Title: Automated machine learning

URL: https://en.wikipedia.org/wiki/Automated_machine_learning

PageID: 55843837

Categories: Category:Artificial intelligence, Category:Computational neuroscience, Category:Machine learning

-----

Supervised learning

Unsupervised learning

Semi-supervised learning

Self-supervised learning

Reinforcement learning

Meta-learning

Online learning

Batch learning

Curriculum learning

Rule-based learning

Neuro-symbolic AI

Neuromorphic engineering

Quantum machine learning

Classification

Generative modeling

Regression

Clustering

Dimensionality reduction

Density estimation

Anomaly detection

Data cleaning

AutoML

Association rules

Semantic analysis

Structured prediction

Feature engineering

Feature learning

Learning to rank

Grammar induction

Ontology learning

Multimodal learning

Apprenticeship learning

Decision trees

Ensembles Bagging Boosting Random forest

Bagging

Boosting

Random forest

k -NN

Linear regression

Naive Bayes

Artificial neural networks

Logistic regression

Perceptron

Relevance vector machine (RVM)

Support vector machine (SVM)

BIRCH

CURE

Hierarchical

k -means

Fuzzy

Expectation–maximization (EM)

DBSCAN

OPTICS

Mean shift

Factor analysis

CCA

ICA

LDA

NMF

PCA

PGD

t-SNE

SDL

Graphical models Bayes net Conditional random field Hidden Markov

Bayes net

Conditional random field

Hidden Markov

RANSAC

k -NN

Local outlier factor
Isolation forest
Autoencoder
Deep learning
Feedforward neural network
Recurrent neural network LSTM GRU ESN reservoir computing
LSTM
GRU
ESN
reservoir computing
Boltzmann machine Restricted
Restricted
GAN
Diffusion model
SOM
Convolutional neural network U-Net LeNet AlexNet DeepDream
U-Net
LeNet
AlexNet
DeepDream
Neural field Neural radiance field Physics-informed neural networks
Neural radiance field
Physics-informed neural networks
Transformer Vision
Vision
Mamba
Spiking neural network
Memtransistor
Electrochemical RAM (ECRAM)
Q-learning
Policy gradient
SARSA
Temporal difference (TD)
Multi-agent Self-play
Self-play
Active learning
Crowdsourcing
Human-in-the-loop

Mechanistic interpretability

RLHF

Coefficient of determination

Confusion matrix

Learning curve

ROC curve

Kernel machines

Bias–variance tradeoff

Computational learning theory

Empirical risk minimization

Occam learning

PAC learning

Statistical learning

VC theory

Topological deep learning

AAAI

ECML PKDD

NeurIPS

ICML

ICLR

IJCAI

ML

JMLR

Glossary of artificial intelligence

List of datasets for machine-learning research List of datasets in computer vision and image processing

List of datasets in computer vision and image processing

Outline of machine learning

v

t

e

Automated machine learning (AutoML) is the process of automating the tasks of applying machine learning to real-world problems. It is the combination of automation and ML.

AutoML potentially includes every stage from beginning with a raw dataset to building a machine learning model ready for deployment. AutoML was proposed as an artificial intelligence -based solution to the growing challenge of applying machine learning. The high degree of automation in AutoML aims to allow non-experts to make use of machine learning models and techniques without requiring them to become experts in machine learning. Automating the process of applying machine learning end-to-end additionally offers the advantages of producing simpler solutions, faster creation of those solutions, and models that often outperform hand-designed models.

Common techniques used in AutoML include hyperparameter optimization , meta-learning and neural architecture search .

Comparison to the standard approach

In a typical machine learning application, practitioners have a set of input data points to be used for training. The raw data may not be in a form that all algorithms can be applied to. To make the data amenable for machine learning, an expert may have to apply appropriate data pre-processing , feature engineering , feature extraction , and feature selection methods. After these steps, practitioners must then perform algorithm selection and hyperparameter optimization to maximize the predictive performance of their model. If deep learning is used, the architecture of the neural network must also be chosen manually by the machine learning expert.

