



Home



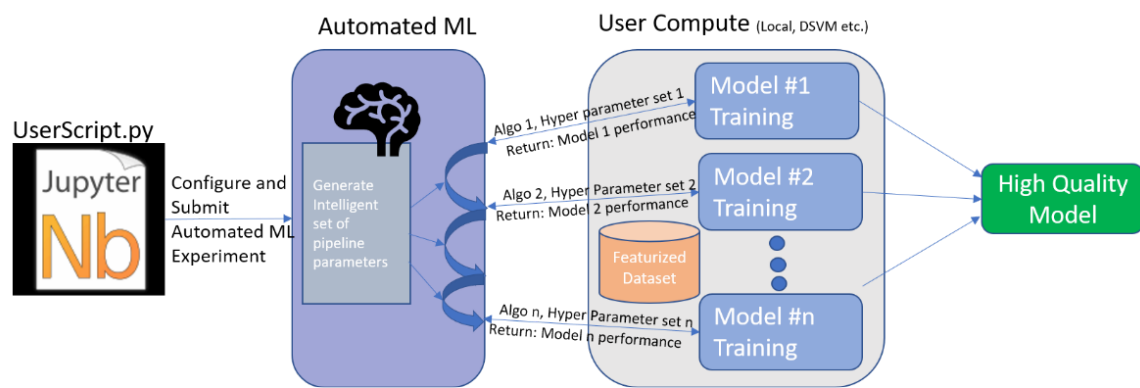
My Network



Jobs

# Top 10 Automated Machine Learning(Auto ML) tools used in 2020-2021

Published on June 6, 2021



**Baishalini Sahu**

Data Scientist || Research And Development || AI ML ||Python||Retail Analytics  
|| Predictive Modelling

6 articles

[+ Follow](#)

## How AutoML developed and its work flow?

- AutoML (automated machine learning) refers to the automated end-to-end process of applying machine learning in real and practical scenarios.
- The boons of machine learning have been leveraged in the industry in the past many years. With its increasing implementation, the ML tools have also evolved with time. Today, people can easily work



been automated enough, people with some knowledge of technology and motivation can work with ML.

- These tools possess the strength to handle the mundane work of collecting data, adding structure and consistency where possible, and then starting the calculation. The modern-day tools can simplify the data gathering process and keeping that information in rows and columns.

### **AutoML has some major advantages:**

- Improve efficiency by automatically running repetitive tasks. This allows data scientists to focus more on problems instead of models.
- Automated ML pipelines also help avoid potential errors caused by manual work.
- AutoML is a big step toward the democratization of machine learning and allows everyone to use ML features.
- Create a model, perform stratified cross validation and evaluate classification metrics
- Automatically tune the hyper-parameters of a classification model
- Analyze model performance using various plots
- Finalize the best model at the end of the experiment
- Make predictions on new / unseen data



## Some important 10 AutoML tools used in today's life:

AutoML has a history of many years. Since last year, many excellent AutoML frameworks have emerged. This article only briefly describes several common frameworks. Subsequent articles will give more information about the use and performance of these frameworks. Below are the five tools that simplify using machine learning algorithms.

### 1. PyCaret

- PyCaret is an open-source, **low-code** machine learning library in Python that aims to reduce the cycle time from hypothesis to insights. It is well suited for **seasoned data scientists** who want to increase the productivity of their ML experiments by using PyCaret in their workflows.



Data  
Preparation



Model  
Training



Hyperparameter  
Tuning



Analysis &  
Interpretability



Model  
Selection

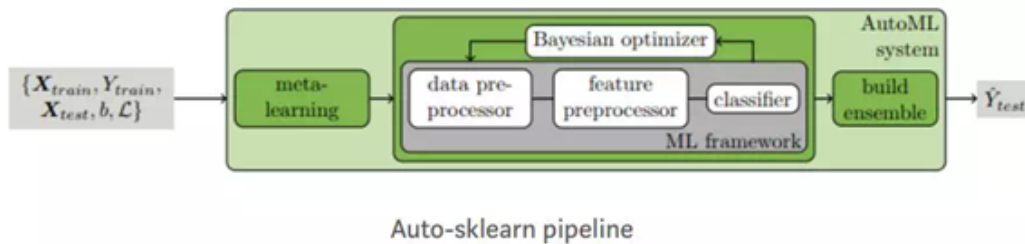


Experiment  
Logging

Resource link: <https://pycaret.org/>



- Auto-SKLearn is an automated machine learning software package built on scikit-learn. Auto-SKLearn frees a machine learning user from **algorithm selection and hyper-parameter tuning**. It includes feature engineering methods such as **One-Hot, digital feature standardization, and PCA**. The model uses SKLearn estimators to process **classification and regression** problems.



