HW 6: ECE 601 Machine Learning for Engineers

Important Notes:

(a) When a HW question asks for writing a code, you would need to include the entire code as well as the output of the program as well as any other analysis requested in the question.

(b) Don't panic about the length of the HW assignment. HW assignments are treated as opportunities for improving learning and understanding, so I might include some extra text to help you better understand the concepts or learn about a point that was not covered during the class. The actual work needed from you is indeed manageable.

(c) Combine your solutions in **one** zip file called `homework6_UMassUSERNAME.zip`.

1. In this problem, we would like to focus on maximum likelihood estimation (MLE). To do so, we first review the definition of the **likelihood** function. Let X_1, X_2, X_3, ..., X_n be a random sample from a distribution with a parameter θ (In general, θ might be a vector, $\theta = (\theta_1, \theta_2, \cdots, \theta_k)$.) Suppose that x_1, x_2, x_3, ..., x_n are the observed values of X_1, X_2, X_3, ..., X_n. If X_i's are discrete random variables, we define the *likelihood* function as the probability of the observed sample sample as a function of θ:

$$\begin{aligned} L(x_1, x_2, \cdots, x_n; \theta) &= P(X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n; \theta) \\ &= P_{X_1 X_2 \cdots X_n}(x_1, x_2, \cdots, x_n; \theta). \end{aligned}$$

If X_1, X_2, X_3, ..., X_n are jointly continuous, we use the joint PDF instead of the joint PMF. Thus, the likelihood is defined by

$$L(x_1, x_2, \cdots, x_n; \theta) = f_{X_1 X_2 \cdots X_n}(x_1, x_2, \cdots, x_n; \theta).$$

Now, let $X_1, \ldots, X_n$ be an i.i.d random sample from a $Poisson(\lambda)$ distribution.

(a) Find the likelihood function, $L(x_1, \ldots, x_n; \lambda)$, using

$$P_{X_i}(x_i; \lambda) = \frac{e^{-\lambda} \lambda^{x_i}}{x_i!}$$

as the PMF.

(b) Find the log likelihood function and use that to obtain the MLE for λ, $\hat{\lambda}_{ML}$. Hint: First find the log-likelihood function, and then take its derivative with respect to λ and set it to 0.

2. You have gained experience applying Convolutional Neural Networks (CNN) to the CIFAR-10 dataset in HW5. The objective of this exercise is to enhance the accuracy of your CNN model that classifies images from the CIFAR-10 dataset. Follow the provided template, entitled `HW6_CNN_CIFAR10.ipynb`. Instructions for each part are provided. Replace the placeholder `'write your code here'` in the template with your code.