Each of these steps may be challenging, resulting in significant hurdles to using machine learning. AutoML aims to simplify these steps for non-experts, and to make it easier for them to use machine learning techniques correctly and effectively.

AutoML plays an important role within the broader approach of automating data science , which also includes challenging tasks such as data engineering, data exploration and model interpretation and prediction.

Targets of automation

Automated machine learning can target various stages of the machine learning process. Steps to automate are:

Data preparation and ingestion (from raw data and miscellaneous formats) Column type detection; e.g., Boolean, discrete numerical, continuous numerical, or text Column intent detection; e.g., target/label, stratification field, numerical feature, categorical text feature, or free text feature Task detection; e.g., binary classification , regression , clustering , or ranking

Column type detection; e.g., Boolean, discrete numerical, continuous numerical, or text

Column intent detection; e.g., target/label, stratification field, numerical feature, categorical text feature, or free text feature

Task detection; e.g., binary classification , regression , clustering , or ranking

Feature engineering Feature selection Feature extraction Meta-learning and transfer learning Detection and handling of skewed data and/or missing values

Feature selection

Feature extraction

Meta-learning and transfer learning

Detection and handling of skewed data and/or missing values

Model selection - choosing which machine learning algorithm to use, often including multiple competing software implementations

Ensembling - a form of consensus where using multiple models often gives better results than any single model

Hyperparameter optimization of the learning algorithm and featurization Neural architecture search

Neural architecture search

Pipeline selection under time, memory, and complexity constraints

Selection of evaluation metrics and validation procedures

Problem checking Leakage detection Misconfiguration detection

Leakage detection

Misconfiguration detection

Analysis of obtained results

Creating user interfaces and visualizations

Challenges and Limitations

There are a number of key challenges being tackled around automated machine learning. A big issue surrounding the field is referred to as "development as a cottage industry". This phrase refers to the issue in machine learning where development relies on manual decisions and biases of experts. This is contrasted to the goal of machine learning which is to create systems that can learn and improve from their own usage and analysis of the data. Basically, it's the struggle between how much experts should get involved in the learning of the systems versus how much freedom they should be giving the machines. However, experts and developers must help create and guide these machines to prepare them for their own learning. To create this system, it requires labor intensive work with knowledge of machine learning algorithms and system design .

Additionally, other challenges include meta-learning and computational resource allocation.

See also

Artificial intelligence

Artificial intelligence and elections

Neural architecture search

Neuroevolution

Self-tuning

Neural Network Intelligence

ModelOps

Hyperparameter optimization

References

Further reading

"Open Source AutoML Tools: AutoGluon, TransmogrifAI, Auto-sklearn, and NNI" . Bizety . 2020-06-16.

Ferreira, Luís, et al. "A comparison of AutoML tools for machine learning, deep learning and XGBoost." 2021 International Joint Conference on Neural Networks (IJCNN). IEEE, 2021. https://repositorium.sdum.uminho.pt/bitstream/1822/74125/1/automl_ijcnn.pdf

Feurer, M., Klein, A., Eggensperger, K., Springenberg, J., Blum, M., & Hutter, F. (2015). Efficient and robust automated machine learning. Advances in neural information processing systems , 28 . https://proceedings.neurips.cc/paper_files/paper/2015/file/11d0e6287202fced83f79975ec59a3a6-Paper.pdf

Differentiable programming

Information geometry

Statistical manifold

Automatic differentiation

Neuromorphic computing

Pattern recognition

Ricci calculus

Computational learning theory

Inductive bias

IPU

TPU

VPU

Memristor

SpiNNaker

TensorFlow

PyTorch

Keras

scikit-learn

Theano

JAX

Flux.jl

MindSpore

Portals Computer programming Technology

Computer programming

Technology