

**Applied Data Science and Practical Machine Learning with AWS SageMaker and AutoML**

**Course Number:** DATA-134WA  
**Duration:** 5 days

**Overview**

This Applied Data Science and Practical Machine Learning with AWS SageMaker and AutoML training course teaches attendees the latest Machine Learning (ML) techniques. Students learn the fundamentals of ML, including exploratory data analysis, model building, and ML explainability. Participants also learn how to use the latest AutoML tools and techniques, such as auto-sklearn, H2O, Auto-Keras, and AWS Auto Pilot. Finally, attendees learn how to use AWS SageMaker to train, evaluate, and deploy models. This data science ML course also includes advanced topics, including neural networks, deep learning, transfer learning, and fine-tuning.

**Prerequisites**

* Proficiency in Python programming
* Understanding of data analysis and manipulation techniques
* Familiarity with Python Pandas or Numpy is recommended
* Basic knowledge of Machine Learning concepts, algorithms, and model evaluation
* Familiarity with AWS and some experience with S3, IAM, and EC2 services

**Materials**

All Data Science and Machine Learning training students receive courseware covering the topics in the class.

**Software Needed on Each Student PC**

* A modern web browser and an Internet connection that allows connections by SSH or Remote Desktop (RDP) into AWS virtual machines.
* Windows, Mac, or Linux with at least 8 GB RAM
* A current version of Anaconda for Python 3.x
* Related lab files that Accelebrate will provide

**Objectives**

* Understand the data science life cycle
* Set up a SageMaker environment
* Train and evaluate ML models using SageMaker
* Deploy ML models
* Work with an AWS AutoML or auto-sklearn environment
* Work with ML models using H2O's automated machine learning
* Understand neural networks and deep learning

**Outline**

* Introduction
* Data Processing Phases and the Data Science Life Cycle
  + Introduction to the data science life cycle
  + Data exploration and visualization
  + Data cleaning and preprocessing
  + Feature engineering
  + Model selection and evaluation
  + Tuning ML: data, parameters, hyperparameters, and artifacts
  + MLI, tuning through data selection/enrichment, analyzing the manifold
  + MLI tools and techniques
* Working with ML Algorithms on SageMaker
  + Introduction to SageMaker
  + Setting up a SageMaker environment
  + Training and evaluating ML models using SageMaker's built-in algorithms
  + Fine-tuning ML models using SageMaker's hyperparameter tuning
* Deploying ML Models as Executable Artifacts
  + Introduction to deploying ML models as executable artifacts
  + Creating and deploying ML models as REST APIs using SageMaker
  + Updating and serving ML models using SageMaker's A/B testing and blue/green deployments
* AWS AutoML and Auto-sklearn
  + Introduction to Canvas and AWS AutoML
  + Costs and examples
  + AutoML as auto-hyperparameter tuning with auto-sklearn
  + Setting up an AWS AutoML or auto-sklearn environment
  + Training and evaluating ML models using AWS AutoML or auto-sklearn
  + Fine-tuning ML models using AWS AutoML or auto-sklearn's hyperparameter tuning
* Automated Machine Learning with H2O
  + Fully automated ML (auto parameter tuning and auto feature engineering)
  + H2O libraries, driverless AI, etc
  + H2O automl vs auto-sklearn (libraries compared/side-by-side0
  + Introduction to H2O and its automated machine learning capabilities
  + Setting up an H2O environment (mention JRE req’s)
  + Training and evaluating ML models using H2O's automated machine learning
  + Fine-tuning ML models using H2O's hyperparameter tuning
* Neural Networks (NN)
  + Neural networks basics and intro
  + NN’s as autoML
  + Common NN topologies and applications (RNN, CNN, LSTM, etc)
  + Thin layer NN, examples, and lab (using TF)
  + Deep Learning
  + Libraries (Keras vs. TF vs. pytorch)
* Conclusion