INDIAN INSTITUTE OF TECHNOLOGY KHARAGPUR	
	Stamp / Signature of the Invigilator

EXAMINATION (End Semester)	SEMESTER (Spring 2025-2026)

Roll Number										Section		Name	
Subject Number	C	S	6	0	0	2	0	Subject Name	*FOUNDATIONS OF ALGORITHM DESIGN AND MACHINE LEARNING*				
Department / Center of the Student		Additional sheets											

Important Instructions and Guidelines for Students

1. You must occupy your seat as per the Examination Schedule/Sitting Plan.
2. Do not keep mobile phones or any similar electronic gadgets with you even in the switched off mode.
3. Loose papers, class notes, books or any such materials must not be in your possession, even if they are irrelevant to the subject you are taking examination.
4. Data book, codes, graph papers, relevant standard tables/charts or any other materials are allowed only when instructed by the paper-setter.
5. Use of instrument box, pencil box and non-programmable calculator is allowed during the examination. However, exchange of these items or any other papers (including question papers) is not permitted.
6. Write on both sides of the answer script and do not tear off any page. **Use last page(s) of the answer script for rough work.** Report to the invigilator if the answer script has torn or distorted page(s).
7. It is your responsibility to ensure that you have signed the Attendance Sheet. Keep your Admit Card/Identity Card on the desk for checking by the invigilator.
8. You may leave the examination hall for wash room or for drinking water for a very short period. Record your absence from the Examination Hall in the register provided. Smoking and the consumption of any kind of beverages are strictly prohibited inside the Examination Hall.
9. Do not leave the Examination Hall without submitting your answer script to the invigilator. **In any case, you are not allowed to take away the answer script with you.** After the completion of the examination, do not leave the seat until the invigilators collect all the answer scripts.
10. During the examination, either inside or outside the Examination Hall, gathering information from any kind of sources or exchanging information with others or any such attempt will be treated as **'unfair means'**. Do not adopt unfair means and do not indulge in unseemly behavior.

Violation of any of the above instructions may lead to severe punishment.

Signature of the Student

To be filled in by the examiner

Question Number	1	2	3	4	5	6	7	8	9	10	Total
Marks Obtained											

Marks obtained (in words)	Signature of the Examiner	Signature of the Scrutineer

Indian Institute of Technology Kharagpur
Department of Computer Science and Engineering

Foundations of Algorithm Design and Machine Learning (CS60020) **Spring 2025-2026**

24-April-2026 **End-Semester Examination** Maximum Marks: 60

Instructions:

- Write your answers in the question paper itself. Be brief and precise. Answer _all_ questions.
- Write the answers only in the respective spaces provided. The last two blank pages may be used for rough work or leftover answers.
- In case you may need more space/pages, please ask for additional sheets in the exam hall and attach the same with this booklet while submitting.
- If you use any algorithm / result / formula covered in the class, just mention it, do not elaborate (unless the same thing has been explicitly asked to answer in the question).

— Question Paper Starts from Next Page —

Q1. [Graph Algorithms: Spanning Tree] **10 marks**

Let $G = (V, E)$ is a weighted undirected graph having all edges of cost 2 except exactly two edges of cost 1. Answer the following questions.

(a) Present an efficient $O(V + E)$ time (or linear) complexity algorithm to find the Minimum Cost Spanning Tree of G. Explain all the steps clearly with comments. **(6)**

Solution:

To be given.

(b) Show the working of your proposed algorithm on a connected weighted undirected graph having at least 10 nodes. **(2)**

Solution:

To be given.

(c) Analyze the time complexity of your proposed algorithm to show why it is $O(V+E)$. **(2)**

Solution:

To be given.

Q2. **[Hardness of Problems]** **10 marks**

Answer the following questions.

(a) When is a problem said to be NP-Complete? Define formally and precisely. **(2)**

Solution:

To be given.

