

Machine Learning

BS/MS (Computer Science)

IQRA UNIVERSITY IU

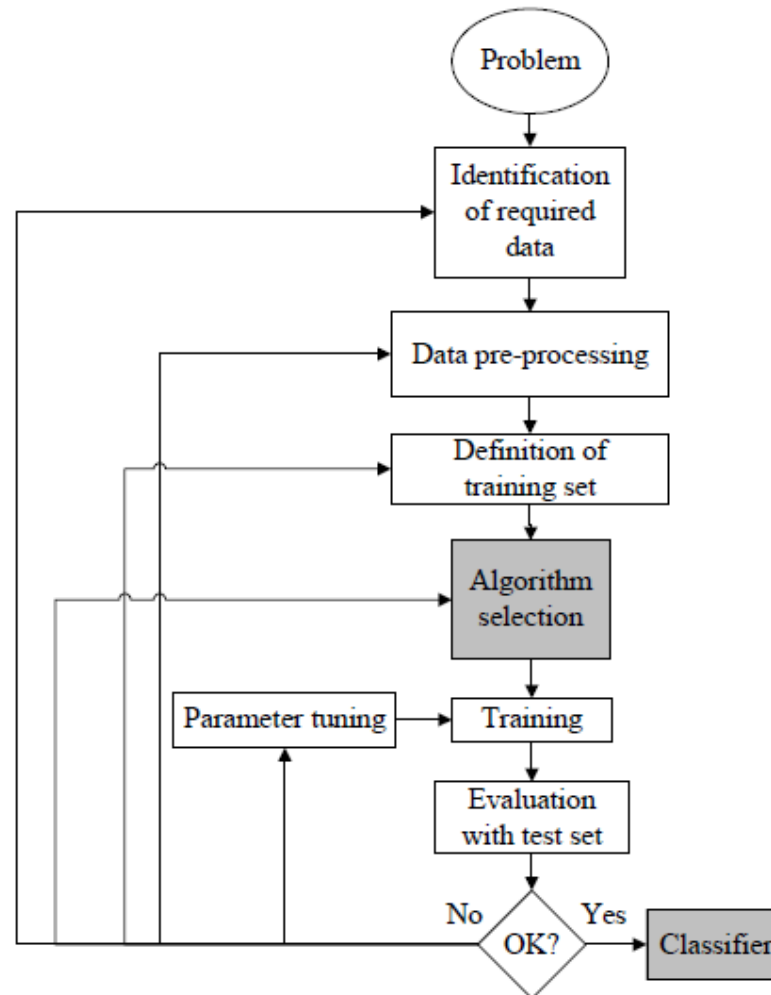
Lecture-02
15-June-2014
Summer Semester

Lecture-02

Designing a Concept Learning System

1. Problem Description
2. Choosing the Training Experience
3. Choosing the Target Function
4. Choosing a Representation for the Target Function
5. Choosing ML Algorithm
6. Final Design

The Process Learning Model (Machine Learning)



Approaches to Build A Model

Following are the two methods, which are commonly used to describe a propose system:

1. **Model-based methods:**

- A mathematical model is used to represent the behavior of the system.
- Practical application is very limited, because it is hard to construct an exact mathematical model for any system

