

## PEC4: Introducción al aprendizaje computacional

### Presentación

Cuarta PEC del curso de Inteligencia Artificial

### Competencias

En esta PEC se trabajarán las siguientes competencias:

#### Competencias de grado:

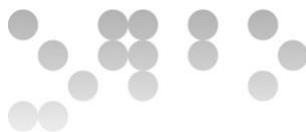
- Capacidad de analizar un problema con el nivel de abstracción adecuado a cada situación y aplicar las habilidades y conocimientos adquiridos para abordarlo y solucionarlo.

#### Competencias específicas:

- Conocer los diferentes aspectos del aprendizaje computacional (aprendizaje supervisado, no supervisado y por refuerzo).
- Conocer los fundamentos teóricos de los métodos más representativos.
- Conocer el tratamiento de los conjuntos de datos para la correcta validación de los sistemas de aprendizaje.

### Objetivos

Esta PEC pretende evaluar diferentes aspectos de aprendizaje supervisado y no supervisado.



## Descripción de la PEC a realizar

### Pregunta 1 (4 puntos)

Un servicio de transportes quiere saber si unas determinadas vías son seguras de circular este invierno en función de la cantidad de nieve caída, el tipo de vía y la elevación del terreno. Para ello, basados en unos datos históricos, quiere implementar un modelo para predecir si la circulación es posible (Y) o no (N) en función de las siguientes características:

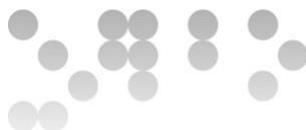
**Tipo de vía:** Autopista (A), Carretera convencional (C)

**Elevación:** Inferior a 300m (<300m) o superior a 300m (>300m)

**Nieve:** Baja (B), Media (M), Alta (A)

La tabla con las observaciones es la siguiente:

	Vía	Elevación	Nieve	Seguro
1	A	<300m	M	Y
2	A	<300m	L	Y
3	C	<300m	M	N
4	C	<300m	H	N
5	A	>300m	H	N
6	A	>300m	L	Y
7	C	>300m	H	N
8	C	>300m	L	Y
9	A	<300m	L	Y
10	A	<300m	M	Y
11	A	>300m	L	Y
12	C	<300m	M	N
13	C	<300m	H	N
14	A	>300m	L	Y
15	C	>300m	L	Y
16	C	>300m	M	N



En este ejercicio, debes construir un árbol de decisión con las tres características que tenemos en la tabla anterior, siendo **Seguro** el objeto de clasificación. Calculad las bondades correspondientes para construir el árbol.

Respuesta:

Realizamos el cálculo de bondad por cada variable descriptora (Nieve, elevación y vía) para hallar la característica de nodo raíz.

Bondad de Vía: 13/16

	Y	N
A	7 {1,2,6,9,10,11,14}	1 {5}
C	2 {8,15}	6 {3,4,7,12,13,16}

Bondad de Elevación: 9/16

	Y	N
<300	4 {1,2,9,10}	4 {3,4,12,13}
>300	5 {6,8,11,14,15}	3 {5,7,16}

Bondad de Nieve: 14/16

	Y	N
M	2 {1, 10}	3 {3, 12, 16}
L	7 {2,6,8,9,11,14,15}	0
H	0	4 {4,5,7,13}

Las bondades de cada atributo fueron calculadas para determinar qué variable separa mejor las clases objetivo (Y/N). En este caso, la variable 'Nieve' resultó ser la que mejor separaba los datos, dado que su bondad fue la más alta al evaluar las diferentes particiones, entonces será nuestro nodo raíz. Por consiguiente, evaluamos las tres ramas que obtenemos de esta partición (M, L y H) para determinar cual es el mejor nodo hijo.