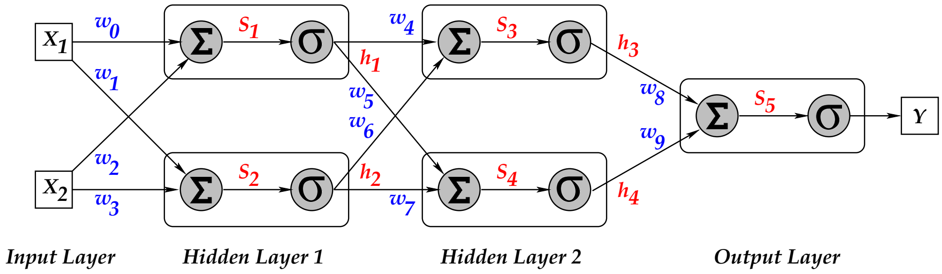
(b) Consider the three-partitioning decision problem of checking whether or not a set of integers S can be split into three parts, one exactly $\frac{1}{3}$, another exactly $\frac{1}{6}$, and a third part exactly $\frac{1}{2}$ of the total sum of elements in S. Using the known NP-Complete problem of Subset Sum, show that this three-partitioning problem is NP-Complete. **(8)**

Solution:

To be given.

Q3. **[Neural Networks]** **10 marks**

Consider the following neural network having two hidden and one output layer. The neurons are indicated using rectangles consisting of a summation unit (indicated using $\sum$) followed by an activation (sigmoid) unit (indicated using σ). Each neuron is provided with only two inputs going into the summation unit and finally it produces one output from the sigmoid unit. The functionalities of each neuron having an input vector $\mathbf{x}$ and the corresponding weight vector $\mathbf{w}$ can therefore be expressed as, $h(\mathbf{x}) = \sigma(s) = \frac{1}{1+e^{-s}} = \frac{1}{1+e^{-\mathbf{w}^T\mathbf{x}}}$. Let the initial weights are given as, $w_k = \begin{cases} +1, & \text{if } k \text{ is even} \\ -1, & \text{if } k \text{ is odd} \end{cases}$.



Further, consider the error to be a squared loss function, that is, $e(\mathbf{w}) = (\hat{Y} - Y)^2$, where $\hat{Y}$ is the predicted output and Y is the given/actual output. Suppose you want to train this neural network over only one training example having $x_1 = x_2 = 1$ and output as $Y = 0.5$. Answer the following.

(a) Perform a forward pass for the input $x_1 = x_2 = 1$ to determine the values of all s_j and h_j (where $1 \le j \le 5$) and the predicted value of the output, i.e. $\hat{Y} = h_5$. Show your calculations. **(5)**

Solution:

$$s_1 = \mathbf{w}^T\mathbf{x} = w_0 x_1 + w_2 x_2 = 2$$

$$h_1 = \sigma(s_1) = \frac{1}{1+e^{-s_1}} \approx 0.88080$$

$$s_2 = \mathbf{w}^T\mathbf{x} = w_1 x_1 + w_3 x_2 = -2$$

$$h_2 = \sigma(s_2) = \frac{1}{1+e^{-s_2}} \approx 0.11920$$

$$s_3 = \mathbf{w}^T\mathbf{x} = w_4 h_1 + w_6 h_2 = 1$$

$$h_3 = \sigma(s_3) = \frac{1}{1+e^{-s_3}} \approx 0.73106$$

$$s_4 = \mathbf{w}^T\mathbf{x} = w_5 h_1 + w_7 h_2 = -1$$

$$h_4 = \sigma(s_4) = \frac{1}{1+e^{-s_4}} \approx 0.26894$$

$$s_5 = \mathbf{w}^T\mathbf{x} = w_8 h_3 + w_9 h_4 \approx 0.46212$$

$$\hat{Y} = h_5 = \sigma(s_5) = \frac{1}{1+e^{-s_5}} \approx 0.61352$$

(b) Now perform a backward pass (using the *backpropagation* approach) and determine the value of the gradient, $\frac{\partial e(\mathbf{W})}{\partial w_0}$. Clearly mention the complete chain rule that you used to compute the gradient and show your calculations. **(4)**

Solution:

It may be noted that,

$$\sigma'(s) = \frac{\partial}{\partial s}\left(\frac{1}{1+e^{-s}}\right) = \frac{e^{-s}}{(1+e^{-s})^2} = \left(\frac{1}{1+e^{-s}}\right)\cdot\left(1-\frac{1}{1+e^{-s}}\right) = \sigma(s)\cdot\Big(1-\sigma(s)\Big)$$

