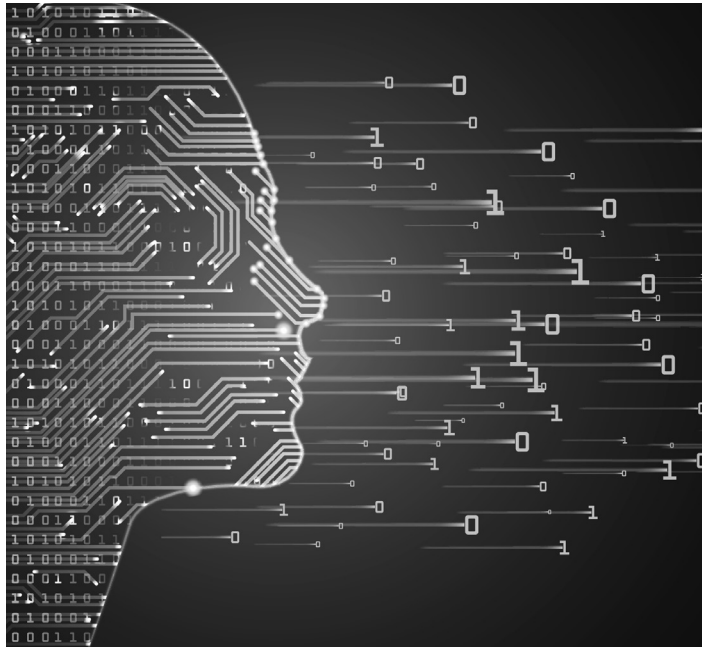
H2O.ai

Machine learning for all: the democratizing of a technology

How the benefits of machine learning can be leveraged by anyone

Opening up the way to enterprise-level machine learning adoption



Data is proliferating exponentially, and with it, the potential to draw business-enhancing, problem-solving insights using artificial intelligence (AI) and machine learning (ML). However, three big challenges—talent, time, and trust—often prevent organizations from adopting enterprise-level AI. Conventional, enterprise-level AI/ML efforts are complex and usually require data scientists. Projects can involve months of data preparation followed by a process of designing, developing, testing, tuning, and deploying a machine learning model. Many available solutions are difficult to trust. They consist either of complex, non-linear models and ensembles that are almost impossible to interpret by humans, or pre-built models that provide no insight into how the predictions are being made.

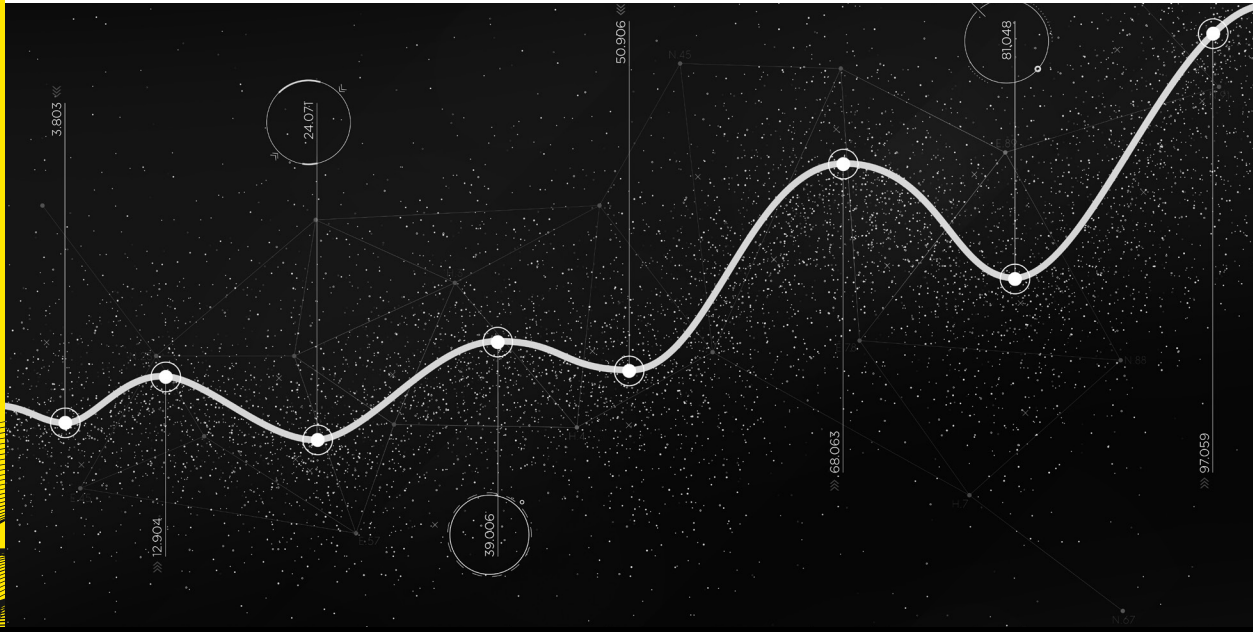
However, solutions are emerging that place the power of advanced machine learning within reach—even for companies with minimal data science experience. The democratization of AI/ML has begun.

