Introduce Machine Learning

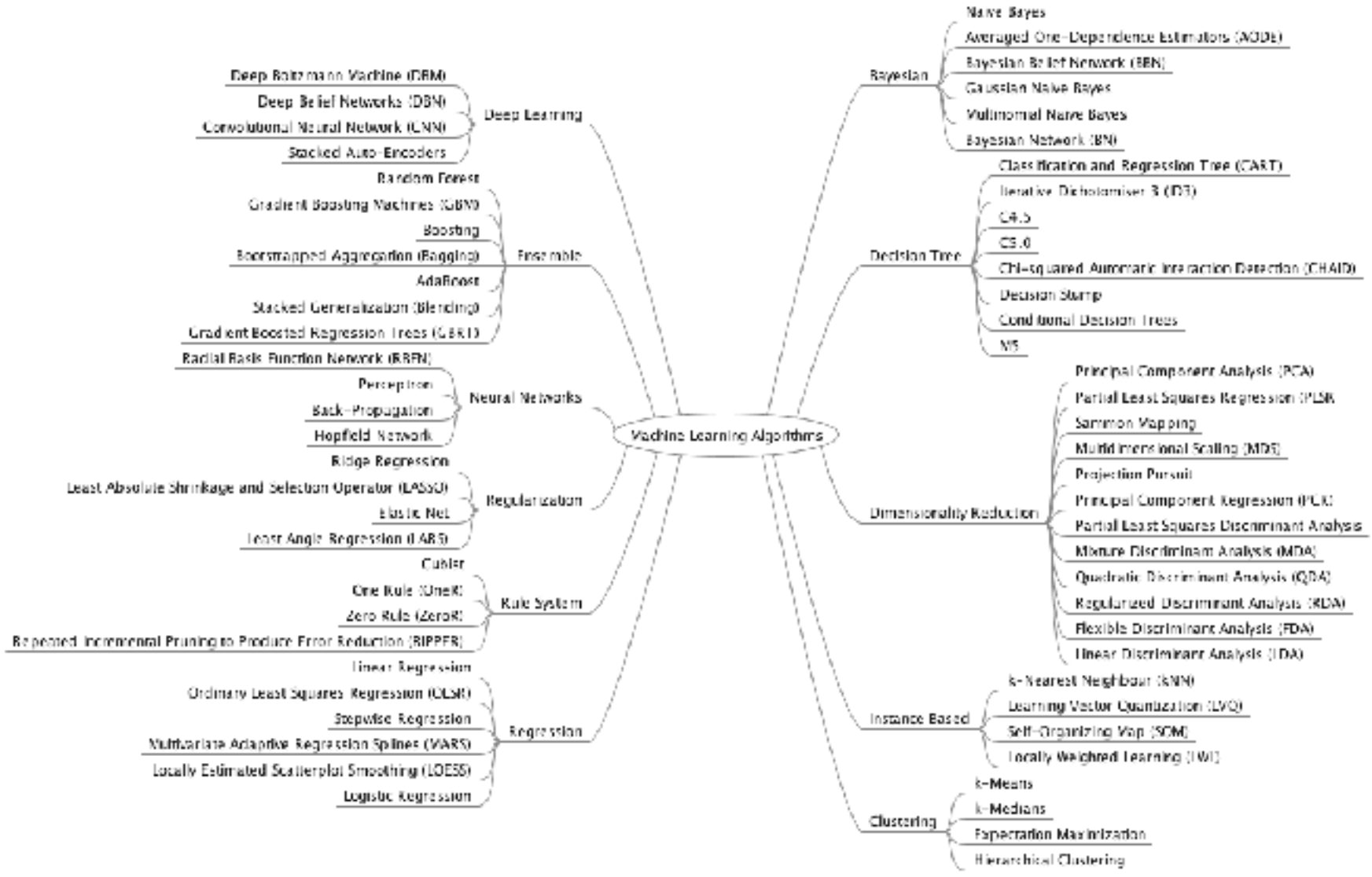
Presented by Tae Geun Kim

- Table of Contents

Types of Machine Learning

Types of Machine Learning

- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning
- Evolutionary Learning


