The Technological Emergence of AutoML: A Survey of Performant Software and Applications in the Context of Industry

Other titles in Foundations and Trends® in Information Systems

Knowledge Paths in Design Science Research
Andreas Drechsler and Alan Hevner
ISBN: 978-1-63828-088-0

The Computer Self-Efficacy Construct: A History of Application in Information Systems Research
George M. Marakas, Miguel Aguirre-Urreta, Amin Shoja, Eunyoung Kim and Shangjun Wang
ISBN: 978-1-63828-080-4

IT Use and Identity: A Review and a Research Agenda
Michelle Carter, Victoria Reibenspiess, Yafang Li and Varun Grover
ISBN: 978-1-63828-072-9

Innovative, Creative and Sustainable Decision-Making in Changing Contexts
Patrick Humphreys and Amy Luk
ISBN: 978-1-63828-010-1

Information Technology Alignment and Innovation: 30 Years of Intersecting Research
Yolande E. Chan, Rashmi Krishnamurthy and Ali S. Ghawe
ISBN: 978-1-68083-816-9

Process Theory: Background, Opportunity, and Challenges
Fred Niederman
ISBN: 978-1-68083-806-0

The Technological Emergence of AutoML: A Survey of Performant Software and Applications in the Context of Industry

Alexander Scriven
University of Technology Sydney
alexander.scriven@uts.edu.au

David Jacob Kedziora
University of Technology Sydney
david.kedziora@uts.edu.au

Katarzyna Musial
University of Technology Sydney
katarzyna.musial-gabrys@uts.edu.au

Bogdan Gabrys
University of Technology Sydney
bogdan.gabrys@uts.edu.au

the essence of knowledge
Boston — Delft

Foundations and Trends® in Information Systems

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is

A. Scriven *et al.*. *The Technological Emergence of AutoML: A Survey of Performant Software and Applications in the Context of Industry*. Foundations and Trends® in Information Systems, vol. 7, no. 1-2, pp. 1–252, 2023.

ISBN: 978-1-63828-323-2

Foundations and Trends® in Information Systems

Volume 7, Issue 1-2, 2023

Editorial Board

Editorial Scope

Topics

Foundations and Trends® in Information Systems publishes survey and tutorial articles in the following topics:

- IS and Individuals
- IS and Groups
- IS and Organizations
- IS and Industries
- IS and Society
- IS Development
- IS Economics
- IS Management
- IS Research Methods

Information for Librarians

Foundations and Trends® in Information Systems, 2023, Volume 7, 4 issues. ISSN paper version 2331-1231. ISSN online version 2331-124X. Also available as a combined paper and online subscription.

Contents

The Technological Emergence of AutoML: A Survey of Performant Software and Applications in the Context of Industry

Alexander Scriven[1], David Jacob Kedziora[1], Katarzyna Musial[1] and Bogdan Gabrys[1]

[1]*Complex Adaptive Systems Lab, Data Science Institute, University of Technology Sydney, Australia; alexander.scriven@uts.edu.au, david.kedziora@uts.edu.au, katarzyna.musial-gabrys@uts.edu.au, bogdan.gabrys@uts.edu.au*

ABSTRACT

With most technical fields, there exists a delay between fundamental academic research and practical industrial uptake. Whilst some sciences have robust and well-established processes for commercialisation, such as the pharmaceutical practice of regimented drug trials, other fields face transitory periods in which fundamental academic advancements diffuse gradually into the space of commerce and industry. For the still relatively young field of Automated/Autonomous Machine Learning (AutoML/AutonoML), that transitory period is under way, spurred on by a burgeoning interest from broader society. Yet, to date, little research has been undertaken to assess the current state of this dissemination and its uptake. Thus, this review makes two primary contributions to knowledge around this topic. Firstly, it provides

---

Alexander Scriven, David Jacob Kedziora, Katarzyna Musial and Bogdan Gabrys (2023), "The Technological Emergence of AutoML: A Survey of Performant Software and Applications in the Context of Industry", Foundations and Trends® in Information Systems: Vol. 7, No. 1-2, pp 1–252. DOI: 10.1561/2900000030.

the most up-to-date and comprehensive survey of existing AutoML tools, both open-source and commercial. Secondly, it motivates and outlines a framework for assessing whether an AutoML solution designed for real-world application is 'performant'; this framework extends beyond the limitations of typical academic criteria, considering a variety of stakeholder needs and the human-computer interactions required to service them. Thus, additionally supported by an extensive assessment and comparison of academic and commercial case-studies, this review evaluates mainstream engagement with AutoML in the early 2020s, identifying obstacles and opportunities for accelerating future uptake.

Keywords: Automated machine learning (AutoML)

1

Introduction

Societal interest in machine learning (ML), especially the subtopic of deep learning (DL), has surged within recent years. This is partially driven by the continuing success of these approaches in many application areas (Grigorescu *et al.*, 2020; Ozbayoglu *et al.*, 2020; Minaee *et al.*, 2021; Piccialli *et al.*, 2021), facilitated by both fundamental advances (Kotsiantis *et al.*, 2006; Chen and Guestrin, 2016; Singh *et al.*, 2016; Choudhary and Gianey, 2017) and the increasing availability of computational resources. Unsurprisingly, on the academic side, the field of artificial intelligence (AI) continues to dominate research outputs, as noted by the 2021 UNESCO Science Report (Schneegans *et al.*, 2021). However, it is the current level of ML engagement in industry that is truly unprecedented. For instance, the 2021 Global AI Adoption Index, commissioned by IBM, found that 80% of 5501 global businesses are either using automation software or planning to within 12 months, and 74% are exploring or deploying AI (IBM, 2021). The Gartner 2019 CIO Agenda survey, with 3000 respondents from across the globe, agrees with this trend, revealing that the proportion of firms deploying AI has increased from 10% in 2015 to 37% in 2019 (Howard and Rowsell-Jones, 2019). Similar conclusions are echoed in the 2020 McKinsey 'State of

AI' report (McKinsey, 2020). Naturally, such a rate of mainstream permeation is also accompanied by intensifying discussions on how to use ML, and AI more broadly, in a socially responsible manner (Makridakis, 2017; Lepri *et al.*, 2018; Jobin *et al.*, 2019; Morley *et al.*, 2020; Toreini *et al.*, 2020).

Nonetheless, despite the growing desire of industry to utilise ML, talent in data science remains scarce (Pompa and Burke, 2017; Basu, Sreeradha, 2018). Both the Gartner and IBM studies agree that lack of expertise creates a barrier to AI adoption (Howard and Rowsell-Jones, 2019; IBM, 2021), especially as, by and large, ML technology still requires specialist skills to implement and employ. Worse yet, in practice, deploying ML solutions for real-world applications requires technical skills beyond the domain of data science. Any shortfall in these broader talents will also adversely affect ML engagement in industry (Australian Computer Society, 2021; Gartner, 2021). So, faced with these realities, a business may ponder: does ML really have to rely so heavily on humans? Enter 'automated machine learning' (AutoML), a research endeavour that has become particularly popular over the last decade (Elshawi *et al.*, 2019; Hutter *et al.*, 2019; Lee *et al.*, 2019; Escalante, 2020; Santu *et al.*, 2020; He *et al.*, 2021; Zöller and Huber, 2021), striving to mechanise as many high-level ML operations as possible. The appeal of this emergent field is multi-faceted, driven by many of the same motivations that inspire automation in general. These include not just democratisation, enabling the broader public to leverage the power of ML approaches, but also efficiency boosts, redistributing the time and effort of existing talent to more valuable functions.

Notably, within the modern era of AutoML, academia has already made much progress. Admittedly, it can be challenging to contain this ever-widening field within a simple overview, and various works lean on taxonomies and categorical systems to aid this (Kedziora *et al.*, 2020; Santu *et al.*, 2020; Dong *et al.*, 2021; Khuat *et al.*, 2021). Consider then a conceptual representation of the processes that are involved in running a real-world ML application, i.e. an ML workflow, as shown in Figure 1.1. With respect to this depiction, the bulk of AutoML research has traditionally focused on automating the model-development phase. Advances in Bayesian optimisation, which continue to be employed (Klein

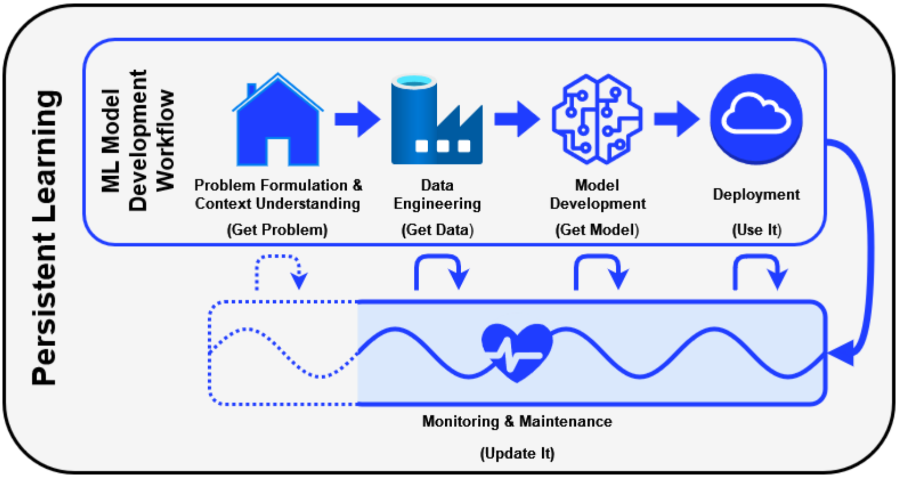

Figure 1.1: Schematic of a general machine learning (ML) workflow, which captures the phases involved in designing, constructing, deploying and maintaining an ML model for a real-world application.

et al., 2017; Frazier, 2018), are frequently credited for jump-starting this process, reducing human involvement in hyperparameter optimisation (HPO) and chipping away at the broader 'combined algorithm selection and HPO' (CASH) problem (Thornton *et al.*, 2013). Since then, this undertaking has evolved in many ways, such as by encapsulating neural architecture search (NAS) (Wistuba *et al.*, 2019; Ren *et al.*, 2021; Dong *et al.*, 2020a; Dong *et al.*, 2020b), which now forms the core of the AutoML-subfield known as 'automated DL' (AutoDL).

However, as previously hinted, the scope of AutoML – AutoDL included (Dong *et al.*, 2021) – has itself gradually expanded to encompass the rest of an ML workflow. For instance, data engineering has received its own fair share of research attention. Some works in this space focus on the initial stage of data preparation (Khayyat *et al.*, 2015), which may involve sampling and cleaning, while others contribute to the topic of automated feature engineering (AutoFE) (Lam *et al.*, 2017), covering both feature generation and selection. Then there are phase-agnostic methodologies, such as meta-learning (Ali *et al.*, 2015; Lemke *et al.*, 2015; Finn *et al.*, 2017; Vanschoren, 2018; Hutter *et al.*, 2019; Lemke and Gabrys, 2010a), that can theoretically be applied anywhere;

these continue to free humans from micro-managing ML systems and supplying domain knowledge. Of course, there is still much further to go. Automating the phase of continuous monitoring and maintenance has recently been highlighted as a crucial prerequisite for truly autonomous machine learning (AutonoML) (Kedziora *et al.*, 2020), where systems persist by adapting ML models to changes in data environment (Kadlec and Gabrys, 2009a; Zliobaite *et al.*, 2012). Progress in this space remains relatively nascent (Lu *et al.*, 2019; Chen *et al.*, 2021a; Gheibi *et al.*, 2021). Additionally, rigorous efforts to survey and benchmark state-of-the-art (SOTA) algorithms and approaches (Escalante, 2020; Romano *et al.*, 2020) are relatively sparse. Nevertheless, the key takeaway from all of this is that, academically, the field of AutoML is rich with activity.

Unfortunately, the translation of pure theory to real-world practice is rarely smooth or one-to-one. That is not to suggest that AutoML has been shunned by industry; to the absolute contrary, a prior dearth of tools to assist with developing ML models – according to the IBM survey (IBM, 2021), one of the top three obstacles for AI uptake – has actually led to an explosion of commercial AutoML services. Alongside numerous open-source packages, these offerings provide businesses plenty of options to choose from, as of the early 2020s, should they wish to apply ML approaches to problems of interest. Yet a healthy scepticism remains warranted, especially where source code is confidential and promotional material is inherently biased. It cannot be assumed that AutoML algorithms and architectures, developed in experimental environments that are well-controlled and sanitised, will deliver optimal outcomes once applied within messy real-world contexts. Certainly, the academic case studies that exist (Orlenko *et al.*, 2018; Padmanabhan *et al.*, 2019; Tsiakmaki *et al.*, 2020), evaluating one or more AutoML solutions within particular industrial domains, are too few in number to make broad claims. So, it is worth asking the question: how are publicly available SOTA AutoML tools and services performing with respect to the demands of industry?

The notion of ‘performant’ ML must be central to any pioneering survey that grapples with this question. In most academic research, performance is usually gauged by purely technical metrics, such as model accuracy and training efficiency. The focus is on how well a computer, in the absence of any human, can generate predictions/prescriptions via ML techniques. On the other hand, industrial contexts are much more human-centric, where stakeholders may have a diversity of interests and obligations; the outcomes and impact of an ML application may only be very loosely correlated with technical performance. Importantly, such matters cannot be ignored by academic AutoML researchers either, as stakeholder requirements can affect the very foundations of algorithms and architectures. For instance, a need for interpretability may force ML model-selection pools to be constrained, a focus on fairness may require mechanisms for bias mitigation, and so on.

Simply put, the technological emergence of AutoML is driven by stakeholder need and the human-computer interaction (HCI) required to service it. Correspondingly, it is impossible to gauge the current state of AutoML technology, especially in terms of whether it can support the needs of industry, without the careful development of an assessment framework anchored by a comprehensive set of HCI-weighted criteria for 'performant ML'. Certainly, the absence of such a systematic appraisal may not only obscure future directions for progress but, if deficiencies are not identified, may also have an eventual chilling effect on technological engagement, especially in the case of unmet expectations.

With all that stated, the primary goal of this review is to present a comprehensive snapshot of how AutoML has permeated into mainstream use within the early 2020s. In contrast to two associated monographs that examined fundamental algorithms and approaches behind AutoML/AutoDL (Kedziora *et al.*, 2020; Dong *et al.*, 2021), this work surveys both their implementation and application in the context of industry. It also defines what a 'performant' AutoML system is – HCI support is valued highly here – and assesses how the current crop of available packages and services, as a whole, lives up to expectation. To do so in a systematic manner, this review is structured as follows. Section 2 begins by elaborating on the notion of an ML workflow, conceptually framing AutoML in terms of the high-level operations required to develop, deploy and maintain an ML model. Section 3 uses this workflow to support the introduction of industry-related stakeholders and their interests/obligations. These requirements are unified into a comprehensive set of criteria, supported by methods of assessment, that determine whether an AutoML system can be considered performant. Section 4 then launches the survey in earnest, assessing the nature and capabilities of existing AutoML technology. This begins with an examination of open-source AutoML packages; some of these are tools dedicated to a singular purpose, e.g. HPO, while others are comprehensive systems that aim to automate a significant portion of an ML workflow. The section additionally investigates AutoML systems that are designed for specific domains, as well as commercial products. Subsequently, Section 5 assesses where AutoML technology has been used and how it has fared. Academic work focusing on real-world applications is surveyed, as are vendor-based case studies. All key findings and assessments are then synthesised in Section 6, with commentary around how mature AutoML technology is, as well as whether there are obstacles and opportunities for future uptake. Finally, Section 7 provides a concluding overview on the technological emergence of AutoML.

Appendices

A

Faded Hyperparameter Optimisation Tools

Table A.1

Name	GitHub	Ref.
Adatune	https://github.com/awslabs/adatune	Amazon Web Services, 2021a
Darts	https://github.com/quark0/darts	Liu, 2021
DeepArchitect	https://github.com/negrinho/deep_architect	Negrinho, 2021
FAR-HO	https://github.com/lucfra/FAR-HO	Franceschi, 2021
GPyOpt	https://github.com/SheffieldML/GPyOpt	The Machine Learning Group at The University of Sheffield, 2021
Optunity	https://github.com/claesenm/optunity	Claesen, 2021
Osprey	https://github.com/msmbuilder/osprey	McGibbon *et al.*, 2021
pyGPGO	https://github.com/josejimenezluna/pyGPGO	Jiménez, 2021
RoBO	https://github.com/automl/RoBO	AutoML Groups Freiburg and Hannover, 2015
sklearn-deap	https://github.com/rsteca/sklearn-deap	rsteca, 2021
Spearmint	https://github.com/HIPS/Spearmint	Harvard Intelligent Probabilistic Systems Group, 2021

B

Faded AutoML Systems

Table B.1

Name	GitHub	Company	Ref.
Adanet	https://github.com/tensorflow/adanet	Google	TensorFlow, 2021
Advisor	https://github.com/tobegit3hub/advisor	Personal	Chen, 2021
Aethos	https://github.com/Ashton-Sidhu/aethos	Personal	Sidhu, 2021
Amla	https://github.com/CiscoAI/amla	Cisco	Cisco AI, 2021
ATM	https://github.com/HDI-Project/ATM	Research	MIT - The Human Data Interaction Project, 2021
auto_ml	https://github.com/ClimbsRocks/auto_ml	Personal	Parry, 2021
Auto-Weka	https://github.com/automl/pyautoweka	Research	AutoML Groups Freiburg and Hannover, 2021c
AutoXGBoost	https://github.com/ja-thomas/autoxgboost	Personal	Thomas, 2021
DeepMining	https://github.com/sds-dubois/DeepMining	Research	Anderson, 2017; Dubois, 2017
FedNas	https://github.com/chaoyanghe/FedNAS	Research	He, 2020; He *et al.*, 2020
HpBandSter	https://github.com/automl/HpBandSter	Research	AutoML Groups Freiburg and Hannover, 2017
MetaQNN	https://github.com/bowenbaker/metaqnn	Research	Baker, 2021; Baker *et al.*, 2017
MLBox	https://github.com/AxeldeRomblay/MLBox	Personal	Romblay, 2017
Recipe	https://github.com/laic-ufmg/Recipe	Research	Laboratório de Inteligência Computacional (Computational Intelligence Lab) at Universidade Federal de Minas Gerais, 2021
tuneRanger	https://github.com/PhilippPro/tuneRanger	Research	Probst, 2021
Xcessiv	https://github.com/reiinakano/xcessiv	Research	Nakano, 2021

C

Insufficiently Detailed Commercial Systems

Table C.1

Name	Website	Ref.
Aible	https://www.aible.com/	*Aible* 2021a
Algolytics	https://algolytics.com/products/abm/	*Algolytics* 2021
DMway	http://dmway.com/	*DMway* 2021
dotData	https://dotdata.com/	*dotData* 2021
Kortical	https://kortical.com/	*Kortical* 2021
neuralstudio.ai	https://neuralstudio.ai/	*neuralstudio.Ai* 2021
OptiScorer	https://optiscorer.com/	*OptiScorer* 2021
Pecan	https://www.pecan.ai/	*Pecan* 2021
Prevision.io	https://prevision.io/	*Prevision.io* 2021
SparkCognition	https://www.sparkcognition.com/products/darwin/	*SparkCognition* 2021
TAZI	https://www.tazi.ai/	*TAZI* 2021
Xpanse	https://xpanse.ai/	*Xpanse* 2021

References

Abdullah-Al-Kafi, M., I. J. Tasnova, M. Wadud Islam, and S. K. Banshal. (2022). "Performances of Different Approaches for Fake News Classification: An Analytical Study". In: *Advanced Network Technologies and Intelligent Computing*. Ed. by I. Woungang, S. K. Dhurandher, K. K. Pattanaik, A. Verma, and P. Verma. *Communications in Computer and Information Science*. Cham: Springer International Publishing. 700–714. DOI: 10.1007/978-3-030-96040-7_53.