- Resource link <https://www.automl.org/automl/auto-sklearn/>
- Auto-SKLearn creates a pipeline and uses Bayes search to optimize that channel. In the ML framework, two components are added for hyperparameter tuning by means of Bayesian reasoning: Meta learning is used to initialize optimizers using **Bayes and evaluate the auto collection construction** of the configuration during the optimization process.
- Auto-SKLearn performs well in medium and small datasets, but it cannot produce modern deep learning systems with the most advanced performance in large datasets.

### 3.MLBox

MLBox is a powerful Automated Machine Learning python library. According to the official document, it provides the features like fast



## hyper-parameter optimization, State-of-the art predictive models for classification and regression (Deep Learning, Stacking, LightGBM,...), Prediction with model interpretation

- MLBox main package contains 3 sub-packages:
- Pre-processing: reading and pre-processing data
- Optimization: testing or optimizing a wide range of learners
- Prediction: predicting the target on a test dataset

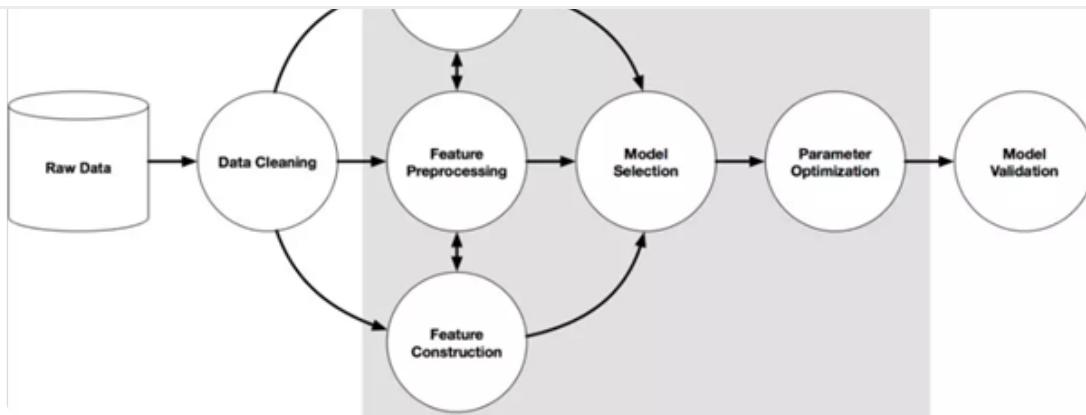


Resource link :

<https://pypi.org/project/mlbox/>

## 4.TPOT

TPOT is a tree-based pipeline optimization tool that uses genetic algorithms to optimize machine learning pipelines. TPOT is built on top of **scikit-learn** and **uses its own regressor and classifier** methods. TPOT explore thousands of possible pipelines and finds the one that best fit the data.



Resource link : <https://wiki.pathmind.com/automl-automated-machine-learning-ai>

## 5.H2O

- H2O is an open source and **distributed in-memory machine learning** platform developed by H2O.ai. H2O supports both R and Python. It supports the most widely used statistical and machine learning algorithms including **gradient boosted machines, generalized linear models, deep learning** and more.
- H2O includes an Automated Machine Learning module and uses its own algorithms to create pipelines. It uses exhaustive search for **feature engineering** methods and model hyper-parameters to optimize pipelines.
- H2O automates some complex data science and machine learning tasks, such as feature engineering, **model validation, model adjustment, model selection, and model deployment**. In addition,