$$\begin{aligned}
\frac{\partial e(\mathbf{W})}{\partial w_0} &= \frac{\partial e(\mathbf{W})}{\partial \hat{Y}}\cdot\frac{\partial \hat{Y}}{\partial s_5}\cdot\left(\frac{\partial s_5}{\partial h_3}\cdot\frac{\partial h_3}{\partial s_3}\cdot\frac{\partial s_3}{\partial h_1} + \frac{\partial s_5}{\partial h_4}\cdot\frac{\partial h_4}{\partial s_4}\cdot\frac{\partial s_4}{\partial h_1}\right)\cdot\frac{\partial h_1}{\partial s_1}\cdot\frac{\partial s_1}{\partial w_0} \\
&= 2(\hat{Y}-Y)\cdot\sigma'(s_5)\cdot\Big(w_8\cdot\sigma'(s_3)\cdot w_4 + w_9\cdot\sigma'(s_4)\cdot w_5\Big)\cdot\sigma'(s_1)\cdot x_1 \\
&\approx 2\times 0.11352\times 0.23711\times\Big(1\times 0.19661\times 1 + (-1)\times 0.19661\times(-1)\Big)\times 0.10499\times 1 \\
&\approx 0.0022235
\end{aligned}$$

(c) Assuming that the learning rate $\eta = 0.5$, what will be the updated weight for w_0 after a round of backpropagation? Clearly mention the weight update rule that you used and your calculations. **(1)**

Solution:

The updated weight for w_0 will be:

$$\begin{aligned}
w_0^{new} &= w_0^{old} \pm \eta\cdot\frac{\partial e(\mathbf{W})}{\partial w_0} \\
&\approx 1 \pm 0.5\times 0.0022235 \approx 1.00111175 \quad or \quad 0.99888825
\end{aligned}$$

Q4. **[Support Vector Machines]** **10 marks**

Consider the training set containing points from three-dimensional space,

$$\mathbf{x_1} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}, \quad \mathbf{x_2} = \begin{bmatrix} 0 \\ 1 \\ 8 \end{bmatrix}, \quad \mathbf{x_3} = \begin{bmatrix} 1 \\ 0 \\ 6 \end{bmatrix}, \quad \mathbf{x_4} = \begin{bmatrix} 1 \\ 1 \\ 7 \end{bmatrix}, \quad \mathbf{x_5} = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}$$

for which the true labels are given respectively as,

$$y_1 = +1, \quad y_2 = +1, \quad y_3 = -1, \quad y_4 = -1, \quad y_5 = +1.$$

After training an SVM classifier with linear kernel (that is, solving the dual optimization problem with these training examples), we find the following Lagrangian coefficients:

$$\alpha_1 = 0, \quad \alpha_2 = 1, \quad \alpha_3 = 0.5, \quad \alpha_4 = 1, \quad \alpha_5 = 0.5.$$

Answer the following questions.

(a) How many support vectors are required and why? **(1)**

Solution:

FOUR – all points (i.e. $\mathbf{x_2}, \mathbf{x_3}, \mathbf{x_4}, \mathbf{x_5}$) those have $\alpha_i > 0$.

(b) What will be the weight vector, $\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$, of the linear separator? Show your calculations. **(2)**

Solution:

The weight vector of the linear separator is found as,

$$\begin{aligned} \mathbf{w} &= \sum_{i=2}^{5} y_i \alpha_i \mathbf{x_i} \\ &= (+1)\cdot 1 \cdot \begin{bmatrix} 0 \\ 1 \\ 8 \end{bmatrix} + (-1)\cdot 0.5 \cdot \begin{bmatrix} 1 \\ 0 \\ 6 \end{bmatrix} + (-1)\cdot 1 \cdot \begin{bmatrix} 1 \\ 1 \\ 7 \end{bmatrix} + (+1)\cdot 0.5 \cdot \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix} \\ &= \begin{bmatrix} -1 \\ 0.5 \\ -0.5 \end{bmatrix} \end{aligned}$$

(c) What will be the bias term, $\mathbf{b}$, of the linear separator? Show your calculations. **(2)**

Solution:

The bias term of the linear separator can be obtained by considering *any one* support vector point (say, $(\mathbf{x_k}, y_k)$, $k \in [2,5]$) using the following equation: $\mathbf{w}^T \mathbf{x_k} + \mathbf{b} = y_k$.