This eBook covers machine learning features and automatic AI solutions, and how organizations can benefit from using them.

Automatic feature engineering

An algorithm is the underlying mathematical logic of an ML solution, and there are different types, depending on the problem being solved. Whatever an organization's ML objectives, a strong automatic ML solution will have access to a broad range of algorithms.

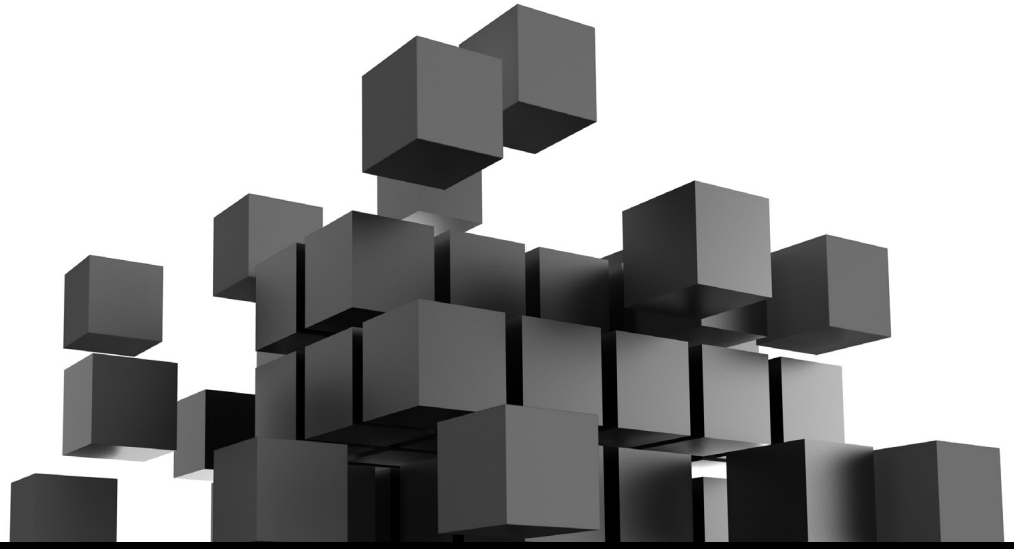
Algorithms need a basis for understanding the data they process, which is accomplished through a process of transforming raw data into features that better represent the underlying structures and patterns in the data, optimizing what the algorithm can learn from the data. This is called feature engineering.



Automatic model building

With a representative feature table, the algorithm is able to build a model of what it expects the data will show. It can then compare the model to the data, learn from what it observes, and modify the model iteratively so that it grows increasingly accurate. Coming up with the right tuning parameters to account for and evaluate all of the possible features and algorithms results in an optimal pipeline that achieves the highest accuracy possible.

The complex and iterative nature of feature engineering puts data scientists in high demand, making them hard to find. Fortunately, machine learning as a technology has reached a point where feature engineering can be done automatically with a high degree of accuracy, allowing organizations to accomplish this important ML workflow without data scientists.