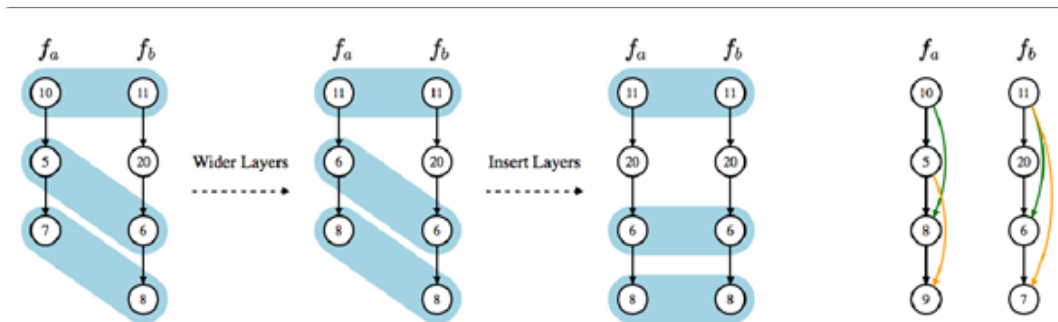
Resource link : <https://www.h2o.ai/products/h2o-automl/>

## 6.Auto-Keras

- Auto-Keras is an open source software library for automated machine learning (AutoML) developed by **DATA Lab**. Built on top of the deep learning framework Keras, Auto-Keras provides functions to automatically search for **architecture and hyper-parameters of deep learning** models.
- Auto-Keras follows the classic Scikit-Learn API design and therefore is easy to use. The current version provides the function to automatically search for hyper-parameters during deep learning.



a set of algorithms that automatically adjust models to replace deep learning engineers/practitioners.

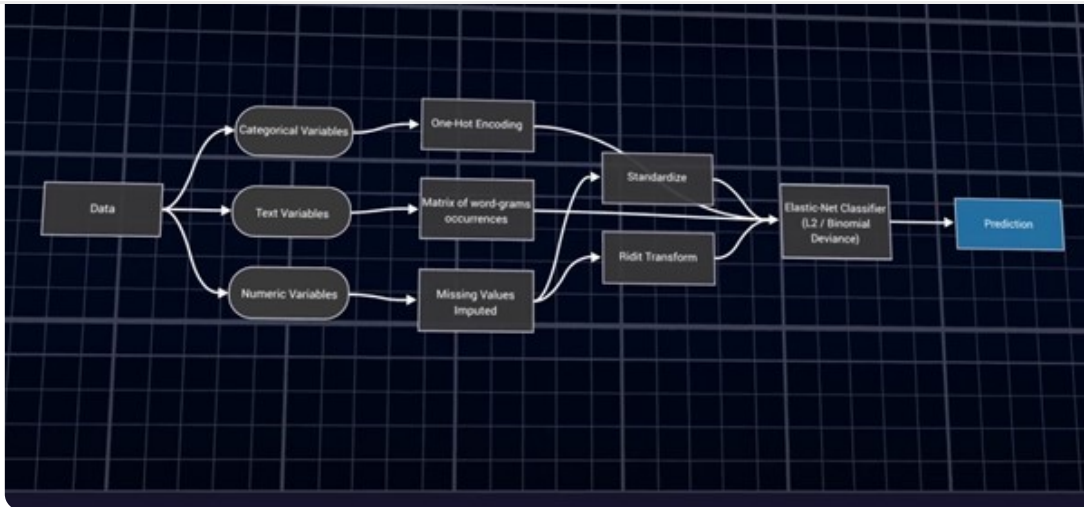


Resource link : <https://autokeras.com/>

## 7.DataRobot

- DataRobot incorporates a variety of regression techniques, ranging from the simplest (linear regression) to complicated **statistical classic regression models, to more complex techniques including gradient boosting and neural networks.**
- The platform can also solve simple binary **classification problems**, as well as highly complex **multiclass classification problems** with up to 100 different categories. Imagine being able to predict which product a customer is likely to purchase next, or why a customer is likely to churn, with a high degree of accuracy.
- With DataRobot it's easy to automate the creation of machine learning models like this – with **unprecedented transparency** so you can understand and trust the predictions they make.





