2025 USA-NA-AIO Round 1, Problem 3, Part 10

USAAIO Mar 2025

Part 10 (5 points, non-coding task)

So far, we have finished preprocessing our dataset. We use this dataset for the purpose of doing **binary classification**.

In all remaining parts in this problem, we will train a logistic regression model with our preprocessed data and test its performance.

Let all training samples be indexed as $0, 1, \cdots, N-1$. For the nth sample, denote by $\mathbf{x}^{(n)} \in \mathbb{R}^{d\times 1}$ a column vector of all features and $y^{(n)} \in \{0, 1\}$ its ground-truth label.

In the logistic regression, denote by $\beta \in \mathbb{R}^{d\times 1}$ the learnable parameters.

Thus, our predicted label is determined according to:

$$y_{predict}^{(n)} = \begin{cases} 1, & \text{with probability } \sigma\left(\mathbf{x}^{(n),\top}\beta\right) \\ 0, & \text{with probability } 1-\sigma\left(\mathbf{x}^{(n),\top}\beta\right) \end{cases}$$

where

$$\sigma(z) = \frac{1}{1+e^{-z}}$$

is the sigmoid function.

Do the following task in this part.

To train our model, we need to solve the following optimization problem:

Skip to main content

$$\min_{\beta} \; L\left(\beta\right).$$

Write down the loss function $L\left(\beta\right)$ in the following form (reasoning is not required):

$$L\left(\beta\right) = \sum_{n=0}^{N-1} \cdots .$$

USAAIO **Mar 2025**

Misplaced '#'

$$\boxed{L\left(\beta\right) = \sum_{n=0}^{N-1} - \left(y^{(n)} \log\left(\sigma\left(\mathbf{x}^{(n),\top}\beta\right)\right) + \left(1 - y^{(n)}\right) \log\left(1 - \sigma\left(\mathbf{x}^{(n),\top}\beta\right)\right)\right).}$$

""" END OF THIS PART """

Related topics

Topic	Replies	Activity
2025 USA-NA-AIO Round 1, Problem 3, Part 17	1	Mar 2025
2025 USA-NA-AIO Round 1, Problem 3, Part 12	1	Mar 2025

Skip to main content

Topic	Replies	Activity
2025 USA-NA-AIO Round 1, Problem 3, Part 13	3	17d
2025 USA-NA-AIO Round 1, Problem 3, Part 16	1	Mar 2025
2025 USA-NA-AIO Round 1, Problem 3, Part 18	1	Mar 2025