



# **Certified Machine Learning Specialist (CMLS)**

## **Course Outline**

**[www.globalicttraining.com](http://www.globalicttraining.com)**

## **DURATION**

- 4 Days

## **COURSE OBJECTIVES**

This specialist course on Machine Learning gives an overview of many concepts, techniques, and algorithms. The course covers basic topics such as classification and linear regression to more advanced topics such as boosting, support vector machines, hidden Markov models, and Bayesian networks in Machine Learning.

This course will introduce professionals to open source Machine Learning tools such as WEKA and Scikit Learn. Professionals attending the course will learn the different algorithms in some of most widely adopted Machine Learning methods such as Supervised Learning, Unsupervised learning and Reinforcement Learning. Please refer to the attached course outline for detailed information of the course.

## **JOB ROLES IN NICF / TARGETED AUDIENCE**

- Data Analyst - Statistics and Mining
- Data Analyst - Text Analytics
- Operations Research Analyst
- Senior Data Analyst- Statistics and Mining

## **PRE-REQUISITES**

Participants are preferred to have experience in software development, business domain or data/business analysis.





Unit 5: **Decision Tree**

- /
- d
- 
- 
- 

Unit 6: **Association Rule Mining**

- 
- 
- Weka
- 
- 

Unit 7: **Association Rule Mining**

- / v š
- , } Á š
- d Z < P Á Á o&v oš
- d Z } u u v > ] v ] v š
- o • • ] ( ) š ] } v Z μ o • v
- š š

Unit 8: **Association Rule Mining**

- ID3 based decision tree algorithm
- Entropy and Information gain
- ID3 implementation using WEKA
- Association rule mining using Frequent Pattern (FP) Growth algorithm
- FP-Tree structure
- FP-Growth Algorithm
- Implementation of FP-Growth using WEKA

## **Unit 9: Machine Learning with Python- Tensor flow, sklearn, keras**

- Introduction to tensor flow
- Tensor flow library
- Sklearn library
- Keras library
- API on Linear regression
- API on logistic regression
- API on random forest
- API on support vector machine (SVM)
- API on naïve bayes
- API on k nearest neighbour (kNN)
- API on Kmean
- API on convolution Neural Network (CNN)
- API on local outlier
- API on deep learning auto decoder
- API on principal component analysis (PCA)

## **Unit 10: Hands on using RapidMiner**

- Data preprocessing
- Correlation
- Association rules
- Kmeans
- Linear regression
- Logistic regression
- Neural network
- Text mining

## **Hands On**

The main objective of this course is to train the professional with open source tools such as Google GO, scikit-learn and Spark MLLIB.

## **WRITTEN ASSESSMENT**

As part of the written examination, each participant will be assessed individually on the last day of the training for their understanding of the subject matter and ability to evaluate, choose and apply them in specific context and also the ability to identify and manage risks. The assessment focuses on higher levels of learning in Bloom's taxonomy: Application, Analysis, Synthesis and Evaluation. This written examination will primarily consist of 40 multiple choice questions spanning various aspects as covered in the program. It is an individual, competency-based assessment.

## **EXAM PREPARATION**

The objective of the certification examination is to evaluate the knowledge and skills acquired by the participants during the course. The weightage in key topics of the course as follows:

- **Introduction and Basic Concepts of machine learning (10%)**
- **Supervised, Unsupervised and Reinforcement learning(10%)**
- **Discriminative and Generative Algorithms (10%)**
- **Neural Network- CNN, RNN, LSTM(10%)**
- **Deep Learning-activation function and learning rate (10%)**
- **Model selection- hyperparameters and optimization techniques (10%)**
- **Machine Learning using WEKA (10%)**
- **Decision Tree and Rule mining using WEKA (10%)**
- **Machine learning with Python- Tensor flow, sklearn, keras (10%)**
- **Hands on using RapidMiner (10%)**

## **Tools Used:**

- **SCIPY**
- **Spark MLLIB**
- **GO**
- **RapidMiner**