CSCI 447 Final Exam Outline

I. Machine Learning Process
- A. Gathering evidence
- B. Modifying the model
- C. Evaluating effectiveness

II. Machine Learning Dimensions
- A. Input
  - 1. Features
  - 2. Data
- B. Algorithms
- C. Output
- D. Evaluation
- E. Other

III. Probability
- A. Random Variables
- B. Joint and Marginal Distributions
- C. Conditional Distributions
- D. Product Rule, Chain Rule, Bayes Rule
- E. Independence

IV. Linear Algebra
- A. What is Linear?
- B. Inputs and Operations
- C. Operations as Inputs

~~V. Python / AWS~~
- ~~A. Data (S3 Buckets)~~
- ~~B. Processing (in AWS)~~
- ~~C. Output (Evaluating Results)~~

VI. Data Preparation
- A. Dataset Issues
  - 1. Missing Data
  - 2. Erroneous Data
  - 3. Data Format
- B. Multiple Source Issues
  - 1. Different Features
  - 2. Different Answers
  - 3. Population Sample

VII. Linear Regression
- A. Linear Models
- B. 1D Ordinary Least Squares (OLS)
- C. Solution of OLS
- D. Interpretation
- E. Anscombe's Quartet
- F. Multivariate OLS
- G. OLS Pros and Cons

VIII. Logistic Regression
- A. Math Behind Logistic Regression

- B. Visualizing Logistic Regression
- C. Loss Function
- D. Batch / Full Regression
- E. Gradient Descent
    1. OLS
    2. Logistic Regression
- F. Comparing OLS and Logistic Regression
- G. Multi-Class Logistic Regression

IX. Summing Up Regression
- A. Assumptions / Extensions of Linear
- B. Beyond Linearity
- C. Extreme Learning Machines
- D. Overfitting
- E. Regularization
- F. Cross Validation

X. Clustering
- A. Hierarchical
- B. K-Means

XI. Nearest Neighbor
- A. K-Nearest Neighbor

XII. Neural Networks
- A. Perceptron
    1. Error Driven Classification
    2. Linear Classifiers
    3. Weight Updates
    4. Improving the Perceptron
- B. Multi-Layer Networks
    1. Overview
    2. Representation
    3. Computing Output – Feed Forward
    4. Vectorizing Across Examples
    5. Activation Functions
    6. Gradient Descent
    7. Backpropagation
- C. Deep Networks
    1. Deep L-Layer Network
    2. Forward Propagation
    3. Matrix Dimensions
    4. Why Deep Representation?
    5. Building Blocks
    6. Backpropagation
    7. Parameters vs. Hyperparameters
    8. Brain Analogy
- D. Convolutional Networks
    1. Overview
    2. Architecture
    3. Intuition
    4. Example

- 5. Visualization
- E. Recurrent Networks
  - 1. Introduction
  - 2. Sequence Data
  - 3. Sequential Memory
  - 4. Recurrent Neural Networks
  - 5. Vanishing Gradient
  - 6. LSTMs and GRUs
- G. Support Vector Machines
  - 1. Optimization Objective
  - 2. Large Margin Intuition
  - 3. Math Behind Large Margin Classification
  - 4. Kernels
  - 5. Using an SVM
- H. Network Considerations
  - 1. Setting up the Network
    - a. Bias / Variance
  - 2. Regularization
    - a. Dropout
  - 3. Optimization

XIII. Bayesian Networks

- A. Overview
  - 1. Representation
  - 2. Conditional Independences
    - a. D-Separation
- B. Inference
  - 1. Probabilistic Inference
    - a. Inference by Enumeration
    - b. Variable Elimination
      - 1. Factors
- C. Learning the Network
  - 1. Learning the Parameters
    - a. Complete Data
    - b. Incomplete Data
  - 2. Learning the Structure

XIV. Genetic Algorithms

- A. Overview
  - 1. Evolutionary Model
  - 2. Local Search
- B. Genetic Operators
  - 1. Selection
  - 2. Crossover
  - 3. Mutation
- C. Development Process

XV. Trees and Forests

- A. Decision Trees
  - 1. Overview
  - 2. Decision Boundaries

- 3. Tree Construction
- 4. Information Gain

B. Random Forests

- 1. Benefits
- 2. Construction
- 3. Operation
- 4. Hyperparameters

XVI. Ensemble Learning

A. Algorithm Comparison / Selection

B. Boosting

- 1. Overview
- 2. Main Ideas
  - a. Use mistakes to train subsequent models
  - b. Use many learners
  - c. Use simple algorithms
  - d. Weight sample contribution to error
  - e. Weight the learners
  - f. Change weights

XVII. Measuring Performance

A. Confusion Matrix

- 1. Precision / Recall
- 2. Sensitivity / Specificity

B. F1 Score

C. Gain / Lift Charts

D. Kolmogorov Smirnov Chart

E. ROC / AUC

F. Regression Metrics

G. Kappa Statistic

XVIII. Ethical Considerations

A. Questions

- 1. Data model is trained on
- 2. Definition of success

B. WMDs

C. Model Audit

- 1. “Do no harm”
- 2. Is it meaningful and fair?
- 3. Do those affected have access to understanding the underlying scoring?