

# **AiU - Certified Machine Learning Engineer (AiU – CMLE)**

**Duration:** 04 days

## **Course Objective**

With the current trends in technology and the growing incorporation of artificial intelligence and machine learning, Artificial Intelligence United (AiU) was created to support the understanding of implementation of these importance advancements. AiU - Certified Machine Learning Engineer (AiU-CMLE) is a 4-day practical certification course, which goes beyond the fundamentals of AI and machine learning, to discover the differences in this new world of expertise.

## **Course Outline**

### **Chapter 1: Data Science Foundations**

This chapter introduces the fundamental principles of data science, focusing on how data is collected, cleaned, processed, and analysed. Participants will explore various types of data, data structures, and the importance of data quality. The module will also cover data wrangling techniques using Python and common libraries such as Pandas and NumPy. Key concepts in exploratory data analysis (EDA) will be discussed, including data visualization techniques using tools like Matplotlib and Seaborn to gain insights from data.

#### **Key Topics:**

- Introduction to Data Science
- Data Collection and Preprocessing
- Exploratory Data Analysis (EDA)
- Data Wrangling and Cleaning
- Data Visualization Techniques
- Introduction to Python for Data Science

### **Chapter 2: Machine Learning Foundations**

This chapter lays the groundwork for understanding machine learning, covering both supervised and unsupervised learning algorithms. Participants will explore key algorithms such as linear regression, decision trees, k-nearest neighbors, and clustering methods. The chapter also introduces the concept of model evaluation, focusing on metrics like accuracy, precision, recall, and F1-score.

#### **Key Topics:**

- Overview of Machine Learning
- Supervised vs. Unsupervised Learning

- Introduction to Key Algorithms (e.g., Linear Regression, Decision Trees, KNN)
- Model Evaluation Metrics
- Introduction to Scikit-Learn for Model Implementation
- Data Splitting (Training, Validation, Testing)

### **Chapter 3: Statistical Model Validation and Testing**

This chapter dives into the statistical methods used for model validation and testing, crucial for ensuring that machine learning models perform well on unseen data. Participants will learn about cross-validation techniques, bias-variance trade-offs, and hyperparameter tuning. The module will also cover statistical hypothesis testing, confidence intervals, and the importance of avoiding overfitting and underfitting.

#### **Key Topics:**

- Cross-Validation Techniques
- Bias-Variance Trade-Off
- Hyperparameter Tuning
- Statistical Hypothesis Testing
- Avoiding Overfitting and Underfitting
- Model Performance Monitoring

### **Chapter 4: Neural Networks and Deep Learning**

This chapter introduces neural networks, the foundation of deep learning. Participants will learn about the architecture of neural networks, including neurons, layers, activation functions, and backpropagation. The module will cover basic network structures such as fully connected networks and explore the training process. Participants will also gain hands-on experience using TensorFlow and Keras for building and training neural networks.

#### **Key Topics:**

- Introduction to Neural Networks
- Neurons, Layers, and Activation Functions
- Forward and Backpropagation
- Network Architectures (e.g., Fully Connected Networks)
- Introduction to TensorFlow and Keras
- Training and Evaluating Neural Networks

### **Chapter 5: Deep Learning and Advanced Data Types**

This chapter explores the application of deep learning to advanced data types, including images, text, and sequences. Participants will learn about convolutional neural networks (CNNs) for image processing, recurrent neural networks (RNNs) for sequential data, and natural language processing (NLP) techniques for text analysis. The module will also discuss transfer learning and how pre-trained models can be fine-tuned for specific tasks.

### **Key Topics:**

- Convolutional Neural Networks (CNNs) for Image Processing
- Recurrent Neural Networks (RNNs) for Sequential Data
- Natural Language Processing (NLP) Techniques
- Transfer Learning and Pre-Trained Models
- Handling Advanced Data Types
- Introduction to Specialized Libraries (e.g., OpenCV, NLTK, SpaCy)

### **Chapter 6: Machine Learning in Production**

The final chapter focuses on the practical aspects of deploying machine learning models into production environments. Participants will learn about the entire machine learning pipeline, from data ingestion to model deployment and monitoring. The module will cover best practices for model versioning, continuous integration/continuous deployment (CI/CD) pipelines, and monitoring models for drift and performance degradation over time.

### **Key Topics:**

- Machine Learning Pipeline Overview
- Model Deployment Strategies
- Continuous Integration/Continuous Deployment (CI/CD) for ML
- Model Versioning and Management
- Monitoring and Managing Model Drift
- Scaling and Optimizing Machine Learning Models in Production

### **Target Audience**

Anybody who is involved in the creation, training, and maintenance of machine learning systems and is looking to certify their knowledge with the global standard of the Artificial Intelligence United (AIU).

### **Requirements**

Although there are generally there are no specific pre-requisites for attending this course, the AiU does recommend that participants hold the ISTQB CTFL (certification for testing/QA), and have some development or code reading skills prior to joining this course. Some experience in statistics and knowledge about Neural Networks is a huge plus.