



## Machine Learning Foundation: Working with Spark and TensorFlow

Explore Spark Essentials, Algorithms, Machine Learning, and Data Mining Concepts, and How TensorFlow Implements Them

[www.triveratech.com](http://www.triveratech.com)

### Course Snapshot

- **Course: TTML6805: Machine Learning Foundation: Working with Spark and TensorFlow**
- **Duration:** 5 days
- **Skill Level & Audience:** Intermediate skills for students new to Machine Learning, Spark & TensorFlow. Targeted for Data Scientists, Software Engineers, and Data Engineers and Developers who have basic Python knowledge.
- **Hands-on Learning:** This course is approximately **50% hands-on lab to 50% lecture ratio**, combining engaging lecture, demos, group activities and discussions with machine-based student labs and exercises. Student machines are required.
- **Delivery Options:** This course is available for **onsite private classroom presentation, live online virtual presentation**, or can be presented in a **flexible blended learning format** for combined onsite and remote attendees. Please also ask about our **Self-Paced / Video / QuickSkills or Mini-Camp Flex Hours / Short Course** options.
- **Public Schedule:** This course is available on our **Public Open Enrollment Schedule** in both **full day format** and **Half Day, Flex-Hours Mini-Camp Format**.
- **Customizable:** This course agenda, topics and labs can be further adjusted to target your specific training skills objectives, tools of choice and learning goals.

*The Machine Learning Foundation Series* is comprised of lab-intensive, hands-on courses that immerse attendees in the cutting-edge world of machine learning. Geared for developers, data analysts and data scientists, this program provides a structured, engaging learning experience for students to progressively attain the technical skills they require to engage in meaningful machine learning projects and activities, right after the training ends.

Our courses are ideally suited for a wide variety of technical learners who need a fast paced, hands-on introduction to the core skills, concepts and technologies related to machine learning. Attendees might include:

- Developers aspiring to be Machine Learning engineers
- Analytics Managers who are leading a team of analysts
- Business Analysts who want to understand data science techniques
- Information Architects who want to gain expertise in Machine Learning algorithms
- Analytics professionals who want to work in machine learning, deep learning or artificial intelligence
- Experienced professionals who would like to harness machine learning in their fields to get more insight about customers

### Overview

Apache Spark, a significant component in the Hadoop Ecosystem, is a cluster computing engine used in Big Data. Building on top of the Hadoop YARN and HDFS ecosystem, offers order-of-magnitude faster processing for many in-memory computing tasks compared to Map/Reduce. It can be programmed in Java, Scala, Python, and R - the favorite languages of Data Scientists - along with SQL-based front ends.

**Machine Learning Foundation: Working with Spark and TensorFlow** is a comprehensive, hands-on machine learning course intended for data scientists and software engineers (with Python experience), new to these technologies and Machine Learning. This course explores popular machine learning algorithms from the ground up. Students will explore Apache Spark essentials, core machine learning concepts, regressions, classifications, clustering and more.

The abundance of data and affordable cloud scale has led to an explosion of interest in Deep Learning. Google has released an excellent library called TensorFlow to open-source, allowing state-of-the-art machine learning done at scale, complete with GPU-based acceleration. Students will explore these skills in an active hands-on manner. The second part of the course introduces students to Deep Learning concepts and how TensorFlow implements them.

### Learning Objectives

This “skills-centric” course is about **50% hands-on lab and 50% lecture**, with extensive practical exercises designed to reinforce fundamental skills, concepts and best practices taught throughout the course. Throughout the program, working in a hands-on

learning environment guided by our expert instructor, students will

- Learn popular machine learning algorithms, their applicability, and limitations
- Practice the application of these methods in the Spark machine learning environment
- Learn practical use cases and limitations of algorithms
- Will explore not just the related APIs, but will also learn the theory behind them
- Work with real world datasets from Uber, Netflix, Walmart, Prosper, etc.

**Need different skills or topics?** If your team requires different topics or tools, additional skills or custom approach, this course may be further adjusted to accommodate. We offer additional big data, analytics. AI, machine learning, programming, Python/R and other related topics that may be blended with this course for a track that best suits your needs. Our team will collaborate with you to understand your needs and will target the course to focus on your specific learning objectives and goals.

### Audience & Pre-Requisites

This is an **intermediate level** course, geared for Data Scientists, Data Analysts and Developers new to Machine Learning, Spark and TensorFlow.

