University of TEHRAN

Machine Learning

Problem Set 11

Hesam Montazeri
Fahimeh Palizban
Zohreh Toghraee

Azar 19, 1398
(December 10, 2019)

Problem 1: Review Questions
Write a summary of the lectures of this week. Write down all formulas we discussed in the lectures and explain in detail each step of derivations. As a guideline, you may consider the following topics:

(a) Decision/regression trees; bagging; random forests

(b) Boosting

Problem 2: Exponential Loss and AdaBoost.M1
Show that AdaBoost.M1 is equivalent to forward stagewise additive modeling using the exponential loss function. Write down the detailed derivations.

Problem 3: Variance reduction of bagging
For B identically distributed random variables with positive pairwise correlation ρ, prove the variance of the average is $\rho\sigma^2 + \frac{1-\rho}{B}\sigma^2$. Discuss how this formula is related to the idea of random forests.

Problem 4: Conceptual questions
[ISL] chapter 8: questions 1, 3, 4, 5; programming questions 8, 10

Submit your solutions (using Easyclass) by Azar 23.