Adebayo, J. (2021). "Fairml". URL: https://github.com/adebayoj/fairml.

Adegboye, A. and M. Kampouridis. (2021). "Machine Learning Classification and Regression Models for Predicting Directional Changes Trend Reversal in FX Markets". *Expert Systems with Applications*. 173(July): 114645. DOI: 10.1016/j.eswa.2021.114645.

Adegboye, A., M. Kampouridis, and F. Otero. (2021). "Improving Trend Reversal Estimation in Forex Markets under a Directional Changes Paradigm with Classification Algorithms". *International Journal of Intelligent Systems*. 36(12): 7609–7640. DOI: 10.1002/int.22601.

Agarwal, A., A. Beygelzimer, M. Dudík, J. Langford, and H. Wallach. (2018). "A Reductions Approach to Fair Classification". In: *International Conference on Machine Learning*. PMLR. 60–69.

Agarwal, A., M. Dudík, and Z. S. Wu. (2019). "Fair Regression: Quantitative Definitions and Reduction-Based Algorithms". In: *International Conference on Machine Learning*. PMLR. 120–129.

Ahmadighohandizi, F. and K. Systä. (2018). "Application Development and Deployment for IoT Devices". In: *Advances in Service-Oriented and Cloud Computing*. Ed. by A. Lazovik and S. Schulte. *Communications in Computer and Information Science*. Cham: Springer International Publishing. 74–85. DOI: 10.1007/978-3-319-72125-5_6.

Ahmed, S., R. Muhammod, Z. H. Khan, S. Adilina, A. Sharma, S. Shatabda, and A. Dehzangi. (2021). "ACP-MHCNN: An Accurate Multi-Headed Deep-Convolutional Neural Network to Predict Anticancer Peptides". *Scientific Reports*. 11(1): 23676. DOI: 10.1038/s41598-021-02703-3.

Ai, S., J. Song, and G. Cai. (2021). "A Real-Time Fault Diagnosis Method for Hypersonic Air Vehicle with Sensor Fault Based on the Auto Temporal Convolutional Network". *Aerospace Science and Technology*. 119(Dec.): 107220. DOI: 10.1016/j.ast.2021.107220.

"Aible". (2021a). URL: https://www.aible.com.

Aible. (2021b). *True AutoML*. URL: https://www.aible.com/blog/true-automl.

Akiba, T., S. Sano, T. Yanase, T. Ohta, and M. Koyama. (2019). "Optuna: A next-Generation Hyperparameter Optimization Framework". In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2623–2631.

Akiba, T., S. Sano, T. Yanase, T. Ohta, and M. Koyama. (2021). "Optuna". URL: https://github.com/optuna/optuna.

Alaa, A. and M. Schaar. (2018). "AutoPrognosis: Automated Clinical Prognostic Modeling via Bayesian Optimization with Structured Kernel Learning". In: *International Conference on Machine Learning*. 139–148.

Alaa, A. M. (2021). "AutoPrognosis". URL: https://github.com/ahmedmalaa/AutoPrognosis.

"Algolytics". (2021). URL: https://algolytics.com/products/abm/.

Ali, A. R., M. Budka, and B. Gabrys. (2015). "A Review of Meta-Level Learning in the Context of Multi-Component, Multi-Level Evolving Prediction Systems". *arXiv preprint arXiv:2007.10818*.

Ali, A. R., B. Gabrys, and M. Budka. (2018). "Cross-Domain Meta-Learning for Time-Series Forecasting". *Procedia Computer Science*. 126: 9–18. DOI: 10.1016/j.procs.2018.07.204.

Alnegheimish, S., N. Alrashed, F. Aleissa, S. Althobaiti, D. Liu, M. Alsaleh, and K. Veeramachaneni. (2020). "Cardea: An Open Automated Machine Learning Framework for Electronic Health Records". In: *2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA)*. IEEE. 536–545.

Alonso-Martín, F., J. J. Gamboa-Montero, J. C. Castillo, Á. Castro-González, and M. Á. Salichs. (2017). "Detecting and Classifying Human Touches in a Social Robot Through Acoustic Sensing and Machine Learning". *Sensors*. 17(5): 1138. DOI: 10.3390/s17051138.

AlternativeTo. (2021). URL: https://alternativeto.net.

"Alteryx". (2021a). URL: https://www.alteryx.com/analytic-process-automation.

Alteryx. (2021b). "Compose". URL: https://github.com/alteryx/compose.

Alteryx. (2021c). "EvalML". URL: https://github.com/alteryx/evalml.

Alteryx. (2021d). "Featuretools". URL: https://github.com/alteryx/featuretools.

Amazon Web Services. (2021a). "Adatune". URL: https://github.com/awslabs/adatune.

Amazon Web Services. (2021b). "AutoGluon". URL: https://github.com/awslabs/autogluon.

Amin, R., C. R. Rahman, S. Ahmed, M. H. R. Sifat, M. N. K. Liton, M. M. Rahman, M. Z. H. Khan, and S. Shatabda. (2020). "iPromoter-BnCNN: A Novel Branched CNN-based Predictor for Identifying and Classifying Sigma Promoters". *Bioinformatics*. 36(19): 4869–4875. DOI: 10.1093/bioinformatics/btaa609.

Anaconda. (2020). "The State of Data Science 2020". *Tech. rep.* URL: https://www.anaconda.com/state-of-data-science-2020.

AnalyticsVidhya. (2021). *AnalyticsVidhya blog posts tagged with 'automl'*. URL: https://www.analyticsvidhya.com/blog/tag/automl/.

Anderson, A. W. (W. (2017). "Deep Mining : Scaling Bayesian Auto-Tuning of Data Science Pipelines". *Thesis*. Massachusetts Institute of Technology.

Angarita-Zapata, J. S., G. Maestre-Gongora, and J. F. Calderín. (2021). "A Bibliometric Analysis and Benchmark of Machine Learning and AutoML in Crash Severity Prediction: The Case Study of Three Colombian Cities". *Sensors*. 21(24): 8401. DOI: 10.3390/s21248401.

Angarita-Zapata, J. S., A. D. Masegosa, and I. Triguero. (2020). "General-Purpose Automated Machine Learning for Transportation: A Case Study of Auto-sklearn for Traffic Forecasting". In: *Information Processing and Management of Uncertainty in Knowledge-Based Systems. Communications in Computer and Information Science*. Cham: Springer International Publishing. 728–744. DOI: 10.1007/978-3-030-50143-3_57.

Angarita-Zapata, J. S., I. Triguero, and A. D. Masegosa. (2018). "A Preliminary Study on Automatic Algorithm Selection for Short-Term Traffic Forecasting". In: *Intelligent Distributed Computing XII*. Ed. by J. Del Ser, E. Osaba, M. N. Bilbao, J. J. Sanchez-Medina, M. Vecchio, and X.-S. Yang. *Studies in Computational Intelligence*. Cham: Springer International Publishing. 204–214. DOI: 10.1007/978-3-319-99626-4_18.

Anwar, T. and H. Anwar. (2021). "Beef Quality Assessment Using AutoML". In: *2021 Mohammad Ali Jinnah University International Conference on Computing (MAJICC)*. 1–4. DOI: 10.1109/MAJICC53071.2021.9526256.

ANZ, I. G. (2019). "IG INDUSTRY SURVEY July 2019 Report". *Tech. rep.* URL: https://www.infogovanz.com/wp-content/uploads/2020/01/IGANZ2019ReportFinal.pdf.

ANZ, I. G. (2021). "IG INDUSTRY REPORT May 2021". *Tech. rep.* URL: https://www.infogovanz.com/wp-content/uploads/2021/05/InfoGov_IndustrySurvey_MAY2021.pdf.

Arya, V., R. K. E. Bellamy, P.-Y. Chen, A. Dhurandhar, M. Hind, S. C. Hoffman, S. Houde, Q. V. Liao, R. Luss, A. Mojsilović, S. Mourad, P. Pedemonte, R. Raghavendra, J. Richards, P. Sattigeri, K. Shanmugam, M. Singh, K. R. Varshney, D. Wei, and Y. Zhang. (2019). "One Explanation Does Not Fit All: A Toolkit and Taxonomy of AI Explainability Techniques". *arXiv preprint arXiv:1909.03012*.

"Auger". (2021a). URL: https://auger.ai/.

Auger. (2021b). *How Does Auger.AI Provide Such Fast And Accurate AutoML? – Auger*. URL: https://auger.ai/how-does-auger-ai-provide-such-fast-and-accurate-automl/.

Australian Computer Society. (2021). “DIGITAL PULSE 2021”. *Tech. rep.* URL: https://www.acs.org.au/insightsandpublications/reports-publications/digital-pulse-2021.html.

AutoML Groups Freiburg and Hannover. (2015). “RoBO”. URL: https://github.com/automl/RoBO.

AutoML Groups Freiburg and Hannover. (2016). “SMAC3”. URL: https://github.com/automl/SMAC3.

AutoML Groups Freiburg and Hannover. (2017). “HpBandSter”. URL: https://github.com/automl/HpBandSter.

AutoML Groups Freiburg and Hannover. (2021a). “Auto-PyTorch”. URL: https://github.com/automl/Auto-PyTorch.

AutoML Groups Freiburg and Hannover. (2021b). “auto-sklearn”. URL: https://github.com/automl/auto-sklearn.

AutoML Groups Freiburg and Hannover. (2021c). “Pyautoweka”. URL: https://github.com/automl/pyautoweka.

AWS. (2021a). *Getting Started with Amazon SageMaker - Amazon Web Services*. URL: https://aws.amazon.com/sagemaker/getting-started/.

AWS. (2021b). *Sagemaker Documentation: Notbooks Sharing*. URL: https://docs.aws.amazon.com/sagemaker/latest/dg/notebooks-sharing.html.

AWS. (2021c). *What Is DevOps? - Amazon Web Services (AWS)*. URL: https://aws.amazon.com/devops/what-is-devops/.

“B2Metric”. (2021). URL: https://b2metric.com.

Baccouri, N. (2020). “Igel”. URL: https://github.com/nidhaloff/igel.

Bai, Y., Y. Li, Y. Shen, M. Yang, W. Zhang, and B. Cui. (2022). “AutoDC: An Automatic Machine Learning Framework for Disease Classification”. *Bioinformatics*. 38(13): 3415–3421. DOI: 10.1093/bioinformatics/btac334.

Baker, B. (2021). “MetaQNN”. URL: https://github.com/bowenbaker/metaqnn.

Baker, B., O. Gupta, N. Naik, and R. Raskar. (2017). “Designing Neural Network Architectures Using Reinforcement Learning”. In: *5th International Conference on Learning Representations, ICLR 2017, Conference Track Proceedings*. OpenReview.net. URL: https://openreview.net/forum?id=S1c2cvqee.

Bakirov, R., D. Fay, and B. Gabrys. (2021). “Automated adaptation strategies for stream learning”. *Machine Learning*: 1–34.

Bakirov, R., B. Gabrys, and D. Fay. (2015). “On Sequences of Different Adaptive Mechanisms in Non-Stationary Regression Problems”. In: *2015 International Joint Conference on Neural Networks (IJCNN)*. IEEE. DOI: 10.1109/ijcnn.2015.7280779.

Bakirov, R., B. Gabrys, and D. Fay. (2017). “Multiple Adaptive Mechanisms for Data-Driven Soft Sensors”. *Computers & Chemical Engineering*. 96: 42–54. DOI: 10.1016/j.compchemeng.2016.08.017.

Bakirov, R., B. Gabrys, and D. Fay. (2018). “Generic Adaptation Strategies for Automated Machine Learning”. *arXiv preprint arXiv:1812.10793v2*.

Balandat, M., B. Karrer, D. R. Jiang, S. Daulton, B. Letham, A. G. Wilson, and E. Bakshy. (2019). “Botorch: Programmable Bayesian Optimization in Pytorch”. *arXiv preprint arXiv:1910.06403*.

Bangash, A. H. (2020). “Leveraging AutoML to Provide NAFLD Screening Diagnosis: Proposed Machine Learning Models”. *medRxiv preprint medRxiv:2020.10.20.20216291v1*.

Bangash, A. H., A. H. Shah, A. Fatima, S. Zehra, S. M. M. Abbas, H. F. Khawaja, M. Ashraf, and A. Baloch. (2021). “Amalgamation of Auto Machine Learning and Ensemble Approaches to Achieve State-of-the-Art Post-Heart Failure Survival Predictions”. *American Heart Journal*. 242(Dec.): 153. DOI: 10.1016/j.ahj.2021.10.021.

Barbadora, L. (2021). *Aktana Acquires Ople.AI’s Machine Learning Automation Technology*. URL: https://www.globenewswire.com/en/news-release/2021/10/27/2321826/0/en/Aktana-Acquires-Ople-AI-s-Machine-Learning-Automation-Technology.html.

Barreiro, E., C. R. Munteanu, M. Cruz-Monteagudo, A. Pazos, and H. González-Díaz. (2018). “Net-Net Auto Machine Learning (AutoML) Prediction of Complex Ecosystems”. *Scientific Reports*. 8(1). DOI: 10.1038/s41598-018-30637-w.

Basgalupp, M. P., R. C. Barros, A. G. de Sá, G. L. Pappa, R. G. Mantovani, A. de Carvalho, and A. A. Freitas. (2020). "An Extensive Experimental Evaluation of Automated Machine Learning Methods for Recommending Classification Algorithms". *Evolutionary Intelligence*: 1–20.

Baškarada, S. and A. Koronios. (2017). "Unicorn Data Scientist: The Rarest of Breeds". *Program: electronic library and information systems*. 51(1): 65–74.

Basu, Sreeradha. (2018). "50,000+ Data Science, AI Jobs Vacant Due to Shortage of Talent [Careers & Companies]: Market Has Twice the Number of Jobs than Jobseekers, Finds a Report". *The Economic Times (Online)*.

Bauer, M., C. van Dinther, and D. Kiefer. (2020). "Machine learning in SME: An empirical study on enablers and success factors". In: *AMCIS 2020 Virtual Conference, August 10-14, 2020*.

Baumeister, F., M. W. Barbosa, and R. R. Gomes. (2020). "What Is Required to Be a Data Scientist?: Analyzing Job Descriptions With Centering Resonance Analysis". *International Journal of Human Capital and Information Technology Professionals (IJHCITP)*. 11(4): 21–40. DOI: 10.4018/IJHCITP.2020100102.

Bayer, P. E., J. Petereit, M. F. Danilevicz, R. Anderson, J. Batley, and D. Edwards. (2021). "The Application of Pangenomics and Machine Learning in Genomic Selection in Plants". *The Plant Genome*. 14(3): e20112. DOI: 10.1002/tpg2.20112.

Beam, A. L., A. K. Manrai, and M. Ghassemi. (2020). "Challenges to the Reproducibility of Machine Learning Models in Health Care". *JAMA*. 323(4): 305–306. DOI: 10.1001/jama.2019.20866.

Bergstra, J., D. Yamins, and D. D. Cox. (2021). "Hyperopt". URL: https://github.com/hyperopt/hyperopt.

Bergstra, J., D. Yamins, and D. Cox. (2013). "Making a Science of Model Search: Hyperparameter Optimization in Hundreds of Dimensions for Vision Architectures". In: *International Conference on Machine Learning*. PMLR. 115–123.

Bertl, M., P. Ross, and D. Draheim. (2021). "Predicting Psychiatric Diseases Using AutoAI: A Performance Analysis Based on Health Insurance Billing Data". In: *Database and Expert Systems Applications*. Ed. by C. Strauss, G. Kotsis, A. M. Tjoa, and I. Khalil. *Lecture Notes in Computer Science*. Cham: Springer International Publishing. 104–111. DOI: 10.1007/978-3-030-86472-9_9.

"Big Squid". (2021). URL: https://bigsquid.ai/kraken.

"BigML". (2021). URL: https://bigml.com.

Bilalli, B., A. Abelló, T. Aluja-Banet, R. F. Munir, and R. Wrembel. (2018). "PRESISTANT: data pre-processing assistant". In: *International Conference on Advanced Information Systems Engineering*. Springer International Publishing. 57–65.

Binder, M., B. Bischl, and A. Hanf. (2020). "mlr3automl". URL: https://github.com/a-hanf/mlr3automl.

Bird, S., M. Dudík, R. Edgar, B. Horn, R. Lutz, V. Milan, M. Sameki, H. Wallach, and K. Walker. (2020). "Fairlearn: A Toolkit for Assessing and Improving Fairness in AI". *Microsoft, Tech. Rep. MSR-TR-2020-32*.

Bischl, B., J. Richter, J. Bossek, D. Horn, J. Thomas, and M. Lang. (2018). "mlrMBO: A Modular Framework for Model-Based Optimization of Expensive Black-Box Functions". *arXiv preprint arXiv:1703.03373*.

Black, R. (2002). *Managing the Testing Process*. John Wiley & Sons.

Boehmke, B. and B. Greenwell. (2019). *Hands-On Machine Learning with R*. Boca Raton: Chapman and Hall/CRC. DOI: 10.1201/9780367816377.

Bond, R., A. Koene, A. Dix, J. Boger, M. D. Mulvenna, M. Galushka, B. W. Bradley, F. Browne, H. Wang, and A. Wong. (2019). "Democratisation of Usable Machine Learning in Computer Vision". *arXiv preprint arXiv:1902.06804v1*.

Bosch, N. (2021). "AutoML Feature Engineering for Student Modeling Yields High Accuracy, but Limited Interpretability". *Journal of Educational Data Mining*. 13(2): 55–79. DOI: 10.5281/zenodo.5275314.

Box, G. E. P., G. M. Jenkins, G. C. Reinsel, and G. M. Ljung. (2015). *Time Series Analysis: Forecasting and Control*. John Wiley & Sons.

Breck, E., S. Cai, E. Nielsen, M. Salib, and D. Sculley. (2017). "The ML Test Score: A Rubric for ML Production Readiness and Technical Debt Reduction". In: *2017 IEEE International Conference on Big Data (Big Data)*. 1123–1132. DOI: 10.1109/BigData.2017.8258038.

Breese, J. S., D. Heckerman, and C. Kadie. (1998). "Empirical Analysis of Predictive Algorithms for Collaborative Filtering". In: *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence. UAI'98*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc. 43–52.

Brindise, L. R. and R. J. Steele. (2018). "Machine Learning-based Predischarge Prediction of Hospital Readmission". In: *2018 International Conference on Computer, Information and Telecommunication Systems (CITS)*. 1–5. DOI: 10.1109/CITS.2018.8440171.

Bulgarevich, D. S., S. Tsukamoto, T. Kasuya, M. Demura, and M. Watanabe. (2019). "Automatic Steel Labeling on Certain Microstructural Constituents with Image Processing and Machine Learning Tools". *Science and Technology of Advanced Materials*. 20(1): 532–542. DOI: 10.1080/14686996.2019.1610668.