**Si Nieve es M:**

Bondad de Vía en Nieve=m: 5/5

	Y	N
A	2{1,10}	0
C	0	3{3,12,16}

Bondad de Elevación en Nieve=m: 3/5

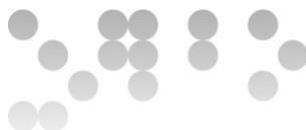
	Y	N
<300	2{1,10}	2{3,12}
>300	0	1{16}

Al tener mayor bondad, el nodo hijo será Vía.

**Si nieve es L:**

Bondad de Vía en Nieve=L: 7/7

	Y	N
A	5 {2,6,9,11,14}	0



C	2 {8,15}	0
---	----------	---

Bondad de Elevación en Nieve=L: 7/7

	Y	N
<300	2 {2,9}	0
>300	5 {6,8,11,14,15}	0

Dado que al realizar el cálculo de bondades y obtener el mismo, escogemos arbitrariamente. En este caso, Vía.

**Si nieve es H:**

Bondad de Vía en Nieve=H: 4/4

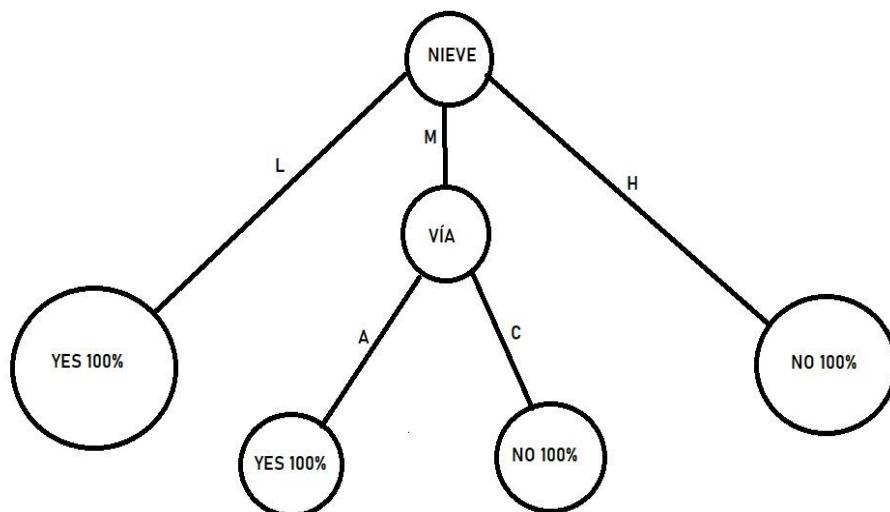
	Y	N
A	0	1{5}
C	0	3{4,7,13}

Bondad de Elevación en Nieve=H: 4/4

	Y	N
<300	0	2{4,13}
>300	0	2{5,7}

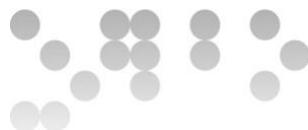
Dado que al realizar el cálculo de bondades y obtener el mismo, escogemos arbitrariamente. En este caso, Vía.

A raíz de todos estos resultados podemos construir el árbol.



El árbol de decisión utiliza 'Nieve' como nodo raíz, ya que es el atributo que mejor separa las clases. En la rama 'Nieve = M', se añade 'Vía' porque genera una división clara. No se utiliza 'Vía' en 'Nieve = L' y 'Nieve = H', ya que no hay una separación significativa en esos casos.