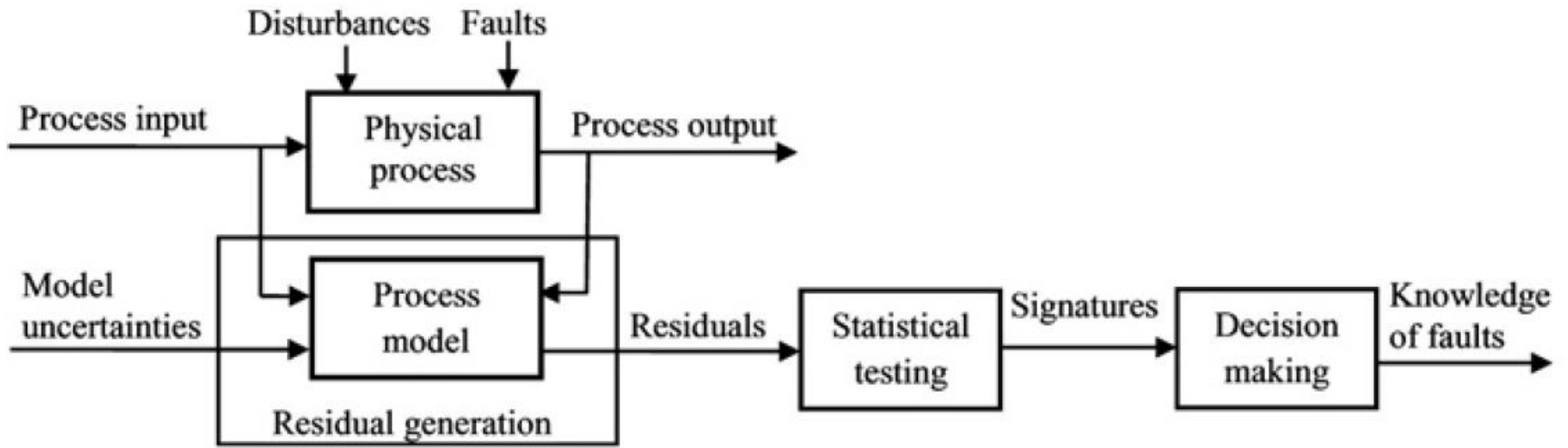
2. **Data-driven methods (multivariate):**

- Data-driven methods rely on relationships between correlated measurements within a system. However, the relationships can be formulated in an implicit way by training an empirical model through analysis of training data obtained during normal operations. The empirical model is then used to estimate true values of new measurements, and results are detected by evaluating the estimation residuals.
- Data-driven methods use multivariate statistical methods and machine learning tools.
- Example of data driven methods are: ANN, SVM etc.

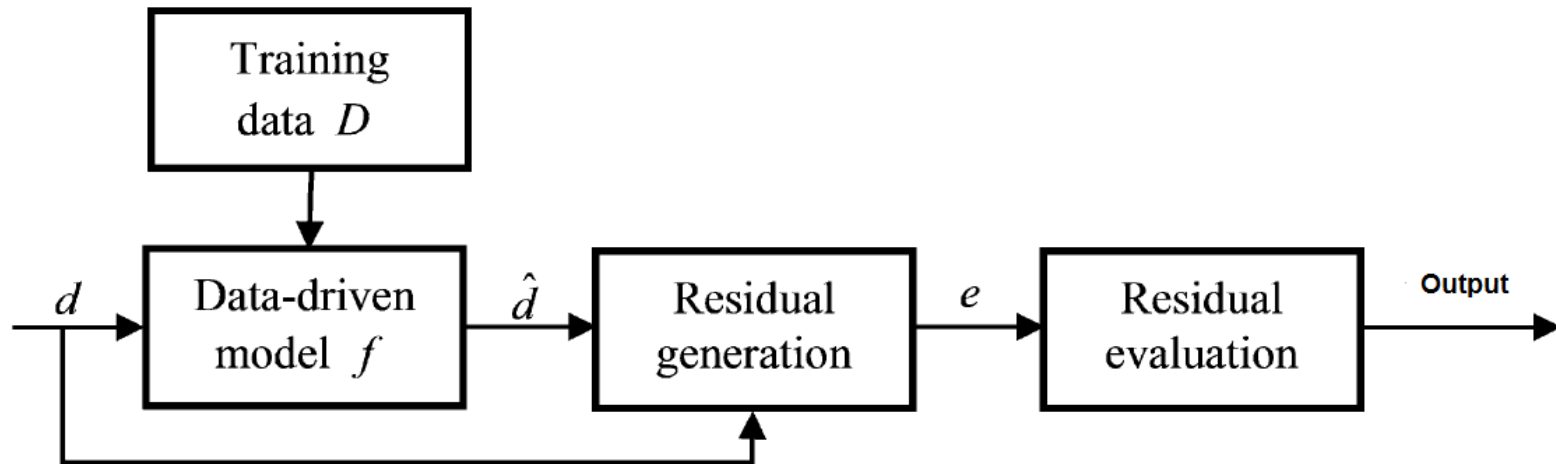
3. **Signal-based methods (univariate):**

- Signal-based methods make decisions by comparing features (e.g., spectrum) extracted from a signal with desired normal baseline values.