Calmon, F., D. Wei, B. Vinzamuri, K. Natesan Ramamurthy, and K. R. Varshney. (2017). "Optimized Pre-Processing for Discrimination Prevention". In: *Advances in Neural Information Processing Systems*. Ed. by I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett. Vol. 30. Curran Associates, Inc. URL: https://proceedings.neurips.cc/paper/2017/file/9a49a25d845a483fae4be7e341368e36-Paper.pdf.

Campero Jurado, I., A. Fedjajevs, J. Vanschoren, and A. Brombacher. (2022). "Interpretable Assessment of ST-Segment Deviation in ECG Time Series". *Sensors*. 22(13): 4919. DOI: 10.3390/s22134919.

Capgemini. (2020). "AI and the Ethical Conundrum". *Tech. rep.* URL: https://www.capgemini.com/wp-content/uploads/2020/10/AI-and-the-Ethical-Conundrum-Report.pdf.

Carper, C., A. McClellan, and C. C. Douglas. (2021). "I-80 Closures: An Autonomous Machine Learning Approach". In: *Computational Science – ICCS 2021*. Ed. by M. Paszynski, D. Kranzlmüller, V. V. Krzhizhanovskaya, J. J. Dongarra, and P. M. Sloot. *Lecture Notes in Computer Science*. Cham: Springer International Publishing. 286–291. DOI: 10.1007/978-3-030-77977-1_22.

Cashman, D., S. R. Humayoun, F. Heimerl, K. Park, S. Das, J. Thompson, B. Saket, A. Mosca, J. Stasko, A. Endert, M. Gleicher, and R. Chang. (2019). "A User-Based Visual Analytics Workflow for Exploratory Model Analysis". *Computer Graphics Forum*. 38(3): 185–199. DOI: 10.1111/cgf.13681.

Castillo, J. C., F. Alonso-Martín, D. Cáceres-Domínguez, M. Malfaz, and M. A. Salichs. (2019). "The Influence of Speed and Position in Dynamic Gesture Recognition for Human-Robot Interaction". *Journal of Sensors*. 2019(Feb.): e7060491. DOI: 10.1155/2019/7060491.

Castro, M. R., B. Van der Heijden, and E. L. Henderson. (2020). "Catalysts in Career Transitions: Academic Researchers Transitioning into Sustainable Careers in Data Science". *Journal of Vocational Behavior*. 122: 103479.

Cegielski, C. G. and L. A. Jones-Farmer. (2016). "Knowledge, Skills, and Abilities for Entry-Level Business Analytics Positions: A Multi-Method Study". *Decision Sciences Journal of Innovative Education*. 14(1): 91–118. DOI: 10.1111/dsji.12086.

Celis, L. E., L. Huang, V. Keswani, and N. K. Vishnoi. (2019). "Classification with Fairness Constraints: A Meta-Algorithm with Provable Guarantees". In: *Proceedings of the Conference on Fairness, Accountability, and Transparency*. 319–328.

Chakraborty, S., S. Shah, K. Soltani, A. Swigart, L. Yang, and K. Buckingham. (2020). "Building an Automated and Self-Aware Anomaly Detection System". In: *2020 IEEE International Conference on Big Data (Big Data)*. IEEE. 1465–1475.

Chan, S., P. Treleaven, and L. Capra. (2013). "Continuous Hyperparameter Optimization for Large-Scale Recommender Systems". In: *2013 IEEE International Conference on Big Data*. IEEE. DOI: 10.1109/bigdata.2013.6691595.

Chaudhary, A., A. Issak, K. Kate, Y. Katsis, A. Valente, D. Wang, A. Evfimievski, S. Gurajada, B. Kawas, C. Malossi, L. Popa, T. Pedapati, H. Samulowitz, M. Wistuba, and Y. Li. (2021). "AutoText: An End-to-End AutoAI Framework for Text". *Proceedings of the AAAI Conference on Artificial Intelligence*. 35(18): 16001–16003.

Chen, D. (2021). "Advisor". URL: https://github.com/tobegit3hub/advisor.

Chen, J., F. Lécué, J. Z. Pan, S. Deng, and H. Chen. (2021a). "Knowledge Graph Embeddings for Dealing with Concept Drift in Machine Learning". *Journal of Web Semantics*. 67(Feb.): 100625. DOI: 10.1016/j.websem.2020.100625.

Chen, T. and C. Guestrin. (2016). "Xgboost: A Scalable Tree Boosting System". In: *Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining*. 785–794.

Chen, Y., Y. Yang, H. Sun, Y. Wang, Y. Xu, W. Shen, R. Zhou, Y. Tong, J. Bai, and R. Zhang. (2020). "AutoADR: Automatic Model Design for Ad Relevance". In: *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*. 2365–2372.

Chen, Z., E. Brophy, and T. Ward. (2021b). "Malware Classification Using Static Disassembly and Machine Learning". *arXiv preprint arXiv:2201.07649*.

Cheng, F., D. Liu, F. Du, Y. Lin, A. Zytek, H. Li, H. Qu, and K. Veeramachaneni. (2022). "VBridge: Connecting the Dots Between Features and Data to Explain Healthcare Models". *IEEE Transactions on Visualization and Computer Graphics*. 28(1): 378–388. DOI: 10.1109/TVCG.2021.3114836.

Chomiak, K. and M. Miktus. (2021). "Harnessing Value from Data Science in Business: Ensuring Explainability and Fairness of Solutions". *arXiv preprint arXiv:2108.07714*.

Choudhary, R. and H. K. Gianey. (2017). "Comprehensive Review On Supervised Machine Learning Algorithms". In: *2017 International Conference on Machine Learning and Data Science (MLDS)*. 37–43. DOI: 10.1109/MLDS.2017.11.

Christ, M., N. Braun, J. Neuffer, and A. W. Kempa-Liehr. (2018). "Time Series FeatuRe Extraction on Basis of Scalable Hypothesis Tests (Tsfresh – A Python Package)". *Neurocomputing*. 307(Sept.): 72–77. DOI: 10.1016/j.neucom.2018.03.067.

Chugh, K. (2020). "Otto". URL: https://github.com/KartikChugh/Otto.

Cimperman, R. (2006). *UAT Defined: A Guide to Practical User Acceptance Testing (Digital Short Cut)*. Pearson Education.

Cisco AI. (2021). "Amla". URL: https://github.com/CiscoAI/amla.

Claesen, M. (2021). "Optunity". URL: https://github.com/claesenm/optunity.

CleverInsight. (2020). "Cognito". URL: https://github.com/CleverInsight/cognito.

"Cloud AutoML (Google)". (2021). URL: https://cloud.google.com/automl.

"cnvrg.Io". (2021). URL: https://cnvrg.io.

Cochran, W. G. (1954). "Some Methods for Strengthening the Common $X2$ Tests". *Biometrics*. 10(4): 417–451. DOI: 10.2307/3001616.

Code of Federal Regulations. (1978). "PART 1607 - UNIFORM GUIDELINES ON EMPLOYEE SELECTION PROCEDURES (1978)". URL: https://www.govinfo.gov/content/pkg/CFR-2014-title29-vol4/xml/CFR-2014-title29-vol4-part1607.xml.

COIN-OR Foundation. (2021). "Rbfopt". URL: https://github.com/coin-or/rbfopt.

"Comet". (2021). URL: https://www.comet.ml/site.

"Compellon". (2021). URL: https://www.compellon.com.

Complex Adaptive Systems (CAS) Lab - University of Technology Sydney. (2022). "AutoWeka4MCPS". URL: https://github.com/UTS-CASLab/autoweka.

Cordeiro, F. M., G. B. França, F. L. de Albuquerque Neto, and I. Gultepe. (2021). "Visibility and Ceiling Nowcasting Using Artificial Intelligence Techniques for Aviation Applications". *Atmosphere*. 12(12): 1657. DOI: 10.3390/atmos12121657.

Corral-Corral, R., J. A. Beltrán, C. A. Brizuela, and G. Del Rio. (2017). "Systematic Identification of Machine-Learning Models Aimed to Classify Critical Residues for Protein Function from Protein Structure". *Molecules*. 22(10): 1673. DOI: 10.3390/molecules22101673.

Costa, A. and G. Nannicini. (2018). "RBFOpt: An Open-Source Library for Black-Box Optimization with Costly Function Evaluations". *Mathematical Programming Computation*. 10(4): 597–629. DOI: 10.1007/s12532-018-0144-7.

Costa, C. and M. Y. Santos. (2017). "The Data Scientist Profile and Its Representativeness in the European E-Competence Framework and the Skills Framework for the Information Age". *International Journal of Information Management*. 37(6): 726–734. DOI: 10.1016/j.ijinfomgt.2017.07.010.

Crankshaw, D., G.-E. Sela, C. Zumar, X. Mo, J. E. Gonzalez, I. Stoica, and A. Tumanov. (2018). "InferLine: ML Inference Pipeline Composition Framework". *arXiv preprint arXiv:1812.01776v1*.

CrowdFlower. (2016). "2016 Data Science Report". *Tech. rep.* URL: https://visit.figure-eight.com/rs/416-ZBE-142/images/CrowdFlower_DataScienceReport_2016.pdf.

Czarnecki, W. M., S. Podlewska, and A. J. Bojarski. (2015). "Robust Optimization of SVM Hyperparameters in the Classification of Bioactive Compounds". *Journal of Cheminformatics*. 7(1): 38. DOI: 10.1186/s13321-015-0088-0.

d'Alessandro, B., C. O'Neil, and T. LaGatta. (2017). "Conscientious Classification: A Data Scientist's Guide to Discrimination-Aware Classification". *Big Data*. 5(2): 120–134. DOI: 10.1089/big.2016.0048.

"D2iQ (Kaptain)". (2021). URL: https://d2iq.com/products/kaptain.

da Ponte Junior, L. A., D. C. Muchaluat Saade, A. Plastino de Carvalho, R. d. C. Alves, L. C. Lima Portugal, L. de Oliveira, and M. Garcia Pereira. (2020). "Identifying Post-Traumatic Stress Symptoms Using Physiological Signals and Data Mining". In: *2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS)*. 233–238. DOI: 10.1109/CBMS49503.2020.00051.

Dafflon, J., W. H. L. Pinaya, F. Turkheimer, J. H. Cole, R. Leech, M. A. Harris, S. R. Cox, H. C. Whalley, A. M. McIntosh, and P. J. Hellyer. (2019). "Analysis of an Automated Machine Learning Approach in Brain Predictive Modelling: A Data-Driven Approach to Predict Brain Age from Cortical Anatomical Measures". *arXiv preprint arXiv:1910.03349*.

Dafflon, J., W. H. L. Pinaya, F. Turkheimer, J. H. Cole, R. Leech, M. A. Harris, S. R. Cox, H. C. Whalley, A. M. McIntosh, and P. J. Hellyer. (2020). "An Automated Machine Learning Approach to Predict Brain Age from Cortical Anatomical Measures". *Human Brain Mapping*. 41(13): 3555–3566. DOI: 10.1002/hbm.25028.

Dai, Z., L. Wang, and S. Yang. (2020). "Data Mining-Based Model Simplification and Optimization of an Electrical Power Generation System". *IEEE Transactions on Transportation Electrification*. 6(4): 1665–1678. DOI: 10.1109/TTE.2020.2995745.

Dallachiesa, M., A. Ebaid, A. Eldawy, A. Elmagarmid, I. F. Ilyas, M. Ouzzani, and N. Tang. (2013). "NADEEF: A Commodity Data Cleaning System". In: *Proceedings of the 2013 ACM SIGMOD International Conference on Management of Data*. 541–552.

Dama International. (2017). *DAMA-DMBOK: Data Management Body of Knowledge*. Second. Technics Publications, LLC.

Data Analytics - Qatar Computing Research Institute, HBKU. (2021). "NADEEF". URL: https://github.com/daqcri/NADEEF.

Data Analytics Lab at Texas A&M University. (2020). "TODS". URL: https://github.com/datamllab/tods.

Data Science for Social Good. (2021). "Aequitas". URL: https://github.com/dssg/aequitas.

"Databricks". (2021a). URL: https://databricks.com/product/automl.

Databricks. (2021b). *Collaborative Notebooks - Databricks*. URL: https://databricks.com/product/collaborative-notebooks.

DataCanvas. (2020). "HyperGBM". URL: https://github.com/DataCanvasIO/HyperGBM.

"Dataiku". (2021). URL: https://www.dataiku.com.

DataQuest. (2021). *Data Analyst Skills – 8 Skills You Need to Get a Job*. URL: https://www.dataquest.io/blog/data-analyst-skills/.

"DataRobot". (2021a). URL: https://www.datarobot.com.

DataRobot. (2021b). *What Is MLOps?* URL: https://www.datarobot.com/lp/what-is-mlops/.

Datarobot Alternatives. (2021). URL: https://www.gartner.com/reviews/market/data-science-machine-learning-platforms/vendor/datarobot/alternatives.

Dattner, B., T. Chamorro-Premuzic, R. Buchband, and L. Schettler. (2019). "The Legal and Ethical Implications of Using AI in Hiring". *Harvard Business Review*. 25.

Davenport, T. H. and D. Patil. (2012). "Data Scientist". *Harvard business review*. 90(5): 70–76.

De Armas Jacomino, L., M. A. Medina-Pérez, R. Monroy, D. Valdes-Ramirez, C. Morell-Pérez, and R. Bello. (2021). "Dwell Time Estimation of Import Containers as an Ordinal Regression Problem". *Applied Sciences*. 11(20): 9380. DOI: 10.3390/app11209380.

De Mauro, A., M. Greco, M. Grimaldi, and P. Ritala. (2018). "Human Resources for Big Data Professions: A Systematic Classification of Job Roles and Required Skill Sets". *Information Processing & Management*. In (Big) Data We Trust: Value Creation in Knowledge Organizations 54(5): 807–817. DOI: 10.1016/j.ipm.2017.05.004.

de Sá, A. G. C., C. G. Pimenta, G. L. Pappa, and A. A. Freitas. (2020). "A Robust Experimental Evaluation of Automated Multi-Label Classification Methods". In: *Proceedings of the 2020 Genetic and Evolutionary Computation Conference. GECCO '20*. New York, NY, USA: Association for Computing Machinery. 175–183. DOI: 10.1145/3377930.3390231.

De Veaux, R. D., M. Agarwal, M. Averett, B. S. Baumer, A. Bray, T. C. Bressoud, L. Bryant, L. Z. Cheng, A. Francis, R. Gould, A. Y. K. Kim, M. Kretchmar, Q. Lu, A. Moskol, D. Nolan, R. Pelayo, S. Raleigh, R. J. Sethi, M. Sondjaja, N. Tiruviluamala, P. X. Uhlig, T. M. Washington, C. L. Wesley, D. White, and P. Ye. (2017). "Curriculum Guidelines for Undergraduate Programs in Data Science". *Annual Review of Statistics and Its Application*. 4: 15–30.

Dean, J., M. Scheffler, T. A. R. Purcell, S. V. Barabash, R. Bhowmik, and T. Bazhirov. (2021). "Interpretable Machine Learning for Materials Design". *arXiv preprint arXiv:2112.00239*.

"Deep Cognition". (2021). URL: https://deepcognition.ai.

Deng, D., R. C. Fernandez, Z. Abedjan, S. Wang, M. Stonebraker, A. K. Elmagarmid, I. F. Ilyas, S. Madden, M. Ouzzani, and N. Tang. (2017). "The Data Civilizer System". In: *8th Biennial Conference on Innovative Data Systems Research, CIDR 2017, Chaminade, CA, USA, January 8-11, 2017, Online Proceedings*. URL: http://cidrdb.org/cidr2017/papers/p44-deng-cidr17.pdf.

"Determined AI". (2021). URL: https://www.determined.ai.

Díaz-Pacheco, A., C. A. Reyes-García, and V. Chicatto-Gasperín. (2021). "Granule-Based Fuzzy Rules to Assist in the Infant-Crying Pattern Recognition Problem". *Sādhanā*. 46(4): 199. DOI: 10.1007/s12046-021-01736-8.

"DMway". (2021). URL: https://dmway.com/.

"Domino". (2021). URL: https://www.dominodatalab.com/product/domino-data-science-platform/.

Dong, X., D. J. Kedziora, K. Musial, and B. Gabrys. (2021). "Automated Deep Learning: Neural Architecture Search Is Not the End". *arXiv preprint arXiv:2112.09245*.

Dong, X., L. Liu, K. Musial, and B. Gabrys. (2020a). "NATS-Bench: Benchmarking NAS Algorithms for Architecture Topology and Size". *arXiv preprint arXiv:2009.00437v4*.

Dong, X., M. Tan, A. W. Yu, D. Peng, B. Gabrys, and Q. V. Le. (2020b). "AutoHAS: Efficient Hyperparameter and Architecture Search". *arXiv preprint arXiv:2006.03656*.

"dotData". (2021). URL: https://dotdata.com.

Drozdal, J., J. Weisz, D. Wang, G. Dass, B. Yao, C. Zhao, M. Muller, L. Ju, and H. Su. (2020). "Trust in Automl: Exploring Information Needs for Establishing Trust in Automated Machine Learning Systems". In: *Proceedings of the 25th International Conference on Intelligent User Interfaces*. 297–307.

Dubois, S. (2017). "DeepMining". URL: https://github.com/sds-dubois/DeepMining.

Dunner, C., T. Parnell, K. Atasu, M. Sifalakis, and H. Pozidis. (2017). "Understanding and Optimizing the Performance of Distributed Machine Learning Applications on Apache Spark". In: *2017 IEEE International Conference on Big Data (Big Data)*. IEEE. DOI: 10.1109/bigdata.2017.8257942.

Ebert, A., M. T. Beck, A. Mattausch, L. Belzner, and C. Linnhoff-Popien. (2017). "Qualitative Assessment of Recurrent Human Motion". In: *2017 25th European Signal Processing Conference (EUSIPCO)*. 306–310. DOI: 10.23919/EUSIPCO.2017.8081218.

Edalati, M., A. S. Imran, Z. Kastrati, and S. M. Daudpota. (2022). "The Potential of Machine Learning Algorithms for Sentiment Classification of Students' Feedback on MOOC". In: *Intelligent Systems and Applications*. Ed. by K. Arai. *Lecture Notes in Networks and Systems*. Cham: Springer International Publishing. 11–22. DOI: 10.1007/978-3-030-82199-9_2.

"Einblick". (2021). URL: https://einblick.ai/product/.

Elshawi, R., M. Maher, and S. Sakr. (2019). "Automated Machine Learning: State-of-The-Art and Open Challenges". *arXiv preprint arXiv:1906.02287*.

Escalante, H. J. (2020). "Automated Machine Learning – a Brief Review at the End of the Early Years". *arXiv preprint arXiv:2008.08516*.

Espinoza Mina, M. A. and D. D. P. Gallegos Barzola. (2019). "Data Scientist: A Systematic Review of the Literature". In: *Technology Trends*. Ed. by M. Botto-Tobar, G. Pizarro, M. Zúñiga-Prieto, M. D'Armas, and M. Zúñiga Sánchez. *Communications in Computer and Information Science*. Cham: Springer International Publishing. 476–487. DOI: 10.1007/978-3-030-05532-5_35.

F. de S. Soares, E., C. A. V. Campos, and S. C. de Lucena. (2019). "Online Travel Mode Detection Method Using Automated Machine Learning and Feature Engineering". *Future Generation Computer Systems*. 101(Dec.): 1201–1212. DOI: 10.1016/j.future.2019.07.056.

Facebook. (2021). "Nevergrad". URL: https://github.com/facebookresearch/nevergrad.

