

CS 383

Midterm Exam

Midterm Exam Rules

- Rules:
 - Close book, however you may bring one piece of paper with notes.
 - I will provide a sheet with necessary rules of calculus
 - You may not use a calculator
 - All computations should be simple enough to do by hand and/or to leave in a formula notation.
 - In class Tuesday February 21st
 - Have entire period to complete (80 minutes)

Midterm Exam Format

- Format:
 1. Definitions
 2. Open ended theory questions
 3. Computational Questions

Studying Tips

- Look over homeworks
- Look over slides, in particular:
 - Examples
 - Derivations
- Refresh yourself on math (if necessary).

Walking through the notes...

- Three types of problems
 - Clustering
 - Regression
 - Classification
- Standardizing data
 - When, how, why
- Curse of Dimensionality
- Feature Selection
 - Exhaustive
 - Information Gain/Entropy

Walking through the notes...

- Feature Projection
 - Projecting points onto a projection matrix
 - PCA
 - Intuition
 - Mechanics
 - LDA
 - Intuition
 - Mechanics
- K-Means
 - Intuition
 - Proof (intra-cluster LSE)

Walking through the notes...

- Expectation Maximization
 - Intuition
 - Mixture Model intuition
- Hierarchical Clustering
 - Building a hierarchical agglomerative cluster tree
 - Intra vs inter cluster distance
 - Link types
 - Purity

Walking through the notes...

- Basic supervised concepts
 - Training, testing, validation sets
 - How are each used and why?
 - Cross validation
 - How and why?
 - Over/under fitting
 - Identifying and solving
 - Bias/variance

Walking through the notes...

- Linear Regression
 - Model form
 - Derivation of closed form solution using LSE
 - Finding closed for solution using LSE
 - Applying model
 - Gradient descent
 - Intuition
 - Derivation for GD for LR using LSE
 - Impact of learning rate
 - Regularization term

Walking through the notes...

- Classification Basics
 - Error types: TP, TN, FP, FN
 - Binary evaluation metrics
 - Precision, recall, f-measure
 - Precision-Recall graph
 - Impact on varying threshold
 - Accuracy metric

Walking through the notes...

- Statistical Classification
 - Computing/using joint distribution tables (JDs)
 - Classifying by inference
 - Use of Bayes' Rule
 - Posterior, Prior, Likelihood, Evidence
 - Naïve Bayes
 - Pros/Cons
 - Intuition/mechanics