Available online www.jsaer.com

***Journal of Scientific and Engineering Research*, 2020, 7(8):36-40**

Review Article

ISSN: 2394-2630
CODEN(USA): JSERBR

Automated Machine Learning

Matthew N. O. Sadiku[1], Uwakwe C. Chukwu[2], Abayomi Ajayi-Majebi[3], Sarhan M. Musa[1]

[1]Roy G. Perry College of Engineering, Prairie View A&M University, Prairie View, TX, USA
[2]Department of Engineering Technology, South Carolina State University, Orangeburg, SC, USA
[3]Department of Manufacturing Engineering, Central State University, P.O. Box 1004, Wilberforce, OH, USA
Email: sadiku@ieee.org; uchukwu@scsu.edu; ajayi-majebi@centralstate.edu; smmusa @pvamu.edu

Abstract Machine learning (ML) is about learning some properties of a data set and applying them to new data. It refers to the automated detection of meaningful patterns in a given data.

Automated machine learning is the process of automating the process of applying machine learning to real-world problems. It provides means for making ML available for non-ML experts. It also offers the advantages of making application of ML easier for non-experts, producing faster creation of solutions, and yielding models that outperform models designed by hand. This paper presents an introduction to the fast-moving field of automated machine learning.

Keywords machine learning, artificial intelligence, automated machine learning

Introduction

Today, we witness an exponential growth in the amount of data generated by various kinds of systems and devices. Many companies and organizations have realized that data analysis is a powerful tool. A crucial step in the data analysis process is the construction and training of a machine learning model [1]. Machine learning techniques and algorithms are playing a crucial role in harnessing the power of massive amounts of data produced daily.

Machine learning is an emerging technology that can facilitate discovery of rules and patterns in a data set. It is the process of applying scientific principles and algorithms to data so that it can be automated to make predictions or decisions. It uses algorithms to parse data, learn from it, and draw conclusions without human intervention. Many industries have realized great benefits from using machine learning to increase reliability, productivity, and the safer operations of their machines.

The rapid growth of ML applications has created a need for off-the-shelf ML methods that can be used easily by non-ML-experts. We refer to the resulting research area as automation of machine learning. Automated machine learning (AutoML) aims to eliminate the need for skilled data scientists to build machine learning models. It basically simplifies machine learning by making the entire process automated. Automated machine learning is regarded as the catalyst for a fundamental change in the way machine learning and data science is approached.

Overview on Machine Learning

Machine learning (ML) is the discipline that gives computers the ability to learn without being explicitly programmed. The term "machine learning" (ML) was initially coined in 1959 by Arthur Samuel, a computer scientist. Machine learning (or statistical learning) is part of artificial intelligence. It assists computers in estimating future events and modelling based on experiences gained from previous information. Machine learning (ML) focuses on how computers "learn" from data. It allows computers to learn from past examples

and to detect hard-to-discern patterns from large data sets. It describes a class of algorithms which learn model parameters from a set of training data with the purpose of accurately predicting outcomes for previously unseen data. ML is a marriage between statistics and computer science [2,3].
As shown in Figure 1, there are two types of learning: supervised learning and unsupervised learning [4].

