Data Science Overview | Technologies, Tools & Modern Roles in the Data-Driven Enterprise

COURSE OVERVIEW:

The Data Science Overview | Technologies, Tools & Modern Roles in the Data-Driven Enterprise is an introductory level course that introduces the entire multi-disciplinary Data Science team to the many evolving and related terms, with focus on Big Data, Data Science, Predictive Analytics, Artificial Intelligence, Data Mining, Data Warehousing. The overview explores the current state of the art and science, the major components of a modern data science infrastructure, team roles and responsibilities, and level-setting realistic possible outcomes for your investment.

This goal of this course is to provide students with a baseline understanding of core concepts that can serve as a platform of knowledge to follow up with more in-depth training and real-world practice.

WHO WILL BENEFIT FROM THIS COURSE?

Attendees should have prior exposure to Enterprise Information Technology. As well as familiarity with Relational Databases.

PREREQUISITES:

This introductory-level / primer course is an overview intended for Business Analysts, Data Analysts, Data Architects, DBAs, Network (Grid) Administrators, Developers or anyone else in the data science realm who need to have a baseline understanding of some of the core areas of modern Data Science technologies, practices and available tools.

COURSE OBJECTIVES:

This course provides a high-level view of a variety of core, current data science-related technologies, strategies, skillsets, initiatives, and supporting tools in common business enterprise practices. This list covers a general range of topics current to the time of course distribution. We will collaborate with your team to refine the level of depth of coverage, understand areas of greater importance to your team, where you would like to add demos, etc.

Students will explore:
- Foundations: Grids & Virtualization; SOA, ESB / EMB, The Cloud
- The Hadoop Ecosystem: HDFS; Resource Navigators, MapReduce, Spark, Distributions
- Big Data, NOSQL, and ETL
- ETL: Exchange, Transform, Load
- Handling Data & a Survey of Useful tools
- Enterprise Integration Patterns and Message Busses
- Developing in Hadoop Ecosystem: R, Python, Java, Scala, Pig, and BPMN
- Artificial Intelligence and Business Systems
- Who's on the Team? Evolving Roles and Functions in Data Science
- Growing your Infrastructure

COURSE OUTLINE:

Please note that this list of topics is based on our standard course offering, evolved from typical industry uses and trends. We'll work with you to tune this course and level of coverage to target the skills you need most. Topics, agenda, and labs are subject to change and may adjust during live delivery based on audience needs and skill-level.

Foundations
- Grids and Virtualization
- Service-Oriented Architecture
- Enterprise Service Bus
- Enterprise Message Bus
- The Cloud

The Hadoop Ecosystem
- HDFS: Hadoop Distributed File System
- Resource Negotiators: YARN, Mesos, and Spark; ZooKeeper
- Hadoop Map/Reduce
- Spark
- Hadoop Ecosystem Distributions: Cloudera, Hortonworks, OpenSource

Big Data, NOSQL, and ETL
- Big Data vs. RDBMS
- NOSQL: Not Only SQL
- Relational Databases: Oracle, MariaDB, DB/2, SQL Server, PostGreSQL
- Key/Value Databases: JBoss Infinispan, Terracotta, Dynamo, Voldemort
- Columnar Databases: Cassandra, HBase, BigTable
- Document Databases: MongoDB, CouchDB/CouchBase
- Graph Databases: Giraph, Neo4J, GraphX
- Apache Hive
- Common Data Formats
- Leveraging SQL and SQL variants

ETL: Exchange, Transform, Load
- Data Ingestion, Transformation, and Loading
- Exporting Data
- Sqoop, Flume, Informatica, and other tools

Enterprise Integration Patterns and Message Busses
- Enterprise Integration Patterns: Apache Camel and Spring Integration
- Enterprise Message Busses: Apache Kafka, ActiveMQ, and other tools

Developing in Hadoop Ecosystem
- Languages: R, Python, Java, Scala, Pig, and BPMN
- Libraries and Frameworks
- Development, Testing, and Deployment

Artificial Intelligence and Business Systems
- Artificial Intelligence: Myths, Legends, and Reality
- The Math
- Statistics
- Probability
- Clustering Algorithms, Mahout, MLLib, SciKit, and Madlib
- Business Rule Systems: Drools, JRules, Pegasus

The Team
- Agile Data Science
- NOSQL Data Architects and Administrators
- Developers
- Grid Administrators
- Business and Data Analysts
- Management
- Evolving your Team
- Growing your Infrastructure

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