Fails, J. A. and D. R. Olsen. (2003). "Interactive Machine Learning". In: *Proceedings of the 8th International Conference on Intelligent User Interfaces. IUI '03*. New York, NY, USA: Association for Computing Machinery. 39–45. DOI: 10.1145/604045.604056.

Fairlearn. (2021). "Fairlearn". URL: https://github.com/fairlearn/fairlearn.

Fakoor, R., J. W. Mueller, N. Erickson, P. Chaudhari, and A. J. Smola. (2020). "Fast, Accurate, and Simple Models for Tabular Data via Augmented Distillation". In: *Advances in Neural Information Processing Systems*. Ed. by H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin. Vol. 33. Curran Associates, Inc. 8671–8681. URL: https://proceedings.neurips.cc/paper/2020/file/62d75fb2e3075506e8837d8f55021ab1-Paper.pdf.

Fati, S. M., A. Muneer, N. A. Akbar, and S. M. Taib. (2021). "A Continuous Cuffless Blood Pressure Estimation Using Tree-Based Pipeline Optimization Tool". *Symmetry*. 13(4): 686. DOI: 10.3390/sym13040686.

Fawcett, T. (2006). "An Introduction to ROC Analysis". *Pattern Recognition Letters*. ROC Analysis in Pattern Recognition 27(8): 861–874. DOI: 10.1016/j.patrec.2005.10.010.

Feldman, M., S. A. Friedler, J. Moeller, C. Scheidegger, and S. Venkatasubramanian. (2015). "Certifying and Removing Disparate Impact". In: *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD '15*. New York, NY, USA: Association for Computing Machinery. 259–268. DOI: 10.1145/2783258.2783311.

Ferreira, L., A. Pilastri, C. Martins, P. Santos, and P. Cortez. (2021). "A Scalable and Automated Machine Learning Framework to Support Risk Management". In: *Agents and Artificial Intelligence*. Ed. by A. P. Rocha, L. Steels, and J. van den Herik. *Lecture Notes in Computer Science*. Cham: Springer International Publishing. 291–307. DOI: 10.1007/978-3-030-71158-0_14.

Finn, C., P. Abbeel, and S. Levine. (2017). "Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks". In: *International Conference on Machine Learning*. PMLR. 1126–1135.

Fleming, S. W., D. C. Garen, A. G. Goodbody, C. S. McCarthy, and L. C. Landers. (2021). "Assessing the New Natural Resources Conservation Service Water Supply Forecast Model for the American West: A Challenging Test of Explainable, Automated, Ensemble Artificial Intelligence". *Journal of Hydrology*. 602(Nov.): 126782. DOI: 10.1016/j.jhydrol.2021.126782.

Fortin, F.-A., F.-M. D. Rainville, M.-A. Gardner, M. Parizeau, and C. Gagné. (2012). "DEAP: Evolutionary Algorithms Made Easy". *Journal of Machine Learning Research*. 13(70): 2171–2175.

Fortin, F.-A., F.-M. D. Rainville, M.-A. Gardner, M. Parizeau, and C. Gagné. (2021). "Deap". URL: https://github.com/DEAP/deap.

Foulds, J. R., R. Islam, K. N. Keya, and S. Pan. (2020). "An Intersectional Definition of Fairness". In: *2020 IEEE 36th International Conference on Data Engineering (ICDE)*. IEEE. 1918–1921.

Fowers, J., K. Ovtcharov, M. Papamichael, T. Massengill, M. Liu, D. Lo, S. Alkalay, M. Haselman, L. Adams, M. Ghandi, S. Heil, P. Patel, A. Sapek, G. Weisz, L. Woods, S. Lanka, S. K. Reinhardt, A. M. Caulfield, E. S. Chung, and D. Burger. (2018). "A Configurable Cloud-Scale DNN Processor for Real-Time AI". In: *2018 ACM/IEEE 45th Annual International Symposium on Computer Architecture (ISCA)*. IEEE. DOI: 10.1109/isca.2018.00012.

Franceschi, L. (2021). "FAR-HO". URL: https://github.com/lucfra/FAR-HO.

Frazier, P. I. (2018). "A Tutorial on Bayesian Optimization". *arXiv preprint arXiv:1807.02811*.

Furnham, A. and H. C. Boo. (2011). "A Literature Review of the Anchoring Effect". *The Journal of Socio-Economics*. 40(1): 35–42. DOI: 10.1016/j.socec.2010.10.008.

Gachechiladze, D., F. Lanubile, N. Novielli, and A. Serebrenik. (2017). "Anger and Its Direction in Collaborative Software Development". In: *2017 IEEE/ACM 39th International Conference on Software Engineering: New Ideas and Emerging Technologies Results Track (ICSE-NIER)*. 11–14. DOI: 10.1109/ICSE-NIER.2017.18.

Galli, S. (2021). "Feature-engine". URL: https://github.com/solegalli/feature%5C_engine.

Garmpis, S., M. Maragoudakis, and A. Garmpis. (2022). "Assisting Educational Analytics with AutoML Functionalities". *Computers*. 11(6): 97. DOI: 10.3390/computers11060097.

Garnica-Caparrós, M. and D. Memmert. (2021). "Understanding Gender Differences in Professional European Football through Machine Learning Interpretability and Match Actions Data". *Scientific Reports*. 11(1): 10805. DOI: 10.1038/s41598-021-90264-w.

Gartner. (2021). “2021-2023 Emerging Technology Roadmap for Large Enterprises”. *Tech. rep.* URL: https://www.gartner.com/en/publications/emerging-technology-roadmap-for-large-enterprises.

Gaurav, S. (2021). *Complete Guide to Feature Engineering: Zero to Hero.* URL: https://www.analyticsvidhya.com/blog/2021/09/complete-guide-to-feature-engineering-zero-to-hero/.

Geman, S., E. Bienenstock, and R. Doursat. (1992). “Neural Networks and the Bias/Variance Dilemma”. *Neural computation.* 4(1): 1–58.

Gerling, A., H. Ziekow, A. Hess, U. Schreier, C. Seiffer, and D. O. Abdeslam. (2022). “Comparison of Algorithms for Error Prediction in Manufacturing with Automl and a Cost-Based Metric”. *Journal of Intelligent Manufacturing.* 33(2): 555–573. DOI: 10.1007/s10845-021-01890-0.

Gesley, J., T. Ahmad, E. Soares, R. Levush, G. Guerra, J. Martin, K. Buchanan, L. Zhang, S. Umeda, A. Grigoryan, N. Boring, E. Hofverberg, G. Sadek, H. Goitom, C. Feikhert-Ahalt, and G. Rodriguez-Ferrand. (2019). “Regulation of Artificial Intelligence in Selected Jurisdictions”. *Tech. rep.* Law Library of Congress, Global Legal Research Directorate. URL: https://tile.loc.gov/storage-services/service/ll/llglrd/2019668143/2019668143.pdf.

Gheibi, O., D. Weyns, and F. Quin. (2021). “Applying Machine Learning in Self-Adaptive Systems: A Systematic Literature Review”. *ACM Transactions on Autonomous and Adaptive Systems (TAAS).* 15(3): 1–37.

Gianfrancesco, M. A., S. Tamang, J. Yazdany, and G. Schmajuk. (2018). “Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data”. *JAMA Internal Medicine.* 178(11): 1544–1547. DOI: 10.1001/jamainternmed.2018.3763.

Gijsbers, P., E. LeDell, J. Thomas, S. Poirier, B. Bischl, and J. Vanschoren. (2019). “An Open Source AutoML Benchmark”. *arXiv preprint arXiv:1907.00909.*

Gijsbers, Pieter. (2021). “GAMA”. URL: https://github.com/PGijsbers/gama.

GmbH, B. Y. (2016). “tsfresh”. URL: https://github.com/blue-yonder/tsfresh.

Grigorescu, S., B. Trasnea, T. Cocias, and G. Macesanu. (2020). "A Survey of Deep Learning Techniques for Autonomous Driving". *Journal of Field Robotics*. 37(3): 362–386. DOI: 10.1002/rob.21918.

Gunawardana, A. and G. Shani. (2009). "A Survey of Accuracy Evaluation Metrics of Recommendation Tasks". *The Journal of Machine Learning Research*. 10(Dec.): 2935–2962.

H2O.ai. (2021). "H2o-3". URL: https://github.com/h2oai/h2o-3.

Haghighatlari, M., C.-Y. Shih, and J. Hachmann. (2019). "Thinking Globally, Acting Locally: On the Issue of Training Set Imbalance and the Case for Local Machine Learning Models in Chemistry". *ChemRxiv preprint ChemRxiv:10.26434/chemrxiv.8796947.v1*.

Haghighatlari, M., G. Vishwakarma, D. Altarawy, R. Subramanian, B. U. Kota, A. Sonpal, S. Setlur, and J. Hachmann. (2020). "ChemML: A Machine Learning and Informatics Program Package for the Analysis, Mining, and Modeling of Chemical and Materials Data". *Wiley Interdisciplinary Reviews: Computational Molecular Science*. 10(4): e1458.

Haibe-Kains, B., G. A. Adam, A. Hosny, F. Khodakarami, T. Shraddha, R. Kusko, S.-A. Sansone, W. Tong, R. D. Wolfinger, C. E. Mason, W. Jones, J. Dopazo, C. Furlanello, L. Waldron, B. Wang, C. McIntosh, A. Goldenberg, A. Kundaje, C. S. Greene, T. Broderick, M. M. Hoffman, J. T. Leek, K. Korthauer, W. Huber, A. Brazma, J. Pineau, R. Tibshirani, T. Hastie, J. P. A. Ioannidis, J. Quackenbush, and H. J. W. L. Aerts. (2020). "Transparency and Reproducibility in Artificial Intelligence". *Nature*. 586(7829): E14–E16.

Halvari, T., J. K. Nurminen, and T. Mikkonen. (2020). "Testing the Robustness of AutoML Systems". *arXiv preprint arXiv:2005.02649*.

Han, T., F. N. B. Gois, R. Oliveira, L. R. Prates, and M. M. d. A. Porto. (2021). "Modeling the Progression of COVID-19 Deaths Using Kalman Filter and AutoML". *Soft Computing*. Jan. DOI: 10.1007/s00500-020-05503-5.

Hansing, T., M. M. Krell, and F. Kirchner. (2016). "hyperSPACE: Automated Optimization of Complex Processing Pipelines for pySPACE". In: *NIPS Workshop on Bayesian Optimization. NIPS Workshop on Bayesian Optimization (BayesOPT2016), December 5-10, Barcelona, Spain*.

Hanussek, M., M. Blohm, and M. Kintz. (2020). “Can AutoML Outperform Humans? An Evaluation on Popular OpenML Datasets Using AutoML Benchmark”. In: *2020 2nd International Conference on Artificial Intelligence, Robotics and Control*. 29–32.

Hardt, M., E. Price, E. Price, and N. Srebro. (2016). “Equality of Opportunity in Supervised Learning”. *Advances in Neural Information Processing Systems*. 29.

Hariharan, A., A. Gupta, and T. Pal. (2020). “Camlpad: Cybersecurity Autonomous Machine Learning Platform for Anomaly Detection”. In: *Future of Information and Communication Conference*. Springer. 705–720.

Harvard Intelligent Probabilistic Systems Group. (2021). “Spearmint”. URL: https://github.com/HIPS/Spearmint.

Hatmaker, T. (2021). *Reddit Is Buying Machine Learning Platform Spell*. TechCrunch. URL: https://social.techcrunch.com/2022/06/16/reddit-spell-machine-learning/.

Hauskrecht, M., I. Batal, M. Valko, S. Visweswaran, G. F. Cooper, and G. Clermont. (2013). “Outlier Detection for Patient Monitoring and Alerting”. *Journal of biomedical informatics*. 46(1): 47–55.

“Hazy”. (2021). URL: https://hazy.com.

He, C. (2020). “FedNas”. URL: https://github.com/chaoyanghe/FedNAS.

He, C., M. Annavaram, and S. Avestimehr. (2020). “FedNAS: Federated Deep Learning via Neural Architecture Search”. *arXiv preprint arXiv:2004.08546*.

He, X., K. Zhao, and X. Chu. (2021). “AutoML: A Survey of the State-of-the-Art”. *Knowledge-Based Systems*. 212(Jan.): 106622. DOI: 10.1016/j.knosys.2020.106622.

He, Y. (2021). “carefree-learn”. URL: https://github.com/carefree0910/carefree-learn.

He, Y. (2022). “carefree-learn-deploy”. URL: https://github.com/carefree0910/carefree-learn-deploy.

Hecht, C., J. Figgener, and D. U. Sauer. (2021). “Predicting Electric Vehicle Charging Station Availability Using Ensemble Machine Learning”. *Energies*. 14(23): 7834.

Herlocker, J. L., J. A. Konstan, L. G. Terveen, and J. T. Riedl. (2004). "Evaluating Collaborative Filtering Recommender Systems". *ACM Transactions on Information Systems*. 22(1): 5–53. DOI: 10.1145/963770.963772.

Hicks, S. C. and R. A. Irizarry. (2018). "A Guide to Teaching Data Science". *The American Statistician*. 72(4): 382–391.

Hodgdon, T., R. E. Thornhill, N. D. James, P. E. Beaulé, A. D. Speirs, and K. S. Rakhra. (2020). "CT Texture Analysis of Acetabular Subchondral Bone Can Discriminate between Normal and Cam-Positive Hips". *European Radiology*. 30(8): 4695–4704. DOI: 10.1007/s00330-020-06781-1.

Hoffman, M., B. Shahriari, and N. Freitas. (2014). "On Correlation and Budget Constraints in Model-Based Bandit Optimization with Application to Automatic Machine Learning". In: *Artificial Intelligence and Statistics*. 365–374.

Homola, D. (2021). "Boruta". URL: https://github.com/scikit-learn-contrib/boruta%5C_py.

Hong, M. S., Y.-H. Lee, J.-M. Kong, O.-J. Kwon, C.-W. Jung, J. Yang, M.-S. Kim, H.-W. Han, S.-M. Nam, and K. O. T. R. S. Group. (2022). "Personalized Prediction of Kidney Function Decline and Network Analysis of the Risk Factors after Kidney Transplantation Using Nationwide Cohort Data". *Journal of Clinical Medicine*. 11(5): 1259. DOI: 10.3390/jcm11051259.

Howard, C. and A. Rowsell-Jones. (2019). "2019 CIO Survey: CIOs Have Awoken to the Importance of AI". *Tech. rep.* Gartner. URL: https://www.gartner.com/document/3897266.

Howard, D., M. M. Maslej, J. Lee, J. Ritchie, G. Woollard, and L. French. (2020). "Transfer Learning for Risk Classification of Social Media Posts: Model Evaluation Study". *Journal of medical Internet research*. 22(5): e15371.

Hu, Y.-J. and S.-W. Huang. (2017). "Challenges of Automated Machine Learning on Causal Impact Analytics for Policy Evaluation". In: *2017 2nd International Conference on Telecommunication and Networks (TEL-NET)*. IEEE. DOI: 10.1109/tel-net.2017.8343571.

Huang, X., Z. Yu, X. Wei, J. Shi, Y. Wang, Z. Wang, J. Chen, S. Bu, L. Li, F. Gao, J. Zhang, and A. Xu. (2021). "Prediction of Vancomycin Dose on High-Dimensional Data Using Machine Learning Techniques". *Expert Review of Clinical Pharmacology*. 14(6): 761–771. DOI: 10.1080/17512433.2021.1911642.

Hutter, F., H. H. Hoos, and K. Leyton-Brown. (2011). "Sequential Model-Based Optimization for General Algorithm Configuration". In: *Learning and Intelligent Optimization*. Ed. by C. A. C. Coello. *Lecture Notes in Computer Science*. Berlin, Heidelberg: Springer. 507–523. DOI: 10.1007/978-3-642-25566-3_40.

Hutter, F., L. Kotthoff, and J. Vanschoren, eds. (2019). *Automated Machine Learning : Methods, Systems, Challenges*. Springer Nature. DOI: 10.1007/978-3-030-05318-5.

IBM. (2021). "Global AI Adoption Index 2021". *Tech. rep.* URL: https://filecache.mediaroom.com/mr5mr_ibmnews/190846/IBM%27s%20Global%20AI%20Adoption%20Index%202021_Executive-Summary.pdf.

"Iguazio". (2021). URL: https://www.iguazio.com/.

Imbus, J. R., R. W. Randle, S. C. Pitt, R. S. Sippel, and D. F. Schneider. (2017). "Machine Learning to Identify Multigland Disease in Primary Hyperparathyroidism". *Journal of Surgical Research*. 219(Nov.): 173–179. DOI: 10.1016/j.jss.2017.05.117.

Ingolfsson, T. M., A. Cossettini, S. Benatti, and L. Benini. (2022). "Energy-Efficient Tree-Based EEG Artifact Detection". *arXiv preprint arXiv:2204.09577*.

Isingizwe, D. F., M. Wang, W. Liu, D. Wang, T. Wu, and J. Li. (2021). "Analyzing Learning-based Encrypted Malware Traffic Classification with AutoML". In: *2021 IEEE 21st International Conference on Communication Technology (ICCT)*. 313–322. DOI: 10.1109/ICCT52962.2021.9658106.

Jain, A. K., M. Dhada, A. K. Parlikad, and B. K. Lad. (2020). "Product Quality Driven Auto-Prognostics: Low-Cost Digital Solution for SMEs". *IFAC-PapersOnLine*. 4th IFAC Workshop on Advanced Maintenance Engineering, Services and Technologies - AMEST 2020 53(3): 78–83. DOI: 10.1016/j.ifacol.2020.11.012.

Jain, S., T. Khanam, A. J. Abedi, and A. A. Khan. (2022). "Efficient Machine Learning for Malnutrition Prediction among Under-Five Children in India". In: *2022 IEEE Delhi Section Conference (DELCON)*. 1–10. DOI: 10.1109/DELCON54057.2022.9753080.

James, G., D. Witten, T. Hastie, and R. Tibshirani. (2013). *An Introduction to Statistical Learning*. Vol. 112. Springer.

Jiménez, J. (2021). "pyGPGO". URL: https://github.com/josejimenezluna/pyGPGO.

Jobin, A., M. Ienca, and E. Vayena. (2019). "The Global Landscape of AI Ethics Guidelines". *Nature Machine Intelligence*. 1(9): 389–399. DOI: 10.1038/s42256-019-0088-2.

Jung, G. and S.-Y. Choi. (2021). "Forecasting Foreign Exchange Volatility Using Deep Learning Autoencoder-LSTM Techniques". *Complexity*. 2021(Mar.): e6647534. DOI: 10.1155/2021/6647534.

Kabathova, J. and M. Drlik. (2021). "Towards Predicting Student's Dropout in University Courses Using Different Machine Learning Techniques". *Applied Sciences*. 11(7): 3130. DOI: 10.3390/app11073130.

Kadlec, P. and B. Gabrys. (2009a). "Architecture for Development of Adaptive On-Line Prediction Models". *Memetic Computing*. 1(4): 241–269. DOI: 10.1007/s12293-009-0017-8.

Kadlec, P. and B. Gabrys. (2009b). "Evolving On-Line Prediction Model Dealing with Industrial Data Sets". In: *2009 IEEE Workshop on Evolving and Self-Developing Intelligent Systems*. IEEE. DOI: 10.1109/esdis.2009.4938995.