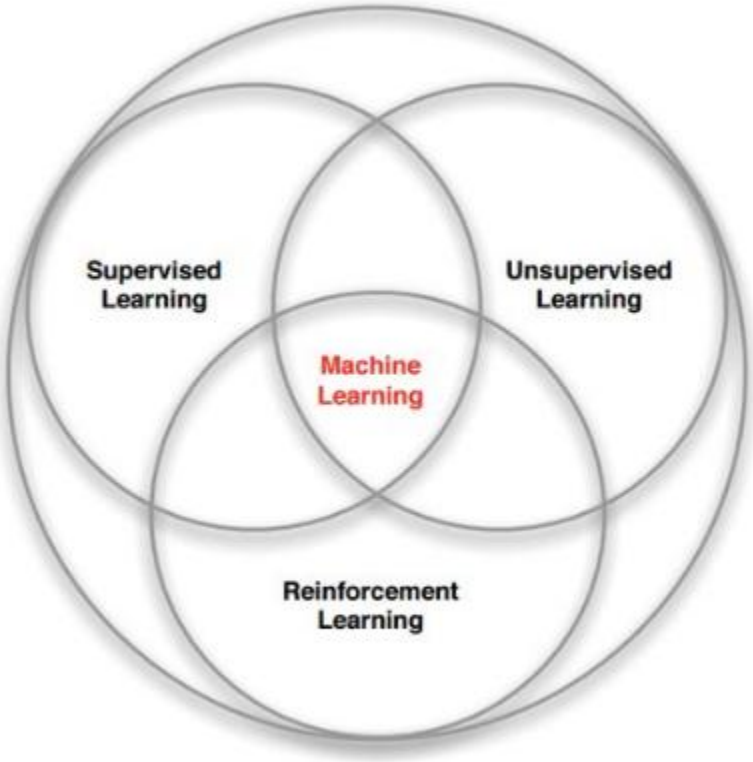

Figure 1: Branches of machine learning [4]

Supervised learning focuses on classification and prediction. It involves building a statistical model for predicting or estimating an outcome based on one or more inputs. It is often used to estimate risk. Supervised ML is where algorithms are given training data. Learning from data is used when there is no theoretical or prior knowledge solution, but data is available to construct an empirical solution. Supervised ML is increasingly being used in medicine such as in cardiac electrophysiology. In unsupervised learning, we are interested in finding naturally occurring patterns within the data. Unlike supervised learning, there is no predicted outcome. Unsupervised learning looks for internal structure in the data. Unsupervised learning algorithms are common in neural network models. A common application of such a process is to explore interrelationships between genetics, biochemistry, histology, and disease states.
The traditional machine learning process typically begins with raw data and ends with a predictive model that can be used to make decisions. The process usually includes the following steps [5]:

1. *Data Gathering* to identify and collect input data.
2. *Data Cleansing* to standardize and clean the raw inputs.
3. *Feature Processing* to transform the input data into formats that can be easily processed to identify the best predictor variables.
4. *Model Training* to train the model, using a wide range of potential algorithms.
5. *Model Validation* to test the model against historical data and assess its performance.
6. *Model Deployment* to load the model into an environment where it can make decisions.

The applications of machine learning are endless, including medicine, machine perception, computer vision, object recognition, natural language processing, cheminformatics, fraud detection, stock market analysis, games,

robotics, health monitoring. Industrial leaders such Google, Amazon, and Microsoft are now offering numerous tools to enable beginners get started with building their own machine learning systems.

Concept of Automatic Machine Learning

The traditional process of manually constructing a machine learning model is very skill-intensive and time consuming. Therefore, applying traditional machine learning methods to real-world problems is time-consuming, resource-intensive, and challenging. It may require experts in the several disciplines such as domain knowledge, data science, engineering, mathematics, statistics, and computer science and takes weeks and months to implement. Automated machine learning changes this situation by incorporating ML best practices from top-ranked skilled data scientists.

Automated machine learning (or autoML) is the process of automating the time consuming, iterative tasks (such as choosing data sources, data prep, and feature selection) of machine learning model development. Instead of searching algorithms and manually tuning hyperparameters, AutoML automatically iterates through various ML algorithms and optimizes hyperparameters. Every machine learning system or model has hyperparameter. The basic task in AutoML is to automatically set these hyperparameters to optimize performance.

The main goal of AutoML is to develop techniques for solving new machine-learning problems automatically, without the need for human intervention. Automated ML tends to democratize the ML model development process by empowering its users, regardless o the background, to identify an end-to-end machine learning pipeline for any problem. A typical automated machine learning is shown in Figure 2 [6].