Machine Learning Algorithms
Bayesian
Naive Bayes
Averaged One-Dependence Estimators (AODE)
Bayesian Belief Network (BBN)
Gaussian Naive Bayes
Multinomial Naive Bayes
Bayesian Network (BN)
Decision Tree
Classification and Regression Tree (CART)
Iterative Dichotomiser 3 (ID3)
C4.5
C5.0
Chi-squared Automatic Interaction Detection (CHAID)
Decision Stump
Conditional Decision Trees
M5
Dimensionality Reduction
Principal Component Analysis (PCA)
Partial Least Squares Regression (PLSR)
Sammon Mapping
Multidimensional Scaling (MDS)
Projection Pursuit
Principal Component Regression (PCR)
Partial Least Squares Discriminant Analysis
Mixture Discriminant Analysis (MDA)
Quadratic Discriminant Analysis (QDA)
Regularized Discriminant Analysis (RDA)
Flexible Discriminant Analysis (FDA)
Linear Discriminant Analysis (LDA)
Instance Based
k-Nearest Neighbour (kNN)
Learning Vector Quantization (LVQ)
Self-Organizing Map (SOM)
Locally Weighted Learning (LWL)
Clustering
k-Means
k-Medians
Expectation Maximization
Hierarchical Clustering
Deep Learning
Deep Boltzmann Machine (DBM)
Deep Belief Networks (DBN)
Convolutional Neural Network (CNN)
Stacked Auto-Encoders
Ensemble
Random Forest
Gradient Boosting Machines (GBM)
Boosting
Bootstrapped Aggregation (Bagging)
AdaBoost
Stacked Generalization (Blending)
Gradient Boosted Regression Trees (GBRT)
Neural Networks
Radial Basis Function Network (RBFN)
Perceptron
Back-Propagation
Hopfield Network
Regularization
Ridge Regression
Least Absolute Shrinkage and Selection Operator (LASSO)
Elastic Net
Least Angle Regression (LARS)
Rule System
Cubist
One Rule (OneR)
Zero Rule (ZeroR)
Repeated Incremental Pruning to Produce Error Reduction (RIPPER)
Regression
Linear Regression
Ordinary Least Squares Regression (OLSR)
Stepwise Regression
Multivariate Adaptive Regression Splines (MARS)
Locally Estimated Scatterplot Smoothing (LOESS)
Logistic Regression

We will divide ML by two ways

- Statistical Learning (Based on Statistics)
- Deep Learning (Based on Induction)

For Statistical Learning,

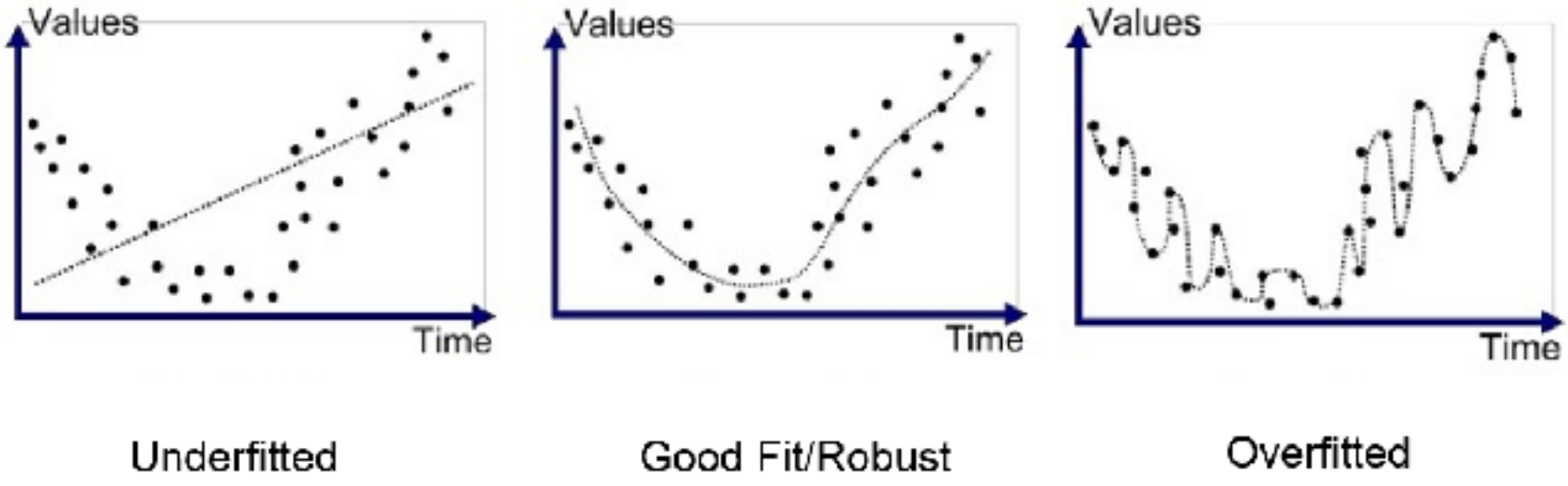
- Statistics
- R (or Scipy or Julia)
- DIY (Do It Yourself with your own languages)