Kadlec, P. and B. Gabrys. (2009c). "Soft Sensor Based on Adaptive Local Learning". In: *Advances in Neuro-Information Processing*. Springer Berlin Heidelberg. 1172–1179. DOI: 10.1007/978-3-642-02490-0_142.

Kadlec, P. and B. Gabrys. (2009d). "Soft Sensors: Where Are We and What Are the Current and Future Challenges?" *IFAC Proceedings Volumes*. 42(19): 572–577. DOI: 10.3182/20090921-3-tr-3005.00098.

Kadlec, P. and B. Gabrys. (2010a). "Adaptive On-Line Prediction Soft Sensing without Historical Data". In: *The 2010 International Joint Conference on Neural Networks (IJCNN)*. IEEE. DOI: 10.1109/ijcnn.2010.5596965.

Kadlec, P. and B. Gabrys. (2010b). "Local Learning-Based Adaptive Soft Sensor for Catalyst Activation Prediction". *AIChE Journal*. 57(5): 1288–1301. DOI: 10.1002/aic.12346.

Kadlec, P., R. Grbić, and B. Gabrys. (2011). "Review of Adaptation Mechanisms for Data-Driven Soft Sensors". *Computers & Chemical Engineering*. 35(1): 1–24. DOI: 10.1016/j.compchemeng.2010.07.034.

Kaggle. (2018). "2018 Kaggle Machine Learning & Data Science Survey". *Tech. rep.* URL: https://www.kaggle.com/kaggle/kaggle-survey-2018.

Kaggle. (2021). "2020 Kaggle Machine Learning & Data Science Survey". *Tech. rep.* URL: https://www.kaggle.com/c/kaggle-survey-2020/.

Kamiran, F. and T. Calders. (2012). "Data Preprocessing Techniques for Classification without Discrimination". *Knowledge and Information Systems*. 33(1): 1–33. DOI: 10.1007/s10115-011-0463-8.

Kamiran, F., A. Karim, and X. Zhang. (2012). "Decision Theory for Discrimination-Aware Classification". In: *2012 IEEE 12th International Conference on Data Mining*. 924–929. DOI: 10.1109/ICDM.2012.45.

Kamishima, T., S. Akaho, H. Asoh, and J. Sakuma. (2012). "Fairness-Aware Classifier with Prejudice Remover Regularizer". In: *Machine Learning and Knowledge Discovery in Databases*. Springer, Berlin, Heidelberg. 35–50. DOI: 10.1007/978-3-642-33486-3_3.

Kanawaday, A. and A. Sane. (2017). "Machine Learning for Predictive Maintenance of Industrial Machines Using IoT Sensor Data". In: *2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS)*. IEEE. 87–90.

Kandasamy, K., K. R. Vysyaraju, W. Neiswanger, B. Paria, C. R. Collins, J. Schneider, B. Poczos, and E. P. Xing. (2020). "Tuning Hyperparameters without Grad Students: Scalable and Robust Bayesian Optimisation with Dragonfly". *Journal of Machine Learning Research*. 21(81): 1–27.

Kandasamy, K., K. R. Vysyaraju, W. Neiswanger, B. Paria, C. R. Collins, J. Schneider, B. Poczos, and E. P. Xing. (2021). "Dragonfly". URL: https://github.com/dragonfly/dragonfly.

Kanter, J. M., O. Gillespie, and K. Veeramachaneni. (2016). "Label, Segment, Featurize: A Cross Domain Framework for Prediction Engineering". In: *2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*. IEEE. 430–439.

Kanter, J. M. and K. Veeramachaneni. (2015). "Deep Feature Synthesis: Towards Automating Data Science Endeavors". In: *2015 IEEE International Conference on Data Science and Advanced Analytics, DSAA 2015, Paris, France, October 19-21, 2015*. IEEE. 1–10.

Katz, G. (2021). "ExploreKit". URL: https://github.com/giladkatz/ExploreKit.

Katz, G., E. C. R. Shin, and D. Song. (2016). "Explorekit: Automatic Feature Generation and Selection". In: *2016 IEEE 16th International Conference on Data Mining (ICDM)*. IEEE. 979–984.

KDnuggets. (2021). URL: https://www.kdnuggets.com/.

Kearns, M., S. Neel, A. Roth, and Z. S. Wu. (2018). "Preventing Fairness Gerrymandering: Auditing and Learning for Subgroup Fairness". In: *International Conference on Machine Learning*. PMLR. 2564–2572.

Kedziora, D. J., K. Musial, and B. Gabrys. (2020). "AutonoML: Towards an Integrated Framework for Autonomous Machine Learning". *arXiv preprint arXiv:2012.12600*.

Kedziora, D. J., T.-D. Nguyen, K. Musial, and B. Gabrys. (2022). "On Taking Advantage of Opportunistic Meta-knowledge to Reduce Configuration Spaces for Automated Machine Learning". *arXiv preprint arXiv:2208.04376*.

Keras. (2021). "AutoKeras". URL: https://github.com/keras-team/autokeras.

Khayyat, Z., I. F. Ilyas, A. Jindal, S. Madden, M. Ouzzani, P. Papotti, J.-A. Quiané-Ruiz, N. Tang, and S. Yin. (2015). "Bigdansing: A System for Big Data Cleansing". In: *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data*. 1215–1230.

Khosravi, K., R. Barzegar, A. Golkarian, G. Busico, E. Cuoco, M. Mastrocicco, N. Colombani, D. Tedesco, M. M. Ntona, and N. Kazakis. (2021). "Predictive Modeling of Selected Trace Elements in Groundwater Using Hybrid Algorithms of Iterative Classifier Optimizer". *Journal of Contaminant Hydrology*. 242(Oct.): 103849. DOI: 10.1016/j.jconhyd.2021.103849.

Khuat, T. T., D. J. Kedziora, K. Musial, and B. Gabrys. (2021). "The Roles and Modes of Human Interactions with Automated Machine Learning Systems". *arXiv preprint arXiv:2205.04139*.

Kiala, Z., J. Odindi, and O. Mutanga. (2022). "Determining the Capability of the Tree-Based Pipeline Optimization Tool (TPOT) in Mapping Parthenium Weed Using Multi-Date Sentinel-2 Image Data". *Remote Sensing*. 14(7): 1687. DOI: 10.3390/rs14071687.

Klein, A., S. Falkner, S. Bartels, P. Hennig, and F. Hutter. (2017). "Fast Bayesian Optimization of Machine Learning Hyperparameters on Large Datasets". In: *Artificial Intelligence and Statistics*. PMLR. 528–536.

"KNIME". (2021). URL: https://www.knime.com.

Knudde, N., J. van der Herten, T. Dhaene, and I. Couckuyt. (2017). "GPflowOpt: A Bayesian Optimization Library Using tensorFlow". *arXiv preprint arXiv:1711.03845*.

Knudde, N., J. van der Herten, T. Dhaene, and I. Couckuyt. (2021). "GPflowOpt". URL: https://github.com/GPflow/GPflowOpt.

Kocbek, S. and B. Gabrys. (2019). "Automated Machine Learning Techniques in Prognostics of Railway Track Defects". In: *2019 International Conference on Data Mining Workshops (ICDMW)*. 777–784. DOI: 10.1109/ICDMW.2019.00115.

Kocbek, S., P. Kocbek, T. Zupanic, G. Štiglic, and B. Gabrys. (2019). "Using (Automated) Machine Learning and Drug Prescription Records to Predict Mortality and Polypharmacy in Older Type 2 Diabetes Mellitus Patients". In: *ICONIP*. DOI: 10.1007/978-3-030-36808-1_68.

Kojima, T., K. Oishi, N. Aoki, Y. Matsubara, T. Uete, Y. Fukushima, G. Inoue, S. Sato, T. Shiraishi, H. Hirooka, and T. Masuda. (2022). "Estimation of Beef Cow Body Condition Score: A Machine Learning Approach Using Three-Dimensional Image Data and a Simple Approach with Heart Girth Measurements". *Livestock Science*. 256(Feb.): 104816. DOI: 10.1016/j.livsci.2021.104816.

Komer B., B. J. and E. C. (2021). "Hyperopt-sklearn". URL: https://github.com/hyperopt/hyperopt-sklearn.

"Kortical". (2021). URL: https://kortical.com.

Kostopoulos, G., T. Panagiotakopoulos, S. Kotsiantis, C. Pierrakeas, and A. Kameas. (2021). "Interpretable Models for Early Prediction of Certification in MOOCs: A Case Study on a MOOC for Smart City Professionals". *IEEE Access*. 9: 165881–165891.

Kotsiantis, S. B., I. D. Zaharakis, and P. E. Pintelas. (2006). "Machine Learning: A Review of Classification and Combining Techniques". *Artificial Intelligence Review*. 26(3): 159–190. DOI: 10.1007/s10462-007-9052-3.

Kraska, T. (2018). "Northstar: An Interactive Data Science System". *Proceedings of the VLDB Endowment*. 11(12): 2150–2164. DOI: 10.14778/3229863.3240493.

Kross, S., R. D. Peng, B. S. Caffo, I. Gooding, and J. T. Leek. (2020). "The Democratization of Data Science Education". *The American Statistician*. 74(1): 1–7.

Kulbach, C., P. Philipp, and S. Thoma. (2020). "Personalized Automated Machine Learning". *ECAI 2020*: 1246–1253. DOI: 10.3233/FAIA200225.

Kursa, M. B. and W. R. Rudnicki. (2010). "Feature Selection with the Boruta Package". *Journal of Statistical Software*. 36(11): 1–13.

Laboratório de Inteligência Computacional (Computational Intelligence Lab) at Universidade Federal de Minas Gerais. (2021). "Recipe". URL: https://github.com/laic-ufmg/Recipe.

Lai, K.-H., D. Zha, G. Wang, J. Xu, Y. Zhao, D. Kumar, Y. Chen, P. Zumkhawaka, M. Wan, D. Martinez, and X. Hu. (2021). "Tods: An Automated Time Series Outlier Detection System". In: *Proceedings of the Aaai Conference on Artificial Intelligence*. Vol. 35. 16060–16062.

Lakkaraju, H., J. Adebayo, and S. Singh. (2020). "Explaining Machine Learning Predictions: State-of-the-art, Challenges, and Opportunities". *NeurIPS Tutorial*.

Lam, H. T., J.-M. Thiebaut, M. Sinn, B. Chen, T. Mai, and O. Alkan. (2017). "One Button Machine for Automating Feature Engineering in Relational Databases". *arXiv preprint arXiv:1706.00327*.

Lamare, J.-B., T. Olatunji, and L. Yao. (2020). "On the Diminishing Return of Labeling Clinical Reports". *arXiv preprint arXiv:2010.14587*.

Larocque-Villiers, J., P. Dumond, and D. Knox. (2021). "Automating Predictive Maintenance Using State-Based Transfer Learning and Ensemble Methods". In: *2021 IEEE International Symposium on Robotic and Sensors Environments (ROSE)*. 1–7. DOI: 10.1109/ROSE52750.2021.9611768.

Larriva-Novo, X., C. Sánchez-Zas, V. A. Villagrá, M. Vega-Barbas, and D. Rivera. (2020). "An Approach for the Application of a Dynamic Multi-Class Classifier for Network Intrusion Detection Systems". *Electronics*. 9(11): 1759. DOI: 10.3390/electronics9111759.

Larsen, E., K. MacVittie, and J. Lilly. (2021a). "A Survey of Machine Learning Algorithms for Detecting Malware in IoT Firmware". *arXiv preprint arXiv:2111.02388*.

Larsen, E., D. Noever, and K. MacVittie. (2021b). "A Survey of Machine Learning Algorithms for Detecting Ransomware Encryption Activity". *arXiv preprint arXiv:2110.07636*.

Lazic, I., F. Hinterwimmer, S. Langer, F. Pohlig, C. Suren, F. Seidl, D. Rückert, R. Burgkart, and R. von Eisenhart-Rothe. (2022). "Prediction of Complications and Surgery Duration in Primary Total Hip Arthroplasty Using Machine Learning: The Necessity of Modified Algorithms and Specific Data". *Journal of Clinical Medicine*. 11(8): 2147. DOI: 10.3390/jcm11082147.

Le, T. T., W. Fu, and J. H. Moore. (2020). "Scaling Tree-Based Automated Machine Learning to Biomedical Big Data with a Feature Set Selector". *Bioinformatics*. 36(1): 250–256.

Leclercq, M., B. Vittrant, M. L. Martin-Magniette, M. P. Scott Boyer, O. Perin, A. Bergeron, Y. Fradet, and A. Droit. (2019). "Large-Scale Automatic Feature Selection for Biomarker Discovery in High-Dimensional OMICs Data". *Frontiers in Genetics*. 10.

Lee, D. J. L., S. Macke, D. Xin, A. Lee, S. Huang, and A. G. Parameswaran. (2019). "A Human-in-the-Loop Perspective on AutoML: Milestones and the Road Ahead." *IEEE Data Eng. Bull.* 42(2): 59–70.

Lee, M. S. A. and J. Singh. (2021). "The Landscape and Gaps in Open Source Fairness Toolkits". In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–13.

Lee, M.-C., J.-C. Lin, and E. G. Gran. (2021). "DistTune: Distributed Fine-Grained Adaptive Traffic Speed Prediction for Growing Transportation Networks". *Transportation Research Record*. 2675(10): 211–227. DOI: 10.1177/03611981211011170.

Lekov, A. (2021). "AutoML Alex". URL: https://github.com/Alex-Lekov/AutoML%5C_Alex.

Lemke, C., M. Budka, and B. Gabrys. (2015). "Metalearning: A Survey of Trends and Technologies". *Artificial Intelligence Review*. 44(1): 117–130. DOI: 10.1007/s10462-013-9406-y.

Lemke, C. and B. Gabrys. (2010a). "Meta-Learning for Time Series Forecasting and Forecast Combination". *Neurocomputing*. 73(10-12): 2006–2016. DOI: 10.1016/j.neucom.2009.09.020.

Lemke, C. and B. Gabrys. (2010b). "Meta-Learning for Time Series Forecasting in the NN GC1 Competition". In: *International Conference on Fuzzy Systems*. IEEE. DOI: 10.1109/fuzzy.2010.5584001.

Lemke, C., S. Riedel, and B. Gabrys. (2012). "Evolving Forecast Combination Structures for Airline Revenue Management". *Journal of Revenue and Pricing Management*. 12(3): 221–234. DOI: 10.1057/rpm.2012.30.

Lepri, B., N. Oliver, E. Letouzé, A. Pentland, and P. Vinck. (2018). "Fair, Transparent, and Accountable Algorithmic Decision-making Processes". *Philosophy & Technology*. 31(4): 611–627. DOI: 10.1007/s13347-017-0279-x.

LF AI Foundation. (2021a). "AIExplainability360". URL: https://github.com/Trusted-AI/AIX360.

LF AI Foundation. (2021b). "AIF360". URL: https://github.com/Trusted-AI/AIF360.

Li, K.-Y., N. G. Burnside, R. S. de Lima, M. V. Peciña, K. Sepp, V. H. Cabral Pinheiro, B. R. C. A. de Lima, M.-D. Yang, A. Vain, and K. Sepp. (2021a). "An Automated Machine Learning Framework in Unmanned Aircraft Systems: New Insights into Agricultural Management Practices Recognition Approaches". *Remote Sensing*. 13(16): 3190. DOI: 10.3390/rs13163190.

Li, K.-Y., R. Sampaio de Lima, N. G. Burnside, E. Vahtmäe, T. Kutser, K. Sepp, V. H. Cabral Pinheiro, M.-D. Yang, A. Vain, and K. Sepp. (2022). "Toward Automated Machine Learning-Based Hyperspectral Image Analysis in Crop Yield and Biomass Estimation". *Remote Sensing*. 14(5): 1114. DOI: 10.3390/rs14051114.

Li, R. Y. M., K. W. Chau, H. C. Y. Li, F. Zeng, B. Tang, and M. Ding. (2021b). "Remote Sensing, Heat Island Effect and Housing Price Prediction via AutoML". In: *Advances in Artificial Intelligence, Software and Systems Engineering*. Ed. by T. Ahram. *Advances in Intelligent Systems and Computing*. Cham: Springer International Publishing. 113–118. DOI: 10.1007/978-3-030-51328-3_17.

Linux Foundation AI & Data. (2021). "Ludwig". URL: https://github.com/ludwig-ai/ludwig.

Liu, G., D. Lu, and J. Lu. (2021a). "Pharm-AutoML: An Open-Source, End-to-End Automated Machine Learning Package for Clinical Outcome Prediction". *CPT: pharmacometrics & systems pharmacology*. 10(5): 478–488.

Liu, H. (2021). "Darts". URL: https://github.com/quark0/darts.

Liu, J., S. Tripathi, U. Kurup, and M. Shah. (2019). "Auptimizer - an Extensible, Open-Source Framework for Hyperparameter Tuning". In: *2019 IEEE International Conference on Big Data (Big Data)*. 339–348. DOI: 10.1109/BigData47090.2019.9006330.

Liu, J., S. Tripathi, U. Kurup, and M. Shah. (2021b). "Auptimizer". URL: https://github.com/LGE-ARC-AdvancedAI/auptimizer.

Liu, P., F. Pan, X. Zhou, S. Li, and L. Jin. (2022a). "CF-DAML: Distributed Automated Machine Learning Based on Collaborative Filtering". *Applied Intelligence*: 1–25.

Liu, P., F. Pan, X. Zhou, S. Li, P. Zeng, S. Liu, and L. Jin. (2022b). "Dsa-PAML: A Parallel Automated Machine Learning System via Dual-Stacked Autoencoder". *Neural Computing and Applications*: 1–22.

Liu, X., P. Li, and Z. Zhu. (2022c). "Bayesian Optimisation-Assisted Neural Network Training Technique for Radio Localisation". *arXiv preprint arXiv:2203.04032*.

Lorraine, J., P. Vicol, and D. Duvenaud. (2020). "Optimizing Millions of Hyperparameters by Implicit Differentiation". In: *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*. Ed. by S. Chiappa and R. Calandra. Vol. 108. *Proceedings of Machine Learning Research*. PMLR. 1540–1552. URL: https://proceedings.mlr.press/v108/lorraine20a.html.

Lu, J., A. Liu, F. Dong, F. Gu, J. Gama, and G. Zhang. (2019). "Learning under Concept Drift: A Review". *IEEE Transactions on Knowledge and Data Engineering*. 31(12): 2346–2363. DOI: 10.1109/TKDE.2018.2876857.

Lu, Y. (2019). *An End-to-End AutoML Solution for Tabular Data at KaggleDays*. URL: https://ai.googleblog.com/2019/05/an-end-to-end-automl-solution-for.html.

Lundberg, S. (2021). "Shap". URL: https://github.com/slundberg/shap.

Lundberg, S. M. and S.-I. Lee. (2017). "A Unified Approach to Interpreting Model Predictions". *Advances in neural information processing systems*. 30.

Lunden, I. (2020). *Intel Acquires SigOpt, a Specialist in Modeling Optimization, to Boost Its AI Business*. URL: https://techcrunch.com/2020/10/29/intel-acquires-sigopt-a-specialist-in-modeling-optimization-to-boost-its-ai-business/.

Luo, C., Z. Zhang, D. Qiao, X. Lai, Y. Li, and S. Wang. (2022). "Life Prediction under Charging Process of Lithium-Ion Batteries Based on AutoML". *Energies*. 15(13): 4594. DOI: 10.3390/en15134594.

