IIT Kanpur Certificate Program on PYTHON for Artificial Intelligence, Machine Learning and Deep Learning

1st to 27th December 2024

Assignment #1

1. Siri, Alexa are examples of
    a. Pattern Recognition
    b. Data Mining
    c. Artificial Intelligence
    d. Linear Regression
2. In test-train split, the fraction of data set aside for testing is typically
    a. $50-60\%$
    b. $80-90\%$
    c. 100%
    d. $10-20\%$
3. Consider the matrix

$$\mathbf{A}=\begin{bmatrix}1 & 4\\ 3 & 2\end{bmatrix}$$

    Its inverse is
    a. $-\frac{1}{10}\begin{bmatrix}2 & -4\\ -3 & 1\end{bmatrix}$
    b. $\begin{bmatrix}2 & -4\\ -3 & 1\end{bmatrix}$
    c. $\frac{1}{10}\begin{bmatrix}1 & -4\\ -3 & 2\end{bmatrix}$
    d. $-\frac{1}{10}\begin{bmatrix}2 & 4\\ 3 & 1\end{bmatrix}$
4. For the linear regression problem $\min\left\|\bar{\mathbf{y}}-\mathbf{X}\,\bar{\mathbf{h}}\right\|^2$, the regression coefficients are given as
    a. $\mathbf{X}^T(\mathbf{X}^T\mathbf{X})^{-1}\bar{\mathbf{y}}$
    b. $(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\bar{\mathbf{y}}$
    c. $\mathbf{X}^T(\mathbf{X}\mathbf{X}^T)^{-1}\bar{\mathbf{y}}$
    d. $(\mathbf{X}\mathbf{X}^T)^{-1}\mathbf{X}^T\bar{\mathbf{y}}$
5. Consider the least squares (LS) problem below

$$\min\left\|\begin{bmatrix}1\\ -2\\ -1\\ 3\end{bmatrix}-\begin{bmatrix}1 & 3\\ 1 & 2\\ 1 & 1\\ 1 & 4\end{bmatrix}\begin{bmatrix}h_0\\ h_1\end{bmatrix}\right\|^2$$

    The regression coefficient vector is given as
    a. $\frac{1}{20}\begin{bmatrix}-70\\ 30\end{bmatrix}$

b. $\begin{bmatrix} -\frac{1}{2} \\ \frac{1}{10} \end{bmatrix}$
c. $\frac{1}{20}\begin{bmatrix} -54 \\ 18 \end{bmatrix}$
d. $\frac{1}{20}\begin{bmatrix} 36 \\ -12 \end{bmatrix}$

6. Logistic regression is well suited for
   a. Linear Approximation
   b. Gaussian clustering
   c. Binary classification
   d. Dimensionality reduction
7. The logistic function is defined as
   a. $\frac{1}{1+e^{z}}$
   b. $\frac{e^{-2z}}{1+e^{-z}}$
   c. $\frac{1}{1+e^{-z}}$
   d. $\frac{1}{1-e^{-z}}$
8. The update rule for logistic regression is given as
   a. $\bar{\mathbf{h}}(k+1) = \bar{\mathbf{h}}(k) + \eta\left(y(k+1) - g\left(\bar{\mathbf{x}}(k+1)\right)\right)\bar{\mathbf{x}}(k+1)$
   b. $\bar{\mathbf{h}}(k+1) = \bar{\mathbf{h}}(k) + \eta\left(g\left(\bar{\mathbf{x}}(k+1)\right) - y(k+1)\right)\bar{\mathbf{x}}(k+1)$
   c. $\bar{\mathbf{h}}(k+1) = \bar{\mathbf{h}}(k) + \eta\left(g\left(\bar{\mathbf{x}}(k+1)\right) - y(k+1)\right)$
   d. $\bar{\mathbf{h}}(k+1) = \bar{\mathbf{h}}(k) + \eta\left(y(k+1) - g\left(\bar{\mathbf{x}}(k+1)\right)\bar{\mathbf{x}}(k+1)\right)$
9. PDF of a Gaussian random vector is given as
   a. $\frac{1}{\sqrt{(2\pi)^n|\mathbf{R}|}} e^{-\frac{1}{2}(\bar{\mathbf{x}}-\bar{\boldsymbol{\mu}})^T\mathbf{R}^{-1}(\bar{\mathbf{x}}-\bar{\boldsymbol{\mu}})}$
   b. $\frac{1}{\sqrt{(2\pi)^n|\mathbf{R}|}} e^{-\frac{1}{2}(\bar{\mathbf{x}}-\bar{\boldsymbol{\mu}})^T\mathbf{R}(\bar{\mathbf{x}}-\bar{\boldsymbol{\mu}})}$
   c. $\frac{1}{\sqrt{(2\pi)^n}} e^{-\frac{1}{2}(\bar{\mathbf{x}}-\bar{\boldsymbol{\mu}})^T\mathbf{R}(\bar{\mathbf{x}}-\bar{\boldsymbol{\mu}})}$
   d. $\frac{1}{|\mathbf{R}|\sqrt{(2\pi)^n}} e^{-\frac{1}{2}(\bar{\mathbf{x}}-\bar{\boldsymbol{\mu}})^T\mathbf{R}^{-1}(\bar{\mathbf{x}}-\bar{\boldsymbol{\mu}})}$
10. Consider the **Gaussian classification** problem with $p_0 = p_1 = \frac{1}{2}$. The classes $\mathcal{C}_0, \mathcal{C}_1$ are distributed as