Specialized machine learning functions that solve business challenges

To address the broad range of objectives that organizations have for ML, specialized technologies are looped in to the machine learning workflow to solve for different types of problems.

For instance, natural language processing (NLP) can process and analyze text-based, unstructured data, which is characteristic of customer feedback use cases. Time series analysis and forecasting can be implemented to evaluate budgets, forecast sales, conduct inventory studies, and improve manufacturing efficiency.

These specialized functions and others like them can be added to automatic machine learning solutions and help organizations increase profits, minimize risk, and achieve their objectives. They also enable the development of an extensible platform that can handle current and future use cases.

Explaining AI results: Machine learning interpretability (MLI)

MLI refers to the ML solution's ability to interpret what the data is saying, usually using non-linear models that can account for patterns and nuances in the data. It aims to help humans understand or explain the results of models with actionable insights. There are many techniques and frameworks that are used for this purpose.

Interpretability is also critical for regulated industries like financial services, insurance, and healthcare to help ensure that the machine learning models are safe, trusted, and pass regulatory oversight.

Some of the more commonly-used interpretability models include K-LIME, Shapley, variable importance, decision tree, and partial dependence.

Any machine learning solution that an organization adopts needs to have strong automatic MLI capabilities.

Deploying a scored machine learning model

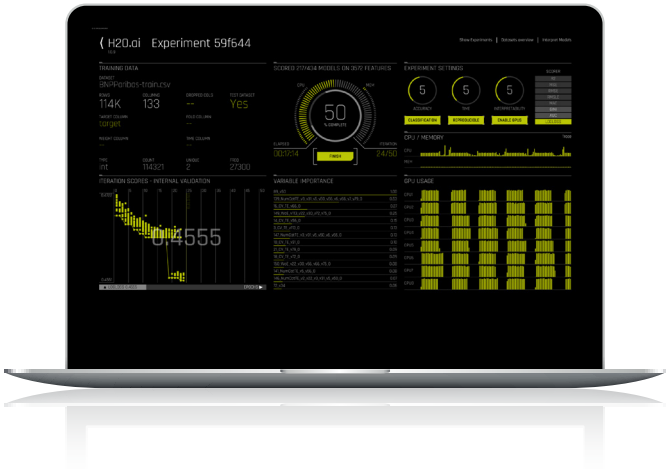
The goal of all machine learning is for the output to be deployed. This final workflow is known as the scoring pipeline, where the machine learning interfaces with a production application.

Prior to being displayed in an application or rendered on a screen, the output of machine learning is fed through an evaluative or predictive model; it is *scored*. Once the output is scored, it is ready to be deployed and put to use by its intended audience.

Using conventional data science deployment methodologies, deploying a scored model can involve considerable time and effort in translating data science output to application- or production-ready artifacts. This is especially true of enterprise-level organizations that need to deploy low-latency scoring engines capable of delivering sub-millisecond inferencing for real-time applications.

Fortunately, with the right machine learning solution, organizations can deploy even complex scored models automatically.

Discover automatic machine learning using H2O.ai



H2O.ai has dedicated itself to democratizing all aspects of AI, including machine learning. H2O Driverless AI is a machine learning solution that automates AI for non-technical users. With it, users can build robust, fast, and accurate machine learning solutions. It also includes visualization and interpretability features that explain the data modeling results in plain English, fostering further adoption and trust in AI.

H2O.ai is an Amazon Web Services (AWS) Machine Learning Competency Partner. Its products work with AWS services to provide such functionality as computer vision using Amazon Rekognition, and customer contact center services using Amazon Connect.

With H2O Driverless AI from H2O.ai, you can:

- Automatically engineer model features.
- Power your efforts with the compute capabilities of Amazon EC2.
- Store your data using Amazon S3 and other AWS storage services.
- Automate model tuning and stacking.
- Deploy models using AWS Lambda.

LEARN MORE ABOUT H2O.AI