## Pregunta 2 (4 puntos)

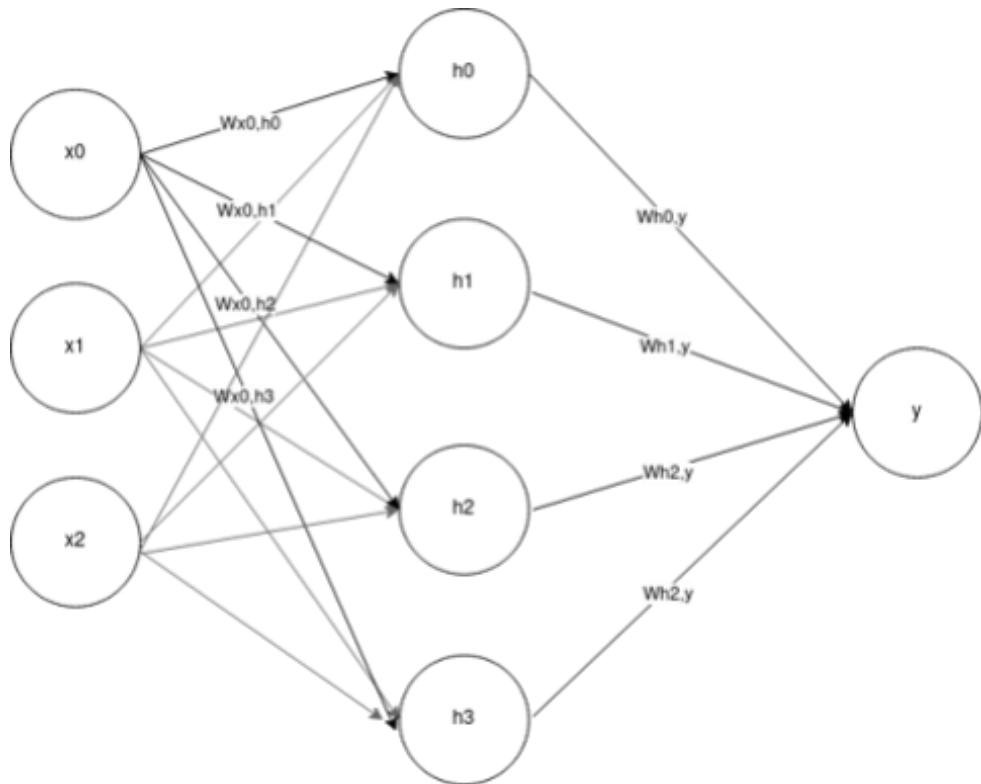


Vamos a asumir ahora que obtenemos datos históricos más completos, siendo estos numéricos, de modo que nuestros datos son los siguientes:

	<b>Sensor1</b>	<b>Sensor2</b>	<b>Sensor3</b>	<b>Clase</b>
1	3	2	9	1
2	6	5	7	0
3	8	1	4	0
4	1	9	2	0
5	4	3	8	1
6	7	6	5	1
7	9	8	3	1
8	2	7	6	0
9	5	4	1	1
10	3	2	9	0
11	8	5	7	0
12	6	1	4	1
13	9	9	2	1
14	1	6	8	0
15	7	3	5	1



Dada la siguiente red neuronal ya entrenada donde las neuronas  $x_0$ ,  $x_1$  y  $x_2$  forman la capa de entrada en la que introducir los datos del Sensor 1, el Sensor 2 y el Sensor 3 respectivamente.



Donde:

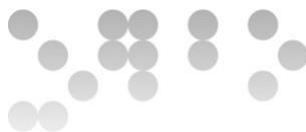
$$h_0 = w_{x_0, h_0} X_0 + w_{x_1, h_0} X_1 + w_{x_2, h_0} X_2$$

$$h_1 = w_{x_0, h_1} X_0 + w_{x_1, h_1} X_1 + w_{x_2, h_1} X_2$$

$$h_2 = w_{x_0, h_2} X_0 + w_{x_1, h_2} X_1 + w_{x_2, h_2} X_2$$

$$h_3 = w_{x_0, h_3} X_0 + w_{x_1, h_3} X_1 + w_{x_2, h_3} X_2$$

$$Y = w_{h_0, y} h_0 + w_{h_1, y} h_1 + w_{h_2, y} h_2 + w_{h_3, y} h_3$$



Los pesos asociados a cada una de las conexiones neuronales están definidos en la siguiente tabla:

<b>W</b>	<b><i>h0</i></b>	<b><i>h1</i></b>	<b><i>h2</i></b>	<b><i>h3</i></b>
<b><i>x</i><sub>0</sub></b>	1	-0,5	0,5	1
<b><i>x</i><sub>1</sub></b>	-1	1	1	-0,5
<b><i>x</i><sub>2</sub></b>	1	-0,5	-0,25	1
<b>y</b>	-1	-0,5	1	-0,25

La función de activación más utilizada en la capa oculta es la ReLU y en la capa de salida se aplica la siguiente función *heaviside*:

$$\text{heaviside}(x) = \begin{cases} 0, & \text{if } x \leq 0 \\ 1, & \text{if } x > 0 \end{cases}$$

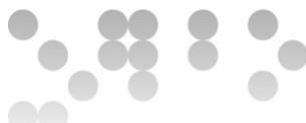
Calcula:

- a) La salida de la red para cada observación.
- b) La matriz de confusión de este clasificador.
- c) La métrica F1. Para calcular la matriz de confusión y la métrica F1 consideraremos 1 como el valor positivo y 0 como el valor negativo.

Respuesta:

Dado los pesos asociados de las conexiones neuronales completamos:

$$\begin{aligned} h_0 &= x_0 - x_1 + x_2 \\ h_1 &= -0.5x_0 + x_1 - 0.5x_2 \\ h_2 &= 0.5x_0 + x_1 - 0.25x_2 \\ h_3 &= x_0 - 0.5x_1 + x_2 \\ y &= -h_0 - 0.5h_1 + h_2 - 0.25h_3 \end{aligned}$$



En la siguiente tabla tenemos los cálculos intermedios necesarios, y la salida de red, Out:

	X0	X1	X2	H0	H1	H2	H3	Y	Out	Clase	F1
1	3	2	9	10	-4	1.25	11	-9.5	0	1	Fn
2	6	5	7	8	-1.5	6.25	10.5	-3.63	0	0	tn
3	8	1	4	11	-5	4	11.5	-7.38	0	0	tn
4	1	9	2	-6	7.5	9	-1.5	11.63	1	0	fp
5	4	3	8	9	-3	3	10.5	-7.13	0	1	Fn
6	7	6	5	6	0	8.25	9	0	0	1	Fn
7	9	8	3	4	2	11.75	8	4.75	1	1	Tp
8	2	7	6	1	3	6.5	4.5	2.88	1	0	fp
9	5	4	1	2	1	6.25	4	2.75	1	1	Tp
10	3	2	9	10	-4	1.25	11	-9.5	0	0	tn
11	8	5	7	10	-2.5	7.25	12.5	-4.63	0	0	tn
12	6	1	4	9	-4	3	9.5	-6.38	0	1	Fn
13	9	9	2	2	3.5	13	6.5	7.63	1	1	Tp
14	1	6	8	3	1.5	4.5	6	-0.75	0	0	tn
15	7	3	5	9	-3	5.25	10.5	-4.88	0	1	Fn

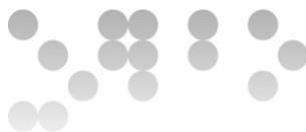
La matriz de confusión es la siguiente:

	Positivo (1)	Negativo (0)
Positivo (1)	3 (TP)	2 (FP)
Negativo (0)	5 (FN)	5 (TN)

A continuación realizamos el cálculo de las medidas:

$$Precisión = \frac{TP}{TP + FP} = 0,6$$

$$Sensibilidad = \frac{TP}{TP + FN} = 0,375$$



$$\text{Exactitud} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{TN} + \text{FP}} = 0,5333 \dots$$

$$F1 = \frac{2 \cdot \text{precisión} \cdot \text{sensibilidad}}{\text{precisión} + \text{sensibilidad}} = 0,4615$$

### Pregunta 3 (2 puntos)

- a) Comenta las diferencias principales entre los dos modelos anteriores, árbol de decisión y red neural.
- b) ¿Qué tipo de paradigma de aprendizaje usan? ¿Y en qué consiste?

