

AutoGluon Training

3 days (21 hours)

Overview

AutoGluon is an AutoML library designed to accelerate the creation of high-performance models on tabular, multimodal, and time-series data. By automating data preprocessing, model selection, and hyperparameter tuning, it helps build robust, production-ready solutions. Our AutoGluon training will enable you to go beyond basic usage and master the essential mechanisms of an automated machine learning pipeline. You'll learn how to prepare your data, configure key predictors, compare candidate models, and leverage model ensembles to improve prediction quality. By the end of the training, you'll be able to design reliable experiments, interpret results, choose the right evaluation metrics, and structure a reproducible workflow with AutoGluon. This training also covers best practices for deployment, performance management, adaptation to business use cases, and the mindset needed to systematically scale your models.

Like all our training courses, this one will introduce you to **the latest stable version** of the technology and its new features.

Objectives

- Understand AutoGluon's position within the AutoML ecosystem.
- Prepare datasets compatible with the main types of predictions.
- Configure and train models for tabular, multimodal, and time-series data.
- Evaluate performance using appropriate metrics and compare the selected approaches.
- Optimize an end-to-end pipeline to achieve reproducible and actionable results.
- Apply best practices to transition from a prototype to a more robust business application.

Target Audience

- Data Scientist
- Data Analyst

- Machine Learning Engineer
- Data Engineer
- Python Developer

Prerequisites

- Proficiency in the fundamentals of Python and data manipulation.
- Fundamental knowledge of machine learning, supervised learning, and model evaluation.
- Experience with Jupyter Notebook or an equivalent IDE.
- Understanding of the concepts of training, validation, and test datasets.
- Prior experience with libraries such as pandas or scikit-learn.

Technical Prerequisites

- A recent computer with sufficient memory to run training and testing runs.
- Python 3.10 through 3.13 installed.
- A development environment with Jupyter Notebook, VS Code, or a compatible IDE.
- Stable internet connection for installing dependencies and accessing datasets.
- A microphone and headset are recommended for remote sessions.

Our AutoGluon Training Program

[Day 1 - Morning] AutoML

Fundamentals

- Understand the principles of AutoML and the added value of AutoGluon.
- Identify the main types of supported tasks, including classification, regression, and time series.
- Explore the library's overall architecture and its main modules.
- Analyze the lifecycle of an experiment, from raw data to a deployable model.
- Identify business use cases where automation accelerates deployment.
- Hands-on workshop: Getting started with AutoGluon, running an initial training session on a simple tabular dataset, and interpreting the results.

[Day 1 - Afternoon]

Preparing Tabular Data

- Structure a dataset for tabular prediction with the correct column types.
- Understand the role of the label, explanatory variables, and identifiers.
- Handling missing values, categorical variables, and numerical variables.
- Set up a consistent separation between training, validation, and testing.
- Explore TabularPredictor's loading and configuration options.
- Hands-on workshop: cleaning a dataset, preparing variables, and training an initial tabular model with evaluation on a test set.

[Day 2 - Morning]

Training and Comparing Models

- Explore the logic behind the automatic selection of candidate models.
- Understand the benefits of ensembles and model stacking for improving performance.
- Configuring the main presets and their impact on training time.
- Interpret scores, model rankings, and validation metrics.
- Identify tuning parameters to balance accuracy, speed, and computational cost.
- Hands-on workshop: comparison of several AutoGluon configurations and analysis of performance differences between models.

[Day 2 - Afternoon]

Optimization and

Interpretability

- Leverage built-in hyperparameter optimization mechanisms.
- Study variable importance methods and prediction explanation techniques.
- Analyze prediction errors to better target improvements.
- Adjust settings according to business constraints and quality objectives.
- Set up a reproducible pipeline to quickly iterate on experiments.
- Hands-on workshop: hyperparameter tuning, interpreting variable importance, and guided improvement of an existing model.

[Day 3 - Morning]

Multimodal and Time Series

- Discover the benefits of AutoGluon Multimodal for combining text, images, and structured data.
- Understand the specific characteristics of data flows for multimodal prediction.
- Discuss time series preparation and long-format constraints.
- Use TimeSeriesPredictor to run probabilistic forecasts.
- Select appropriate metrics based on the forecasting context.
- Hands-on workshop: Implementing a multimodal or time-series use case, including training and comparing predictions.

[Day 3 - Afternoon] Deployment and

Best Practices

- Organize an end-to-end workflow for deployment to production.
- Save, reload, and reuse trained models.
- Establish a process for validating and monitoring performance.
- Identify the limitations of the AutoML framework based on project constraints.
- Apply best practices for documentation, reproducibility, and maintenance.
- Hands-on workshop: building a complete mini-project, from data preparation to delivering a model ready for production.

Target Audience

This training is intended for both individuals and companies—large or small—that wish to train their teams in a new advanced IT technology or to acquire specific business knowledge or modern methods.

Entry-Level Assessment

The pre-training assessment complies with Qualiopi quality standards. Upon final registration, the learner receives a self-assessment questionnaire that allows us to evaluate their estimated proficiency in various types of technologies, as well as their expectations and personal goals for the upcoming training, within the limits imposed by the selected format. This questionnaire also allows us to anticipate certain connection or internal security issues within the company (intra-company or virtual classroom) that could pose challenges for monitoring and ensuring the smooth running of the training session.

Teaching Methods

Hands-On Training: 60% hands-on, 40% theory. Training materials distributed in digital format to all participants.

Organization

The course alternates between theoretical input from the instructor—supported by examples and reflection sessions—and group work.

Assessment

At the end of the session, a multiple-choice quiz is administered to verify that participants have successfully acquired the skills.

Certification

A certificate will be issued to each trainee who has completed the entire training program.