$$k = 2, \text{ i.e., taking } (\mathbf{x_2}, y_2) \implies \mathbf{b} = (+1) - \begin{bmatrix} -1 & 0.5 & -0.5 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ 8 \end{bmatrix} = 4.5$$

$$k = 3, \text{ i.e., taking } (\mathbf{x_3}, y_3) \implies \mathbf{b} = (-1) - \begin{bmatrix} -1 & 0.5 & -0.5 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 6 \end{bmatrix} = 3$$

$$k = 4, \text{ i.e., taking } (\mathbf{x_4}, y_4) \implies \mathbf{b} = (-1) - \begin{bmatrix} -1 & 0.5 & -0.5 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 7 \end{bmatrix} = 3$$

$$k = 5, \text{ i.e., taking } (\mathbf{x_5}, y_5) \implies \mathbf{b} = (+1) - \begin{bmatrix} -1 & 0.5 & -0.5 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix} = 3$$

(d) Express the equation of the linear separator. Show your calculations. **(2)**

Solution:

The equation the linear separator is given as,

$$\mathbf{w}^T\mathbf{x}+\mathbf{b}=0 \implies \begin{bmatrix}-1 & 0.5 & -0.5\end{bmatrix}\cdot\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}+3=0 \text{ (taking } b=3)$$
$$\implies 2x_1-x_2+x_3-6=0$$

Alternatively:

$$\mathbf{w}^T\mathbf{x}+\mathbf{b}=0 \implies \begin{bmatrix}-1 & 0.5 & -0.5\end{bmatrix}\cdot\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}+4.5=0 \text{ (taking } b=4.5)$$
$$\implies 2x_1-x_2+x_3-9=0$$

(d) Derive the value of the margin (that is, the total width of the separation region). **(2)**

Solution:

The margin is twice the distance between any support vector and the boundary line, which can be computed as follows:

$$margin = \frac{2}{||\,\mathbf{w}\,||} = \frac{2}{\sqrt{(-1)^2+(0.5)^2+(-0.5)^2}} \approx 1.633$$

(e) How does the SVM classify a new point, $\mathbf{x}_{new}=\begin{bmatrix}1\\1\\8\end{bmatrix}$? **(1)**

Solution:

Since we find that,

$$\mathbf{w}^T\mathbf{x}_{new}+\mathbf{b} = \begin{bmatrix}-1 & 0.5 & -0.5\end{bmatrix}\cdot\begin{bmatrix}1\\1\\8\end{bmatrix}+3 = 0.5 \quad >0,$$

the point $\mathbf{x}_{new}$ will be classified as, $\hat{y}_{new}=+1$.

Alternatively:

$$\mathbf{w}^T\mathbf{x}_{new}+\mathbf{b} = \begin{bmatrix}-1 & 0.5 & -0.5\end{bmatrix}\cdot\begin{bmatrix}1\\1\\8\end{bmatrix}+4.5 = 0,$$

the point $\mathbf{x}_{new}$ will be classified as, $\hat{y}_{new}=\pm1$.

Q5. **[Learning Theory]** **10 marks**

For this problem, assume that all examples are points in two-dimensional space, i.e. $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$ ($\mathbb{R}$ denotes the set of reals and $\mathbb{R}^2 \equiv \mathbb{R} \times \mathbb{R}$).

(a) Given two examples, $\mathbf{x}$, $\mathbf{z} \in \mathbb{R}^2$, let us define $K(\mathbf{x},\mathbf{z}) = \alpha(\mathbf{x}^T\mathbf{z})^3 + \beta(\mathbf{x}^T\mathbf{z})^2$, where α and β are constants. Prove that $K(\mathbf{x},\mathbf{z})$ is a valid kernel function. **(5)**

Solution:

There are three parts to this proof as ilustrated below.

- *Proving $K_1(\mathbf{x},\mathbf{z}) = (\mathbf{x}^T\mathbf{z})^3$ is a valid kernel*:

$$\begin{aligned} (\mathbf{x}^T\mathbf{z})^3 &= (x_1z_1 + x_2z_2)^3 \\ &= x_1^3z_1^3 + 3x_1^2x_2z_1^2z_2 + 3x_1x_2^2z_1z_2^2 + x_2^3z_2^3 \\ &= \phi_1(\mathbf{x})^T\phi_1(\mathbf{z}), \qquad \text{where } \phi_1(\mathbf{x}) = \begin{bmatrix} x_1^3 \\ \sqrt{3}x_1^2x_2 \\ \sqrt{3}x_1x_2^2 \\ x_2^3 \end{bmatrix} \text{ and } \phi_1(\mathbf{z}) = \begin{bmatrix} z_1^3 \\ \sqrt{3}z_1^2z_2 \\ \sqrt{3}z_1z_2^2 \\ z_2^3 \end{bmatrix} \end{aligned}$$

