AutoML | Automated Machine Learning

AutoML enables everyone to build the machine learning models and make use of its power without having expertise in machine learning.

In recent years, Machine Learning has evolved very rapidly and has become one of the most popular and demanding technology in current times. It is currently being used in every field, making it more valuable. But there are two biggest barriers to making efficient use of machine learning (classical & deep learning): skills and **computing resources.** However, computing resources can be made available by spending a good amount of money, but the availability of skills to solve the machine learning problem is still difficult. It means it is not available for those with limited machine learning knowledge. To solve this problem, Automated Machine Learning (AutoML) came into existence. In this topic, we will understand what AuotML is and how it affects the world?

What is AutoML?

Automated Machine Learning or AutoML is a way to automate the time-consuming and iterative tasks involved in the machine learning model development process. It provides various methods to make machine learning available for people with limited knowledge of Machine Learning. It aims to reduce the need for skilled people to build the ML model. It also helps to improve efficiency and to accelerate the research on Machine learning.

To better understand automated machine learning, we must know the life cycle of a data science or ML project. A typical lifecycle of a data science project contains the following phases:

- o **Data Cleaning**
- o **Feature Selection/Feature Engineering**
- o **Model Selection**
- o **Parameter Optimization**
- o **Model Validation.**

Despite advancements in technology, these processes still require manual effort, making them time-consuming and demanding for non-experts. The rapid growth of ML applications has generated a demand for automating these processes, enabling easier usage without expert knowledge. AutoML emerged to automate the entire process from data cleaning to parameter optimization, saving time and delivering excellent performance.

AutoML Platforms

AutoML has evolved before many years, but in the last few years, it has gained popularity. There are several platforms or frameworks that have emerged. These platforms enable the user to train the model using drag & drop design tools.

1. Google Cloud AutoML

Google has launched several AutoML products for building our own custom machine learning models as per the business needs, and it also allows us to integrate these models into our applications or websites. Google has created the following product:

- **AutoML Natural Language**
- **AutoML Tables**
- **AutoML translation**
- **AutoML Video Intelligence**
- **AutoML Vision**

The above products provide various tools to train the model for specific use cases with limited machine learning expertise. For cloud AutoML, we don't need to have knowledge of transfer learning or how to create a neural network, as it provides the out-of-box for deep learning models.

2. Microsoft Azure AutoML

Microsoft Azure AutoML, released in 2018, simplifies machine learning model building for non-experts by providing a transparent model selection process and automating key steps such as data preprocessing, feature engineering, and hyperparameter tuning. It enables users to easily experiment with different algorithms and configurations, deploy models as web services, and monitor their performance.

3. H2O.ai

H2O is an open-source platform that enables the user to create ML models. It can be used for automating the machine learning workflow, such as automatic training and tuning of many models within a user-specified time limit. Although H2O AutoML can make the development of ML models easy for the non-experts still, a good knowledge of data science is required to build the high-performing ML models.

4. TPOT

TPOT(Tree-based Pipeline Optimization) can be considered as a Data science assistant for developers. It is a Python packaged Automated Machine Learning tool, which uses genetic

programming to optimize the machine learning pipelines. It is built on the top of the scikit-learn, so it will be easy for the developers to work with it (if they are aware of scikit learn). It automates all the tedious parts of the ML lifecycle by exploring thousands of possible processes to find the best one for the particular requirement. After finishing the search, it provides us with the Python code for the best pipeline.

5. DataRobot

DataRobot is one of the best AutoML tools platforms. It provides complete automation by automating the ML pipeline and supports all the steps required for the preparation, building, deployment, monitoring, and maintaining the powerful AI applications.

6. Auto-Sklearn

Auto-Sklearn is an open-source library built on the top of scikit learn. It automatically does algorithm selection and parameter tuning for a machine learning model. It provides out-of-the-box features of supervised learning.
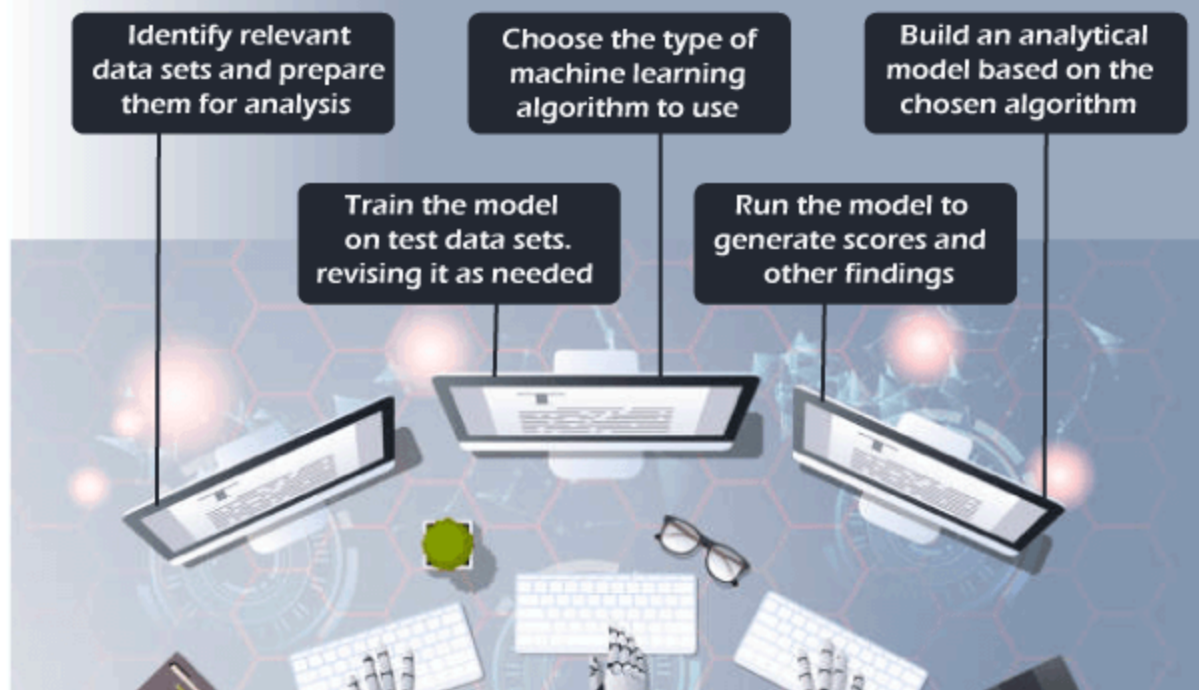
7. MLBox