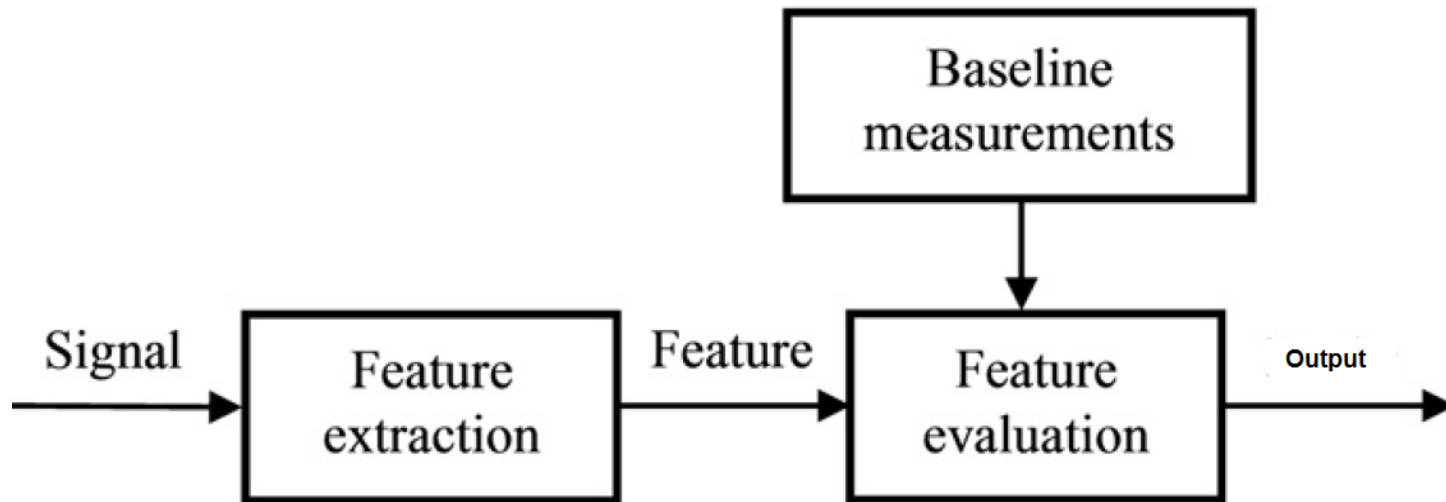
Mathematical Model based Method



Data-Driven based Method



Signal based Method



Data Driven Methods (Machine Learning)

- There are two general classes of problems in Machine learning:
 1. Supervised Learning
 2. Unsupervised Learning

Supervised Machine Learning

- In case of supervised machine learning algorithms, the classes are predetermined.
- Classes can be considered as a finite set (which may be discrete or continuous), which are previously known.
- The data in the supervised learning can be represented as a pairs of $\{X.Y\}$ where Y s are the actual labels of different data elements in X

Supervised Machine Learning

- Supervised machine learning techniques are further divided into two different categories:
 1. Regression: In this case, ML algorithms predict continuous or numerical values.
 2. Classification: In this case, ML algorithms are used to classify each instance into a predefined discrete value or label like, *true* or *false*.
- Regression further divided into two different types of regression methods:
 1. Parametric Regression: It is used when we have a data set with known distribution.
 2. Non Parametric Regression: It is used when the distribution of data is unknown

