Practical Machine Learning with R

Instructor

Matthew Renze
Twitter: @matthewrenze
Email: info@matthewrenze.com
Web: http://www.matthewrenze.com

Course Description

R is a very popular open-source programming language for machine learning. Its interactive programming environment and powerful data analysis capabilities make R an ideal tool for machine learning.

This workshop will provide an introduction to the R programming language using RStudio. In addition, we will demonstrate how we can use R to train a series of machine learning models. We'll cover supervised and unsupervised learning in the form of classification, regression, and clustering. We'll also learn about ensemble learning, deep learning, and reinforcement learning. Finally, we'll learn how to deploy these models to production.

Prerequisites

Please bring your own Windows laptop and complete Lab 0 to install all of the necessary software before the workshop begins.

Module Descriptions

1. **Introduction** – introduce machine learning and the R programming language
2. **Classification** – learn how to predict categorical outcomes
3. **Regression** – learn how predict numeric outcomes
4. **Clustering** – learn how to predict groups of data based on similarity
5. **ML in Practice** – learn how to prep, tune, and evaluate machine learning models
6. **Ensemble Learning** – learn how to train and test ensembles of models
7. **Deep Learning** – learn how to train and test deep neural network models
8. **Reinforcement Learning** – learn how to train and test reinforcement learning agents
9. **ML in Production**– learn how to deploy machine learning solutions into production
10. **Conclusion** – learn where to go next for additional training and resources

Learning Objectives

When students are finished with this workshop, they should understand the following:

Introduction

- What machine learning is, why it is important, and how the machine learning process works
- What R is and why it has become so popular for machine learning
- How to create data types, data structures, subset data tables, and find help on R topics

Classification

- What classification is, how it works, and applications for classification
- What are k-nearest neighbors, decision trees, and neural networks
- How to train, test, and make predictions with a classifier

Regression

- What regression is, how it works, and applications for regression
- What are linear regression, multiple regression, and neural network regression
- How to train, test, and make predictions with regression

Clustering

- What clustering is, how it works, and applications for clustering
- What are k-Means and hierarchical clustering
- How to group similar data points using clustering algorithms

ML in Practice

- What the machine learning process is
- What are overfitting and the curse of dimensionality
- How to prep data, tune hyperparameters, and evaluate model performance

Ensemble Learning

- What ensemble learning is, why it's important, and how it works
- What a random forest classifier is and how it works
- How to train, test, and predict with ensemble classifiers

Deep Learning

- What deep learning is, why it's important, and how it works
- What a deep neural network is and how it works
- How to train, test, and predict with a deep neural network

Reinforcement Learning

- What reinforcement learning is, why it's important, and how it works
- What are agents, environments, states, actions, and rewards
- How to train, test, and predict with a reinforcement learning agent

ML in Production

- How to deploy machine learning models into production with R

Course Outline

Introduction

Lecture

- What is machine learning?
- What is R?

Lab

- Installation and setup
- Hello World
- Working with data types
- Working with data structures
- Working with data frames

Classification

Lecture

- Classification
- K-nearest neighbors
- Decision tree classifier
- Neural network classifier

Lab

- Predicting categories with k-nearest neighbors
- Predicting categories with a decision tree
- Predicting categories with a neural network

Regression

Lecture

- Regression
- Simple linear regression
- Multiple linear regression
- Neural network regression

Lab

- Predicting values with simple linear regression
- Predicting values with multiple linear regression
- Predicting values with a neural network

Clustering

Lecture

- Clustering
- k-Means clustering
- Hierarchical clustering

Lab

- Grouping data with k-means clustering
- Grouping data with hierarchical clustering

Machine Learning in Practice

Lecture

- The machine learning process
- Overfitting, underfitting, and regularization
- The curse of dimensionality

Lab

- Data preparation
- Hyperparameter tuning and model selection
- Evaluating models

Ensemble Learning

Lecture

- What is ensemble learning?
- Random forest classifier

Lab

- Classification with random forests

Deep Learning

Lecture

- What is deep learning?
- Deep neural networks

Lab

- Handwritten digit recognition with a deep neural network

Reinforcement Learning

Lecture

- What is reinforcement learning?
- Deep neural networks

Lab

- Handwritten digit recognition with a deep neural network

Machine Learning in Production

Lecture

- Deploying ML in production with R

Lab

- Deploying a ML model to production with Shiny

Conclusion

Lecture

- Where to go next
- Course summary