- *Proving $K_2(\mathbf{x},\mathbf{z}) = (\mathbf{x}^T\mathbf{z})^2$ is a valid kernel*:

$$\begin{aligned} (\mathbf{x}^T\mathbf{z})^2 &= (x_1z_1 + x_2z_2)^2 \\ &= x_1^2z_1^2 + 2x_1x_2z_1z_2 + x_2^2z_2^2 \\ &= \phi_2(\mathbf{x})^T\phi_2(\mathbf{z}), \qquad \text{where } \phi_2(\mathbf{x}) = \begin{bmatrix} x_1^2 \\ \sqrt{2}x_1x_2 \\ x_2^2 \end{bmatrix} \text{ and } \phi_2(\mathbf{z}) = \begin{bmatrix} z_1^2 \\ \sqrt{2}z_1z_2 \\ z_2^2 \end{bmatrix} \end{aligned}$$

- *Proving the linearity of kernel functions*:

  We have proven in the above arguments that K_1 and K_2 all are valid kernel functions. Hence, there exists two functions ϕ_1 and ϕ_2 such that,

$$\begin{aligned} K_1(\mathbf{x},\mathbf{z}) &= \phi_1(\mathbf{x})^T\phi_1(\mathbf{z}) \\ K_2(\mathbf{x},\mathbf{z}) &= \phi_2(\mathbf{x})^T\phi_2(\mathbf{z}) \end{aligned}$$

  Define, $$\phi(\mathbf{x}) = \begin{bmatrix} \sqrt{\alpha}\phi_1(\mathbf{x}) \\ \sqrt{\beta}\phi_2(\mathbf{x}) \end{bmatrix} \quad \text{and} \quad \phi(\mathbf{z}) = \begin{bmatrix} \sqrt{\alpha}\phi_1(\mathbf{z}) \\ \sqrt{\beta}\phi_2(\mathbf{z}) \end{bmatrix}.$$

  It follows that,

$$\begin{aligned} K(\mathbf{x},\mathbf{z}) &= \phi(\mathbf{x})^T\phi(\mathbf{z}) \\ &= \alpha\phi_1(\mathbf{x})^T\phi_1(\mathbf{z}) + \beta\phi_2(\mathbf{x})^T\phi_2(\mathbf{z}) \\ &= \alpha K_1(\mathbf{x},\mathbf{z}) + \beta K_2(\mathbf{x},\mathbf{z}) \\ &= \alpha(\mathbf{x}^T\mathbf{z})^3 + \beta(\mathbf{x}^T\mathbf{z})^2 \end{aligned}$$

  Therefore, K is a valid kernel.

(b) The hypothesis space H is the region between two axis-parallel lines, either $(x = a,\ x = b)$ or $(y = a,\ y = b)$ for $a < b$. That is, each hypothesis $h \in H$ is defined by two numbers, $a, b \in \mathbb{R}$ and another Boolean indicator (X or Y, respectively) that determines whether the lines are parallel to the x-axis or the y-axis. An example (x, y) gets a *positive* label for the hypothesis (X, a, b) if and only if $a \leq x \leq b$. Similarly, an example (x, y) gets a *positive* label for the hypothesis (Y, a, b) if and only if $a \leq y \leq b$. Determine the VC-dimension of H. Justify your answer. **(5)**

Solution:

For hypothesis spaces with lines parallel to x or y-axis, VC-dimension is 4.

It is trivial to give a set of *four* points that there is some hypothesis in the hypothesis space that can correctly assign all possible label assignments to these points.

Given any set of *five* points, then there exist a minimal rectangle that contains all these five points. At least two of them are on the edges of the rectangle. Label the points that are interior to the rectangle as negative if there are any. If all these five points are on the edges of the rectangle, then there must exist two points that on the same edge, in which case label any one of these two points as negative. For this labeling, these five points cannot be correctly classified using a member of our hypothesis space, because the negative point must lie in the region between two parallel lines, otherwise contradicting that either the negative point is the interior point of the rectangle or the negative point and a positive point are on a same line. Therefore, we proved that for any set of five points, the points cannot be shattered using hypothesis in H.

