

Certified Professional Machine Learning & Data Science Foundation (CP-ML&DS Foundation) Certification Course



<http://cpmlddf.devopsppalliance.org/>
<http://devopsppalliance.org/cp-mlddf.html>



info@devopsppalliance.org



[/devopspp](https://twitter.com/devopspp)



[/devopspp](https://facebook.com/devopspp)

<http://devopsppalliance.org/>

+91-22-2778-2378

© DevOps++ Alliance 2018-2019



What is CP-ML&DS Foundation course?

CP-ML&DS Foundation stands for “Certified Professional – Machine Learning and Data Science Foundation” certification prepared and honored by “Agile Testing Alliance”.

The course is applicable for all roles and knowledge, experience & certification is consciously designed for all those who want to learning practical Machine learning and Data Science

How is it useful?

Machine learning is based on algorithms that can learn from data without relying on rules-based programming

As ever more of the analog world gets digitized, our ability to learn from data by developing and testing algorithms will only become more important for what are now seen as traditional businesses. Google chief economist Hal Varian calls this “computer kaizen.” For “just as mass production changed the way products were assembled and continuous improvement changed how manufacturing was done,” he says, “so continuous [and often automatic] experimentation will improve the way we optimize business processes in our organizations.

This is where it is clear, that we are into a computer kaizen world. A world where Machine learning and Data science algorithms are driving this self-learning and continuous learning process and bringing about a massive change in almost every industry. We have machine learning and data science now being used in banking, insurance, health care, sports, manufacturing, smart cities, lot Solutioning, Automotive, Aviation, Shipping and every other industry for that matter.

Machine learning and Data science need thus has increased multifold in past few years and would keep on increasing. At the same time there is a dearth of experienced professionals who know Machine learning and Data Science. There is no dearth of machine learning and data science programs where folks would need to spend infinite amount of time and efforts to acquire this knowledge. The challenge is for working professionals to spend so much time. Most often this rigor is lost over few weeks or months.

This program solves this issue addresses two basic needs

- a) Practical tool-based Machine learning and Data Science exposure for every working professional
- b) Allow working professionals to acquire this knowledge in the most agile manner

Am I Eligible?

There are no pre-requisites for this certification program except having some prior knowledge of any programming language and basics of mathematics and statistics. Program is Python driven and having prior knowledge of Python would be an advantage.

Duration?

CP-ML&DS Foundation duration is 3 full days. It is designed specifically for corporates and working professionals alike. If you are a corporate and can't dedicate full day for training, then you can opt for either 6 half days course or 3 full days course which is followed by theory and practical exams.

Learning Objectives and Main Topics

Module 1: The World of Machine Learning

Topics:

- A. What is Machine Learning?
- B. Types of Machine Learning
 - Supervised Learning
 - Unsupervised Learning
 - Reinforcement Learning
- C. Steps in Machine Learning
 - Training
 - Testing
 - Inference
- D. Applications of Machine Learning in various industries
- E. Linear Algebra, Calculus and Statistics refresher

Learning Objectives:

At the end of this module students should

- Tell the definition of Machine Learning
- Give the types of Machine Learning and the names of algorithms used
- Be able to tell the real-world applications of Machine Learning
- Be able to answer questions related to Linear Algebra, Calculus and Statistics

Module 2: Setting up Environment for Machine Learning

Topics:

- A. Anaconda environment setup
- B. Installing libraries required for Machine Learning
- C. Introduction to Jupyter notebook
- D. Python Data types
 - Operators
 - Condition and Loops

- Data structures – List, Dictionary and Tuples
 - Writing functions
 - File handling
 - Object oriented programming
 - Exception Handling
- E. Introduction to basic libraries of Python Numpy
- Pandas
 - Matplotlib

Learning Objectives:

At the end of this module students should

- Setup an environment in Anaconda for Machine Learning
- Independently install various libraries required for Machine Learning
- Solve small programming assignments in python
- Create Vectors and Matrices in Numpy. Able to perform addition and matrix multiplication using Numpy
- Upload a csv in Pandas. Perform statistics operations such as finding mean and standard deviation of a column
- Visualize the distribution of data using Matplotlib

Application Activity:

Students will implement important functions of Machine Learning such as Sigmoid and Softmax in Numpy

Project: Students will create a jupyter notebook where they will apply statistics and visualization using Pandas and Matplotlib on the Boston Housing Dataset

Module 3: Exploratory Data Analysis

Topics:

- A. Types of data
- B. Using summary statistics to understand data
- C. Using boxplot to visualize data
- D. Pre-processing Handling missing data

- Removing outliers
- Dealing with duplicate values
- Normalization

E. Visualizing relationships between columns using scatter plots

Learning Objectives:

At the end of this module students should

- Understand data that is used for Machine Learning
- Apply all the steps of pre-processing
- Explain summary statistics for a data frame
- Understand relationship between columns using scatter plots

Application Activity:

Project: Explore the 1974 Motor Trend US magazine dataset

Module 4: Simple Linear Regression

Topics:

- A. Understanding the linear equation Features
 - Weights
 - Target
- B. Learning a model for $y=wx+b$ equation Training a model in python from scratch
 - Defining and visualizing the cost function
 - Understanding gradient descent
- C. Testing a linear model Root Mean Square Error (RMSE)
- D. Using scikit-learn for Machine Learning Installing the library
 - Using scikit-learn for training $y=wx+b$
 - Saving a model
 - Infer on the saved model
- E. Finding the time required to repair a computer given the number of parts requiring replacement

Learning Objectives:

At the end of this module students

- Will understand basic concepts of linear regression.
- Able to apply linear regression on a dataset with one feature.
- Find accuracy of a linear model
- Will be able to use scikit-learn library for applying linear regression

Application Activity:

Project: Predict the distance traveled given speed of the car

Module 5: Multiple Linear Regression

Topics:

- A. Understand the concepts of multiple linear regression Multiple features
 - Multiple weights
 - Single target
- B. Data pre-processing. Converting categorical features to multiple features
 - Normalizing features
- C. Finding the best features for training a model Find correlation between features
 - Understand which features are highly correlated with the target
- D. Learning a model for $y=Wx+b$ equation. Understanding how a model gets trained with multiple features
 - Fitting a multiple linear regression model
- E. Testing a model RMSE
 - Improving accuracy
- F. Predict housing prices applying multiple linear regression

Learning Objectives:

- At the end of this module students
- Will understand how to solve multiple linear regression use case
- Will be able to apply multiple linear regression on different datasets
- Select best features for higher accuracy

Application Activity:

Students will present five different problems where linear regression can be applied

Project: Predict the power output of a power generation plant given environmental readings from different sensors

Module 6: Classification using Logistic Regression

Topics:

- A. Understand classification problems
- B. Logistic regression in detail
- C. The sigmoid function
 - Cost function for classification
 - Finding derivative of sigmoid function
- D. Training a logistic regression model Training logistic regression using scikit-learn
- E. Evaluating a model Accuracy
 - Confusion Matrix
 - Precision
 - Recall
 - Sensitivity
 - Specificity
 - F1 score
- F. Predict whether the income of an employee exceeds \$50K per year

Learning Objectives:

At the end of this module students

- Will be able to apply logistic regression to different datasets using scikit-learn
- Understand on what kind of problems logistic regression can be applied

Application Activity:

Students will present five different problems where logistic regression can be applied

Project: Apply logistic regression to predict whether a client of a bank will subscribe to a term deposit given campaign marketing data