Luo, G. (2016). "PredicT-ML: A Tool for Automating Machine Learning Model Building with Big Clinical Data". *Health Information Science and Systems*. 4(1). DOI: 10.1186/s13755-016-0018-1.

Luo, G., S. He, B. L. Stone, F. L. Nkoy, and M. D. Johnson. (2020). "Developing a Model to Predict Hospital Encounters for Asthma in Asthmatic Patients: Secondary Analysis". *JMIR Medical Informatics*. 8(1): e16080. DOI: 10.2196/16080.

Luo, G., B. L. Stone, M. D. Johnson, P. Tarczy-Hornoch, A. B. Wilcox, S. D. Mooney, X. Sheng, P. J. Haug, and F. L. Nkoy. (2017). "Automating Construction of Machine Learning Models with Clinical Big Data: Proposal Rationale and Methods". *JMIR Research Protocols*. 6(8): e175. DOI: 10.2196/resprot.7757.

Luo, G., B. L. Stone, X. Sheng, S. He, C. Koebnick, and F. L. Nkoy. (2021). "Using Computational Methods to Improve Integrated Disease Management for Asthma and Chronic Obstructive Pulmonary Disease: Protocol for a Secondary Analysis". *JMIR Research Protocols*. 10(5): e27065. DOI: 10.2196/27065.

Lv, Z., D. Wang, H. Ding, B. Zhong, and L. Xu. (2020). "Escherichia Coli DNA N-4-Methycytosine Site Prediction Accuracy Improved by Light Gradient Boosting Machine Feature Selection Technology". *IEEE Access*. 8: 14851–14859. DOI: 10.1109/ACCESS.2020.2966576.

Maass, K., A. Aravkin, and M. Kim. (2022). "A Hyperparameter-Tuning Approach to Automated Inverse Planning". *Medical Physics*. 49(5): 3405–3415. DOI: 10.1002/mp.15557.

Madrid, J. G., H. J. Escalante, E. F. Morales, W.-W. Tu, Y. Yu, L. Sun-Hosoya, I. Guyon, and M. Sebag. (2019). "Towards AutoML in the Presence of Drift: First Results". *arXiv preprint arXiv:1907.10772*.

Makarious, M. B., H. L. Leonard, D. Vitale, H. Iwaki, D. Saffo, L. Sargent, A. Dadu, E. S. Castaño, J. F. Carter, M. Maleknia, J. A. Botia, C. Blauwendraat, R. H. Campbell, S. H. Hashemi, A. B. Singleton, M. A. Nalls, and F. Faghri. (2021). "GenoML: Automated Machine Learning for Genomics". *arXiv preprint arXiv:2103.03221*.

Makridakis, S. (2017). "The Forthcoming Artificial Intelligence (AI) Revolution: Its Impact on Society and Firms". *Futures*. 90(June): 46–60. DOI: 10.1016/j.futures.2017.03.006.

Manduchi, E., J. D. Romano, and J. H. Moore. (2021). "The Promise of Automated Machine Learning for the Genetic Analysis of Complex Traits". *Human Genetics*. Oct. DOI: 10.1007/s00439-021-02393-x.

Mangalath Ravindran, S., S. K. Moorakkal Bhaskaran, S. K. Ambat, K. Balakrishnan, and M. Manguttathil Gopalakrishnan. (2022). "An Automated Machine Learning Methodology for the Improved Prediction of Reference Evapotranspiration Using Minimal Input Parameters". *Hydrological Processes*. 36(5): e14571. DOI: 10.1002/hyp.14571.

Mantel, N. and W. Haenszel. (1959). "Statistical Aspects of the Analysis of Data From Retrospective Studies of Disease". *JNCI: Journal of the National Cancer Institute*. 22(4): 719–748. DOI: 10.1093/jnci/22.4.719.

Martinez-Cantin, R. (2014). "BayesOpt: A Bayesian Optimization Library for Nonlinear Optimization, Experimental Design and Bandits." *Journal of Machine Learning Research*. 15(1): 3735–3739.

Martinez-Cantin, R. (2021). "Bayesopt". URL: https://github.com/rmcantin/bayesopt.

Masrom, S., T. Mohd, N. S. Jamil, A. S. A. Rahman, and N. Baharun. (2019). "Automated Machine Learning Based on Genetic Programming: A Case Study on a Real House Pricing Dataset". In: *2019 1st International Conference on Artificial Intelligence and Data Sciences (AiDAS)*. 48–52. DOI: 10.1109/AiDAS47888.2019.8970916.

Mathew, S. K. and Y. Zhang. (2020). "Acoustic-Based Engine Fault Diagnosis Using WPT, PCA and Bayesian Optimization". *Applied Sciences*. 10(19): 6890. DOI: 10.3390/app10196890.

Matos, C. E. F., J. C. Souza, J. O. B. Diniz, G. B. Junior, A. C. de Paiva, J. D. S. de Almeida, S. V. da Rocha, and A. C. Silva. (2019). "Diagnosis of Breast Tissue in Mammography Images Based Local Feature Descriptors". *Multimedia Tools and Applications*. 78(10): 12961–12986. DOI: 10.1007/s11042-018-6390-x.

Mazaev, G., A. P. Ompusunggu, G. Tod, G. Crevecoeur, and S. Van Hoecke. (2020). "Data-Driven Prognostics of Alternating Current Solenoid Valves". In: *2020 Prognostics and Health Management Conference (PHM-Besançon)*. 109–115. DOI: 10.1109/PHM-Besancon49106.2020.00024.

Mazza-Anthony, C. (2021). *A Five-Step Guide for Conducting Exploratory Data Analysis*. Shopify. URL: https://shopify.engineering/conducting-exploratory-data-analysis.

McGibbon, R. T., C. X. Hernández, M. P. Harrigan, S. Kearnes, M. M. Sultan, S. Jastrzebski, B. E. Husic, and V. S. Pande. (2021). "Osprey". URL: https://github.com/msmbuilder/osprey.

McKinsey. (2020). "Global Survey: The State of AI in 2020 | McKinsey". *Tech. rep.* URL: https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/global-survey-the-state-of-ai-in-2020.

Mehrabi, N., F. Morstatter, N. Saxena, K. Lerman, and A. Galstyan. (2021). "A Survey on Bias and Fairness in Machine Learning". *ACM Computing Surveys (CSUR)*. 54(6): 1–35.

Mei, N., M. D. Grossberg, K. Ng, K. T. Navarro, and T. M. Ellmore. (2017). "Identifying Sleep Spindles with Multichannel EEG and Classification Optimization". *Computers in biology and medicine*. 89: 441–453.

Meiners, M., M. Kuhn, and J. Franke. (2021). "Manufacturing Process Curve Monitoring with Deep Learning". *Manufacturing Letters*. 30(Oct.): 15–18. DOI: 10.1016/j.mfglet.2021.09.006.

Mendes, H. A. (2021). "On AutoMLs for Short-Term Solar Radiation Forecasting in Brazilian Northeast". In: *2021 International Conference on Engineering and Emerging Technologies (ICEET)*. 1–6. DOI: 10.1109/ICEET53442.2021.9659788.

Mendikute, J., J. Plazaola, M. Baskaran, E. Zugasti, L. Aretxabaleta, and J. Aurrekoetxea. (2021). "Impregnation Quality Diagnosis in Resin Transfer Moulding by Machine Learning". *Composites Part B: Engineering*. 221(Sept.): 108973. DOI: 10.1016/j.compositesb.2021.108973.

Merritt, R. (2021). *What Is MLOps?* NVIDIA. URL: https://blogs.nvidia.com/blog/2020/09/03/what-is-mlops/.

Michael, E., H. Ma, H. Li, and S. Qi. (2022). "An Optimized Framework for Breast Cancer Classification Using Machine Learning". *BioMed Research International*. 2022(Feb.): e8482022. DOI: 10.1155/2022/8482022.

Microsoft. (2020). "FLAML". URL: https://github.com/microsoft/FLAML.

"Microsoft Azure Automl". (2021). URL: https://azure.microsoft.com/en-us/services/machine-learning/automatedml.

Mikalef, P., M. N. Giannakos, I. O. Pappas, and J. Krogstie. (2018). "The Human Side of Big Data: Understanding the Skills of the Data Scientist in Education and Industry". In: *2018 IEEE Global Engineering Education Conference (EDUCON)*. 503–512. DOI: 10.1109/EDUCON.2018.8363273.

Mills, S., S. Duranton, M. Santinelli, G. Hua, E. Baltassis, S. Thiel, and O. Muehlstein. (2021). "Are You Overestimating Your Responsible AI Maturity?" *Tech. rep.* BCG. URL: https://www.bcg.com/publications/2021/the-four-stages-of-responsible-ai-maturity.

Minaee, S., Y. Y. Boykov, F. Porikli, A. J. Plaza, N. Kehtarnavaz, and D. Terzopoulos. (2021). "Image Segmentation Using Deep Learning: A Survey". *IEEE Transactions on Pattern Analysis and Machine Intelligence*: 1–1. DOI: 10.1109/TPAMI.2021.3059968.

MindsDB Inc. (2022). "Lightwood". URL: https://github.com/mindsdb/lightwood.

MIT. (2021a). "Data To AI Lab at The Massachusetts Institute of Technology. BTB". URL: https://github.com/MLBazaar/BTB.

MIT. (2021b). "Data To AI Lab at The Massachusetts Institute of Technology. Cardea". URL: https://github.com/MLBazaar/Cardea.

MIT - The Human Data Interaction Project. (2021). "ATM". URL: https://github.com/HDI-Project/ATM.

Mitchell, S., E. Potash, S. Barocas, A. D'Amour, and K. Lum. (2018). "Prediction-Based Decisions and Fairness: A Catalogue of Choices, Assumptions, and Definitions". *arXiv preprint arXiv:1811.07867*.

MLJAR. (2021). "Mljar". URL: https://github.com/mljar/mljar-supervised.

Mo, Y., D. Zhao, J. Du, M. Syal, A. Aziz, and H. Li. (2020). "Automated Staff Assignment for Building Maintenance Using Natural Language Processing". *Automation in Construction*. 113(May): 103150. DOI: 10.1016/j.autcon.2020.103150.

Mohapatra, A., S. Pattnaik, B. K. Pattanayak, S. Patnaik, and S. R. Laha. (2022). "Software Quality Prediction Using Machine Learning". In: *Advances in Data Science and Management*. Ed. by S. Borah, S. K. Mishra, B. K. Mishra, V. E. Balas, and Z. Polkowski. *Lecture Notes on Data Engineering and Communications Technologies*. Singapore: Springer Nature. 137–146. DOI: 10.1007/978-981-16-5685-9_14.

Morley, J., L. Floridi, L. Kinsey, and A. Elhalal. (2020). "From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices". *Science and Engineering Ethics*. 26(4): 2141–2168. DOI: 10.1007/s11948-019-00165-5.

"MOSTLY AI". (2021). URL: https://mostly.ai.

Muhammod, R., S. Ahmed, D. Md Farid, S. Shatabda, A. Sharma, and A. Dehzangi. (2019). "PyFeat: A Python-based Effective Feature Generation Tool for DNA, RNA and Protein Sequences". *Bioinformatics*. 35(19): 3831–3833. DOI: 10.1093/bioinformatics/btz165.

Mulpuru, V. and N. Mishra. (2021). "In Silico Prediction of Fraction Unbound in Human Plasma from Chemical Fingerprint Using Automated Machine Learning". *ACS omega*. 6(10): 6791–6797.

Muniz Do Nascimento, W. and L. Gomes-Jr. (2022). "Enabling Low-Cost Automatic Water Leakage Detection: A Semi-Supervised, autoML-based Approach". *Urban Water Journal*. 0(0): 1–11. DOI: 10.1080/1573062X.2022.2056710.

"MyDataModels". (2021). URL: https://www.mydatamodels.com.

Nakano, R. (2021). "Xcessiv". URL: https://github.com/reiinakano/xcessiv.

National Disability Insurance Scheme (NDIS). (2021). "Personalised Budgets - Proposal for a New NDIS Budget Model Technical Information Paper". *Tech. rep.* URL: https://www.ndis.gov.au/media/3124/download.

Negrinho, R. (2021). "DeepArchitect". URL: https://github.com/negrinho/deep%5C_architect.

"Neptune.Ai". (2021). URL: https://neptune.ai/.

"neuralstudio.Ai". (2021). URL: https://neuralstudio.ai.

Ngai, E. W. T., Y. Hu, Y. H. Wong, Y. Chen, and X. Sun. (2011). "The Application of Data Mining Techniques in Financial Fraud Detection: A Classification Framework and an Academic Review of Literature". *Decision Support Systems*. On Quantitative Methods for Detection of Financial Fraud 50(3): 559–569. DOI: 10.1016/j.dss.2010.08.006.

Nguyen, T.-D., D. J. Kedziora, K. Musial, and B. Gabrys. (2021a). "Exploring Opportunistic Meta-knowledge to Reduce Search Spaces for Automated Machine Learning". In: *2021 International Joint Conference on Neural Networks (IJCNN)*. 1–8. DOI: 10.1109/IJCNN52387.2021.9533431.

Nguyen, T.-D., K. Musial, and B. Gabrys. (2021b). "AutoWeka4MCPS-AVATAR: Accelerating automated machine learning pipeline composition and optimisation". *Expert Systems with Applications*. 185: 115643. DOI: https://doi.org/10.1016/j.eswa.2021.115643.

Nguyen, T.-D., T. T. Sang Nguyen, and N.-T. Le. (2018). "Calibration of Conductivity Sensor Using Combined Algorithm Selection and Hyperparameter Optimization: A Case Study". In: *2018 International Conference on Advanced Technologies for Communications (ATC)*. 296–300. DOI: 10.1109/ATC.2018.8587559.

Nogueira, F. (2021). "Bayesian Optimization". URL: https://github.com/fmfn/BayesianOptimization.

Novillo Rangone, G., C. Pizarro, and G. Montejano. (2022). "Automation of an Educational Data Mining Model Applying Interpretable Machine Learning and Auto Machine Learning". In: *Communication and Smart Technologies*. Ed. by Á. Rocha, D. Barredo, P. C. López-López, and I. Puentes-Rivera. *Smart Innovation, Systems and Technologies*. Singapore: Springer. 22–30. DOI: 10.1007/978-981-16-5792-4_3.

"Number Theory". (2021). URL: http://numbertheory.ai.

Nyabako, T., B. M. Mvumi, T. Stathers, S. Mlambo, and M. Mubayiwa. (2020). "Predicting Prostephanus Truncatus (Horn) (Coleoptera: Bostrichidae) Populations and Associated Grain Damage in Smallholder Farmers' Maize Stores: A Machine Learning Approach". *Journal of Stored Products Research*. 87(May): 101592. DOI: 10.1016/j.jspr.2020.101592.

Office of the Australian Information Commissioner. (2018). "Guide to Data Analytics and the Australian Privacy Principles". *Tech. rep.* URL: https://www.oaic.gov.au/privacy/guidance-and-advice/guide-to-data-analytics-and-the-australian-privacy-principles.

Office of the Prime Minister and Cabinet. (2021). "Budget 2021-22 Fact Sheets - Artificial Intelligence". *Tech. rep.* URL: https://web.archive.org/web/20211027035957/https://digitaleconomy.pmc.gov.au/fact-sheets/artificial-intelligence.

Olson, R. S., R. J. Urbanowicz, P. C. Andrews, N. A. Lavender, L. C. Kidd, and J. H. Moore. (2016). "Automating Biomedical Data Science through Tree-Based Pipeline Optimization". In: *Applications of Evolutionary Computation*. Springer International Publishing. 123–137. DOI: 10.1007/978-3-319-31204-0_9.

"OptiScorer". (2021). URL: https://optiscorer.com.

Orlenko, A., J. H. Moore, P. Orzechowski, R. S. Olson, J. Cairns, P. J. Caraballo, R. M. Weinshilboum, L. Wang, and M. K. Breitenstein. (2018). "Considerations for Automated Machine Learning in Clinical Metabolic Profiling: Altered Homocysteine Plasma Concentration Associated with Metformin Exposure". In: *Pacific Symposium on Biocomputing 2018*. World Scientific. 460–471.

Ozbayoglu, A. M., M. U. Gudelek, and O. B. Sezer. (2020). "Deep Learning for Financial Applications : A Survey". *Applied Soft Computing*. 93(Aug.): 106384. DOI: 10.1016/j.asoc.2020.106384.

Padmanabhan, M., P. Yuan, G. Chada, and H. V. Nguyen. (2019). "Physician-Friendly Machine Learning: A Case Study with Cardiovascular Disease Risk Prediction". *Journal of Clinical Medicine*. 8(7): 1050. DOI: 10.3390/jcm8071050.

Pahlberg, T., M. Thurley, D. Popovic, and O. Hagman. (2018). "Crack Detection in Oak Flooring Lamellae Using Ultrasound-Excited Thermography". *Infrared Physics & Technology*. 88(Jan.): 57–69. DOI: 10.1016/j.infrared.2017.11.007.

Palacios Salinas, N. R., M. Baratchi, J. N. van Rijn, and A. Vollrath. (2021). "Automated Machine Learning for Satellite Data: Integrating Remote Sensing Pre-trained Models into AutoML Systems". In: *Machine Learning and Knowledge Discovery in Databases. Applied Data Science Track*. Ed. by Y. Dong, N. Kourtellis, B. Hammer, and J. A. Lozano. *Lecture Notes in Computer Science*. Cham: Springer International Publishing. 447–462. DOI: 10.1007/978-3-030-86517-7_28.

Parker, A. (2010). *The Spawn of Craigslist*. URL: https://thegongshow.tumblr.com/post/345941486/the-spawn-of-craigslist-like-most-vcs-that-focus.

Parry, P. (2021). "auto_ml". URL: https://github.com/ClimbsRocks/auto%5C_ml.

Pastor, K., M. Ilić, J. Kojić, M. Ačanski, and D. Vujić. (2022). "Classification of Cereal Flour by Gas Chromatography – Mass Spectrometry (GC-MS) Liposoluble Fingerprints and Automated Machine Learning". *Analytical Letters*. 55(14): 2220–2226. DOI: 10.1080/00032719.2022.2050921.

Patidar, K. (2020). "Predicting Spin-Symmetry Breaking in Organic Photovoltaic Compounds Using a Data Mining Approach". *PhD thesis*. State University of New York at Buffalo.

Payton, N. (2020). "Insights from Interviewing Dozens of Modelers". *Tech. rep.* SigOpt. URL: https://sigopt.com/blog/insights-from-interviewing-dozens-of-modelers/.

"Pecan". (2021). URL: https://www.pecan.ai.

Peleg, Y. (2021). "HungaBunga". URL: https://github.com/ypeleg/HungaBunga.

Pellegrini, M. (2021). "Accurate Prediction of Breast Cancer Survival through Coherent Voting Networks with Gene Expression Profiling". *Scientific Reports*. 11(1): 14645. DOI: 10.1038/s41598-021-94243-z.

Peng, R. (2015). "The Reproducibility Crisis in Science: A Statistical Counterattack". *Significance*. 12(3): 30–32. DOI: 10.1111/j.1740-9713.2015.00827.x.

People+AI Research (PAIR) Initiative at Google. (2021). "What-If-Tool". URL: https://github.com/PAIR-code/what-if-tool.

Piatetsky, G. (2019). “KDnuggets Software Poll 2019”. *Tech. rep.* KGnuggets. URL: https://www.kdnuggets.com/2019/05/poll-top-data-science-machine-learning-platforms.html.

