ECE 492-45 Introduction to Machine Learning
2019 Fall Exam 2
Instructor: Dr. Chau-Wai Wong

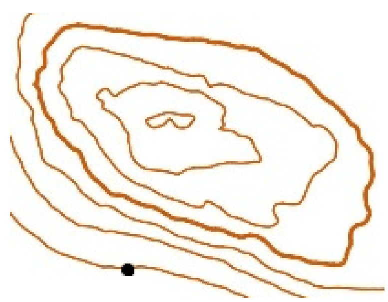
This is a closed-book exam. You may use a scientific calculator with cleared memory, but not a smart phone or computer. You should answer ***all four*** problems.

Problem 1 (25 pts)

(a) State one advantage of cross-validation over the validation-set approach.

(b) Given a random sample $\{1, 2, -1, 0\}$, generate a bootstrap samples of the same size. What is the chance that a bootstrap sample $\{0, 0, 0, 0\}$ is obtained? Justify your answer.

(c) A set of level curves are shown as follows. Use the dot as the starting point, draw a trajectory of gradient descent steps. Annotate each descent step using a line segment with an arrow at the end. Use a tangent line to assist you to determine the negative gradient direction if needed. You may vary the descent step size.



(d) A convolution neural network takes as input RGB pictures of size $512 \times 512 \times 3$. After passing through one convolutional layer, the output data is of size $512 \times 512 \times 10$. Explain what has happened to the third dimension.

Problem 2 (25 pts) Data from class i with prior probability π_i follow Gaussian distribution with density function:

$$f_i(x) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(x-\mu_i)^2}{2\sigma_i^2}}, \quad i = 1, 2. \tag{1}$$

Show that the LDA decision threshold between class 1 and class 2 is one of the roots of the following quadratic equation:

$$\left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}\right) x^2 - 2\left(\frac{\mu_1}{\sigma_1^2} - \frac{\mu_2}{\sigma_2^2}\right) x + \left(\frac{\mu_1^2}{\sigma_1^2} - \frac{\mu_2^2}{\sigma_2^2}\right) + 2\ln\left(\frac{\pi_2\sigma_1}{\pi_1\sigma_2}\right) = 0. \tag{2}$$

Problem 3 (25 pts) Two classes' PDFs are unit-variance Gaussian with mean 0 and 3, respectively.

You are given values of cumulative distribution function (CDF) for standard Gaussian evaluated at a few locations, namely, $\Phi(0) = 0.5$, $\Phi(-1) = 0.16$, $\Phi(-2) = 0.02$, $\Phi(-3) \approx 0$. These Φ values will be useful when calculating error rates for part (b). Note that $\Phi(x) = F[X \leq x]$, where $X \sim \mathcal{N}(0, 1)$.

(a) Draw the PDFs of two classes in the same plot. The x-axis must range from -3 to 6. Clearly label the mean for each class. Illustrate graphically $\Phi(0) = 0.5$ and $\Phi(-1) = 0.16$.

(b) Illustrate graphically the false positive rate (FPR) and the false negative rate (FNR) on the plot drawn for (a) using a specific threshold η that can reasonably convey the definitions. For example, do not use an η that is too small, too larger, or sits at the middle of two distributions. Calculate the numerical values of (FPR, FNR) pairs for decision threshold $\eta = -3, -2, -1, 0, 1, 2, 3, 4, 5, 6$.

(c) Draw ROC curve using FPR as the horizontal axis and FNR as the vertical axis. Limit the range of both axes to be $[0, 1]$.

Problem 4 (25 pts) Response $Y_i \sim \mathrm{B}(n, p_i)$ is a binomial random variable in which n is known. The (conditional) PDF is shown as follows:

$$\mathbb{P}[Y_i = k | \underset{\sim}{X}_i = \underset{\sim}{x}_i] = \binom{n}{k} p_i^k (1 - p_i)^{n-k}, \quad k \in \{0, 1, \ldots, n\}. \tag{3}$$

(a) Explain why the linear regression may not the best fit to find the relation between Y_i and a set of predictors $X_{i,1}, \ldots, X_{i,q}$.

(b) One proposes to link the conditional mean μ_i and the predictors $\underset{\sim}{x}_i$ using a generalized linear model shown as follows:

$$g(\mu_i) = \underset{\sim}{\beta}^T \underset{\sim}{x}_i \tag{4}$$

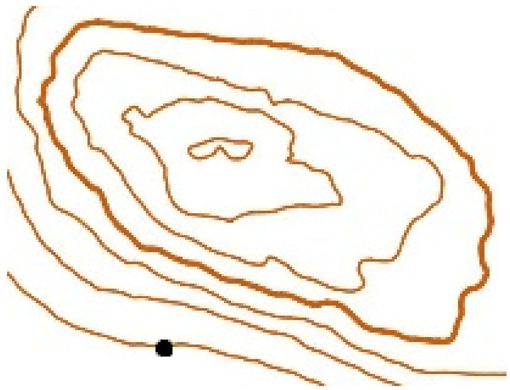
where $g(u) = \log(\frac{u}{n-u})$ and $\mu_i = \mathbb{E}[Y_i | \underset{\sim}{X}_i = \underset{\sim}{x}_i] = np_i$. From the variable transformation viewpoint, show that $g(\cdot)$ matches the ranges for the two sides of Eq. (4).

(c) Rewrite the PDF into an exponential family form shown as follows:

$$f_Y(y; \theta) = \exp\left(\frac{y\theta - b(\theta)}{a(\phi)} + c(y, \phi)\right), \tag{5}$$

where θ is the natural parameter. Show that $g(\cdot)$ in (b) is the canonical link function when taking μ_i as the input.

Name:
Student ID:
Answer to Problem 1:



Answer to Problem 1 (cont'd):

Name:

Answer to Problem 2:

Answer to Problem 2 (cont'd):

Name:

Answer to Problem 3:

Answer to Problem 3 (cont'd):

Name:

Answer to Problem 4:

Answer to Problem 4 (cont'd):