For Deep Learning,

- Tensorflow
- Torch / PyTorch
- MXNet or Other frameworks

Scoring Machine Learning Algorithm

1. Overfitting



Underfitted | Good Fit/Robust | Overfitted

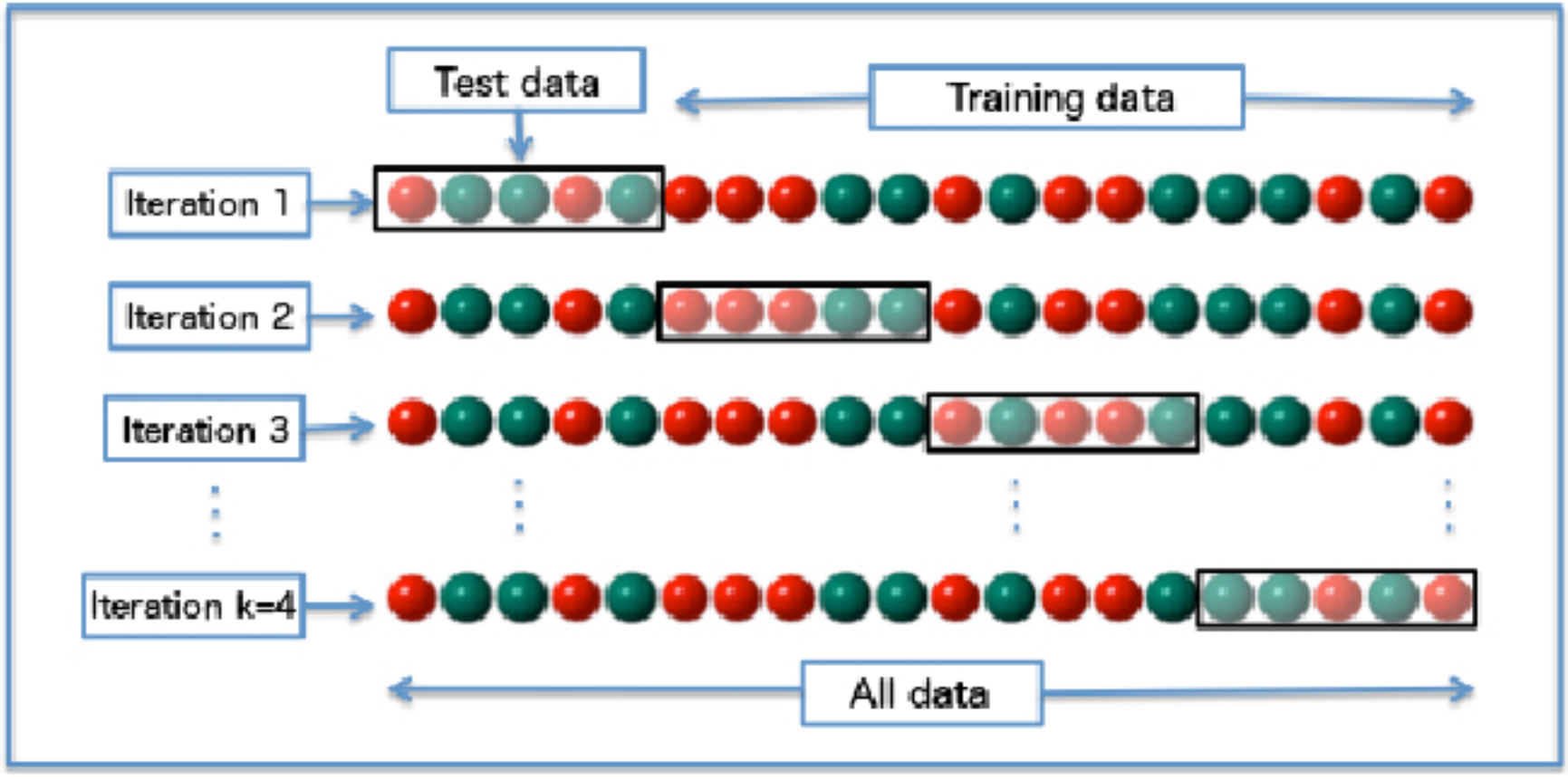
We need third data sets - **validation set**. This procedure called **Cross Validation**.

But in some cases, we can't get enough labeled data. So, we need **semi-supervised learning**.

There are some ways :

