

**AI/ML Model Development, Deployment, and Monitoring Training**

**Course Number:** PYTH-282WA  
**Duration:** 2 days

**Overview**

This AI/ML Model Development, Deployment, and Monitoring course teaches attendees how to operationalize machine learning solutions. Students learn how to design and build effective models tailored to their organization’s business challenges, seamlessly integrate them into existing systems, and implement robust monitoring strategies. Participants master selecting the right tools, interpreting data, detecting and addressing model drift, and safeguarding privacy and security, all while ensuring models consistently deliver on intended business objectives.

**Prerequisites**

* A foundational understanding of predictive machine learning concepts, such as supervised and unsupervised learning, neural networks, and evaluation metrics (e.g., precision, recall, F1).
* Python programming experience, including data structures, functions, and control flow.
* Experience with Python data processing libraries, such as Pandas and NumPy.

**Materials**

All AI and ML training students receive comprehensive courseware.

**Software Needed on Each Student PC**

Detailed setup will be provided upon request.

**Objectives**

* An understanding of the Machine Learning Operations (MLOps) workflow and pipeline.
* Best practices for data preprocessing and transformation using Airflow, Spark, and Pandas.
* Exposure to modern ML development workflows using SparkML, scikit-Learn, Hugging Face, and TensorFlow.
* Strategies for deploying and monitoring ML Models using MLFlow, TensorFlow Service, Prometheus, and Grafana.
* Best practices for organizing, running, and tracking Agile AI/ML projects.

**Outline**

* Introduction to Machine Learning
  + Understanding the basics of machine learning applications, deployment, and pipelines
  + Overview of common machine learning tools
* Data Preprocessing
  + Introduction to data preprocessing
  + Preprocessing tools such as Airflow, Spark, and Pandas
  + ML/AI Tools Examples
  + In this chapter, examples of ML/AI tools could include:
    - Airflow: Apache Airflow is a platform to programmatically author, schedule, and monitor workflows. It allows for data preprocessing tasks to be executed in a scalable manner.
    - Spark: Apache Spark is a powerful open-source unified analytics engine that provides support for both batch and streaming data processing. Spark can be used for efficient data preprocessing tasks.
    - Pandas: Pandas is a popular open-source data manipulation and analysis tool that provides data structures and functions to efficiently clean and preprocess data.
  + Preprocessing data using Pandas and Spark pipelines
* Developing ML Models
  + Overview of ML at scale using Spark ML pipelines
  + Building traditional ML models with Scikit-learn pipelines
  + Introduction to NLP/NN using Hugging Face pipelines
  + Deep dive into LLM and NN with TensorFlow pipelines
  + ML/AI Tools Examples
  + In this chapter, examples of ML/AI tools could include:
    - Spark ML: Apache Spark MLlib is a scalable machine learning library that provides easy-to-use APIs for building scalable machine learning pipelines.
    - Scikit-learn: Scikit-learn is a popular machine learning library in Python that provides a wide range of tools for building traditional machine learning models.
    - Hugging Face Transformers: Hugging Face Transformers is a popular open-source library that provides state-of-the-art models for Natural Language Processing (NLP) tasks.
    - TensorFlow: TensorFlow is an open-source machine learning framework developed by Google that provides tools and libraries for building and training neural network models.
  + Building a machine learning model using Spark ML pipelines and TensorFlow pipelines
* Deployment and Monitoring
  + Strategies for monitoring model drift
  + Deploying machine learning models in production
  + Tools for model deployment and monitoring
  + ML/AI Tools Examples
  + In this chapter, examples of ML/AI tools could include:
    - MLflow: MLflow is an open-source platform for the end-to-end machine learning lifecycle. It provides tools for tracking experiments, packaging code, and deploying models.
    - TensorFlow Serving: TensorFlow Serving is a flexible, high-performance serving system for machine learning models designed for production environments.
    - Prometheus and Grafana: Tools like Prometheus and Grafana can be used for monitoring model performance and detecting model drift in real-time.
  + Deploying a machine learning model and setting up monitoring for model drift
* Updating and Fine-Tuning Models
  + Green/blue deployments for model updates
  + Techniques for fine-tuning models in live applications
  + Methods for continuous model improvement
  + ML/AI Tools Examples
  + In this chapter, examples of ML/AI tools could include:
    - Kubeflow: Kubeflow is an open-source platform built on Kubernetes that provides a com prehensive solution for managing end-to-end machine learning workflows, including model training, deployment, and monitoring.
    - Amazon SageMaker: Amazon SageMaker is a fully managed service that provides tools to build, train, and deploy machine learning models at scale. It also offers capabilities for model tuning and continuous deployment.
    - GitLab CI/CD: GitLab's CI/CD pipelines can be used for automating model updates and deployments, enabling green/blue deployments and continuous integration of model changes.
  + Implementing green/blue deployments for model updates
* Project Management Tools
  + Overview of project management tools for machine learning projects
  + Best practices for managing machine learning projects
  + Tools for collaboration, tracking, and documentation
  + ML/AI Tools Examples
  + In this chapter, examples of ML/AI tools could include:
    - Jira: Jira is a popular project management tool that can be used to track tasks, collaborate with team members, and manage project documentation in machine learning projects.
    - Confluence: Confluence is a collaboration tool that allows team members to create, share, and collaborate on project documentation, keeping all project-related information in one place.
    - Trello: Trello is a simple and flexible project management tool that can be used to organize tasks, assign responsibilities, and track progress in machine learning projects.
* Full ML Project Pipelines
  + Utilizing MLflow for end-to-end ML project management
  + Introduction to Amazon SageMaker for developing, training, and deploying ML models
  + Setting up a complete ML project pipeline from data preprocessing to model deployment
  + ML/AI Tools Examples
  + In this chapter, examples of ML/AI tools could include:
    - MLflow: MLflow can be used to manage the end-to-end machine learning workflow, including data preprocessing, model training, deployment, and monitoring, providing a comprehensive solution for ML project pipelines.
    - Amazon SageMaker: Amazon SageMaker offers a complete set of tools for building, training, and deploying machine learning models at scale, making it an ideal choice for developing and managing ML project pipelines in production environments.
    - AWS Step Functions: AWS Step Functions can be used to orchestrate and automate the various steps in an ML project pipeline, enabling seamless integration of data preprocessing, model training, and deployment tasks.
* Conclusion