



Introduction to Machine Learning using Python API for SAS

Nadeem Chaudhry

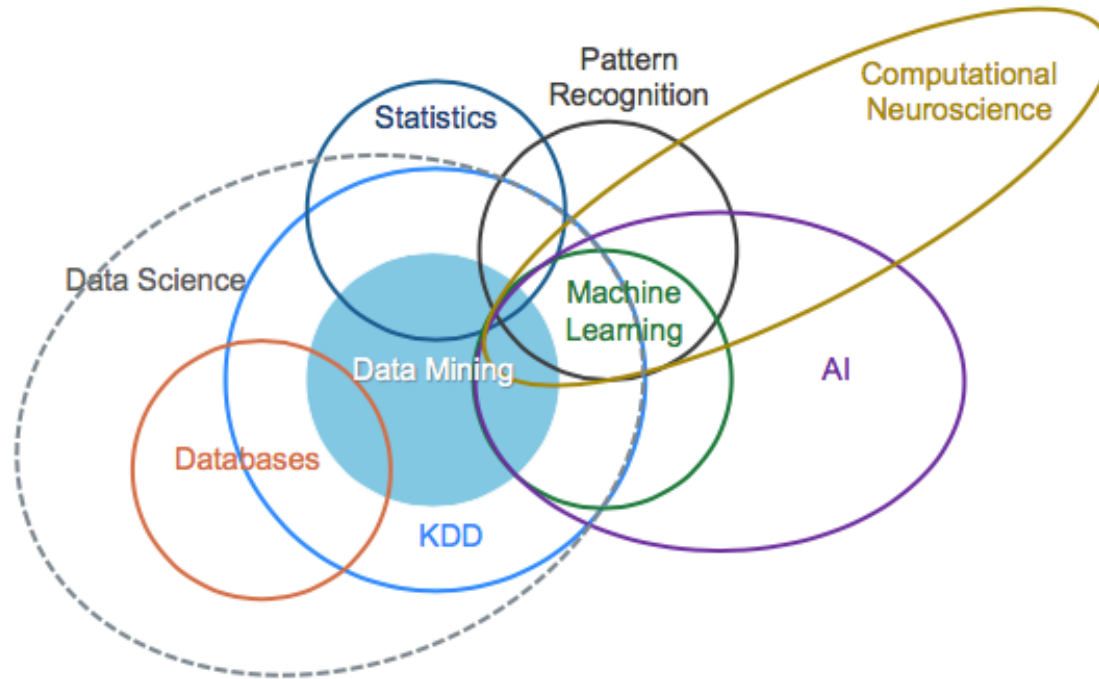
Sr. Solutions Architect, Data Science

<https://www.linkedin.com/in/nadeemc/>

Training Agenda

- Machine Learning Concepts
- Python API for SAS
 - Human Resource – Kaggle Dataset
 - Build & compare Machine Learning model
 - Hands-on exercise
- Machine Learning Algorithms

Many Disciplines of Data Mining (Source – KDD)



Types of Machine Learning Algorithms

- Supervised Learning
 - Make predictions – draw inference from a set of labeled example
- Unsupervised Learning
 - Discover intrinsic patterns in unlabeled data
- Semi-supervised Learning

Types of Machine Learning Algorithms

- Supervised Learning

- Make predictions – draw inference from a set of labeled example

- Business Problems

- Which customers are likely to attrite in the next 30 days?
- Which transaction is likely to be fraudulent?

- Unsupervised Learning

- Discover intrinsic patterns in unlabeled data

- Business Problems

- Can we find clusters or segments in customer shopping behavior?
- Detect anomalous behavior?

Training Data

inputs						label
■	■	■	■	■	■	■
■	■	■	■	■	■	■
■	■	■	■	■	■	■
■	■	■	■	■	■	■
■	■	■	■	■	■	■

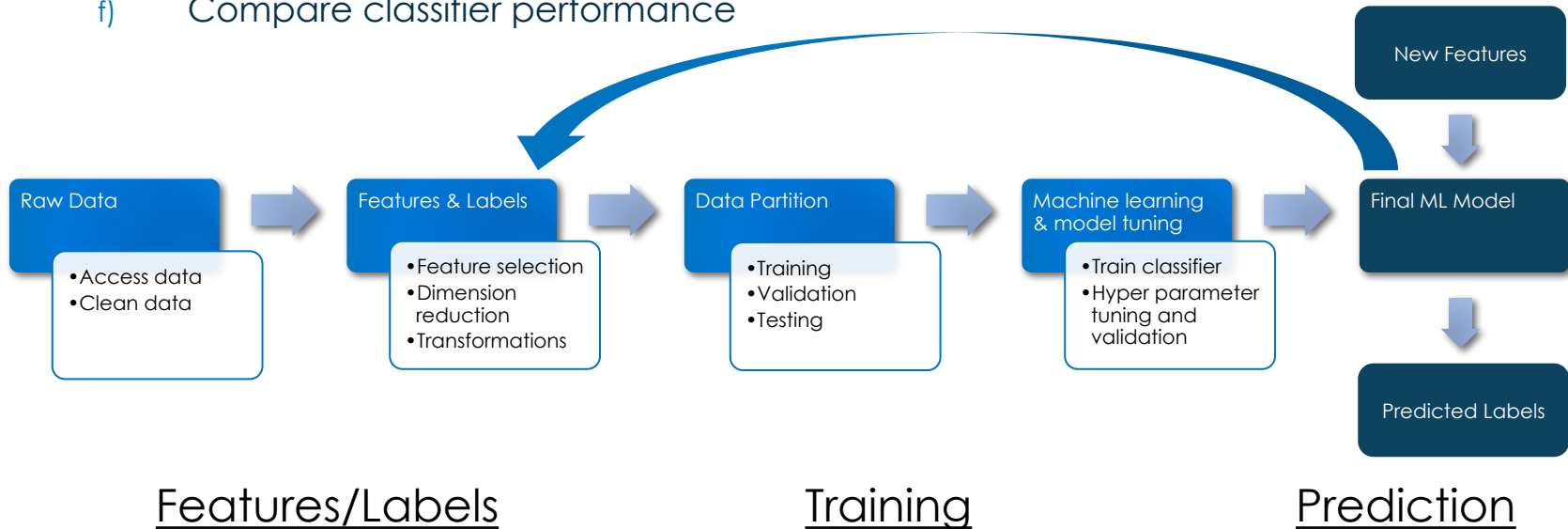
inputs					
■	■	■	■	■	■
■	■	■	■	■	■
■	■	■	■	■	■
■	■	■	■	■	■
■	■	■	■	■	■

Supervised Learning

- **Classification:** When data are being used to predict a categorical variable, supervised learning is also called classification.
 - assigning a label or indicator, either dog or cat to an image.
 - With two labels, this is called binary classification.
 - more than two categories, the problems are called multi-class classification.
- **Regression:** When predicting continuous values, the problems become a regression problem.

Machine Learning Workflow

- a) Upload data to the CAS server
- b) Data Manipulation & Exploration
- c) Build Classifiers
- d) Score Data
- e) Assess Classification Models
- f) Compare classifier performance



Machine Learning Algorithms

