CptS 437 — Exam-01 Notes

ML Overview

- **features**:

- **training data**: a random sample of input/output pairs drawn from a distribution $\mathcal{D}$.

Decision Tree

- Entropy & Information Gain:

$$E = -\sum_{i}^{N} p_i log_2 p_i \quad \text{where} \quad i \in \{+, -\}$$

$$G(T, A) = E(T) - \sum_{v \in A} \frac{|S_v|}{|S|} E(S_v)$$

Bayes Optimal Classfier

- Representation bias: making decision based incorrect perceptions of similar situations.

- Inductive bias: input is only dependent on ouput label and an assumed normal distribution.

- No free lunch theorem: implies that there is no single best optimization algorithm, because all algorithms perform equally when performance is averged across all possible problems.

Geometric Algorithms

- k-nearest neighbors:

- distance function:

$$d(a, b) = \left[\sum_{d=1}^{D} (a_d - b_d)^2 \right]^{1/2}$$

- normalization: to make data consistent in some way.

- feature normalization: adjust all examples consistently for a feature.

- example normalization: example is ajusted individually.

- inductive bias: assumes that nearby points have the same label

- decision boundaries: colored differently for positive and negative.

- boundaries for specific classifiers:

- features as dimensions:

K-Means Clustering

- algorithm: is an unsupervised learning algorithm, page 37.
- iterative refinement: starts with a guess and gradually build up its quality.
- convergence: guaranteed to converge and converge quickly.
- computational complexity:
- optimality:
- sum of squared error:
- elbow method:
- the curse of dimensionality:
- variations on clustering, density and hierachy

Perceptron

- algorithm, weighted sum of inputs, neuron, activation function, bias term
- impact of weights, training, modify weights
- number of iterations, order of training instances, overfit, convergence
- voting percetron, averaged perceptron
- decision boundary, linear separability
- data preparation, imbalanced class distribution