**Pre-Requisites:** Students should have attended or have incoming skills equivalent to those in this course:

- Strong basic Python Skills. Attendees without Python background may view labs as follow along exercises or team with others to complete them.
- Good foundational mathematics in Linear Algebra and Probability
- Basic Linux skills, including familiarity with command-line options such as ls, cd, cp, and su

**Take Before:** Attending students should have incoming skills equivalent to those in the course(s) below, or should have attended these as a pre-requisite:

- TTDS6600 Data Science Primer – Overview
- TTPS4800 Introduction to Python

**Take Next / Follow-on Courses:** This course is a core component of our **AI & Machine Learning Skills Path**, designed to trainer participants of all skill levels in modern AI, Machine Learning and Analytics skills across the enterprise. We offer courses in next level AI and Machine Learning, Deep Learning, Natural Language Processing, Applied Machine Learning (Chatbots, Intelligent Web) and many more related titles. Please contact us for details and next step recommendations based on your specific roles and goals.

**Enhanced Learning Support:** Please ask about our **Pre-Training Class Prep & Primer offerings, Skills Gap Assessment Services, Case Studies, Knowledge Check Quizzes, Skills Immersion Programs & Camps, Collaborative Mentoring Services and Extended Learning Support** services.

### Course Topics / Agenda

*Please note that this list of topics is based on our standard course offering, evolved from typical industry uses and trends. We will work with you to tune this course and level of coverage to target the skills you need most. Course agenda, topics and labs are subject to adjust during live delivery in response to student skill level, interests and participation.*

#### Part 1: Introduction to Machine Learning

##### 1. Machine Learning (ML) Overview

- Machine Learning landscape
- Machine Learning applications
- Understanding ML algorithms & models

##### 2. ML in Python and Spark

- Spark ML Overview
- Introduction to Jupyter

notebooks

- Lab: Working with Jupyter + Python + Spark
- Lab: Spark ML utilities

##### 3. Machine Learning Concepts

- Statistics Primer
- Covariance, Correlation, Covariance Matrix
- Errors, Residuals
- Overfitting / Underfitting

- Cross-validation, bootstrapping
- Confusion Matrix
- ROC curve, Area Under Curve (AUC)
- Lab: Basic stats

##### 4. Feature Engineering (FE)

- Preparing data for ML
- Extracting features, enhancing data
- Data cleanup

- Visualizing Data
  - Lab: data cleanup
  - Lab: visualizing data
- 5. Linear regression**
- Simple Linear Regression
  - Multiple Linear Regression
  - Running LR
  - Evaluating LR model performance
  - Lab
  - Use case: House price estimates
- 6. Logistic Regression**
- Understanding Logistic Regression
  - Calculating Logistic Regression
  - Evaluating model performance
  - Lab: Use case: credit card application, college admissions
- 7. Classification: SVM (Supervised Vector Machines)**
- SVM concepts and theory
  - SVM with kernel
  - Lab: Use case: Customer churn data
- 8. Classification: Decision Trees & Random Forests**
- Theory behind trees
  - Classification and Regression Trees (CART)
  - Random Forest concepts
  - Labs: Use case: predicting loan defaults, estimating election contributions
- 9. Classification: Naive Bayes**
- Theory
  - Lab
  - Use case: spam filtering
- 10. Clustering (K-Means)**
- Theory behind K-Means
  - Running K-Means algorithm
  - Estimating the performance
  - Lab: Use case: grouping cars data, grouping shopping data
- 11. Principal Component Analysis (PCA)**
- Understanding PCA concepts
  - PCA applications
- Running a PCA algorithm
  - Evaluating results
  - Lab: Use case: analyzing retail shopping data
- 12. Recommendations (Collaborative filtering)**
- Recommender systems overview
  - Collaborative Filtering concepts
  - Lab: Use case: movie recommendations, music recommendations
- 13. Performance**
- Best practices for scaling and optimizing Apache Spark
  - Memory caching
  - Testing and validation
- Part Two: Introduction to Deep Learning with TensorFlow**
- 14. Machine Learning Quick Review**
- Understanding Machine Learning
  - Supervised versus Unsupervised Learning
  - Regression
  - Classification
  - Clustering
- 15. Introducing Tensorflow**
- Tensorflow intro
  - Tensorflow Features
  - Tensorflow Versions
  - GPU and TPU scalability
  - Lab: Setting up and Running Tensorflow
- 16. The Tensor: The Basic Unit of Tensorflow**
- Introducing Tensors
  - Tensorflow Execution Model
  - Lab: Learning about Tensors
- 17. Single Layer Linear Perceptron Classifier With TensorFlow**
- Introducing Perceptrons
  - Linear Separability and Xor Problem
  - Activation Functions
  - Softmax output
  - Backpropagation, loss functions, and Gradient Descent
- Lab: Single-Layer Perceptron in Tensorflow
- 18. Hidden Layers: Intro to Deep Learning**
- Hidden Layers as a solution to XOR problem
  - Distributed Training with Tensorflow
  - Vanishing Gradient Problem and ReLU
  - Loss Functions
  - Lab: Feedforward Neural Network Classifier in Tensorflow
- 19. High level Tensorflow: tf.learn**
- Using high level tensorflow
  - Developing a model with tf.learn
  - Lab: Developing a tf.learn model
- 20. Convolutional Neural Networks in Tensorflow**
- Introducing CNNs
  - CNNs in Tensorflow
  - Lab : CNN apps
- 21. Introducing Keras**
- What is Keras?
  - Using Keras with a Tensorflow Backend
  - Lab: Example with a Keras
- 22. Recurrent Neural Networks in Tensorflow**
- Introducing RNNs
  - RNNs in Tensorflow
  - Lab: RNN
- 23. Long Short-Term Memory (LSTM) in Tensorflow**
- Introducing RNNs
  - RNNs in Tensorflow
  - Lab: RNN
- 24. Conclusion**
- Summarize features and advantages of Tensorflow
  - Summarize Deep Learning and How Tensorflow can help
  - Next steps

