

## COSC 522 – Homework #4 (Due 10/26)

---

**Problem 1:** (35 pts.) On Gradient Descent (GD). Find the global minimum of  $f(t) = 50 * \sin(t) + t^2$  over  $-10 \leq t \leq 10$ . This problem intends to give you a hands-on experience on how gradient descent works and how it can get trapped at the local minima.

- a) (5) Plot this function. Visualize the multiple local minima and the global minimum.
- b) (25) Implement gradient descent in Python to find the local minimum.
  - i) (15) Pick a starting point at  $t=7$ . What's the minimum? Show the convergence path. Experimenting with different learning rate.
  - ii) (10) Pick a starting point at  $t=1$ . What's the minimum? Show the convergence path. Experimenting with different learning rate.
  - iii) (5) Comment on the results from the above experiments

**Problem 2:** (30 pts.) On classifier fusion. Assume there are two classifiers (L1, L2), performing classification tasks on three objects ( $\omega_1, \omega_2, \omega_3$ ). Assume there are 30 samples in the training data for each category. Following are the confusion matrices generated from the training data.

L1	$\omega_1$	$\omega_2$	$\omega_3$
$\omega_1$	20	5	5
$\omega_2$	3	24	3
$\omega_3$	0	9	21

L2	$\omega_1$	$\omega_2$	$\omega_3$
$\omega_1$	25	2	3
$\omega_2$	3	22	5
$\omega_3$	5	6	19

In the confusion matrix, each row represents the ground truth label and each column represents the label given by the classifier.

- (a) (20 pts.) Derive a lookup table that includes the fused result from all possible combinations of labels from the two classifiers using Naïve Bayesian (show details and justify your selection of fused labels)

L <sub>1</sub> , L <sub>2</sub>	Fused label
1, 1	
1, 2	
1, 3	
2, 1	
2, 2	
2, 3	
3, 1	
3, 2	
3, 3	

- (b) (10 pts.) How to generate the lookup table using BKS from the above confusion matrices, if possible at all? If not possible, then what additional information would you need to generate the lookup table for BKS? What is the difference between Naïve Bayesian and BKS?

**Problem 3:** (35 pts.) Generate the ROC curve for the 3 cases of MPP classifier, applied on the pima dataset (pima.tr and pima.te). Note: 1) No need to draw a smooth curve. Just generate 20 discrete points based on different prior probability ratios ranging from 0 to inf. 2) Make sure the dataset is standardized.

**Bonus:** (+10 pts) Add kNN with k=5 to the ROC curve in Problem 3. You have to think about how to incorporate prior probability ratios into kNN which assumes the prior probability determined by the sample size in the training set.