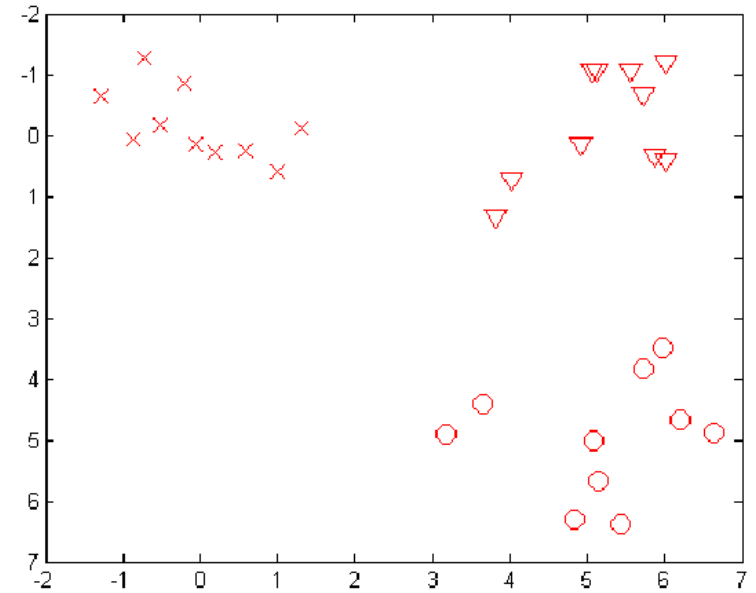
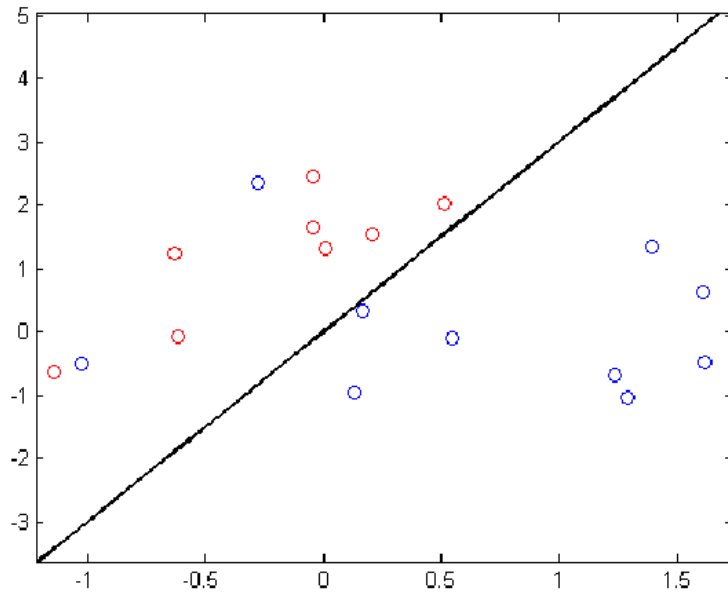
Unsupervised Machine Learning

- Unsupervised learners are not provided with classifications.
- In fact, the basic task of unsupervised learning is to develop classification labels automatically.
- Unsupervised algorithms seek out a similarity between pieces of data in order to determine whether they can be characterized as forming a group. These groups are termed clusters, and there is a whole family of clustering machine learning techniques.

Unsupervised Machine Learning

- Clustering
 - k-means clustering
 - mixture models
 - hierarchical clustering
- Self-Organizing Map
- Expectation Maximization Algorithm

Supervised and Unsupervised Classification



Some Issues in Machine Learning

- Which algorithm can approximate function well and when?
- How does number of training examples influence accuracy?
- How does noisy data influence accuracy?
- What are the theoretical limits of learnability?
- How can prior knowledge of learner can help?

LAB Work

Machine Learning Tools

- Introduction to MATLAB
- Introduction to WEKA

Project Work

List of Projects

1. Machine learning application to predict fault prone program source files
2. Machine learning application for bug triaging (text mining)
3. Machine Learning Application in Text Mining
4. Machine learning application for Image Processing
5. Machine learning application for voice recognition and signal processing
6. Machine learning application in Robotics
7. Machine Learning Application for Autopilot system
8. Machine learning application in Networking (Network Intrusion Detection System)