University of TEHRAN

# Machine Learning

## Problem Set 10

Hesam Montazeri
Fahimeh Palizban
Zohreh Toghraee

Azar 12, 1398
(December 3, 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) Statistical hypothesis testing; p-value

(b) Statistical testing for comparing machine learners

(c) Feature selection methods

**Problem 2: Common statistical hypothesis tests**
For each of the following tests: one-sample t-test; two-sample t-test, paired t-test, Wilcoxon signed-rank test, Mann-Whitney-Wilcoxon, Pearson's chi-squared test, Fisher's Exact Test, Analysis of variance (the F-test), hypergeometric test, Pearson correlation and Spearman correlation test

(a) Explain the aim of the test in your own words.

(b) Define the null and alternative hypotheses.

(c) Explain the main assumptions of the test.

(d) Write down the test statistic and null distribution.

(e) Give an example application of the test in bioinformatics (you may give the same example for several tests, if applicable).

(f) How do you perform this test using a built-in function in the programming language of your choice?

**Important:** In addition, draw your own flowchart on how to choose a statistical test from the above list.

Submit your solutions (using Easyclass) by Azar 16 for the review part and Azar 24 for the problem 2.