Final exam

Milos Hauskrecht
milos@cs.pitt.edu
5329 Sennott Square, x4-8845

people.cs.pitt.edu/~milos/

Final exam

When: **April 26, 2019 at 8:00am**

Exam:
- **closed book**
- **cumulative**

Calculators:
- **Bring your own calculator**
- **No cell phones**

Exam: study material

- **Lecture notes**
- **Textbook: Bishop**
- **Homework assignments**
- **TA's recitation notes:**
  - **http://people.cs.pitt.edu/~jlee/teaching/cs1675/**

+ Basic algebra, calculus, probabilities

- **Reviews on January 15 and January 22**

Matlab coding and programs

Are there any Matlab programming questions in the exam?

- No, no Matlab code during the exam

Probability distributions

Do I need to know the exact formulas of all distribution models we covered?

- Yes, for Bernoulli
- No for others, but you need to know when the formula is given to you:
  - What are the parameters
  - Ranges of values the specific distribution is defined on
    - Say Gamma is on non-negative reals, Beta is on [0,1] interval
  - How to use it calculate the probability of a specific instance

Derivations in the notes/book

Do I need to know how to replicate the derivations?

- No, for very long ones, e.g. gradient descend for the neural networks, but you need to understand the principles of what has been done
- Yes, for short ones, such as ML estimates for the sequence of Bernoulli trials.

Understand the concepts/terminology and methods

Be able to describe basic concepts used throughout the course. Examples:

- Cross-validation
- Gradient-descend
- Overfitting
- Error function
- Maximum likelihood
- Support vectors
- Regularization penalty
- Impurity measure
- Distance metric
- Similarity
- Linkage distance
- Model bias and model variance
- Filter methods
- PCA
- Bootstrap
- Exploration-exploitation dilemma
- Boltzman exploration
- **Etc …**

Questions (examples):

$$AB = \begin{bmatrix} 3 & -4 \\ 10 & 9 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 6 & 5 & 4 \end{bmatrix}$$

Questions (examples):

Bernoulli trials. Assume the probability of head is 0.6. Assume we have observed the following sequence of coin flips: Tail-Head-Tail. What is the probability of seeing this sequence of outcomes.

Questions (examples):

Log function.

- Draw a graph of a log function.
- Describe the property of the function (monotonicity, trend)?
- Assume a function f that is restricted to positive values.
  - Argue that:
    finding the value u^* that maximizes $f(u)$
    can be found by maximizing log $f(u)$

Questions (examples):

Regression. Assume you have a dataset D that consists of (**x**,y) pairs. You believe the relation between x and y could be modeled using a combination y = a sin(x) + bcos(x) + c, where a, b, c are parameters. Explain briefly how could you find the best model using the linear regression solver.

Questions (examples):

Support vector machines. The solution of the SVM is defined in terms of Lagrange parameters alpha. There is one alpha α_j for every training instance $(\mathbf{x}_j, y_j)$. Answer the following questions:

- What is the meaning of $\alpha_j > 0$?
- What is the meaning of $\alpha_j = 0$?
- How are weights **w** (representing the discriminant functions and decision boundaries) defined in terms of α_j and $(\mathbf{x}_j, y_j)$ for training instances? Give an expression.
- Are all training instances important?

Questions (examples):

BBN.

Assume a binary classification problem with 3 binary input variables x1, x2, x3. Assume you choose to define the classifier for the problem using the Naïve Bayes model.

Part a. Draw the BBN graph corresponding to the Naïve Bayes model.

Part b. How many parameters are needed to define the model.

Part c. Write an expression for calculating

P(y=1| x1=a1, x2=a2, x3=a3) in terms of parameters of the Naive Bayes model.

Questions (examples):

Clustering.

Explain how is the linkage distance used in hierarchical clustering? What does it measure?

How is the min linkage distance defined?

Questions (examples):

Reinforcement learning.

Let R(*x*,*a*) defines an expected one step reward for performing an action *a* in state x. Explain two solutions the expectation can be estimated from state, action, reward trajectories in reinforcement learning.

Questions (examples):

True/false questions with explanation.

Please note justification/explanation is needed in addition to marking True/False

The greedy wrapper method always finds the optimal set of inputs.

True/false