

MACHINE LEARNING USING ZEPPELIN AND SCIKIT-LEARN

COURSE OVERVIEW:

Machine learning applications have been written since the late 1950s (see Perceptron, 1958). The term “machine learning” was accredited to Arthur Samuel in 1959. With the recent interest in Deep Learning (see Geoffrey Hinton et al., 2006) machine learning techniques have become valuable tools for the data scientist.

Machine learning is a field of computer science that gives computer systems the ability to progressively improve performance on a specific task with data, without being explicitly programmed. Many of the tasks performed by data scientists such as computational statistics, which focuses on prediction-making using computers, usually engage machine learning algorithms.

Machine learning is often associated with data mining. Data mining is often accomplished by data scientists through the application of exploratory data analysis (EDA) and unsupervised learning. Machine learning is used in the credit card industry, for example, to evince and establish baseline behavioral profiles for various entities and then find meaningful anomalies like the fraudulent use of credit cards or identities.

Machine learning algorithms are used to devise complex models and algorithms that lend themselves to prediction. In commercial use, this is known as “predictive analytics.” Analytical models allow data scientists to “produce reliable, repeatable decisions and results” and uncover “hidden insights” through learning from historical relationships and trends in the data.

This course instructs the student in key concepts and fundamental practices of machine learning (through lecture and labs using the Scikit-Learn libraries) that are relevant to the activities of a data scientist.

50% Lecture 50% Hands-on Labs

WHO WILL BENEFIT FROM THIS COURSE?

Individuals who are new to the application of Machine Learning. The goal of this course is to provide the concepts and the tools a data scientist needs to implement programs that are capable of “learning” from data. Applications will be written in the Python programming language using the Apache Zeppelin environment.

PREREQUISITES:

Experience with the Python programming language, the Zeppelin IDE and exposure to EDA statistics is a prerequisite. It is suggested that a student new to programming and new to Zeppelin take the course “Introduction to Python using Zeppelin.” Either experience with programming EDA statistics using Zeppelin or the completion of the course “Statistics for Data Science using Zeppelin” is a prerequisite for this course.

COURSE OUTLINE:

- Day 1: Fundamental Machine Learning and Classification
- Day 2: Training Models and Support Vector Machines
- Day 3: Decision Trees and Ensemble learning and Random Forests
- Day 4: Dimensionality Reduction

SUNSET LEARNING INSTITUTE (SLI) DIFFERENTIATORS:

Sunset Learning Institute (SLI) has been an innovative leader in developing and delivering authorized technical training since 1996. Our goal is to help our customers optimize their cloud technology investments by providing convenient, high quality technical training that our customers can rely on. We empower students to master their desired technologies for their unique environments.

What sets SLI apart is not only our immense selection of trainings options, but our convenient and consistent delivery system. No matter how complex your environment is or where you are located, SLI is sure to have a training solution that you can count on!

Premiere World Class Instruction Team

- All SLI instructors have a four-year technical degree, instructor level certifications and field consulting work experience.
- Sunset Learning has won numerous Instructor Excellence and Instructor Quality Distinction awards since 2012

Enhanced Learning Experience

- The goal of our instructors during class is ensure students understand the material, guide them through our labs and encourage questions and interactive discussions.

Convenient and Reliable Training Experience

- You have the option to attend classes at any of our established training facilities or from the convenience of your home or office with the use of our HD-ILT network (High Definition Instructor Led Training)
- All Sunset Learning Institute classes are guaranteed to run – you can count on us to deliver the training you need when you need it!

Outstanding Customer Service

- Dedicated account manager to suggest the optimal learning path for you and your team
- Enthusiastic Student Services team available to answer any questions and ensure a quality training experience