Advanced Machine Learning is an eight-week, part-time, online bootcamp that offers a deeper understanding of advanced machine learning techniques and how to handle unstructured data. In this course, attendees explore decision trees, neural networks, clustering, KMeans, time series, signal processing and more. At the end of this course, students will be able to handle structured data and apply these techniques to large volumes of real-world unstructured data, while harnessing the power of raw data to solve pressing business problems.

Who should attend?
Data analysts, economists, researchers, software or data engineers who want to gain hands-on experience applying machine learning techniques to real-world datasets

Key Skills Covered
NumPy, SciPy, Pandas, probability, anomaly detection, recommendation engines, classification, transformers, preprocessing, unsupervised learning, decision trees

Prerequisites
To achieve the greatest benefit from this course, attendees must take Essential Data Tools and Practical Machine Learning or possess the following skills prior to attending:
- Intermediate Python
- Linear algebra
- Statistical modeling

Register for Advanced Machine Learning or learn more about other courses in our data curriculum by visiting pragmaticinstitute.com or calling 480.515.1411.
What Students Learn

Over the course of eight weeks, attendees get hands-on experience with advanced machine learning techniques that can be used the very next day.

### WEEK 1

**STUDENTS REFRESH AND EXPAND** on their knowledge of Python and statistics to prepare for advanced machine learning techniques.

- **Python refresher**
  - NumPy, SciPy and Matplotlib
  - Pandas
  - Functions
  - Object-oriented programming
  - Error handling

- **Statistics refresher**
  - Probability
  - Statistics of random variables
  - Distributions
  - Estimating statistical parameters
  - Bootstrapping

### WEEK 2

**STUDENTS EXPLORE MACHINE LEARNING** concepts, advancing skills learned in previous courses.

- Overview of machine learning
  - Predictive modeling framework
  - Scaling, normalization and standardization
  - Feature importance

### WEEK 3

**STUDENTS BEGIN BUILDING REGRESSION**, classification and signal models using real-world data.

- Bias, variance and overfitting linear and logistic regression
  - Linear regression
  - Regularization
  - Combining signals
  - Logistic regression
  - Multiclass classification problems

### WEEK 4

**STUDENTS CREATE CUSTOM MODELS** and programs using Python and Scikit-Learn.

- Scikit-Learn workflow
  - Writing custom estimators and transformers
  - Pipelines
  - Feature unions
  - Stochastic gradient descent
STUDENTS EXPAND THEIR KNOWLEDGE of clustering, reduction and other models to help them better understand and manage unsupervised learning programs

- Unsupervised learning
  - K-Means clustering
  - Metrics for clustering
  - Dimension reduction
  - Matrix factorization
  - Principal components analysis
  - Anomaly detection

STUDENTS LEARN ADVANCED ALGORITHMS and decision-making tools that can be used for a variety of data types.

- Advanced algorithms
  - Decision trees and random forests
  - Support vector machines (SVM)
  - K nearest neighbors (KNN)
  - Naive bayes
  - Ensemble methods

  - Decision making
    - Comparing machine learning algorithms
    - Decision-making flowchart
    - Unbalanced classes

STUDENTS DIVE INTO NATURAL LANGUAGE processes and time series modeling for more advanced program training.

- Natural language processing
  - Text as a “bag of words”
  - TF-IDF
  - Document similarity metrics
  - Word2Vec

  - Time series
    - Trends in time series data
    - Cross-validation for time series
    - Modeling drift
    - Modeling seasonality
    - Modeling “noise”
    - Using external data sources as features
    - More advanced time series modeling frameworks

STUDENTS BUILD RECOMMENDATION ENGINES and explore deep learning with the help of TensorFlow.

- Recommendation engines
  - Feature engineering
  - Nearest neighbors
  - Dimensionality reduction
  - Recommendation for a user
  - Cooperative learning
  - Regression of ratings

  - Deep learning and TensorFlow
    - TensorFlow
    - Neural networks

Register for Advanced Machine Learning or learn more about other courses in our data curriculum by visiting pragmaticinstitute.com or calling 480.515.1411.