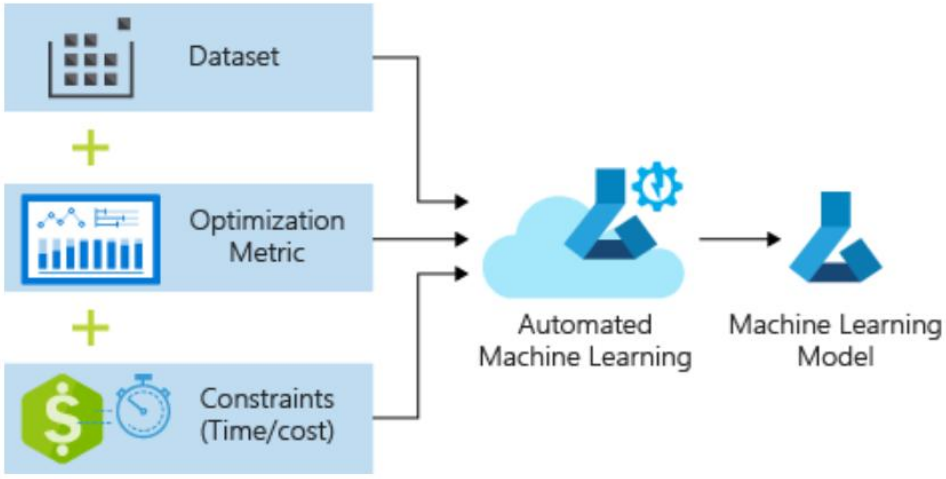

Figure 2: A typical automated machine learning [6]

Automated ML has emerged to minimize the design complexity of a complete ML solution, which usually includes data preprocessing, feature engineering, model selection and ensemble, fine-tuning of hyperparameters, etc. The various stages of the automated machine learning process is shown in Figure 3 [7]. These stages are automated [8]:

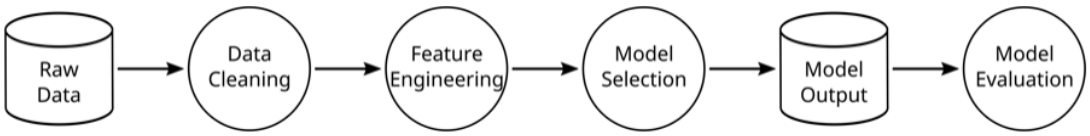

Figure 3: Stages of the automated machine learning process [7].

- *Automated feature engineering:* A feature is an individual measurable property or characteristic of a phenomenon being observed. Feature engineering refers to the process of finding the best set of variables and data encoding for input to the model training process.
- *Automated model selection:* To be effective in practice, AutoML systems need to automatically choose a good algorithm for a given data set. Given the ever increasing number of ML models being

developed, model selection is becoming increasingly important. Automating the selection has long been one of the goals of the machine.

- *Automatic deep transfer learning:* When training models, rather than starting from scratch AutoML implements automatic deep transfer learning.

Several off-the-shelf packages have been developed for automated machine learning [9]. For example, DataRobot claims to have invented AutoML and has a strong reputation in the market.

Applications

ML is already a powerful tool that can give your business decision-making processes a massive upgrade. The technology is here, it is already proving itself in the market, and it is increasingly being built with business in mind.

As the name suggests, automated machine learning is machine learning practice at which human involvements are minimized. The user provides the data, and the AutoML system automatically determines the approach that performs best for this particular application. Automated machine learning makes it possible for businesses in every industry (such as healthcare, banking, marketing) to leverage machine learning. New ways of applying ML for mundane, repetitive, and complex tasks is increasing: from using ML to reply to e-mail messages automatically to self-driving cars and using ML for sustainable mobility.

Benefits and Challenges

AutoML are emerging approaches that are making great strides and they automatically search for possible solutions from a large space of possible kinds of models. It makes data science accessible to organizations and companies of all sizes. This helps them to easily retrieve relevant information from the huge data they possess and generate hundreds of models automatically. AutoML improves efficiency of ML, accelerates research on ML, and helps in making intelligent predictions. It makes the process of machine learning simple and repeatable. It enables users to implement machine learning solutions with ease and frees up data scientists to focus on more complex problems. AutoML compensates for the lack of ML experts. It greatly eliminates or reduces bias and errors that occur when a human being is designing the ML models.