Resource link : <https://www.datarobot.com/>

## 8.BigML

- BigML's AutoML is an Automated Machine Learning tool for BigML. The first version of AutoML helps automate the complete **Machine Learning pipeline**, not only the model selection. To boot, **it's pretty easy to execute**.
- The user needs to give it training and validation datasets and it will give back a Fusion with the best possible models using the least possible number of features. **BigML's AutoML performs three main operations: Feature Generation, Feature Selection, and Model Selection.**



Home

My Network

Jobs



Resource link : <https://bigml.com/tools>

## 9.RapidMiner

- RapidMiner's automated machine learning can **exponentially reduce the time and effort** required to create predictive models for all businesses and organizations regardless of size, resources or industry.
- With its Auto Model, it's possible to build predictive models in just 5 clicks. There's no need for technical expertise. All users need to do is upload his data and specify the outcomes he wants, then Auto Model will produce the **high-value insights** he needs. RapidMiner

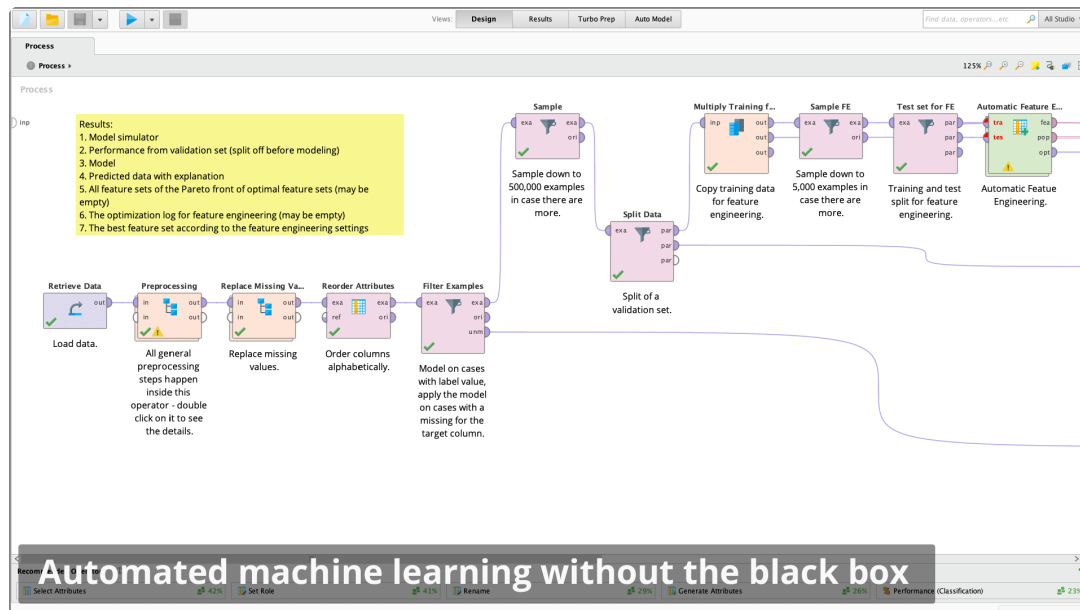


Home

My Network

Jobs

## Turbo Prep and Model Ops in RapidMiner Studio Enterprise.



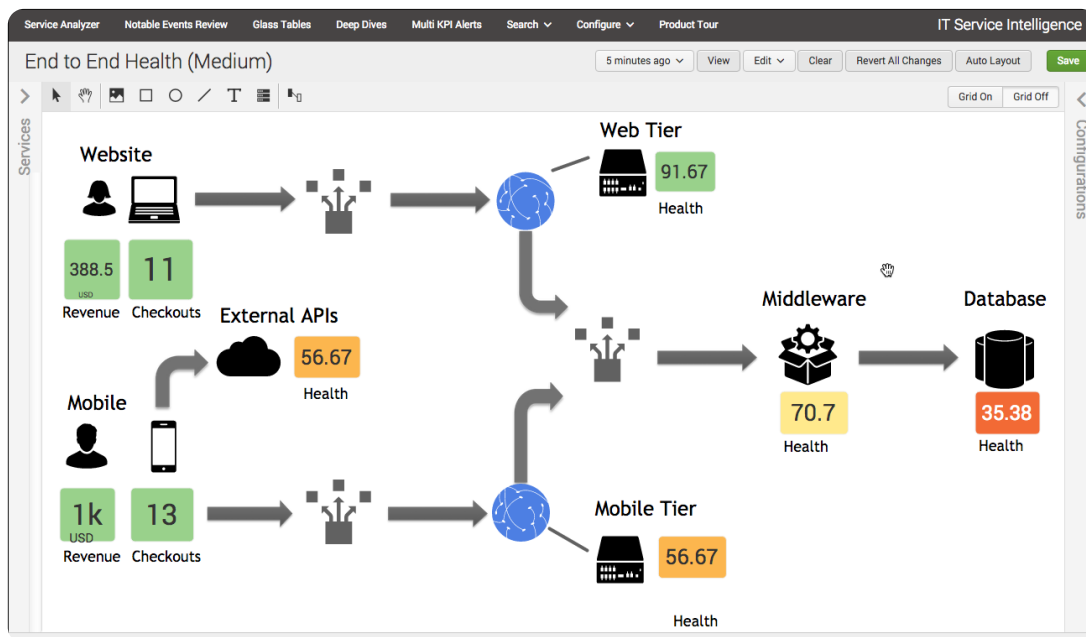
Resource Link: <https://rapidminer.com/>