**Student Materials:** Each participant will receive a **Student Guide** with course notes, code samples, software tutorials, step-by-step written lab instructions, diagrams and related reference materials and resource links. Students will also receive the project files (or code, if applicable) and solutions required for the hands-on work.

**Hands-On Setup Made Simple!** Our dedicated tech team will work with you to ensure our 'easy-access' cloud-based course environment is accessible, fully-tested and verified as ready to go well in advance of the course start date, ensuring a smooth start to class and effective learning experience for all participants. Please inquire for details and options.

### For More Information

All courses can be presented **onsite** or **online**, or in a **combined / flex / blended learning format**, tailored to target your specific audience, needs and learning goals. We also offer focused, flexible **short courses, self-paced learning options, recorded sessions** and more. We train beginner to advanced skills in all areas we cover, and offer **New Hire / Cohort Training, Boot Camps, Skills Immersion Programs, Reskilling Programs, Skills Migration & Transition Programs**, and more. We collaborate with you to ensure all courses are truly targeted to meet your specific needs and learning skills, maximizing your valuable training time, as well as your important budget.

Please also visit our extensive **Public Training Schedule** for training for smaller groups or individuals. Please contact us for course details, **Corporate Rates** and **Special Discount Offers**.

**For more information** about our dedicated training services, collaborative coaching services, courseware licensing and development services, public course schedule, training management services, partner programs, or to see our complete list of course offerings and special offers please visit us at [www.triveratech.com](http://www.triveratech.com), email [Info@triveratech.com](mailto:Info@triveratech.com) or call us toll free at **844-475-4559**. Our pricing and services are always satisfaction guaranteed.

---

**TRIVERA TECHNOLOGIES • Collaborative IT Training, Coaching & Skills Development Solutions**  
[www.triveratech.com](http://www.triveratech.com) • toll free +1-844-475-4559 • [Info@triveratech.com](mailto:Info@triveratech.com) • Twitter TriveraTech

ONSITE, ONLINE & BLENDED TRAINING SOLUTIONS • PUBLIC / OPEN ENROLLMENT COURSES • COURSEWARE LICENSING & DEVELOPMENT MENTORING • ASSESSMENTS • LEARNING PLAN DEVELOPMENT • SKILLS IMMERSION PROGRAMS / RESKILLING / NEW HIRE / BOOT CAMPS PARTNER & RESELLER PROGRAMS • CORPORATE TRAINING MANAGEMENT • VENDOR MANAGEMENT SERVICES

Trivera Technologies is a Woman-Owned Small-Business Firm

## Explore Trivera's Ways to Learn...