Q6. **[Boosting: AdaBoost]** **10 marks**

Consider 10 labeled data $(x_1, x_2) \in \mathbb{R}^2$ (i is the example index) given the the table (right). In this problem, we study how AdaBoost algorithm performs on a very simple classification problem. We shall use two rounds of AdaBoost to learn a hypothesis for this data set. In round number t, AdaBoost chooses a weak learner that minimizes the weighted error ε_t. We shall use *decision stumps* as our weak learner / hypothesis which use axis-parallel lines of the form (i) Label $+$ if $x > a$, else $-$ or (ii) Label $+$ if $y > b$, else $-$, for some integers a, b (either one of the two forms, not a disjunction of the two). Answer the following.

i	x_1	x_2	*Label*
0	11	3	$-$
1	10	1	$-$
2	4	4	$-$
3	12	10	$+$
4	2	4	$-$
5	10	5	$+$
6	8	8	$-$
7	6	5	$+$
8	7	7	$+$
9	7	8	$+$

(a) What is the weight assigned to each data point according to the initial data weight distribution D_1? **(1)**

Solution:

The initial weight distribution is uniform, therefore all data points get the same weight as,

$$W_i^{(0)} = \frac{1}{10} \; (0 \leq i \leq 9).$$

(b) Which is the hypothesis (decision stump) h_1 that minimizes the weighted error in `Round-1` of AdaBoost, using the distribution D_1 computed in part (a)? **(1)**

Solution:

Analyzing the various decision stumps possible over the given data set, it becomes obvious that the best option is, $h_1 : x_2 > 4$.

(c) What is the weighted error (ε_1) and the weight (α_1) assigned to h_1 computed in part (b)? Show your calculations. **(1+1)**

Solution:

h_1 wrongly classifies only one point ($i = 6$), and hence the weighted error, $\varepsilon_1 = \frac{1}{10}$.

Weight assigned to h_1 is given by, $\alpha_1 = \ln\left(\sqrt{\frac{1-\varepsilon_1}{\varepsilon_1}}\right) = \ln 3 \approx 1.09861$.

(d) Next, we proceed to `Round-2` of AdaBoost. We begin by recomputing data weights depending on the error of h_1 and whether a point was (mis)classified by h_1. What will be weight assigned to each data point after boosting is performed, creating modified weight distribution D_2? Show your approach. *Do not forget to normalize the new data weights so that they sum to* 1. **(2)**

Solution:

Firstly, the normalization factor, $Z = 2\sqrt{\varepsilon_1(1-\varepsilon_1)} = \frac{3\sqrt{2}}{5} \approx 0.84853.$

After boosting, the updated weights will be:

- For correctly classified data-points (that is, $0 \leq i \leq 9$ and $i \neq 6$):

$$W_i^{(1)} = \frac{W_i^{(0)} e^{-\alpha_1}}{Z} = \frac{W_i^{(0)}}{2(1-\varepsilon_1)} = \frac{1}{18} \approx 0.05556$$

- For correctly classified data-points (that is, $i = 6$):

$$W_6^{(1)} = \frac{W_i^{(0)} e^{\alpha_1}}{Z} = \frac{W_i^{(0)}}{2\varepsilon_1} = \frac{1}{2} = 0.5$$

(e) Which is the hypothesis (decision stump) h_2 that minimizes the weighted error in `Round-2` of AdaBoost, using the distribution D_2 computed in part (d)? **(1)**

Solution:

The best decision stump option is, $h_2 : x_2 > 9.$

(f) What is the weighted error (ε_2) and the weight (α_2) assigned to h_2 computed in part (e)? Show your calculations. **(1+1)**

Solution:

Since h_2 wrongly classifies four points ($i = \{5,7,8,9\}$), the weighted error $\varepsilon_2 = \frac{4}{18} \approx 0.22222.$

Weight assigned to h_2 is given by, $\alpha_2 = \ln\left(\sqrt{\frac{1-\varepsilon_2}{\varepsilon_2}}\right) = \ln\left(\sqrt{3.5}\right) \approx 0.62638.$

(g) Now that we have completed two rounds of AdaBoost, it is time to create the final output hypothesis. What is the final weighted hypothesis after two rounds of AdaBoost? **(1)**

Solution:

$$h_{final} = sign\left[\alpha_1 h_1 + \alpha_2 h_2\right] = sign\left[1.09861\ h_1 + 0.62638\ h_2\right]$$

— Question Paper Ends Here —