Assignment 1
CSE975 - Topics in Machine Learning

Submission Deadline: **20th September 2018**

September 9, 2018

1 Instructions

- Submit your writeup in pdf and all source code (without datasets) in a zip file (with proper documentation).
- Write a script for each programming exercise so that the TAs can easily run and verify your results. Make sure your code runs!
- We will be checking code for plagiarism. Any plagiarism detected will result in all involved students getting a zero for this assignment evaluation.
- For the 3rd Question, Please submit the scanned copy of handwritten answer.

2 Problem 1 [3 Marks]

Consider 2 arm bandit problem, where each arm's reward distribution is Bernoulli. Consider the following 3 problem variants. Implement UCB, Exp3, ϵ-Greedy (with $\epsilon = 0.01, 0.1$) algorithms.

Problem	Arm 1	Arm 2
P1	0.9	0.6
P2	0.9	0.8
P3	0.55	0.45

- Choose the horizon T as 10000.
- For each algorithm, repeat the experiment $N = 100$ times.
- Store the number of times an algorithm plays the optimal arm, for each round $t = 1, \ldots, T$ (averaged over N runs).
- Store the regret in each round $t = 1, \ldots, T$ (averaged over N runs).
- Plot the percentage of times optimal arm played against the rounds $t = 1, ..., T$.
- For each plot, add standard error bars.

Explain the results obtained on each problem instance with each algorithm and correlate the results to the theoretical findings.

3 Problem 2 [4 Marks]

Dr. Vishnu is in-charge of product acquisitions at SmartWheelChair Inc. Due to being registered on a number of suspicious websites, his email is flooded with mails that are both legitimate and Spam. He needs your help to sort his mailbox so that he can keep his job.

3.1 Part 1

- Download the spambase dataset:link.
- Implement the Perceptron and Winnow (with different values of η) algorithms on the dataset features to classify Spam/Not Spam.
- Choose the horizon T as 10000.
- For each algorithm, repeat the experiment $N = 100$ times.
- Store the rate of misclassification an each round $t = 1, \ldots, T$ (averaged over N runs).
- Plot misclassification rate against time t.
- For each plot, add standard error bars.
- Why do you observe this relative performance of the algorithms?

3.2 Part 2

- Download the SMS spam collection dataset:link.
- Build a bag-of-words representation (link) of each sample in the dataset with appropriate text preprocessing (remove stop words, lower-case etc.).
- Implement the Perceptron and Winnow algorithms to classify Spam/Not Spam.
- Choose the horizon T as 10000.
- For each algorithm, repeat the experiment $N = 100$ times.
- Store the rate of misclassification an each round $t = 1, \ldots, T$ (averaged over N runs).
- Plot misclassification rate against time t
- For each plot, add standard error bars.
- Why do you observe this relative performance of the algorithms?

3.3 Part 3

Modify the Spam-base dataset by flipping the labels of 10% randomly chosen points in the training set. Generate the same plots as earlier.

4 Problem 3 [3 Marks]

Consider the ϵ-greedy bandit algorithm as follows.
For $t = 1, 2, ..., T$, repeat

1. Let i^t be the arm with the highest sample mean so far, i.e.,

$$i^t = \arg \max_{k \in \{1,\ldots,n\}} \hat{\mu}_k^{t-1}$$

   where $\hat{\mu}_k^{t-1}$ is the average reward obtained for ark k upto time $t-1$.

2. With probability $1 - \epsilon_t$, play arm i^t and with probability ϵ_t, play a random arm.

For a **two-armed** bandit problem, show that the pseudo regret incurred by the ϵ-greedy algorithm, with $\epsilon_t = t^{-\frac{1}{3}}$, satisfies

$$\hat{R}_T \leq cT^{\frac{2}{3}} (\log T)^{\frac{1}{3}}$$

for some universal constant c.