$$\mathcal{C}_0 \sim N\left(\begin{bmatrix} 1 \\ -2 \end{bmatrix}, \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}\right), \mathcal{C}_1 \sim N\left(\begin{bmatrix} -1 \\ 2 \end{bmatrix}, \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}\right)$$

    The discriminant function to choose $\mathcal{C}_0$ is given as
    a. $8x_1 - 6x_2 < 0$
    b. $6x_1 - 8x_2 \geq 0$
    c. $2x_1 - 3x_2 \geq 3$
    d. $3x_1 - 2x_2 < 3$
11. The **SVC** problem is given as

a. $\min\|\bar{\mathbf{a}}\|$
$\bar{\mathbf{a}}^T\bar{\mathbf{x}}_i + b \geq 1, \bar{\mathbf{x}}_i \in \mathcal{C}_0$
$\bar{\mathbf{a}}^T\bar{\mathbf{x}}_i + b \leq -1, \bar{\mathbf{x}}_i \in \mathcal{C}_1$

b. $\max\|\bar{\mathbf{a}}\|$
$\bar{\mathbf{a}}^T\bar{\mathbf{x}}_i + b \geq 1, \bar{\mathbf{x}}_i \in \mathcal{C}_0$
$\bar{\mathbf{a}}^T\bar{\mathbf{x}}_i + b \leq -1, \bar{\mathbf{x}}_i \in \mathcal{C}_1$

c. $\min\|\bar{\mathbf{a}}\|$
$\bar{\mathbf{a}}^T\bar{\mathbf{x}}_i + b \geq -1, \bar{\mathbf{x}}_i \in \mathcal{C}_0$
$\bar{\mathbf{a}}^T\bar{\mathbf{x}}_i + b \leq 1, \bar{\mathbf{x}}_i \in \mathcal{C}_1$

d. $\min\|\bar{\mathbf{a}}\|$
$\bar{\mathbf{a}}^T\bar{\mathbf{x}}_i + b = -1, \bar{\mathbf{x}}_i \in \mathcal{C}_0$
$\bar{\mathbf{a}}^T\bar{\mathbf{x}}_i + b = 1, \bar{\mathbf{x}}_i \in \mathcal{C}_1$

12. Using the Naïve Bayes assumption, $p(\bar{\mathbf{x}}|y) = p(x_1, x_2, \ldots, x_N|y)$ can be expressed as

a. $p(x_1|y) + p(x_2|y) + \cdots + p(x_N|y)$
b. $p(x_1) \times p(x_2) \times \ldots \times p(x_N)$
c. $p(x_1|y) \times p(x_2|y) \times \ldots \times p(x_N|y)$
d. $p(y|x_1) \times p(y|x_2) \times \ldots \times p(y|x_N)$

13. The posterior probability $p(y = 1|\bar{\mathbf{x}})$ is given as

a. $\frac{p(\bar{\mathbf{x}}|y=1)\times p(\bar{\mathbf{x}})}{p(y=1)}$
b. $\frac{p(\bar{\mathbf{x}}|y=1)\times p(y=1)}{p(\bar{\mathbf{x}})}$
c. $\frac{p(\bar{\mathbf{x}})\times p(y=1)}{p(\bar{\mathbf{x}}|y=1)}$
d. $\frac{p(\bar{\mathbf{x}}|y=1)}{p(\bar{\mathbf{x}})\times p(y=1)}$

14. Consider the data table given below

Y \ X_1	counts 0	1
0	3	10
1	4	13

The quantity $P(Y = 0|X_1 = 0)$ can be evaluated as

a. $\frac{3}{13}$
b. $\frac{7}{23}$
c. $\frac{13}{13}$
d. $\frac{3}{7}$

15. The K −means cost-function to minimize is given as

a. $\min\sum_{i=1}^{K}\sum_{j=1}^{M}\alpha_i(j)\|\bar{\mathbf{x}}(j) - \bar{\boldsymbol{\mu}}_i\|$
b. $\min\sum_{i=1}^{K}\sum_{j=1}^{M}\alpha_i(j)\|\bar{\mathbf{x}}(j) - \bar{\boldsymbol{\mu}}_i\|^2$
c. $\min\sum_{i=1}^{K}\sum_{j=1}^{M}\|\bar{\mathbf{x}}(j) - \bar{\boldsymbol{\mu}}_i\|$
d. $\min\sum_{i=1}^{K}\sum_{j=1}^{M}\|\bar{\mathbf{x}}(j) - \bar{\boldsymbol{\mu}}_i\|^2$