## 10.Splunk

- Splunk's original version started off as a tool for searching through the **voluminous log files created by modern web applications**. Since then it has grown to analyze all forms of data, especially time-series and others produced in sequence.
- The latest newest versions of Splunk includes apps that **integrate the data sources with machine learning tools like TensorFlow**



and generating predictions for future values.



Resource Link : [https://www.splunk.com/en\\_us/software/splunk-enterprise/machine-learning.html](https://www.splunk.com/en_us/software/splunk-enterprise/machine-learning.html)

## Conclusion:

Data scientists can accelerate ML development by using AutoML to implement really efficient machine learning. The essence of AutoML is to automate repetitive tasks such as pipeline creation and hyper-parameter tuning so that data scientists can spend more time on business problems on hand in practical scenarios. AutoML also allows everyone instead a small group of people to use the machine learning technology.



Home

My Network

Jobs

part of machine learning in the future.

Thank you By Baishalini Sahu

[Report this](#)

Published by

**Baishalini Sahu**

Data Scientist || Research And Development || AI ML  
||Python||Retail Analytics || Predictive Modelling  
Published • 5mo

6

articles

[+ Follow](#)

[#lowcode](#) [#automl](#) [#automatedmachinelearning](#) [#Autodeeplearning](#) [#pycaret](#)  
[#datarobot](#) [#h2o](#) [#splunk](#) [#Autokeras](#) [#Autosklearn](#) [#BigML](#) [#artificialintelligence](#)  
[#innovation](#) [#2021tech](#)



Like



Comment



Share



51 • 3 comments

Reactions



+39

3 Comments

Most relevant ▼



Add a comment...

**Anna Guseva** • 3rd+

1d ...

Marketing Content Manager at GBKSOFT - Web & Mobile App Development

Machine Learning is probably one of the most influential technologies in 2021. I think it can really transform so many businesses and automate numerous

[Home](#)[My Network](#)[Jobs](#)

Earlier updates of all models in the future. Here is more info about it.

[Like](#) | [Reply](#)**Dyutimoy Mohanta** • 3rd+

5mo ...

Electronics & Communication student. Trained in VLSI (verilog). KNOWS SQL & Computer networking concepts. A Data Enthusiastic lover knows statistics for machine learning, python for data science, & basics of power BI

Tq for sharing this

[Like](#) · [2](#) | [Reply](#)[Load more comments](#)**Baishalini Sahu**

Data Scientist || Research And Development || AI ML ||Python||Retail Analytics || Predictive Modelling

[+ Follow](#)

## More from Baishalini Sahu



**Artificial Intelligence powered for Retail in 2021: Real-World Use Cases**

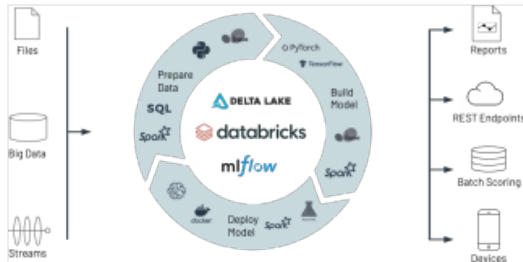
Baishalini Sahu on LinkedIn



**What evaluation approaches**

[Home](#)[My Network](#)[Jobs](#)

Baishalini Sahu on LinkedIn



### Databricks with Machine Learning flow all in one solution #2021

Baishalini Sahu on LinkedIn



### Most useful 2021 python libraries for data science and web development

Baishalini Sahu on LinkedIn

[See all 6 articles](#)