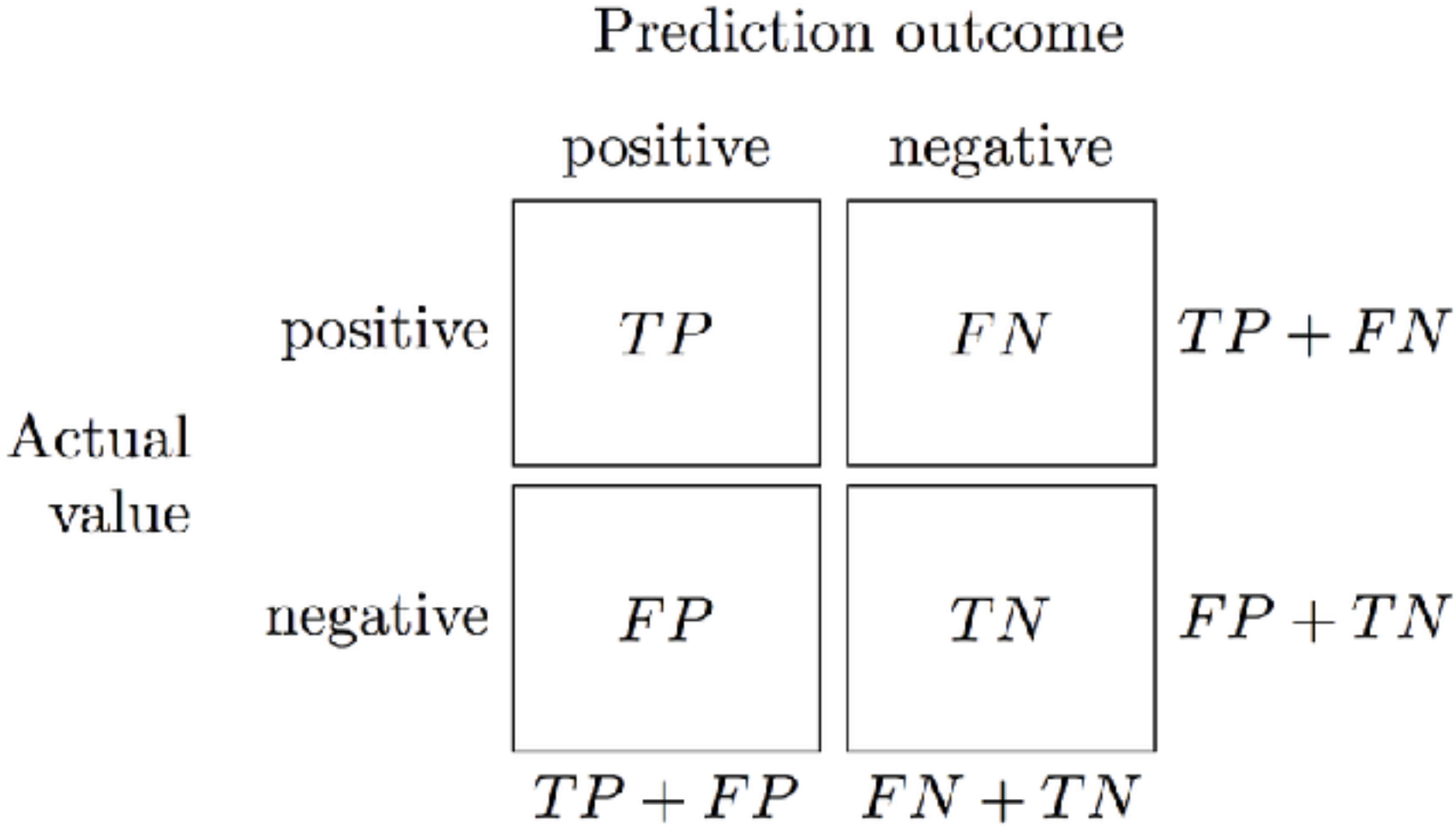
- Leave Some Out (Leave one out)
- Multifold cross validation

Example : Multifold cross validation



2. Confusion Matrix

Actual \ Predicted	Cat	Dog	Rabbit
Cat	5	3	0
Dog	2	3	1
Rabbit	0	2	11


Prediction outcome
positive
negative
Actual value
positive
negative
TP
FN
FP
TN
$TP + FN$
$FP + TN$
$TP + FP$
$FN + TN$

Predicted \ Actual	Cat	Non-cat
Cat	5 TP	3 FN
Non-cat	2 FP	17 TN

- Notations

$$P = TP + FN,\ N = FP + TN$$

- Accuracy

$$ACC = \frac{TP + TN}{P + N}$$

- Sensitivity, Recall, True positive rate

$$TPR = \frac{TP}{P}$$

- Specificity, True negative rate

$$TNR = \frac{TN}{N}$$

- Precision, Positive predictive value

$$PPV = \frac{TP}{TP + FP}$$

- High Recall $\rightarrow$ the class is correctly recognized (small FN)
- High Precision $\rightarrow$ an example labeled as positive is indeed positive (small FP)
- High Recall, Low Precision $\rightarrow$ Miss a lot of positive examples, but those we predict as positive are indeed positive (low FP)

- F-measure

$$F = \frac{1}{\alpha \frac{1}{PPV} + (1-\alpha)\frac{1}{TPR}} = \frac{(\beta^2+1)PPV \times TPR}{\beta^2 PPV + TPR}$$

- F_1-measure ($\beta = 1$)

$$F_1 = 2 \times \frac{TPR \times PPV}{TPR + PPV}$$

- $F_1 \rightarrow 1$: Best!
- $F_1 \rightarrow 0$: Worst!