

# Math 342W/642/742W

*Recitation – Day #17 (4.10.25)*

## **I. Trees**

- (i) What is a *tree*?
  
  
  
  
  
- (ii) What is the tree-method are we interested in for machine learning?
  
  
  
  
  
- (iii) What tree-based algorithm will we be implementing?
  
  
  
  
  
- (iv) What are the two types of trees we will be considering?
  
  
  
  
  
- (v) What is being done to the predictor/feature space with this tree-based method?
  
  
  
  
  
- (vi) What are the advantages of tree-based models over the linear based models we have seen?
  
  
  
  
  
- (vii) What are the disadvantages of tree-based models when compared with linear based models?

## II. Regression Trees

- (i) What is the candidate set of functions  $\mathcal{H}$  for regression trees? Compare that with the candidate set for the linear regression model and the logistic regression model.
  
  
  
  
  
  
  
  
  
  
- (ii) How are the “*splits*” of the training data made?
  
  
  
  
  
  
  
  
  
  
- (iii) After each split is made, what is computed?
  
  
  
  
  
  
  
  
  
  
- (iv) What are loss/objective function that we are trying to minimize to find the “best split”?
  
  
  
  
  
  
  
  
  
  
- (v) What type of algorithm is CART described as?
  
  
  
  
  
  
  
  
  
  
- (vi) Give a pictorial example of a regression tree and a partitioned feature space.