

Mis à jour le 26/08/2026

S'inscrire

Formation AutoGluon

3 jours (21 heures)

Présentation

AutoGluon est une bibliothèque AutoML conçue pour accélérer la création de modèles performants sur des données tabulaires, multimodales et de séries temporelles. Avec l'automatisation du prétraitement, la sélection de modèles et l'optimisation d'hyperparamètres, elle aide à bâtir des solutions robustes et prêtes pour la production. Notre formation AutoGluon vous permettra d'aller au-delà des usages de base pour maîtriser les mécanismes essentiels d'un pipeline de machine learning automatisé. Vous apprendrez à préparer vos données, à configurer les principaux predictors, à comparer les modèles candidats et à exploiter les ensembles de modèles pour améliorer la qualité des prédictions. À l'issue de la formation, vous serez en mesure de concevoir des expérimentations fiables, d'interpréter les résultats, de choisir les bons indicateurs d'évaluation et de structurer un flux de travail reproductible avec AutoGluon. Cette formation aborde également les bonnes pratiques de mise en production, la gestion des performances, l'adaptation aux cas d'usage métiers et les réflexes nécessaires pour industrialiser vos modèles avec méthode.

Comme toutes nos formations, celle-ci vous présentera **la dernière version stable** de la technologie et ses nouveautés.

Objectifs

- Comprendre le positionnement de AutoGluon dans l'écosystème AutoML.
- Préparer des jeux de données compatibles avec les principaux types de prédiction.
- Configurer et entraîner des modèles pour les données tabulaires, multimodales et time series.
- Évaluer les performances avec des métriques adaptées et comparer les approches retenues.
- Optimiser un pipeline de bout en bout pour obtenir des résultats reproductibles et exploitables.
- Appliquer des bonnes pratiques pour passer d'un prototype à un usage métier plus robuste.

Public visé

- Data Scientist
- Data Analyst

- Machine Learning Engineer
- Ingénieur Data
- Développeur Python

Pré-requis

- Maîtrise des bases de Python et de la manipulation de données.
- Connaissances fondamentales en machine learning, apprentissage supervisé et évaluation de modèles.
- Pratique d'un environnement Jupyter Notebook ou d'un IDE équivalent.
- Compréhension des notions de jeux d'entraînement, validation et test.
- Expérience préalable avec des bibliothèques comme pandas ou scikit-learn.

Pré-requis techniques

- Ordinateur récent avec suffisamment de mémoire pour exécuter des entraînements de test.
- Installation de Python 3.10 à Python 3.13.
- Environnement de travail avec Jupyter Notebook, VS Code ou un IDE compatible.
- Connexion réseau stable pour l'installation des dépendances et l'accès aux jeux de données.
- Micro et casque recommandés pour les sessions à distance.

Programme de notre formation AutoGluon

[Jour 1 - Matin]

Fondamentaux de l'AutoML

- Comprendre les principes de l'AutoML et la valeur ajoutée d'AutoGluon.
- Identifier les grands types de tâches prises en charge, dont la classification, la régression et les séries temporelles.
- Découvrir l'architecture générale de la bibliothèque et ses principaux modules.
- Analyser le cycle de vie d'une expérimentation, de la donnée brute au modèle exploitable.
- Repérer les cas d'usage métiers où l'automatisation accélère la mise en production.
- Atelier pratique : prise en main d'AutoGluon, lancement d'un premier entraînement sur un jeu de données tabulaire simple et lecture des résultats.

[Jour 1 - Après-midi]

Préparation des données tabulaires

- Structurer un jeu de données pour la prédiction tabulaire avec les bons types de colonnes.
- Comprendre le rôle du label, des variables explicatives et des identifiants.
- Gérer les valeurs manquantes, les variables catégorielles et les variables numériques.
- Mettre en place une séparation cohérente entre apprentissage, validation et test.
- Découvrir les options de chargement et de configuration de TabularPredictor.
- Atelier pratique : nettoyage d'un dataset, préparation des variables et entraînement d'un premier modèle tabulaire avec évaluation sur un jeu de test.

