Machine Learning with MATLAB

Antti Löytynoja
Application Engineer
Goals

- Overview of machine learning
- Machine learning models & techniques available in MATLAB
- MATLAB as an interactive environment for evaluating and choosing the best algorithm
What is Machine Learning?

- Algorithms and techniques used for **data analytics** *(think data analysis)*
  - Obtain valuable information from the data

- Why is it called “learning”?  
  - Systems learn from initial training data  
  - Use resulting model (or knowledge) to predict outcomes or classes of new samples
Machine Learning
Characteristics and Examples

- **Characteristics**
  - Lots of data (many variables)
  - System too complex to know the governing equation (e.g., black-box modeling)

- **Examples**
  - Pattern recognition (speech, images)
  - Financial algorithms (credit rating, algorithmic trading)
  - Energy forecasting (load, price)
  - Biology (tumour detection, drug discovery)
Challenges – Machine Learning

- Significant technical expertise required
- No “one size fits all” solution
- Locked into Black Box solutions
- Time required to conduct the analysis
Overview – Machine Learning

Type of Learning

Unsupervised Learning
- Group and interpret data based only on input data

Supervised Learning
- Develop predictive model based on both input and output data

Categories of Algorithms

Clustering

Classification

Regression
Unsupervised Learning

Clustering

- k-Means, Fuzzy C-Means
- Hierarchical
- Neural Networks
- Gaussian Mixture
- Hidden Markov Model
Supervised Learning

Regression

- Neural Networks
- Decision Trees
- Ensemble Methods
- Non-linear Reg. (GLM, Logistic)
- Linear Regression

Classification

- Support Vector Machines
- Discriminant Analysis
- Naive Bayes
- Nearest Neighbor
Supervised Learning - Workflow

1. Import Data
2. Explore Data
3. Prepare Data

- Speed up Computations

- Select Model
- Train the Model
  - Known data
  - Known responses
- Measure Accuracy

- Use for Prediction
  - Model
  - Predicted Responses
  - New Data
Demo – Bank Marketing Campaign

- **Goal:**
  - Predict if customer would subscribe to bank term deposit based on different attributes

- **Approach:**
  - Import historical data
  - Divide data into training and testing sets
  - Train a classifier using different models
  - Measure accuracy and compare models

Data set downloaded from UCI Machine Learning repository
http://archive.ics.uci.edu/ml/datasets/Bank+Marketing
Demo – Bank Marketing Campaign

- Numerous predictive models with rich documentation
  - Also available: decision trees, neural networks, naïve Bayes etc.
- Interactive visualizations and apps to aid discovery
- Quick prototyping; Focus on modeling not programming

There’s more…

- Methods to simplify model
- Integrate algorithms into enterprise applications
Clustering
What MATLAB has to offer

- Numerous clustering functions with rich documentation
  - Hierarchical, k-means, Gaussian Mixture, Hidden Markov…
- Interactive visualizations to aid discovery
- Automatically determine the correct number of clusters (R2013b): `evalclusters`
- Viewable source; not a black box
- Rapid exploration & development
Learn More: Machine Learning with MATLAB

Data Driven Fitting with MATLAB

Classification with MATLAB

Regression with MATLAB

Multivariate Classification in the Life Sciences

Electricity Load and Price Forecasting

Credit Risk Modeling with MATLAB

Questions and answers