Piatetsky, G. (2020). “AutoML Poll Results: If You Try It, You’ll like It More”. *Tech. rep.* KGnuggets. URL: https://www.kdnuggets.com/automl-poll-results-if-you-try-it-youll-like-it-more.html/.

Piccialli, F., V. D. Somma, F. Giampaolo, S. Cuomo, and G. Fortino. (2021). “A Survey on Deep Learning in Medicine: Why, How and When?” *Information Fusion*. 66(Feb.): 111–137. DOI: 10.1016/j.inffus.2020.09.006.

Pineda-Jaramillo, J. and Ó. Arbeláez-Arenas. (2022). “Assessing the Performance of Gradient-Boosting Models for Predicting the Travel Mode Choice Using Household Survey Data”. *Journal of Urban Planning and Development*. 148(2): 04022007.

Pleiss, G., M. Raghavan, F. Wu, J. Kleinberg, and K. Q. Weinberger. (2017). “On Fairness and Calibration”. *Advances in neural information processing systems*. 30.

Polonskaia, I. S., I. R. Aliev, and N. O. Nikitin. (2021). “Automated Evolutionary Design of CNN Classifiers for Object Recognition on Satellite Images”. *Procedia Computer Science*. 10th International Young Scientists Conference in Computational Science, YSC2021, 28 June – 2 July, 2021 193(Jan.): 210–219. DOI: 10.1016/j.procs.2021.10.021.

Pompa, C. and T. Burke. (2017). “Data Science and Analytics Skills Shortage: Equipping the APEC Workforce with the Competencies Demanded by Employers”. *APEC Human Resource Development Working Group*.

Powers, D. M. W. (2020). “Evaluation: From Precision, Recall and F-measure to ROC, Informedness, Markedness and Correlation”. *arXiv preprint arXiv:2010.16061*.

“Prevision.io”. (2021). URL: https://prevision.io.

Probst, P. (2021). “tuneRanger”. URL: https://github.com/PhilippPro/tuneRanger.

Prochnow, S. J., N. S. Raterman, M. Swenberg, L. Reddy, I. Smith, M. Romanyuk, and T. Fernandez. (2022). "A Subsurface Machine Learning Approach at Hydrocarbon Production Recovery & Resource Estimates for Unconventional Reservoir Systems: Making Subsurface Predictions from Multimensional Data Analysis". *Journal of Petroleum Science and Engineering*. 215(Aug.): 110598. DOI: 10.1016/j.petrol.2022.110598.

Purwanto, R., A. Pal, A. Blair, and S. Jha. (2021). "Man versus Machine: AutoML and Human Experts' Role in Phishing Detection". *arXiv preprint arXiv:2108.12193*.

Puspus, Xavier. (2021). "Parity-Fairness". URL: https://github.com/xmpuspus/parity-fairness.

PyCaret. (2021). "PyCaret". URL: https://github.com/pycaret/pycaret.

Pymetrics. (2021). "Audit-Ai". URL: https://github.com/pymetrics/audit-ai.

Pytorch. (2021). "Botorch". URL: https://github.com/pytorch/botorch.

Qatar Computing Research Institute. (2020). "Data_civilizer_system". URL: https://github.com/qcri/data%5C_civilizer%5C_system.

Qi, W., C. Xu, and X. Xu. (2021). "AutoGluon: A Revolutionary Framework for Landslide Hazard Analysis". *Natural Hazards Research*. 1(3): 103–108. DOI: 10.1016/j.nhres.2021.07.002.

Qlik. (2021). *Qlik Acquires Big Squid to Expand Its Industry Leading Augmented Analytics Capabilities with No-Code Automated Machine Learning*. URL: https://www.qlik.com/us/company/press-room/press-releases/qlik-acquires-big-squid-to-expand-its-industry-leading-augmented-analytics-capabilities.

Qolomany, B., I. Mohammed, A. Al-Fuqaha, M. Guizani, and J. Qadir. (2021). "Trust-Based Cloud Machine Learning Model Selection for Industrial IoT and Smart City Services". *IEEE Internet of Things Journal*. 8(4): 2943–2958. DOI: 10.1109/JIOT.2020.3022323.

"Qubole". (2021). URL: https://www.qubole.com/platform/open-data-lake-platform/machine-learning/.

Radzi, S. F. M., M. K. A. Karim, M. I. Saripan, M. A. A. Rahman, I. N. C. Isa, and M. J. Ibahim. (2021). "Hyperparameter Tuning and Pipeline Optimization via Grid Search Method and Tree-Based AutoML in Breast Cancer Prediction". *Journal of Personalized Medicine*. 11(10): 978. DOI: 10.3390/jpm11100978.

"RapidMiner". (2021). URL: https://rapidminer.com.

RapidMiner. (2019). *Automated Machine Learning*. URL: https://rapidminer.com/glossary/automated-machine-learning/.

Ray. (2021). "Tune-Sklearn". URL: https://github.com/ray-project/tune-sklearn.

Real, E., C. Liang, D. So, and Q. Le. (2020). "Automl-Zero: Evolving Machine Learning Algorithms from Scratch". In: *International Conference on Machine Learning*. PMLR. 8007–8019.

Real, E., C. Liang, D. R. So, and Q. V. Le. (2021). "Automl-Zero". URL: https://github.com/google-research/google-research/tree/master/automl%5C_zero.

Ren, P., Y. Xiao, X. Chang, P.-Y. Huang, Z. Li, X. Chen, and X. Wang. (2021). "A Comprehensive Survey of Neural Architecture Search: Challenges and Solutions". *ACM Computing Surveys (CSUR)*. 54(4): 1–34.

Rençberoğlu, E. (2021). *Fundamental Techniques of Feature Engineering for Machine Learning*. URL: https://towardsdatascience.com/feature-engineering-for-machine-learning-3a5e293a5114.

Resmini, R., L. Faria da Silva, P. R. T. Medeiros, A. S. Araujo, D. C. Muchaluat-Saade, and A. Conci. (2021). "A Hybrid Methodology for Breast Screening and Cancer Diagnosis Using Thermography". *Computers in Biology and Medicine*. 135(Aug.): 104553. DOI: 10.1016/j.compbiomed.2021.104553.

Riad, R., M. Lunven, H. Titeux, X.-N. Cao, J. Hamet Bagnou, L. Lemoine, J. Montillot, A. Sliwinski, K. Youssov, L. Cleret de Langavant, E. Dupoux, and A.-C. Bachoud-Lévi. (2022). "Predicting Clinical Scores in Huntington's Disease: A Lightweight Speech Test". *Journal of Neurology*. May. DOI: 10.1007/s00415-022-11148-1.

Ribeiro, M. T., S. Singh, and C. Guestrin. (2016). “"Why Should I Trust You?": Explaining the Predictions of Any Classifier”. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, August 13-17, 2016*. 1135–1144.

Ribeiro, M. T. C. (2021). “Lime”. URL: https://github.com/marcotcr/lime.

Ritter, C., T. Wollmann, P. Bernhard, M. Gunkel, D. M. Braun, J.-Y. Lee, J. Meiners, R. Simon, G. Sauter, H. Erfle, K. Rippe, R. Bartenschlager, and K. Rohr. (2019). “Hyperparameter Optimization for Image Analysis: Application to Prostate Tissue Images and Live Cell Data of Virus-Infected Cells”. *International Journal of Computer Assisted Radiology and Surgery*. 14(11): 1847–1857. DOI: 10.1007/s11548-019-02010-3.

Roberts, M., D. Driggs, M. Thorpe, J. Gilbey, M. Yeung, S. Ursprung, A. I. Aviles-Rivero, C. Etmann, C. McCague, L. Beer, J. R. Weir-McCall, Z. Teng, E. Gkrania-Klotsas, J. H. F. Rudd, E. Sala, and C.-B. Schönlieb. (2021). “Common Pitfalls and Recommendations for Using Machine Learning to Detect and Prognosticate for COVID-19 Using Chest Radiographs and CT Scans”. *Nature Machine Intelligence*. 3(3): 199–217. DOI: 10.1038/s42256-021-00307-0.

Robnik-Šikonja, M., M. Radović, S. Đorović, B. Anđelković-Ćirković, and N. Filipović. (2018). “Modeling Ischemia with Finite Elements and Automated Machine Learning”. *Journal of Computational Science*. 29(Nov.): 99–106. DOI: 10.1016/j.jocs.2018.09.017.

Romano, J. D., T. T. Le, W. Fu, and J. H. Moore. (2020). “Is Deep Learning Necessary for Simple Classification Tasks?” *arXiv preprint arXiv:2006.06730*.

Romblay, A. de. (2017). “MLBox”. URL: https://github.com/AxeldeRomblay/MLBox.

rsteca. (2021). “sklearn-deap”. URL: https://github.com/rsteca/sklearn-deap.

Rudrabhatla, C. K. (2020). "Comparison of Zero Downtime Based Deployment Techniques in Public Cloud Infrastructure". In: *2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*. 1082–1086. DOI: 10.1109/I-SMAC49090.2020.9243605.

Ruta, D., B. Gabrys, and C. Lemke. (2011). "A Generic Multilevel Architecture for Time Series Prediction". *IEEE Transactions on Knowledge and Data Engineering*. 23(3): 350–359. DOI: 10.1109/tkde.2010.137.

Rye, I., A. Vik, M. Kocinski, A. S. Lundervold, and A. J. Lundervold. (2022). "Predicting Conversion to Alzheimer's Disease in Individuals with Mild Cognitive Impairment Using Clinically Transferable Features". *Scientific Reports*. 12(1): 15566. DOI: 10.1038/s41598-022-18805-5.

Ryoo, J., S. Rizvi, W. Aiken, and J. Kissell. (2014). "Cloud Security Auditing: Challenges and Emerging Approaches". *IEEE Security Privacy*. 12(6): 68–74. DOI: 10.1109/MSP.2013.132.

Sadley, D. (2021). *NDIS 'Robo-Plans' Test Algorithmic Transparency*. URL: https://www.innovationaus.com/ndis-robo-plans-test-algorithmic-transparency/.

Saeed, W. (2021). "Comparison of Automated Machine Learning Tools for SMS Spam Message Filtering". In: *Advances in Cyber Security*. Ed. by N. Abdullah, S. Manickam, and M. Anbar. *Communications in Computer and Information Science*. Singapore: Springer. 307–316. DOI: 10.1007/978-981-16-8059-5_18.

"SageMaker (AWS)". (2021). URL: https://aws.amazon.com/sagemaker/.

Saha, B. C., J. A. Dhanraj, M. Sujatha, R. Vallikannu, M. Alanazi, A. Almadhor, R. Sathyamurthy, K. G. Erko, and V. Sugumaran. (2022). "Investigating Rotor Conditions on Wind Turbines Using Integrating Tree Classifiers". *International Journal of Photoenergy*. 2022(June): e5389574. DOI: 10.1155/2022/5389574.

Saleiro, P., B. Kuester, L. Hinkson, J. London, A. Stevens, A. Anisfeld, K. T. Rodolfa, and R. Ghani. (2019). "Aequitas: A Bias and Fairness Audit Toolkit". *arXiv preprint arXiv:1811.05577*.

Salvador, M. M. (2017). "Automatic and Adaptive Preprocessing for the Development of Predictive Models." *PhD thesis*. Bournemouth University.

Salvador, M. M., M. Budka, and B. Gabrys. (2016a). "Adapting Multicomponent Predictive Systems Using Hybrid Adaptation Strategies with Auto-Weka in Process Industry". In: *Workshop on Automatic Machine Learning*. 48–57.

Salvador, M. M., M. Budka, and B. Gabrys. (2016b). "Towards Automatic Composition of Multicomponent Predictive Systems". In: *Hybrid Artificial Intelligent Systems*. Springer International Publishing. 27–39.

Salvador, M. M., M. Budka, and B. Gabrys. (2019). "Automatic Composition and Optimization of Multicomponent Predictive Systems With an Extended Auto-WEKA". *IEEE Transactions on Automation Science and Engineering*. 16(2): 946–959. DOI: 10.1109/TASE.2018.2876430. arXiv: 1612.08789.

Samala, R. K., H.-P. Chan, L. Hadjiiski, and S. Koneru. (2020). "Hazards of Data Leakage in Machine Learning: A Study on Classification of Breast Cancer Using Deep Neural Networks". In: *Medical Imaging 2020: Computer-Aided Diagnosis*. Vol. 11314. International Society for Optics and Photonics. 1131416. DOI: 10.1117/12.2549313.

Sánchez-Pozo, N. N., J. S. Mejía-Ordóñez, D. C. Chamorro, D. Mayorca-Torres, and D. H. Peluffo-Ordóñez. (2021). "Predicting High School Students' Academic Performance: A Comparative Study of Supervised Machine Learning Techniques". In: *2021 Machine Learning-Driven Digital Technologies for Educational Innovation Workshop*. IEEE. 1–6.

Santu, S. K. K., M. M. Hassan, M. J. Smith, L. Xu, C. Zhai, and K. Veeramachaneni. (2020). "A Level-wise Taxonomic Perspective on Automated Machine Learning to Date and Beyond: Challenges and Opportunities". *arXiv preprint arXiv:2010.10777*.

Schaer, R., H. Müller, and A. Depeursinge. (2016). "Optimized Distributed Hyperparameter Search and Simulation for Lung Texture Classification in CT Using Hadoop". *Journal of Imaging*. 2(2): 19. DOI: 10.3390/jimaging2020019.

Schneegans, S., J. Lewis, and T. Straza. (2021). "UNESCO Science Report 2021: The Race against Time for Smarter Development". *Tech. rep.* UNESCO. URL: https://unesdoc.unesco.org/ark:/48223/pf0000377433.

Schwen, L. O., D. Schacherer, C. Geißler, and A. Homeyer. (2022). "Evaluating Generic AutoML Tools for Computational Pathology". *Informatics in Medicine Unlocked*. 29(Jan.): 100853. DOI: 10.1016/j.imu.2022.100853.

Scikit-Optimize. (2021). "Scikit-Optimize". URL: https://github.com/scikit-optimize/scikit-optimize.

Scottfree Analytics LLC. (2021). "AlphaPy". URL: https://github.com/ScottfreeLLC/AlphaPy.

Sculley, D., G. Holt, D. Golovin, E. Davydov, T. Phillips, D. Ebner, V. Chaudhary, M. Young, J.-F. Crespo, and D. Dennison. (2015). "Hidden Technical Debt in Machine Learning Systems". *Advances in Neural Information Processing Systems*. 28.

Sehnke, F., M. D. Felder, and A. K. Kaifel. (2012). "Learn-o-Matic: A Fully Automated Machine Learning Suite for Profile Retrieval Applications". In: *Proceedings of the 2012 EUMETSAT Meteorological Satellite Conference*.

Seo, K., B. Chung, H. P. Panchaseelan, T. Kim, H. Park, B. Oh, M. Chun, S. Won, D. Kim, J. Beom, D. Jeon, and J. Yang. (2021). "Forecasting the Walking Assistance Rehabilitation Level of Stroke Patients Using Artificial Intelligence". *Diagnostics*. 11(6): 1096. DOI: 10.3390/diagnostics11061096.

Seshadri, R. (2021a). "Auto_TS". URL: https://github.com/AutoViML/Auto%5C_TS.

Seshadri, R. (2021b). "Auto_ViML". URL: https://github.com/AutoViML/Auto%5C_ViML.

Shah, P. (2020). "Libra". URL: https://github.com/Palashio/libra.

Shearer, C. (2000). "The CRISP-DM Model: The New Blueprint for Data Mining". *Journal of data warehousing*. 5(4): 13–22.

Shi, Y., R. Biswas, M. Noori, M. Kilberry, J. Oram, J. Mays, S. Kharude, D. Rao, and X. Chen. (2021). "Predicting Road Accident Risk Using Geospatial Data and Machine Learning (Demo Paper)". In: *Proceedings of the 29th International Conference on Advances in Geographic Information Systems. SIGSPATIAL '21*. New York, NY, USA: Association for Computing Machinery. 512–515. DOI: 10.1145/3474717.3484253.

Shin, T. (2021). *An Extensive Step by Step Guide to Exploratory Data Analysis*. URL: https://towardsdatascience.com/an-extensive-guide-to-exploratory-data-analysis-ddd99a03199e.

Shu, R., T. Xia, J. Chen, L. Williams, and T. Menzies. (2021). "How to Better Distinguish Security Bug Reports (Using Dual Hyperparameter Optimization)". *Empirical Software Engineering*. 26(3): 53. DOI: 10.1007/s10664-020-09906-8.

Shubhangi, V., E. Brethenoux, F. Choudhary, and J. Hare. (2020). "Use Gartner's 3-Stage MLOps Framework to Successfully Operationalize Machine Learning Projects". *Tech. rep.* URL: https://www.gartner.com/document/3987104.

Sidhu, A. (2021). "Aethos". URL: https://github.com/Ashton-Sidhu/aethos.

Sills, M. R., M. Ozkaynak, and H. Jang. (2021). "Predicting Hospitalization of Pediatric Asthma Patients in Emergency Departments Using Machine Learning". *International Journal of Medical Informatics*. 151(July): 104468. DOI: 10.1016/j.ijmedinf.2021.104468.

Silva, L., F. Seixas, C. Fontes, D. Muchaluat-Saade, and A. Conci. (2020). "A Computational Method for Breast Abnormality Detection Using Thermographs". In: *2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS)*. 469–474. DOI: 10.1109/CBMS49503.2020.00095.

Silva, L. F., A. A. S. M. D. Santos, R. S. Bravo, A. C. Silva, D. C. Muchaluat-Saade, and A. Conci. (2016). "Hybrid Analysis for Indicating Patients with Breast Cancer Using Temperature Time Series". *Computer Methods and Programs in Biomedicine*. 130(July): 142–153. DOI: 10.1016/j.cmpb.2016.03.002.

Singh, A., N. Thakur, and A. Sharma. (2016). “A Review of Supervised Machine Learning Algorithms”. In: *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*. 1310–1315.

Singh, D., P. K. Pant, H. Pant, and D. C. Dobhal. (2021). “Robust Automated Machine Learning (AutoML) System for Early Stage Hepatic Disease Detection”. In: *Intelligent Data Communication Technologies and Internet of Things*. Ed. by J. Hemanth, R. Bestak, and J. I.-Z. Chen. *Lecture Notes on Data Engineering and Communications Technologies*. Singapore: Springer. 65–76. DOI: 10.1007/978-981-15-9509-7_6.

Sklearn documentation - List of all Estimators. (2021). URL: https://scikit-learn.org/stable/modules/generated/sklearn.utils.all_estimators.html#sklearn-utils-all-estimators.

Sleimi, A., N. Sannier, M. Sabetzadeh, L. Briand, M. Ceci, and J. Dann. (2021). “An Automated Framework for the Extraction of Semantic Legal Metadata from Legal Texts”. *Empirical Software Engineering*. 26(3): 43. DOI: 10.1007/s10664-020-09933-5.

Smith, M. J., C. Sala, J. M. Kanter, and K. Veeramachaneni. (2020). “The Machine Learning Bazaar: Harnessing the ML Ecosystem for Effective System Development”. *Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data*. June: 785–800. DOI: 10.1145/3318464.3386146. arXiv: 1905.08942.

Society For Human Resource Management. (2016). “Jobs of the Future: Data Analysis Skills”. *Tech. rep.* URL: https://www.shrm.org/hr-today/trends-and-forecasting/research-and-surveys/documents/data-analysis-skills.pdf.

