**AutoML - Automating Machine Learning /Deep Learning tasks & phases**

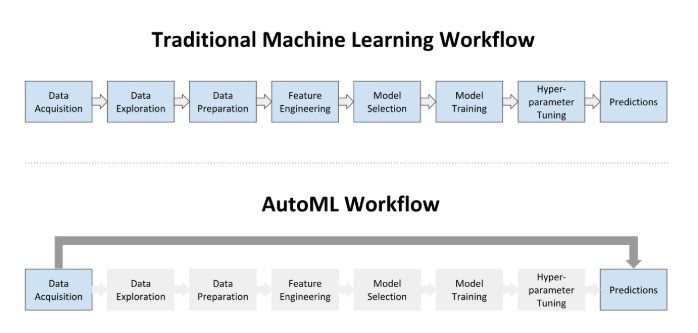
**Why AutoML?**

As we all know the demand for machine learning systems has soared over the past few years. This is because of the success of Machine Learning techniques in a wide range of applications.

AutoML is fundamentally changing the face of ML-based solutions today by enabling people from diverse backgrounds to use machine learning models to address complex scenarios.

However, even with a clear indication that machine learning can provide a boost to certain businesses, a lot of companies today struggle to deploy ML models.

This is because there is a shortage of experienced and seasoned data scientists in the industry. In a way, the demand for machine learning experts has outpaced the supply. Secondly, a lot of machine learning steps require more experience than knowledge, especially when deciding which models to train and how to evaluate them. Such gaps are pretty apparent today, and a lot of efforts are being taken to address these issues. Automated Machine learning may be an answer to such impediments.



**Diagram 1.1 :-Traditional ML Vs AutoML**

**What is AutoML?**

Automated machine learning (AutoML) is the process of automating the end-to-end process of applying machine learning to real-world problems. AutoML tends to automate the maximum number of steps in an ML pipeline — with a minimum amount of human effort — without compromising the model’s performance.

Automated machine learning can be thought of as the standard machine learning process with the automation of some of the steps involved. AutoML very broadly includes:

* Automating certain parts of data preparation, e.g. imputation, standardization, feature selection, etc.
* Being able to generate various models automatically, e.g. random grid search, Bayesian Hyperparameter Optimization, etc.
* Getting the best model out of all the generated models, which most of the time is an Ensemble, e.g. ensemble selection, stacking, etc.

**H2O’s Automatic Machine Learning (AutoML)**

**Introduction to H2O’s AutoML:-**

[**H2O**](https://www.h2o.ai/products/h2o/#overview) is a fully open-source, distributed in-memory machine learning platform with linear scalability. H2O supports the most widely used statistical & machine learning algorithms, including gradient boosted machines, generalized linear models, deep learning, and many more.

H2O also has an industry-leading AutoML functionality (available in H2O ≥3.14) that automates the process of building a large number of models, to find the “**best**” model without any prior knowledge or effort by the Data Scientist. H2O AutoML can be used for automating the machine learning workflow, which includes automatic training and tuning of many models within a user-specified time-limit.

**Features of H2O AutoML**

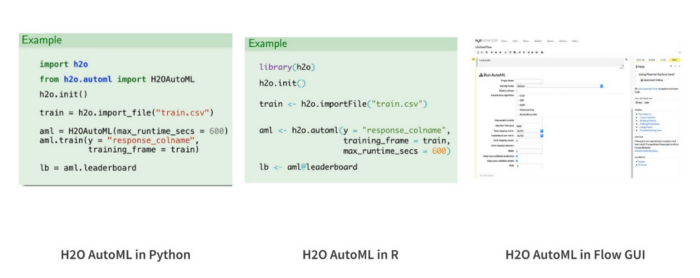
* Open-source, distributed (multi-core + multi-node) implementations of cutting edge ML algorithms.
* Availability of core algorithms in high-performance Java. including APIs in R, Python, Scala, web GUI.



* Easily deployable models to production as pure Java code.
* Seamlessly works on Hadoop, Spark, AWS, your laptop, etc.

## **AutoML Interface**

H2O AutoML has an R and Python interface along with a web GUI called Flow. The H2O AutoML interface is designed to have as few parameters as possible so that all the user needs to do is to point to their dataset, identify the response column and optionally specify a time constraint or limit on the number of total models trained.



H2O AutoML is available in R, Python, and a web GUI.