

- **Lesson 1 - Introduction to Sci-Kit**

- Introduction
- Installing scikit-learn
- Data manipulation: from Pandas to scikit-learn
- Creating synthetic data

- **Supervised Methods**

- Naive Bayes: Bernoulli - Multinomial
- Detecting spam in real SMS Kaggle data
- Linear Support Vector Machines (SVM): SVM and LinearSVC
- Linear Support Vector Machines (SVM): NuSVM
- Logistic regression
- Predicting if income >50k using real US Census Data
- Isotonic regression
- Linear regression - Lasso - Ridge
- Lasso - Ridge
- Decision trees
- Introduction to ensemble methods
- Averaging ensemble methods - Part 1: Bagging
- Averaging ensemble methods - Part 2: Random forests
- Digit Classification via Random Forests
- Boosting ensemble methods

- Grid Search Cross Validation
- Predicting real house prices in the US using ExtraTreesRegressor

- **Unsupervised Methods**

- Density Estimation
- Principal Components
- Principal Components
- K-means
- DBScan
- Clustering
- Clustering and PCA on real countries data from Kaggle
- Outlier detection
- Novelty detection

