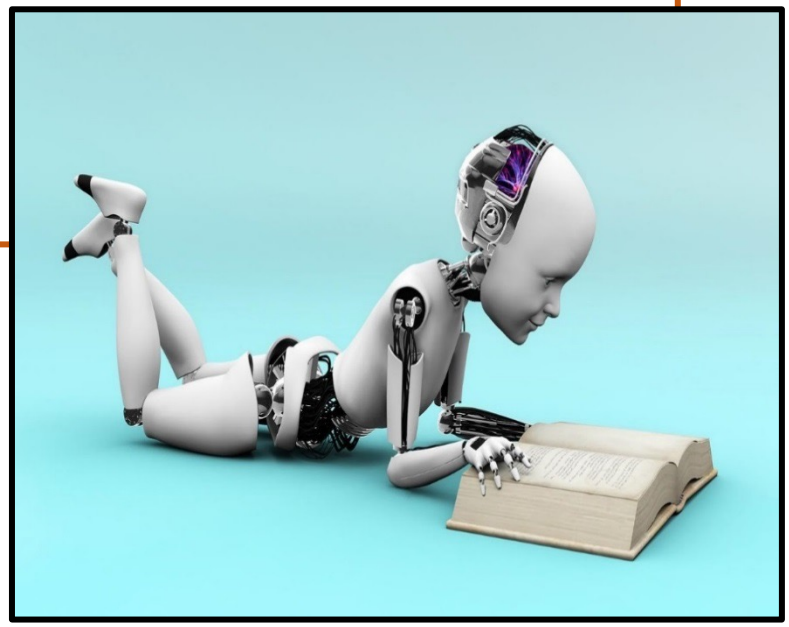
Machine Learning from Data, Summary of topics

Zohar Yakhini & Ariel Shamir

Riechman University

2022/TASHPAB



Reichman University

Types of Learning Tasks

- Regression
  - Given training data $\{x_i, y_i\}$, learn a function f to be used to predict the value y for a new data point x: $\hat{y} = f(x)$
- Classification
  - Given training data $\{x_i, t_i\}$ where $t_i \in \{0,1\}$, learn a mechanism that determines, for a new data point x, its label $t(x)$
  - Also applies to multiple categories
- Parameter Estimation
  - Given $\{x_i\}$ find a PDF that best explains the data
- Unsupervised Learning
  - Given $\{x_i\}$ find regularities, such as clustering

Techniques/topics we have learned and their properties	Numerical or categorical features/attributes?	Algorithm Type	Related techniques and concepts; Comments
Linear regression	Numerical only	Regression	Normalization, Gradient descent, pseudo-inverse Ridge, LASSO
Decision trees	Any type of attributes. May need to scale/normalize	Both regression and Classification	Entropy, Gini
Bayes classifiers (full and naïve)	Both (Irises and playing tennis)	Classification	Probability distributions, prior and posterior probabilities
MLE		Estimation and inference	Poisson distribution, Binomial distribution, Normal distribution, **EM algorithm**
KNN	Mostly numerical but can be adapted	Both regression and Classification	
Perceptron and dual perceptron	Numerical	Classification	Mapping into higher dimensions, Cover's Thm, Kernels
SVM (soft margins, hard margins)	Numerical	Classification	Lagrange multipliers, kernels, slack variables

Techniques/topics we have learned and their properties	Numerical or categorical features/attributes?	Algorithm Type	Related techniques and concepts; Comments
Logistic regression	Numerical	Classification	Gradient descent
Sample complexity		Classification	Bounds on sample complexity, PAC learning
VC dimension		Classification	
Clustering	Numerical	Unsupervised	K-means, Naïve cluster growing, hierarchical clustering
Performance evaluation	Both	Both regression and Classification	Confusion matrix, Cost function, TPR/FPR and ROC curves, PR curves, conf intervals
PCA, LDA	Numerical	Unsupervised & Supervised	Dimensionality Reduction
tSNE, MDS	Numerical	Unsupervised & Supervised	Only introduced very briefly

General techniques and principles

- Gradient descent
- Pseudo-inverse
- Stochastic, batch, mini-batch approaches (to running through training data)
- Cross validation
- Split to training and test data
- Multidimensional probability distributions; the covariance matrix
- Confusion matrix, ROC curves, PR curves, Conf Intervals

General techniques and principles

- Vector geometry in $\mathbf{R}^n$
- MLE, EM algorithm
- Confidence intervals for the total error
- Histograms for representing distributions
- Entropy, Gini index
- Conditional independence
- Cover's Thm
- Total variance principle

General techniques and principles

- Lagrange multipliers for solving constrained optimization problems
- Kernels
- Sample complexity
- VC dimensions
- Feature selection techniques and how to work with them
- Dimensionality reduction & representation learning

Happy Summer!!

Reichman University