Developing an AutoML model can be an expensive task. Lack of interpretability hinders wide industry adoption of AutoML. Since AutoML is fairly new in the machine learning community, one must exercise caution while applying the current AutoML solutions. Automated machine learning has not been able automated all tasks, which means that some human experts are still required to help in building ML solutions.

Conclusion

Machine learning involves using the past experience to optimize the performance of a given algorithm. ML algorithms facilitate computers to learn from past experience.

It has a rapidly increasing presence across industries. The huge impact of ML on our everyday lives will continue to increase with the introduction of automated machine learning (AutoML), which makes it possible for organizations and companies of all sizes to benefit from using machine learning. AutoML model development is increasingly becoming an established research area. For more information on automated machine learning, one should consult books in [10-13].

References

[1]. L. Tuggener et al., "Automated machine learning in practice: State of the art and recent results," *Proceedings of the 6th Swiss Conference on Data Science*, 2019, pp. 31-36.

[2]. M. N. O. Sadiku, S. M. Musa, and O. S. Musa, "Machine learning, " *International Research Journal of Advanced Engineering and Science*, vol. 2, no. 4, 2017, pp. 79-81.

[3]. M. N. O. Sadiku, S. M. Musa, and O. S. Musa, "Machine learning in chemical industry," *International Journal of Advances in Scientific Research and Engineering*, vol. 3, no. 10, November 2017, pp. 12-15.

[4]. R. Mosic, "Man group hedge fund uses machine / Reinforcement learning for financial trading," July 2017, https://medium.com/@ranko.mosic/why-reinforcement-learning-is-a-perfect-match-for-financial-securities-trading-9462907069ba
[5]. "How automated machine learning (AutoML) is leveling the playing field," May 2019, https://digifi.io/blog/introduction-to-automated-machine-learning-automl/
[6]. "Tutorial: Use automated machine learning to predict taxi fares," August 2019, https://docs.microsoft.com/en-us/azure/machine-learning/service/tutorial-auto-train-models
[7]. D. Vatterott, "Looking towards the future of automated machine-learning," November 2018, http://www.pybloggers.com/2018/11/looking-towards-the-future-of-automated-machine-learning/
[8]. "Automated machine learning\," *Wikipedia*, the free encyclopedia https://en.wikipedia.org/wiki/Automated_machine_learning
[9]. "AUTOML," https://www.ml4aad.org/automl/
[10]. K. Whitford, *Automated Machine Learning Pipeline with Mesos.* Packt Publishing, 2017.
[11]. F. Hutter, L. Kotthoff, and J. Vanschoren (eds.), *Automated Machine Learning: Methods, Systems, Challenges*. Springer, 2019.
[12]. S. Das and U. M. Cakmak. *Hands-On Automated Machine Learning*. Packt Publishing, 2018.
[13]. D. Mukunthu, P. Shah, and W. H. Tok, *Practical Automated Machine Learning on Azure: Using Azure Machine Learning to Quickly Build AI Solutions.* O'Reilly Media, 2019.

About the Authors

Matthew N.O. Sadiku is a professor emeritus in the Department of Electrical and Computer Engineering at Prairie View A&M University, Prairie View, Texas. He is the author of several books and papers. His areas of research interest include computational electromagnetics and computer networks. He is a fellow of IEEE.

Uwakwe C. Chukwu is an associate professor in the Department of Industrial & Electrical Engineering Technology of South Carolina State University. He has published several books and papers. His research interests are power systems, smart grid, V2G, energy scavenging, renewable energies, and microgrids.

Abayomi Ajayi-Majebi is a professor in the Department of Manufacturing Engineering at Central State University in Wilberforce, Ohio. In 2015 he was honored by the White House as a Champion of Change for his significant contributions to the engineering education of minority students. He is a senior member of both the Society of Manufacturing Engineers and the American Society for Quality.

Sarhan M. Musa is a professor in the Department Electrical and Computer Engineering at Prairie View A&M University, Texas. He has been the director of Prairie View Networking Academy, Texas, since 2004. He is an LTD Sprint and Boeing Welliver Fellow. His areas of research interest include computational electromagnetics and computer networks.