“SparkCognition”. (2021). URL: https://www.sparkcognition.com/products/darwin/.

Sparks, E. R., A. Talwalkar, D. Haas, M. J. Franklin, M. I. Jordan, and T. Kraska. (2015). “Automating Model Search for Large Scale Machine Learning”. In: *Proceedings of the Sixth ACM Symposium on Cloud Computing - SoCC 15*. ACM Press. DOI: 10.1145/2806777.2806945.

Sparks, E. R., S. Venkataraman, T. Kaftan, M. J. Franklin, and B. Recht. (2017). "KeystoneML: Optimizing Pipelines for Large-Scale Advanced Analytics". In: *2017 IEEE 33rd International Conference on Data Engineering (ICDE)*. IEEE. DOI: 10.1109/icde.2017.109.

Speicher, T., H. Heidari, N. Grgic-Hlaca, K. P. Gummadi, A. Singla, A. Weller, and M. B. Zafar. (2018). "A Unified Approach to Quantifying Algorithmic Unfairness: Measuring Individual &Group Unfairness via Inequality Indices". In: *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. KDD '18*. New York, NY, USA: Association for Computing Machinery. 2239–2248. DOI: 10.1145/3219819.3220046.

"Spell". (2021). URL: https://spell.ml/platform.

Spruit, M. and M. Lytras. (2018). "Applied Data Science in Patient-Centric Healthcare: Adaptive Analytic Systems for Empowering Physicians and Patients". *Telematics and Informatics*. 35(4): 643–653. DOI: 10.1016/j.tele.2018.04.002.

Srivastava, S., R. N. K., R. Srinivasan, N. K. M. Nambison, and S. S. Gorthi. (2021). "Diagnosis of Sickle Cell Anemia Using AutoML on UV-Vis Absorbance Spectroscopy Data". *arXiv preprint arXiv:2111.12711*.

Stackoverflow. (2021). "2021 Developer Survey". *Tech. rep.* URL: https://insights.stackoverflow.com/survey/2021#technology.

Standards Australia. (2020). "An Artificial Intelligence Standards Roadmap: Making Australia's Voice Heard". *Tech. rep.* URL: https://www.standards.org.au/getmedia/ede81912-55a2-4d8e-849f-9844993c3b9d/R_1515-An-Artificial-Intelligence-Standards-Roadmap-soft.pdf.aspx.

Stanton, B. and J. Theodore. (2021). "Trust and Artificial Intelligence". *Tech. rep.* National Institute of Standards and Technology, U.S. Department of Commerce. URL: https://nvlpubs.nist.gov/nistpubs/ir/2021/NIST.IR.8332-draft.pdf.

Studer, S., T. B. Bui, C. Drescher, A. Hanuschkin, L. Winkler, S. Peters, and K.-R. Müller. (2021). "Towards CRISP-ML (Q): A Machine Learning Process Model with Quality Assurance Methodology". *Machine Learning and Knowledge Extraction*. 3(2): 392–413.

Sun, A. Y., B. R. Scanlon, H. Save, and A. Rateb. (2021). "Reconstruction of GRACE Total Water Storage Through Automated Machine Learning". *Water Resources Research*. 57(2): e2020WR028666. DOI: 10.1029/2020WR028666.

Suresh, H. and J. V. Guttag. (2020). "A Framework for Understanding Unintended Consequences of Machine Learning". *arXiv preprint arXiv:1901.10002*.

Taino, D. F., M. G. Ribeiro, G. F. Roberto, G. F. D. Zafalon, M. Z. do Nascimento, T. A. A. Tosta, A. S. Martins, and L. A. Neves. (2021). "Analysis of Cancer in Histological Images: Employing an Approach Based on Genetic Algorithm". *Pattern Analysis and Applications*. 24(2): 483–496. DOI: 10.1007/s10044-020-00931-3.

Tan, J., S. Xie, W. Wu, P. Qin, and T. Ouyang. (2022). "Evaluating and Optimizing the Cold Energy Efficiency of Power Generation and Wastewater Treatment in LNG-fired Power Plant Based on Data-Driven Approach". *Journal of Cleaner Production*. 334(Feb.): 130149. DOI: 10.1016/j.jclepro.2021.130149.

Taylor, S. J. and B. Letham. (2018). "Forecasting at Scale". *The American Statistician*. 72(1): 37–45.

"TAZI". (2021). URL: https://www.tazi.ai/.

TensorFlow. (2021). "Adanet". URL: https://github.com/tensorflow/adanet.

The Artificial Intelligence Innovation (A2I) research laboratory at Cedars-Sinai Medical Center. (2021). "TPOT". URL: https://github.com/EpistasisLab/tpot.

The European Commission. (2021). "Proposal for a Regulation Laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act)". *Tech. rep.* URL: https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence-artificial-intelligence.

The Institute for Ethical Machine Learning. (2021). *Awesome-Production-Machine-Learning*. URL: https://github.com/EthicalML/awesome-production-machine-learning.

The Machine Learning Group at The University of Sheffield. (2021). "GPyOpt". URL: https://github.com/SheffieldML/GPyOpt.

The mlr-org. (2021). "mlrMBO". URL: https://github.com/mlr-org/mlrMBO.

The University at Buffalo. (2021). "The Hachmann Lab. ChemML". URL: https://github.com/hachmannlab/chemml.

The White House. (2021). *The Biden Administration Launches the National Artificial Intelligence Research Resource Task Force*. URL: https://www.whitehouse.gov/ostp/news-upyears/2021/06/10/the-biden-administration-launches-the-national-artificial-intelligence-research-resource-task-force/.

Thomas, J. (2021). "AutoXGBoost". URL: https://github.com/ja-thomas/autoxgboost.

Thompson, N. C., K. Greenewald, K. Lee, and G. F. Manso. (2021). "Deep Learning's Diminishing Returns: The Cost of Improvement Is Becoming Unsustainable". *IEEE Spectrum*. 58(10): 50–55. DOI: 10.1109/MSPEC.2021.9563954.

Thornton, C., F. Hutter, H. H. Hoos, and K. Leyton-Brown. (2013). "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*. 847–855.

Tighe, D., T. Lewis-Morris, and A. Freitas. (2019). "Machine Learning Methods Applied to Audit of Surgical Outcomes after Treatment for Cancer of the Head and Neck". *British Journal of Oral and Maxillofacial Surgery*. 57(8): 771–777. DOI: 10.1016/j.bjoms.2019.05.026.

"TIMi". (2021). URL: https://timi.eu/timi/timi-modeler/.

Tiobe. (2021). *Tiobe Index*. URL: https://www.tiobe.com/tiobe-index/.

Top 10 DataRobot Alternatives & Competitors. (2021). URL: https://www.g2.com/products/datarobot/competitors/alternatives.

Toreini, E., M. Aitken, K. Coopamootoo, K. Elliott, C. G. Zelaya, and A. van Moorsel. (2020). "The Relationship between Trust in AI and Trustworthy Machine Learning Technologies". In: *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. 272–283.

Truong, A., A. Walters, J. Goodsitt, K. Hines, C. B. Bruss, and R. Farivar. (2019). "Towards Automated Machine Learning: Evaluation and Comparison of AutoML Approaches and Tools". *2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI)*. Nov.: 1471–1479. DOI: 10.1109/ICTAI.2019.00209. arXiv: 1908.05557.

Tsai, C.-W. and Z.-Y. Fang. (2021). "An Effective Hyperparameter Optimization Algorithm for DNN to Predict Passengers at a Metro Station". *ACM Transactions on Internet Technology*. 21(2): 32:1–32:24. DOI: 10.1145/3410156.

Tsakonas, A. and B. Gabrys. (2012). "GRADIENT: Grammar-driven Genetic Programming Framework for Building Multi-Component, Hierarchical Predictive Systems". *Expert Systems with Applications*. 39(18): 13253–13266. DOI: 10.1016/j.eswa.2012.05.076.

Tsamardinos, I., P. Charonyktakis, G. Papoutsoglou, G. Borboudakis, K. Lakiotaki, J. C. Zenklusen, H. Juhl, E. Chatzaki, and V. Lagani. (2022). "Just Add Data: Automated Predictive Modeling for Knowledge Discovery and Feature Selection". *NPJ precision oncology*. 6(1): 1–17.

Tsiakmaki, M., G. Kostopoulos, S. Kotsiantis, and O. Ragos. (2020). "Implementing AutoML in Educational Data Mining for Prediction Tasks". *Applied Sciences*. 10(1): 90. DOI: 10.3390/app10010090.

Ukanwa, K. and R. T. Rust. (2021). "Algorithmic Discrimination in Service". *USC Marshall School of Business Research Paper*.

Valdes-Ramirez, D., M. A. Medina-Pérez, and R. Monroy. (2021). "An Ensemble of Fingerprint Matching Algorithms Based on Cylinder Codes and Mtriplets for Latent Fingerprint Identification". *Pattern Analysis and Applications*. 24(2): 433–444. DOI: 10.1007/s10044-020-00911-7.

van Eeden, W. A., C. Luo, A. M. van Hemert, I. V. E. Carlier, B. W. Penninx, K. J. Wardenaar, H. Hoos, and E. J. Giltay. (2021). "Predicting the 9-Year Course of Mood and Anxiety Disorders with Automated Machine Learning: A Comparison between Auto-Sklearn, Naïve Bayes Classifier, and Traditional Logistic Regression". *Psychiatry Research*. 299(May): 113823. DOI: 10.1016/j.psychres.2021.113823.

Vanawat, Nimit. (2021). *How To Perform Exploratory Data Analysis -A Guide for Beginners.* URL: https://www.analyticsvidhya.com/blog/2021/08/how-to-perform-exploratory-data-analysis-a-guide-for-beginners/.

VanderPlas, J. (2016). *Python Data Science Handbook: Essential Tools for Working with Data.* " O'Reilly Media, Inc."

Vanschoren, J. (2018). "Meta-Learning: A Survey". *arXiv preprint arXiv:1810.03548.*

Vega, A. de la, D. García-Saiz, M. Zorrilla, and P. Sánchez. (2019). "How Far Are We from Data Mining Democratisation? A Systematic Review". *arXiv preprint arXiv:1903.08431v1.*

Venkata, D., P. Senthil Kumar, L. Y. Venkataramana, G. Prasannamedha, S. Harshana, S. Jahnavi Srividya, K. Harrinei, and S. Indraganti. (2021). "Automating Water Quality Analysis Using ML and Auto ML Techniques". *Environmental Research.* 202(Nov.): 111720. DOI: 10.1016/j.envres.2021.111720.

Vishwakarma, G., M. Haghighatlari, and J. Hachmann. (2019). "Towards Autonomous Machine Learning in Chemistry via Evolutionary Algorithms". *ChemRxiv preprint ChemRxiv:10.26434/chemrxiv.9782387.v1.*

"Viya (SAS)". (2021). URL: https://www.sas.com/en%5C_us/software/visual-data-mining-machine-learning.html.

Wang, C., T. Bäck, H. H. Hoos, M. Baratchi, S. Limmer, and M. Olhofer. (2019a). "Automated Machine Learning for Short-term Electric Load Forecasting". In: *2019 IEEE Symposium Series on Computational Intelligence (SSCI).* 314–321. DOI: 10.1109/SSCI44817.2019.9002839.

Wang, D., J. D. Weisz, M. Muller, P. Ram, W. Geyer, C. Dugan, Y. Tausczik, H. Samulowitz, and A. Gray. (2019b). "Human-AI Collaboration in Data Science: Exploring Data Scientists' Perceptions of Automated AI". *Proceedings of the ACM on Human-Computer Interaction.* 3(CSCW): 1–24. DOI: 10.1145/3359313.

Wang, T., X. Wu, and T. He. (2019c). "Trustable and Automated Machine Learning Running with Blockchain and Its Applications". *arXiv preprint arXiv:1908.05725.*

Wang, Z., J. Poon, S. Wang, S. Sun, and S. Poon. (2021). "A Novel Method for Clinical Risk Prediction with Low-Quality Data". *Artificial Intelligence in Medicine*. 114(Apr.): 102052. DOI: 10.1016/j.artmed.2021.102052.

Warmerdam, V. D. (2021). "Scikit-Fairness". URL: https://github.com/koaning/scikit-fairness.

"Watson Studio (IBM)". (2021). URL: https://www.ibm.com/au-en/cloud/watson-studio/autoai.

Webb, G. I., R. Hyde, H. Cao, H. L. Nguyen, and F. Petitjean. (2016). "Characterizing Concept Drift". *Data Mining and Knowledge Discovery*. 30(4): 964–994. DOI: 10.1007/s10618-015-0448-4.

Wei, W. (2021). *Awesome-AutoML*. URL: https://github.com/windmaple/awesome-AutoML.

West, J. (2018). "Teaching Data Science: An Objective Approach to Curriculum Validation". *Computer Science Education*. 28(2): 136–157.

Wickham, H. and G. Grolemund. (2017). *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. 1st. O'Reilly Media, Inc.

Wikipedia. (2021). *Automated machine learning*. URL: https://en.wikipedia.org/w/index.php?title=Automated_machine_learning&oldid=1054510848.

Williams, L. (2021). *Clearsense Acquires Compellon*. URL: .https://clearsense.com/insights/clearsense-acquires-plug-and-play-ai-analytics-firm/.

Williams, M. L., W. P. James, and M. T. Rose. (2019). "Variable Segmentation and Ensemble Classifiers for Predicting Dairy Cow Behaviour". *Biosystems Engineering*. 178(Feb.): 156–167. DOI: 10.1016/j.biosystemseng.2018.11.011.

Wistuba, M., A. Rawat, and T. Pedapati. (2019). "A Survey on Neural Architecture Search". *arXiv preprint arXiv:1905.01392v2*.

Wixom, B., T. Ariyachandra, D. Douglas, M. Goul, B. Gupta, L. Iyer, U. Kulkarni, J. Mooney, G. Phillips-Wren, and O. Turetken. (2014). "The Current State of Business Intelligence in Academia: The Arrival of Big Data". *Communications of the Association for Information Systems*. 34(1). DOI: 10.17705/1CAIS.03401.

"Xpanse". (2021). URL: https://xpanse.ai/.

Yang, C., Y. Akimoto, D. W. Kim, and M. Udell. (2021). "OBOE". URL: https://github.com/udellgroup/oboe.

Yang, L. and A. Shami. (2020). "On Hyperparameter Optimization of Machine Learning Algorithms: Theory and Practice". *Neurocomputing*. July. DOI: 10.1016/j.neucom.2020.07.061.

Yang, Y., J. Mao, R. Nguyen, A. Tohmeh, and H.-G. Yeh. (2022). "Feature Construction and Selection for PV Solar Power Modeling". *arXiv preprint arXiv:2202.06226*.

Yao, Q., M. Wang, Y. Chen, W. Dai, Y.-F. Li, W.-W. Tu, Q. Yang, and Y. Yu. (2018). "Taking Human out of Learning Applications: A Survey on Automated Machine Learning". *arXiv preprint arXiv:1810.13306*.

Zemel, R., Y. Wu, K. Swersky, T. Pitassi, and C. Dwork. (2013). "Learning Fair Representations". In: *Proceedings of the 30th International Conference on Machine Learning*. PMLR. 325–333.

Zeng, S., M. Arjomandi, Y. Tong, Z. C. Liao, and G. Luo. (2022). "Developing a Machine Learning Model to Predict Severe Chronic Obstructive Pulmonary Disease Exacerbations: Retrospective Cohort Study". *Journal of Medical Internet Research*. 24(1): e28953. DOI: 10.2196/28953.

Zhang, B. H., B. Lemoine, and M. Mitchell. (2018a). "Mitigating Unwanted Biases with Adversarial Learning". In: *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*. 335–340.

Zhang, C. and Z. Ye. (2020). "Water Pipe Failure Prediction Using AutoML". *Facilities*. 39(1/2): 36–49. DOI: 10.1108/F-08-2019-0084.

Zhang, T., J. Su, C. Liu, and W.-H. Chen. (2021). "State and Parameter Estimation of the AquaCrop Model for Winter Wheat Using Sensitivity Informed Particle Filter". *Computers and Electronics in Agriculture*. 180(Jan.): 105909. DOI: 10.1016/j.compag.2020.105909.

Zhang, W., P. Ge, W. Jin, and J. Guo. (2018b). "Radar Signal Recognition Based on TPOT and LIME". In: *2018 37th Chinese Control Conference (CCC)*. 4158–4163. DOI: 10.23919/ChiCC.2018.8483165.

Zhang, X., Y. Li, and Z. Li. (2022). "Comparative Research of Hyper-Parameters Mathematical Optimization Algorithms for Automatic Machine Learning in New Generation Mobile Network". *Mobile Networks and Applications*. 27(3): 928–935. DOI: 10.1007/s11036-022-01913-x.

Zhang, Y., M. T. Bahadori, H. Su, and J. Sun. (2016). "FLASH: Fast Bayesian Optimization for Data Analytic Pipelines". In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD 16*. ACM Press. DOI: 10.1145/2939672.2939829.

Zhang, Z. and D. B. Neill. (2017). "Identifying Significant Predictive Bias in Classifiers". *arXiv preprint arXiv:1611.08292*.

Zhao, Y., Z. Nasrullah, and Z. Li. (2019). "PyOD: A Python Toolbox for Scalable Outlier Detection". *Journal of Machine Learning Research*. 20(96): 1–7.

Zhong, J., W. Guo, and Z. Wang. (2016). "Study on Network Failure Prediction Based on Alarm Logs". In: *2016 3rd MEC International Conference on Big Data and Smart City (ICBDSC)*. 1–7. DOI: 10.1109/ICBDSC.2016.7460337.

Zhou, B., E. Erell, I. Hough, A. Shtein, A. C. Just, V. Novack, J. Rosenblatt, and I. Kloog. (2020). "Estimation of Hourly near Surface Air Temperature Across Israel Using an Ensemble Model". *Remote Sensing*. 12(11): 1741. DOI: 10.3390/rs12111741.

Zhou, L., Z. Wang, X. Tian, and L. Peng. (2021). "LPI-deepGBDT: A Multiple-Layer Deep Framework Based on Gradient Boosting Decision Trees for lncRNA–Protein Interaction Identification". *BMC Bioinformatics*. 22(1): 479. DOI: 10.1186/s12859-021-04399-8.

Zillow. (2020). "Luminaire". URL: https://github.com/zillow/luminaire.

Zliobaite, I., A. Bifet, M. Gaber, B. Gabrys, J. Gama, L. Minku, and K. Musial. (2012). "Next Challenges for Adaptive Learning Systems". *ACM SIGKDD Explorations Newsletter*. 14(1): 48. DOI: 10.1145/2408736.2408746.

Zliobaite, I. and B. Gabrys. (2014). "Adaptive Preprocessing for Streaming Data". *IEEE Transactions on Knowledge and Data Engineering*. 26(2): 309–321. DOI: 10.1109/tkde.2012.147.

Žliobaitė, I., M. Budka, and F. Stahl. (2015). "Towards Cost-Sensitive Adaptation: When Is It Worth Updating Your Predictive Model?" *Neurocomputing*. 150: 240–249. DOI: 10.1016/j.neucom.2014.05.084.

Zöller, M.-A. and M. F. Huber. (2021). "Benchmark and Survey of Automated Machine Learning Frameworks". *Journal of Artificial Intelligence Research*. 70(Jan.): 409–472. DOI: 10.1613/jair.1.11854.