The Machine Learning Pipeline on AWS

Overview

This course explores how to use the machine learning (ML) pipeline to solve a real business problem in a project-based learning environment. Students will learn about each phase of the pipeline from instructor presentations and demonstrations and then apply that knowledge to complete a project solving one of three business problems: fraud detection, recommendation engines, or flight delays. By the end of the course, students will have successfully built, trained, evaluated, tuned, and deployed an ML model using Amazon SageMaker that solves their selected business problem.

Prerequisite Comments

We recommend that attendees of this course have:
Basic knowledge of Python programming language
Basic understanding of AWS Cloud infrastructure (Amazon S3 and Amazon CloudWatch)
Basic experience working in a Jupyter notebook environment

Target Audience

This course is intended for:
Developers
Solutions Architects
Data Engineers
Anyone with little to no experience with ML and wants to learn about the ML pipeline using Amazon SageMaker

Course Objectives

In this course, you will learn to:
Select and justify the appropriate ML approach for a given business problem
Use the ML pipeline to solve a specific business problem
Train, evaluate, deploy, and tune an ML model using Amazon SageMaker
Describe some of the best practices for designing scalable, cost-optimized, and secure ML pipelines in AWS
Apply machine learning to a real-life business problem after the course is complete

Course Outline

1 - Module 0: Introduction

Pre-assessment

2 - Module 1: Introduction to Machine Learning and the ML Pipeline

Overview of machine learning, including use cases, types of machine learning, and key concepts
Overview of the ML pipeline
Introduction to course projects and approach

3 - Module 2: Introduction to Amazon SageMaker

Introduction to Amazon SageMaker
Demo: Amazon SageMaker and Jupyter notebooks
Hands-on: Amazon SageMaker and Jupyter notebooks

4 - Module 3: Problem Formulation

Overview of problem formulation and deciding if ML is the right solution
Converting a business problem into an ML problem
Demo: Amazon SageMaker Ground Truth
Hands-on: Amazon SageMaker Ground Truth
Practice problem formulation
Formulate problems for projects

5 - Module 4: Preprocessing

Overview of data collection and integration, and techniques for data preprocessing and visualization
Practice preprocessing
Preprocess project data
Class discussion about projects

6 - Module 5: Model Training

Choosing the right algorithm
Formatting and splitting your data for training
Loss functions and gradient descent for improving your model
Demo: Create a training job in Amazon SageMaker

7 - Module 6: Model Evaluation

How to evaluate classification models
How to evaluate regression models
Practice model training and evaluation
Train and evaluate project models
Initial project presentations

8 - Module 7: Feature Engineering and Model Tuning

Feature extraction, selection, creation, and transformation
Hyperparameter tuning
Demo: SageMaker hyperparameter optimization
Practice feature engineering and model tuning
Apply feature engineering and model tuning to projects
Final project presentations

9 - Module 8: Deployment

How to deploy, inference, and monitor your model on Amazon SageMaker
Deploying ML at the edge
Demo: Creating an Amazon SageMaker endpoint
Post-assessment
Course wrap-up