[Jour 2 - Matin]

Entraînement et comparaison des modèles

- Explorer la logique de sélection automatique des modèles candidats.
- Comprendre l'intérêt des ensembles et du stacking pour améliorer la performance.
- Paramétrer les principaux presets et leurs impacts sur le temps d'entraînement.
- Interpréter les scores, les classements de modèles et les métriques de validation.
- Identifier les leviers de réglage pour équilibrer précision, vitesse et coût calculatoire.
- Atelier pratique : comparaison de plusieurs configurations AutoGluon et analyse des écarts de performance entre modèles.

[Jour 2 - Après-midi]

Optimisation et interprétabilité

- Exploiter les mécanismes d'optimisation d'hyperparamètres intégrés.
- Étudier les méthodes d'importance des variables et d'explication des prédictions.
- Analyser les erreurs de prédiction pour mieux cibler les améliorations.
- Adapter les réglages selon les contraintes métier et les objectifs de qualité.
- Préparer un pipeline reproductible pour itérer rapidement sur les expérimentations.
- Atelier pratique : ajustement des hyperparamètres, lecture des importances de variables et amélioration guidée d'un modèle existant.

[Jour 3 - Matin]

Multimodal et séries temporelles

- Découvrir les apports de AutoGluon Multimodal pour combiner texte, image et données structurées.
- Comprendre les spécificités des flux de données pour la prédiction multimodale.
- Aborder la préparation des séries temporelles et les contraintes de format long.
- Utiliser TimeSeriesPredictor pour lancer des prévisions probabilistes.
- Choisir des métriques adaptées selon le contexte de prévision.
- Atelier pratique : mise en œuvre d'un cas d'usage multimodal ou temporel, avec entraînement et comparaison des prédictions.

[Jour 3 - Après-midi]

Industrialisation et bonnes pratiques

- Organiser un flux de travail de bout en bout pour le passage en production.
- Sauvegarder, recharger et réutiliser les modèles entraînés.
- Mettre en place une démarche de validation et de suivi des performances.
- Identifier les limites du cadre AutoML selon les contraintes de projet.
- Appliquer les bonnes pratiques de documentation, de reproductibilité et de maintenance.
- Atelier pratique : construction d'un mini-projet complet, depuis la préparation des données jusqu'à la restitution d'un modèle prêt à être industrialisé.

Sociétés concernées

Cette formation s'adresse à la fois aux particuliers ainsi qu'aux entreprises, petites ou grandes, souhaitant former ses équipes à une nouvelle technologie informatique avancée ou bien à acquérir des connaissances métiers spécifiques ou des méthodes modernes.

Positionnement à l'entrée en formation

Le positionnement à l'entrée en formation respecte les critères qualité Qualiopi. Dès son inscription définitive, l'apprenant reçoit un questionnaire d'auto-évaluation nous permettant d'apprécier son niveau estimé sur différents types de technologies, ses attentes et objectifs personnels quant à la formation à venir, dans les limites imposées par le format sélectionné. Ce questionnaire nous permet également d'anticiper certaines difficultés de connexion ou de sécurité interne en entreprise (intraentreprise ou classe virtuelle) qui pourraient être problématiques pour le suivi et le bon déroulement de la session de formation.

Méthodes pédagogiques

Stage Pratique : 60% Pratique, 40% Théorie. Support de la formation distribué au format numérique à tous les participants.

Organisation

Le cours alterne les apports théoriques du formateur soutenus par des exemples et des séances de réflexions, et de travail en groupe.

Validation

À la fin de la session, un questionnaire à choix multiples permet de vérifier l'acquisition correcte des compétences.

Sanction

Une attestation sera remise à chaque stagiaire qui aura suivi la totalité de la formation.