MLBox also provides the powerful Python Library for automated Machine Learning. It provides a range of features and functionalities to automate various aspects of the ML workflow, making it easier for users to develop machine learning models efficiently.

How does Automated Machine Learning Work?

Automated machine learning or AutoML is an open-source library that automates each step of the machine learning lifecycle, including preparing a dataset to deploy an ML model. It works in a completely different way than the traditional machine learning method, where we need to develop the model manually, and each step is handled separately.

The automated machine learning process

Identify relevant data sets and prepare them for analysis

Choose the type of machine learning algorithm to use

Build an analytical model based on the chosen algorithm

Train the model on test data sets. revising it as needed

Run the model to generate scores and other findings

AutoML automatically selects and locates the optimal and most suitable algorithm as per our problem or given task. It performs by following the two basic concepts:

o **Neural Architecture Search:** It helps in automating the design of neural networks. It enables AutoML models to discover new architectures as per the problem requirement.

o **Transfer Learning:** With the help of transfer learning, previously trained models can apply their logic to new datasets that they have learned. It enables AutoML models to apply available architectures to the new problems.

With AutoML, a Machine learning enthusiast can use Machine learning or deep learning models by using Python language. Moreover, below are the steps that are automated by AutoML that occur in the Machine learning lifecycle or learning process:

o **Raw data processing**
o **Feature engineering**
o **Model selection**
o **Hyperparameter optimization and parameter optimization**
o **Deployment with consideration for business and technology constraints**
o **Evaluation metric selection**
o **Monitoring and problem checking**

o **Result Analysis**

Pros of AutoML

o **Performance:** AutoML performs most of the steps automatically and gives a great performance.
o **Efficiency:** It provides good efficiency by speeding up the machine learning process and by reducing the training time required to train the models.
o **Cost Savings:** As it saves time and the learning process of machine learning models, hence also reduces the cost of developing an ML model.
o **Accessibility:** AutoML enables those with little background in the area to use the potential of ML models by making machine learning accessible to them.
o **Democratization of ML:** AutoML democratises machine learning by making it easier for anybody to use, hence maximising its advantages.

Cons of AutoML

o **Lack of Human Expertise:** AutoML can be considered as a substitute for human knowledge, but human oversight, interpretation, and decision-making are still required.
o **Limited Customization:** Limited customization possibilities on some AutoML systems may make it difficult to fine-tune models to meet particular needs.
o **Dependency on Data Quality:** The accuracy and relevancy of the supplied data are crucial to AutoML. The quality and performance of the generated models may be impacted by biassed, noisy, or missing data.
o **Complexity of Implementation:** Even while AutoML makes many parts of machine learning simpler, incorporating AutoML frameworks into current processes may need more time and technical know-how.
o **Lack of Platform Maturity:** Since AutoML is still a relatively young and developing area, certain platforms could still be in the works and be in need of improvements.

Applications of AutoML

AutoML shares common use cases with traditional machine learning. Some of these include:

o **Image Recognition:** AutoML is also used in image recognition for Facial Recognition.
o **Risk Assessment:** For banking, finance, and insurance, it can be used for Risk Assessment and management.

- o **Cybersecurity:** In the cybersecurity field, it can be used for risk monitoring, assessment, and testing.
- o **Customer Support:** Customer support where can be used for sentiment analysis in chatbots and to increase the efficiency of the customer support team.
- o **Malware & Spam:** To detect malware and spam, AutoML can generate adaptive cyberthreats.
- o **Agriculture:** In the Agriculture field, it can be used to accelerate the quality testing process.
- o **Marketing:** In the Marketing field, AutoML is employed to predict analytics and improve engagement rates. Moreover, it can also be used to enhance the efficiency of behavioral marketing campaigns on social media.
- o **Entertainment:** In the entertainment field, it can be used as the content selection engine.
- o **Retail:** In Retail, AutoML can be used to improve profits and reduce the inventory carry.

Conclusion:

AutoML has taken huge steps in democratizing AI via mechanizing and working on the cycle. It permits people with restricted AI aptitude to tackle the force of ML models. The article gives a prologue to AutoML, talks about well known stages and instruments, makes sense of its functioning standards, and investigates its stars, cons, and applications. By ceaselessly remaining refreshed with the most recent progressions in AutoML, people can completely use its true capacity for different use cases across various ventures.