

# Mid-Term Review

Lecturer: Ji Liu

# Overview

- Midterm: 10 questions (Basic questions 80% + tricky questions 20%)
- Time: 75 mins (use your time wisely)
- You can expect to receive >80 points in exam if you well understand all contents in this slides
- Tricky questions are still covered by this slides, but require deeper understanding on all algorithms

# Probability

- Joint Probability
- Conditional Probability
- Marginal Probability
- Bayes Rule
- Independence
- Conditional Independence

# Probability

- Basically, given the joint probability, you should know how to compute the conditional probability, marginal probability, and know how to verify the independence and conditional independence
- Understand Bayes rule and know how to use Bayes rule to compute the probability you desire

# Naive Bayes

- Key assumption: Conditional independence
- How to predict the label?

# Naive Bayes

- Given the training samples, know how to use the Naive Bayes rule to predict the labels. Basically, you know the procedure of Naive Bayes classifier
- Understand the key assumption of Naive Bayes and its limitation

# K-Means

- Algorithm procedure (given you the data points, know how to use this algorithm to generate the clustering result)
- Distortion function (understand the geometric meaning of this distortion function)

# Hierarchical Clustering

- Algorithm procedure (given you the data points, know how to use this algorithm to generate the clustering result)
- Cluster distance (understand several commonly used cluster distance and know which one to use in different scenarios)



# KNN

- Algorithm procedure (given you the data points, know how to use this algorithm to generate the clustering result)
- Decision boundary
- Accuracy evaluation (know how to evaluate accuracy given the testing points)

# SVM

- Understand how to obtain the SVM formulation (understand the meanings of each component:  $w'w$ ,  $C$ , constraint, slack variable, optimization variables)
- Slack variables (Why do we need slack variables? What does it mean if slack variable  $>0$  or  $=0$ ?)

# SVM

- Kernel approach (know why do we use kernel approach)
- Linear separable / non-separable (understand the definition)
- Prediction (know how to use  $w$  and  $b$  to predict the label for testing sample)
- Accuracy evaluation (know how to evaluate the accuracy on given a data set)

# Decision Tree

- Entropy (given a distribution, how to compute the entropy; what does entropy standard for?)
- How to construct a decision tree? (Given you a set of training samples, know how to construct a decision tree)
- Accuracy evaluation (given you a data set, know how to use the decision tree to evaluate the prediction accuracy)
- How to prune a decision tree? (Know how to prune an internal node (change a internal node to a leaf node); know how to evaluate the accuracy on the tuning data set on the pruned decision tree.)