Test for CSC 2515
Overview

• The test will be marked out of 100

• PART A: 8 easy questions
  – Worth 5 points each.
  – Each question should take ~3 minutes.
  – You should answer all of them

• Typical easy question:
  – a) Write down the softmax function
  – b) What is the purpose of the denominator?
Overview (cont.)

- Second Half: Four multi-part questions, you select 3

- Each question should take ~10 minutes.

- Typical question (one part):

  Under what conditions is the algorithm for K-means clustering equivalent to the algorithm for fitting a mixture of Gaussians?
Summary

• What to study?
  – Material covered in lectures and tutorial
  – Use the books/readings as back-up, to help understand the methods and derivations

• The exam is closed book and notes
  – Do not focus on memorizing formulas, but instead main ideas and methods
Topics to Study

• Week 1: Linear Regression
  • When is minimizing the squared error equivalent to Maximum Likelihood Learning?
  • Online vs. Batch learning
  • Regularized least squares

• Week 2: Classification
  • Linear Classifier: defn, utility
  • Logistic Regression: form, probabilistic interpretation
  • Loss functions
  • Regularizer: form, rationale
  • Use of validation set

• Week 3: Nonparametric methods
  • kNN: decision boundaries; compare to parametric methods
  • Decision trees: information gain, decision boundaries

• Week 4: Generative classifiers
  • Compare to discriminative classifiers
  • Underlying assumptions
  • Forms of GBC
  • Naïve Bayes: basic assumptions, application
• Week 5: Neural Networks
  • Backprop: when applicable (constraints on cost, activation functions)?
  • Weight constraints: how and why
  • Methods to prevent over-fitting
  • Alternative activation, loss functions

• Week 6: Clustering & EM
  • Hard vs. soft k-means
  • Mixture models
  • EM: definition of steps in algorithm
  • is convergence guaranteed?
  • what are responsibilities?
  • understand (but not memorize) eqns, objective
  • what does it mean that this is a generative model?

• Week 7: Continuous Latent Variable Models
  • PCA: motivation, basic algorithm
  • Probabilistic vs. standard PCA, and compare to full Gaussian
  • what is the objective function(s)?
  • what is a principal component?
  • PCA vs. clustering?
  • How does PCA compare to autoencoders
Topics to Study (cont.)

- **Week 8: Support Vector Machines**
  - what is the kernel trick?
  - when can the kernel trick be applied?
  - what is its purpose
  - how is an SVM similar and different than a linear classifier?
  - what is a support vector?
  - What is the objective function?

- **Week 10: Combining Models**
  - When is ensemble a win?
  - Bagging vs. Boosting
  - Boosting: main steps in algorithm; 2 different types of weights, rationale
  - Decision trees: algorithm; choosing nodes via MI; decision boundaries, stumps
  - Mixture of experts

- **Week 11: Reinforcement Learning**
  - Compare to other forms of learning
  - Q learning algorithm: updates, objective
  - Exploration/exploitation
Topics *NOT* to Study

- Free energy
- Structured prediction
- Non-deterministic Q-learning
- Derivations