Indian Statistical Institute
Mid-Semester Examination : 2025-2026

M. Tech. (CrS) 2nd year

Subject: Machine Learning for Security

Date: 11/09/2025 Maximum Marks: 40 Time: 2 Hours

Note: Notations used are as explained in the class.

1. We write the poisson distribution parameterized by λ as

$$p(y;\lambda) = \frac{e^{-\lambda}\lambda^y}{y!}.$$

Show that the poisson distribution is in the exponential family with parameter η, and clearly state what are $b(y)$, η and $a(\eta)$. Using the Generalized Linear Model (GLM) on poisson distribution find the hypothesis function (also known as canonical response function). [10]

2. Let a sequence of examples $(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \cdots, (x^{(m)}, y^{(m)})$ be given. Suppose that $||x^{(i)}||_2 \leq D$ for all i, and further that there exists a unit-length vector u ($||u||_2 = 1$) such that $y^{(i)}(u^T x^{(i)}) \geq \gamma$ for all examples in the sequence (i.e., $u^T x^{(i)} \geq \gamma$ if $y^{(i)} = 1$, and $u^T x^{(i)} \leq -\gamma$ if $y^{(i)} = -1$, so that u separates the data with a margin of at least γ). Then the total number of mistakes that the perceptron algorithm makes on this sequence is at most $(D/\gamma)^2$. [10]

3. Consider using a logistic regression model $h_\theta(x) = g(\theta^T x)$ where g is the sigmoid function, and let a training set $\{x^{(i)}, y^{(i)} : i = 1, \ldots, m\}$ be given as usual. The maximum likelihood estimate of parameters θ is given by

$$\theta_{\text{ML}} = \arg\max_\theta \prod_{i=1}^m p(y^{(i)}|x^{(i)};\theta).$$

If we wanted to regularize logistic regression, then we might put a Bayesian prior on the parameters. Suppose we choose the prior $\theta \sim \mathcal{N}(0, \tau^2 I)$ (here, $\tau > 0$, and I is the $n+1$ by $n+1$ identity matrix), and then found the MAP estimate of θ as:

$$\theta_{\text{MAP}} = \arg\max_\theta p(\theta) \prod_{i=1}^m p(y^{(i)}|x^{(i)}, \theta).$$

Prove that $||\theta_{\text{MAP}}||_2 \leq ||\theta_{\text{ML}}||_2$. [10]

4. Suppose there are m training examples $(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \cdots, (x^{(m)}, y^{(m)})$. Consider the cost function for linear regression:

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^m (h_\theta(x^{(i)}) - y^{(i)})^2$$

where $h_\theta(x) = \theta^T x$. Find the Hessian of this cost function and prove that J is convex. [10]