Kim. A streaming ensemble algorithm (sea) for large-scale classification. In *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 377–382, 2001.
- [144] Q. Sun and B. Pfahringer. Pairwise meta-rules for better meta-learning-based algorithm ranking. *Machine learning*, 93(1):141–161, 2013.

- [145] Q. Sun, B. Pfahringer, and M. Mayo. Towards a Framework for Designing Full Model Selection and Optimization Systems. In *Multiple Classifier Systems*, pages 259–270. Springer, 2013.
- [146] A. S. Szalay, J. Gray, A. R. Thakar, P. Z. Kunszt, T. Malik, J. Raddick, C. Stoughton, and J. vandenBerg. The sdss skyserver: public access to the sloan digital sky server data. In *Proceedings of the 2002 ACM SIGMOD international conference on Management of data*, pages 570–581. ACM, 2002.
- [147] C. Thornton, F. Hutter, H. Hoos, and K. Leyton-Brown. Auto-WEKA: Combined selection and Hyperparameter Optimization of Classification Algorithms. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 847–855. ACM, 2013.
- [148] L. Torgo. *Data Mining with R: Learning with Case Studies*. Chapman & Hall/CRC, 1st edition, 2010. ISBN 1439810184, 9781439810187.
- [149] L. G. Valiant. A theory of the learnable. *Communications of the ACM*, 27(11): 1134–1142, 1984.
- [150] J. N. Van Rijn and J. Vanschoren. OpenML: An Open Science Platform for Machine Learning. In *Proceedings of the 22th Belgian-Dutch Conference on Machine Learning (BeNeLearn 2013)*, 1 page, 2013.
- [151] J. N. Van Rijn, V. Umaashankar, S. Fischer, B. Bischl, L. Torgo, B. Gao, P. Winter, B. Wiswedel, M. R. Berthold, and J. Vanschoren. A RapidMiner extension for Open Machine Learning. In *RapidMiner Community Meeting and Conference*, pages 59–70, 2013.
- [152] J. Vanschoren and H. Blockeel. Towards understanding learning behavior. In *Proceedings of the annual machine learning conference of Belgium and the Netherlands*, pages 89–96, 2006.
- [153] J. Vanschoren, H. Blockeel, B. Pfahringer, and G. Holmes. Experiment databases. A new way to share, organize and learn from experiments. *Machine Learning*, 87(2):127–158, 2012.
- [154] J. Vanschoren, J. N. van Rijn, B. Bischl, and L. Torgo. OpenML: networked science in machine learning. *ACM SIGKDD Explorations Newsletter*, 15(2):49–60, 2014.
- [155] J. Vanschoren, J. N. van Rijn, and B. Bischl. Taking machine learning research online with OpenML. In *Proceedings of the 4th International Workshop on Big*

- Data, Streams and Heterogeneous Source Mining: Algorithms, Systems, Programming Models and Applications*, pages 1–4. JLMR.org, 2015.
- [156] J. Vanschoren, J. N. van Rijn, B. Bischl, G. Casalicchio, M. Lang, and M. Feurer. OpenML: a Networked Science Platform for Machine Learning. In *ICML 2015 MLOSS Workshop*, 3 pages, 2015.
- [157] V. Vidulin, M. Bohanec, and M. Gams. Combining human analysis and machine data mining to obtain credible data relations. *Information Sciences*, 288:254–278, 2014.
- [158] R. Vilalta and Y. Drissi. A perspective view and survey of meta-learning. *Artificial Intelligence Review*, 18(2):77–95, 2002.
- [159] R. Vilalta, C. Giraud-Carrier, and P. Brazdil. *Meta-Learning*, pages 731–748. Springer US, Boston, MA, 2005.
- [160] M. Wainberg, B. Alipanahi, and B. J. Frey. Are random forests truly the best classifiers? *Journal of Machine Learning Research*, 17(110):1–5, 2016.
- [161] H. Wang, W. Fan, P. S. Yu, and J. Han. Mining Concept-Drifting Data Streams using Ensemble Classifiers. In *KDD*, pages 226–235, 2003.
- [162] M. Wojnarski, S. Stawicki, and P. Wojnarowski. Tunedit.org: System for automated evaluation of algorithms in repeatable experiments. *Lecture Notes in Computer Science*, 6086:20–29, 2010.
- [163] D. H. Wolpert. Stacked generalization. *Neural networks*, 5(2):241–259, 1992.
- [164] M. N. Zarmehri and C. Soares. Using Metalearning for Prediction of Taxi Trip Duration Using Different Granularity Levels. In *Advances in Intelligent Data Analysis XIV*, pages 205–216. Springer, 2015.
- [165] P. Zhang, B. J. Gao, X. Zhu, and L. Guo. Enabling Fast Lazy Learning for Data Streams. In *Data Mining (ICDM), 2011 IEEE 11th International Conference on*, pages 932–941. IEEE, 2011.