CS 771A: Introduction to Machine Learning			**Quiz I** (14 Aug 2019)
Name	SAMPLE SOLUTIONS		**30 marks**
Roll No		Dept.	Page **1** of **2**

Instructions:
1. This question paper contains 1 page (2 sides of paper). Please verify.
2. Write your name, roll number, department above in **block letters neatly with ink**.
3. Write your final answers neatly **with a blue/black pen**. Pencil marks may get smudged.
4. Don't overwrite/scratch answers especially in MCQ. We will entertain no requests for leniency.
5. Do not rush to fill in answers. You have enough time to solve this quiz.

Q1. Write **T** or **F** for True/False (write **only in the box on the right hand side**) (**5x2=10 marks**)

1	Classifying natural numbers into prime vs non-prime is a good example of a problem that can be solved using machine learning (e.g. by binary classification)	F
2	Learning with prototypes cannot be used if there are more than two classes	F
3	If $f, g: \mathbb{R} \to \mathbb{R}$ are two convex functions (not necessarily differentiable), then the average function $h(x) \triangleq \left(f(x) + g(x)\right)/2$ must always be convex as well	T
4	Let $f: \mathbb{R} \to \mathbb{R}$ be a doubly differentiable function (i.e. first and second derivatives exist). If $f'(x^0) = 0$ at $x^0 \in \mathbb{R}$, then it must always be true that $f''(x^0) = 0$	F
5	The boundary of the unit 2D circle i.e. $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ is a convex set	F

Q2. Fill the circle (**don't tick**) next to all the correct options (**many may be correct).**(**4x3=12 marks**)

2.1 Suppose $f, g: \mathbb{R} \to \mathbb{R}$ are two convex functions (not necessarily differentiable) and we define $p(x) = f(x) + g(x)$ and $q(x) = f(x) - g(x)$. Two claims are made about these functions

Claim 1: $p(x) + q(x)$ must always be convex

Claim 2: $q(x)$ must always be convex

A	Claim 1 is TRUE, claim 2 is FALSE	●
B	Claim 1 is FALSE, claim 2 is TRUE	○
C	Both claims are TRUE	○
D	Both claims are FALSE	○

2.2 Let $f(x) = \sin(x)$. Which of the following statements is true about the function $f(x)$?

A	$f(x)$ has more than one local minima	●
B	$f''''(x) = f(x)$	●
C	$f(x)$ is a convex function	○
D	$f(x)$ is a concave function	○

2.3 Which of the following statements is true about the kNN algorithm?

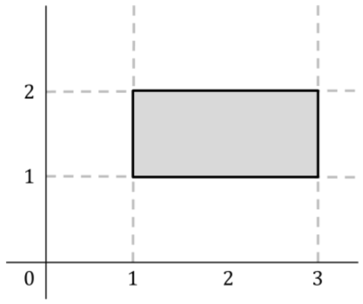
A	When used for binary classification, the kNN algorithm always produces decision boundaries that are linear (i.e. a line or a hyperplane)	○
B	The kNN algorithm can be used to solve regression problems	●
C	The value k in kNN must always be a positive integer	●
D	There exists no dataset, nor any value of k, for which kNN has linear decision boundary	○

2.4 Which of the following statements is true?

A	In held-out validation, the validation set is a subset of the test set	○
B	In held-out validation, the validation set is a subset of the training set	●
C	Using cross validation is more expensive than held-out validation	●
D	Using cross validation is less expensive than held-out validation	○

Q3 Consider the following optimization problem. **(2+2=4 marks)**

$$\max_{(x,y)\in\mathbb{R}^2} x^2 + y^2$$
$$x \in [1,3]$$
$$y \in [1,2]$$



At which point in $\mathbb{R}^2$ is the solution to this optimization problem achieved? Write your answer in space below.

$(3, 2)$

At which point in $\mathbb{R}^2$ would the solution have been achieved, had the objective function instead been $(-x^2 - y^2)$ (but constraints remained the same)?

$(1, 1)$

Q4. Let $\mathbf{a} \in \mathbb{R}^d$ be a constant vector and $b \in \mathbb{R}$ be a constant scalar. For $\mathbf{x} \in \mathbb{R}^d$, let us define the function $f(\mathbf{x}) = \ln(1 + \exp(-b \cdot \mathbf{a}^\top \mathbf{x}))$ (where $\ln$ is the natural logarithm). Find $\nabla f(\mathbf{x})$ and briefly show all major steps in your derivation. Write only in the space provided. **(4 marks)**

We have $f(t) = \ln(t)$ where $t(s) = 1 + \exp(-s)$ where $s(\mathbf{x}) = b \cdot \mathbf{a}^\top \mathbf{x}$. We have $f'(t) = \frac{1}{t}$, $t'(s) = -\exp(-s)$, and $\nabla s(\mathbf{x}) = b \cdot \mathbf{a}$. Thus, applying the chain rule gives us

$$\nabla f(\mathbf{x}) = \frac{-b \cdot \exp(-b \cdot \mathbf{a}^\top \mathbf{x})}{1 + \exp(-b \cdot \mathbf{a}^\top \mathbf{x})} \cdot \mathbf{a}$$

------------------------------------ END OF QUIZ ------------------------------------