

General Notes

CS, ML and Stats

Patrick Daly

1. Computer Science
 - (a) Algorithms
 - (b) Data Structures
 - (c) Linux
2. Machine Learning (combine with stats?)
 - (a) Supervised
 - i. Ordinary Least Squares
 - ii. Logistic Regression
 - iii. Linear Discriminant Analysis
 - iv. Support Vector Machines
 - v. K-Nearest Neighbors
 - vi. Gaussian Process
 - vii. Decision Trees
 - viii. Random Forest
 - ix. Naive Bayes
 - x. Deep Learning
 - A. Deep Feedforward Networks
 - B. Convolutional Networks
 - C. Recurrent Networks
 - D. Long Short-Term Memory (LSTM)
 - E. Autoencoders
 - (b) Unsupervised
 - i. Gaussian Mixture Models
 - ii. K-Means
 - iii. Density-Based Spatial Clustering of Applications with Noise (DBSCAN)
 - iv. Spectral Clustering
 - v. Hierarchical Clustering
 - vi. Factor Analysis
 - vii. Independent Component Analysis (ICA)
 - viii. Principal Component Analysis
 - ix. Non-Negative Matrix Factorization (NMF)

- x. Latent Dirichlet Allocation (LDA)
- xi. PCA
- xii. Outlier Detection?

3. Linear Algebra

- (a) Norms
 - i. Euclidean or Frobenius (L^2)
 - ii. Manhattan (L^1)
 - iii. Infinity (L^∞)
 - iv. Nuclear (L^0)
 - v. Spectral
- (b) Special Matrices & Vectors
 - i. Diagonal
 - ii. Symmetric
 - iii. Positive Definite
 - iv. Positive Semi-Definite
 - v. Negative Definite
 - vi. Negative Semi-Definite
- (c) Eigendecomposition
- (d) Singular Value Decomposition (SVD)
- (e) Principal Component Analysis (PCA)
- (f) Independent Component Analysis (ICA)
- (g) Canonical Component Analysis (CCA)
- (h) Factor Analysis

4. Statistics

- (a) Probability Theory
- (b) Distributions
 - i. Discrete
 - A. Binomial
 - B. Geometric
 - C. Uniform
 - D. Exponential

- E. Poisson
- ii. Continuous
 - A. Normal
 - B. Log-Normal
 - C. Laplace
 - D. Gamma
- (c) Combinatorics