AI & ML Course
Quiz 2(Mar 11, 2024)

Time: 30 minutes

Instructions

- Answer all questions
- See upload instructions in the form

Question:	1	2	Total
Points:	10	10	20
Score:			

1. Consider the problem of Binary Classification using the following Soft margin SVM on a Dataset of total 100 observations, equally divided between the positive and negative classes.

$$min_{\mathbf{w},b,\xi} \frac{1}{2}\|\mathbf{w}\|^2 + 2\sum_{i=1}^{N} \xi_i$$

$$y_i \left(\mathbf{w}^\top \mathbf{x}^{(i)} + b\right) \geq 1 - \xi_i, \quad \xi_i \geq 0, i \in \{1, \ldots, 100\}$$

Following information is available. At optimality,

- $$\mathbf{w}^* = -2\sum_{i=1}^{3} \mathbf{x}^{(i)} + r\sum_{i=11}^{16} \mathbf{x}^{(i)}$$

  The value of r is not known.
- The training set error(expressed as percentage) of the optimal classifier is 2%.
- It is given that $\xi = \beta\mathbf{e}, \quad \beta\sum_{i=1}^{N} e_i = 10$ where $\mathbf{e} \in \{0,1\}^N$ and β is an unknown scalar

Based on the information provided, answer the following questions

(a) (2 points) What is the value of r _____

(b) (2 points) What percentage of positive class examples are mis-classified _____

(c) (2 points) What percentage of negative class examples are mis-classified _____

(d) (2 points) What is the cardinality of the set $M = \{i|\xi_i > 1, i \in [100]\}$ _____

(e) (2 points) What is the value of β _____

2. A dataset $\mathcal{D} = \{(y_i, x_i)|i \in [n]\}$ was generated as follows

$$y_i = 2x_i + \epsilon_i, \quad \epsilon_i \sim N(0,1)$$

where $x_i \in \mathbb{R}$. Consider implementing Ridge regression

$$w_{RR} = \text{argmin}_{\ w} w^2 + \frac{1}{2}\sum_{i=1}^{n}(y_i - wx_i)^2$$

on $\mathcal{D}$. Following information is available $\frac{1}{n}\sum_{i=1}^{n} x_i y_i = 0.1, \quad \frac{1}{n}\sum_{i=1}^{n} x_i^2 = 2, \quad n = 100$ Answer the following.

(a) (2 points) What is the value of w_{RR}? ◯ $\frac{50}{100}$ ◯ $\frac{0.1}{4}$ ◯ $\frac{5}{101}$ ◯ 2

(b) (4 points) What is $E(w_{RR})$ ◯ $\frac{200}{101}$ ◯ $\frac{100}{101}$ ◯ $\frac{1}{101}$ ◯ 2

(c) (4 points) What is the squared bias of the prediction of the ridge regression model on a test point x. ◯ $\frac{4}{202^2}x^2$ ◯ $\frac{4}{101^2}x^2$ ◯ $\frac{8}{202^2}x^2$ ◯ x^2