

El examen contiene 2 páginas (incluyendo esta) y 6 preguntas. El total de puntaje es 20.

Tabla de puntaje (uso del profesor)

Question	Points	Score
1	4	
2	4	
3	3	
4	2	
5	4	
6	3	
Total:	20	

1. (4 points) Dada la variable aleatoria discreta X que representa el comportamiento de los datos de entrenamiento para un problema de clasificación con función de probabilidad definida por

$$p(x) = \begin{cases} (\frac{1}{2})^x & \text{para } x = 1, 2, 3, \dots \\ 0 & \text{en otro caso.} \end{cases}$$

Hallar la esperanza $E(X)$ de la variable aleatoria en cuestión.

2. (4 points) Henry llega típicamente tarde a la escuela. Si Henry llega tarde, denotamos a esto como $H = \text{tarde}$, en otro caso, $H = \text{notarde}$. Cuando su madre pregunta si él llegó tarde o no a la escuela, él nunca admite que llegó tarde. La respuesta que Henry da R_H queda representada como

$$p(R_H = \text{notarde} | H = \text{notarde}) = 1, \quad p(R_H = \text{tarde} | H = \text{tarde}) = 0.$$

Dado que $R_H = \text{notarde}$, cuál es la probabilidad que Henry llegó tarde, es decir, $p(H = \text{tarde} | R_H = \text{notarde})$.

3. (3 points) Sabemos de la regla de entrenamiento Delta para entrenar un Perceptron que el error cuadrático medio es

$$E(\vec{w}) = \frac{1}{2} \sum_{d \in D} (t_d - o_d)^2,$$

Day	Outlook	Temperature	Humidity	Wind	PlayTennis
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rain	Mild	High	Weak	Yes
5	Rain	Cool	Normal	Weak	Yes
6	Rain	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Strong	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	Yes
10	Rain	Mild	Normal	Weak	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes
13	Overcast	Hot	Normal	Weak	Yes
14	Rain	Mild	High	Strong	No

Table 1: Datos de entrenamiento PlayTennis.

donde D es el conjunto de datos de entrenamiento y la forma de actualización del vector \vec{w} esta dada por

$$\vec{w} \leftarrow \vec{w} + \Delta \vec{w},$$

donde $\Delta \vec{w} = -\eta \nabla E(\vec{w})$. Deducir la expresión de actualización para cada componente del vector \vec{w} .

4. (2 points) Hallar la entropía de un mensaje M de longitud 1 caracter, considerando el conjunto de caracteres ASCII y suponiendo una equiprobabilidad en sus 256 caracteres utilizando la formula general de la entropía para n estados.
5. (4 points) Hallar el árbol de decisión resultante haciendo uso del algoritmo ID3 y utilizando los datos de entrenamiento de la Tabla 1. ¿Cuáles son las variables aleatorias resultantes en los nodos del árbol? ¿Por qué algunas son descartadas?
6. (3 points) Demostrar la ida del siguiente teorema:
Sea \mathbb{H}_0 cualquier subespacio de \mathbb{C}^E , el espacio de funciones complejas definidas en E , en el cual definimos un producto interno $\langle \cdot, \cdot \rangle_{\mathbb{H}_0}$, con la norma generada $\|\cdot\|_{\mathbb{H}_0}$. Si existe un espacio de Hilbert \mathbb{H} tal que
 - a) $\mathbb{H}_0 \subset \mathbb{H} \subset \mathbb{C}^E$ y la topología definida en \mathbb{H}_0 inducida por el producto interno $\langle \cdot, \cdot \rangle_{\mathbb{H}_0}$ coincide con la topología inducida sobre \mathbb{H}_0 por \mathbb{H} ,
 - b) \mathbb{H} tiene un kernel reproductivo K ,

entonces se cumple que

- c) las evaluaciones funcionales $(e_t)_{t \in E}$ son continuas en \mathbb{H}_0 ,
- d) cualquier sucesión de cauchy (f_n) en \mathbb{H}_0 puntualmente convergente a 0 converge también a 0 en el sentido de norma.