Respuesta:

a) El árbol de decisión se representa como una jerarquía sencilla de nodos (variables) conectados por aristas (valores), lo que lo hace fácil de interpretar visualmente. Por otro lado, la red neuronal tiene una estructura más compleja, formada por múltiples capas de nodos interconectados y esto dificulta su comprensión directa.

Aunque los árboles de decisión pueden manejar relaciones no lineales, su capacidad es limitada, especialmente cuando se usan árboles poco profundos. Por otro lado, la red neuronal es mucho más efectiva en modelar relaciones no lineales complejas gracias a sus funciones de activación y transformaciones.

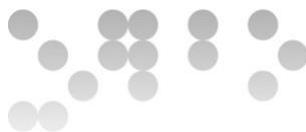
El árbol de decisión tiene un proceso de entrenamiento rápido, ya que segmenta los datos en particiones basadas en las combinaciones de las variables. Por otro lado, la red neuronal necesita un entrenamiento más largo y complejo, utilizando métodos numéricos avanzados sobre todo cuando se trata de redes profundas.

Los árboles de decisión pueden manejar bien tamaños pequeños de datos de entrada. Por otro lado, la red neuronal requiere un gran volumen de datos para entrenar de manera adecuada y modelar patrones complejos de comportamiento.

b) El tipo de paradigma de aprendizaje que utilizan tanto el árbol de decisión como la red neuronal es el aprendizaje supervisado. Este enfoque consiste en entrenar el modelo con un conjunto de datos en el que las entradas y las salidas ya son conocidas.

En el aprendizaje supervisado, el modelo recibe ejemplos de datos con etiquetas o valores asociados, lo que se denomina conjunto de entrenamiento. El objetivo es que el modelo aprenda a hacer predicciones precisas. Para ello, debe minimizar la diferencia entre las salidas predichas y las reales, que son conocidas a partir de los datos de entrenamiento.

Con el tiempo, el modelo ajusta y refina sus predicciones de manera iterativa, reduciendo el error con cada ciclo de entrenamiento. Este proceso permite que tanto los árboles de decisión como las redes neuronales aprendan patrones y relaciones en los datos. De esta manera, pueden generalizar y realizar predicciones sobre datos nuevos y desconocidos.



## Recursos

Para hacer esta PEC el material imprescindible es el módulo 5.

## Criterios de valoración

Las puntuaciones se muestran en cada pregunta del enunciado.

## Formato y fecha de entrega

Para dudas y aclaraciones sobre el enunciado, diríguirlos al consultor responsable del aula.

Hay que entregar la solución en un archivo PDF usando una de las plantillas entregadas conjuntamente con este enunciado. Adjuntar el fichero a un mensaje en el apartado Entrega y Registro de EC (REC).

El nombre del archivo debe ser *Apellidos\_Nombre\_IA\_PEC4* con la extensión .pdf (PDF).

La fecha límite de entrega es el: **31/12/2024** (a las 24 horas).

Razonad la respuesta en todos los ejercicios. Las respuestas sin justificación no recibirán puntuación.

### Nota: Propiedad intelectual

A menudo es inevitable, al producir una obra multimedia, hacer uso de recursos creados por terceras personas. Es por tanto comprensible hacerlo en el marco de una práctica de los estudios de Informática, siempre que se documente claramente y no suponga plagio en la práctica.

Por lo tanto, al presentar una práctica que haga uso de recursos ajenos, se presentará junto con ella un documento en el que se detallen todos ellos, especificando el nombre de cada recurso, su autor, el lugar donde se obtuvo y el su estatus legal: si la obra está protegida por copyright o se acoge a alguna otra licencia de uso (Creative Commons, licencia GNU, GPL ...).

El estudiante deberá asegurarse de que la licencia que sea no impide específicamente su uso en el marco de la práctica. En caso de no encontrar la información correspondiente deberá asumir que la obra está protegida por copyright.

Deberán, además, adjuntar los archivos originales cuando las obras utilizadas sean digitales, y su código fuente esté corresponde.