CLARUSWAY

# Lesson Plan

# Machine Learning

Machine learning is a data analytics technique, getting computers to learn and also act like a human. Machine learning algorithms use computational methods to “learn” by feeding data and information. And it is also known as a field of data analytics to make predictions depending on trends and insights of the data. In this section you will learn the basics of machine learning and algorithms. At the end of this you will be able to build and evaluate your own machine learning models.

**Custodian :** Orion-Instructor (orion@clarusway.com)

**In-class Sessions :** 24 Sessions / 72 hours

**Lab Sessions :** 5 Labs / 5 hours

## Prerequisites

- Basic Math
- Inferential and descriptive statistics
- Pandas, Numpy, Seaborn, Matplotlib
- Python Experience

## Course Outline

**1. Introduction**

**1.1. Introduction to Machine Learning**

1.1.1. Why Machine Learning?

1.1.2. Types of Machine Learning Algorithms

**2. Supervised Learning**

**2.1. Supervised Machine Learning Process**

2.1.1 Supervised Machine Learning

**2.2. Linear Regression**

2.2.1. Linear Regression Theory

2.2.2. Linear Regression with Python

**2.2.3 Polynomial Regression**
2.2.3.1. Polynomial Regression Theory
2.2.3.2. Polynomial Regression with Python
**2.2.4. Bias-Variance Trade-Off**
2.2.4.1. Introduction Bias Variance Trade-Off
2.2.4.2. Underfitting and Overfitting Problems(Choosing Degree of Polynomial)
**2.2.5. Regularization**
2.2.5.1. Regularization overview
2.2.5.2. Feature Scaling
2.2.5.3. Cross-Validation and Grid Search
2.2.5.3.1. Introduction to cross-validation
2.2.5.3.2. Test | Train Split
2.2.5.3.3. Test | Validation | Train Split
2.2.5.3.4. Cross_val_score method
2.2.5.3.5. Cross_validate method
2.2.5.3.6. Grid Search
2.2.5.4. Ridge Regression
2.2.5.5. Lasso Regression
2.2.5.6. Elastic Net

**2.3. Logistic Regression**

2.3.1. Introduction to Logistic Regression
2.3.2. Logistic Regression Theory - Linear to Logistic
2.3.3 Logistic Regression Theory - Maximum Likelihood
2.3.4 Logistic Regression with Python-Part 1
**2.3.5. Classification Metrics**
2.3.5.1. Confusion Matrix and Accuracy
2.3.5.2. Precision, Recall, F1-Score
2.3.5.3. ROC Curves
2.3.6. Logistic Regression with Python - Part 2
2.3.7. Multi-class Classification with Python

**2.4. K Nearest Neighbors**
2.4.1. KNN Theory
2.4.2. KNN with Python

**2.5. Support Vector Machines**
2.5.1. SVM Theory
2.5.2. Support Vector Machines with Python

**2.6. Decision Trees**
2.6.1. Decision Tree Theory
2.6.2. Decision Trees with Python

**2.7. Random Forest**
2.7.1. The Bagging Method as one of the Ensemble Methods
2.7.2. Random Forest Theory
2.7.3. Random Forest Classifier with Python
2.7.4. Random Forest Regressor with Python

**2.8. Boosting Methods**
2.8.1. Boosting Theory
2.8.2. AdaBoost Theory
2.8.3. AdaBoost with Python
2.8.4. Gradient Boosting Theory
2.8.5. Gradient Boosting with Python
2.8.6. XGBoost Theory
2.8.7. XGBoost with Python

3. **Unsupervised Learning**

**3.1. Unsupervised Learning Overview**
3.1.1. Unsupervised Learning

**3.2. K Means Clustering**
3.1.1. K Means Algorithm Theory
3.1.2. K Means with Python

**3.2. Hierarchical Clustering**
3.2.1. Hierarchical Clustering Theory
3.2.2. Hierarchical Clustering with Python

**3.3. Principal Component Analysis**
3.3.1. Principal Component Analysis
3.3.2. PCA with Python

## Materials & Resources

- Clarusway Learning Management System (LMS)
- Scikit-Learn Documentation
- Pandas Documentation
- Numpy Documentation
- Seaborn Documentation
- Matplotlib Documentation
- Yellowbrick Documentation
- Portilla Udemy Course
- Wikipedia
- StatQuest with Josh Starmer

## Tools and Software

- Zoom, Slack, Kahoot, Peardeck Applications
- Jupyter Notebook / Google Colab
- Scikit-learn, Numpy, Pandas, Matplotlib, Seaborn,Yellowbrick

## Assignments & Projects

### Assignments

1. Assignment-1 (Car Price)
2. Assignment-2 (New York Payroll(overfit-underfit))
3. Assignment-3 (Heart Stroke Prediction)
4. Assignment-4 (Unbalanced Classification)
5. Assignment-5 (Soldiers' Races Prediction)
6. Assignment-6 (World Military Power Clustering)

### Projects

1. Autoscout Price Prediction
2. Tree Cover Types Multi-classification
3. Mall Customer Segmentation