

# Test: Prueba de Contenido General Recuperativa.

Profesor: Rodolfo Anibal Lobo

Noviembre 2023

## Instrucciones

Lea atentamente las preguntas y conteste de manera breve y clara sus respuestas. Este archivo contiene preguntas extras. La prueba solo tendrá 3 preguntas cada una valiendo 2 puntos.

1. Defina Machine Learning o Aprendizado de Máquina. Mencione 3 ejemplos de aplicación general (fuera del mundo del audio) y una aplicación dentro del mundo del audio. De ser necesario referencie su respuesta.
2. Defina aprendizaje supervisado, no supervisado y semi supervisado. De un ejemplo de uso en cada caso.
3. Supongamos que tenemos un pokedex programado en `python` ¿Cuáles serían los 3 pasos para poder consultar por el pokemón llamado *Onix*?. La clase utilizada para crear el pokedex es la siguiente:

```
1 import requests
2 class Pokedex:
3     def __init__(self):
4         # Variables que utilizaremos dentro de toda la clase
5         self.base_url =
6             "https://pokeapi.co/api/v2/pokemon/{"
7         self.name = None
8         self.type = None
9         self.attack = None
10    def search(self, name: str):
11        # Dando formato al nombre del pokemon en la URL
12        self.name = name.lower()
13        url = self.base_url.format(self.name)
14        # Manejo de excepciones: intento preguntarle a la API
15        try:
16            response = requests.get(url)
17        except:
18            print("Algo extraño ocurrió")
19        # Observando la respuesta de la API
20        if response.status_code == 200:
21            data = response.json()
22            self.type = data['types'][0]['type']['name']
23            self.attack = data['moves'][0]['move']['name']
24        else:
25            print(f"No pudimos encontrar datos para
26                {self.name}")
```

```

25     # Pantalla del pokedex
26     def mostrar_informacion(self):
27         print(f"Nombre: {self.name}")
28         print(f"Tipo: {self.type}")
29         print(f"Ataque: {self.attack}")

```

4. Dados los siguientes grupos de códigos escribe en cada caso que obtendremos por salida:

```

1 numbers = [1, 2, 3, 4, 5]
2 output = [x**2 for x in numbers]
3 print(output)

```

```

1 class Animal:
2     def __init__(self, name):
3         self.name = name
4
5     def speak(self):
6         return "Hola, soy un animal llamado " + self.name
7
8 animal = Animal('Firulais')
9 print(animal.speak())

```

```

1 datos = {'a': 1, 'b': 2, 'c': 3}
2 del datos['b']
3 print(datos)

```

```

1 def suma(a, b=2, c=3):
2     return a + b + c
3
4 resultado = suma(1, c=1)
5 print(resultado)

```

```

1 import numpy as np
2
3 A = np.array([[1, 2], [3, 4], [5, 6]])
4 v = np.array([1, 1])
5
6 resultado = A@v
7 print(resultado)
8 print("Shape de A:", A.shape)
9 print("Shape de v:", v.shape)
10 print("Shape del resultado:", resultado.shape)

```

5. Dado el cálculo de  $\text{accuracy} = \frac{\text{total de aciertos}}{\text{total de aciertos} + \text{total de errores}}$  en un problema de clasificación. Analicemos el siguiente problema:

Supongamos que estás trabajando en la clasificación de un conjunto de datos con tres posibles etiquetas de clase: A, B y C. Se te proporcionan dos conjuntos de datos pequeños, uno balanceado y otro desbalanceado, junto con las predicciones de un modelo. Tu tarea es calcular la precisión (accuracy) del modelo para cada conjunto de datos y discutir por qué el **accuracy** podría o no ser una buena métrica en estos escenarios.

Observación	Clase Verdadera	Predicción del Modelo
1	A	A
2	A	B
3	B	B
4	B	A
5	C	C
6	C	C

Tabela 1: Dataset Balanceado

Observación	Clase Verdadera	Predicción del Modelo
1	A	A
2	A	C
3	B	B
4	C	C
5	C	C
6	C	C

Tabela 2: Dataset Desbalanceado

- Calcula la accuracy del modelo para el dataset balanceado.
  - Calcula la accuracy del modelo para el dataset desbalanceado.
  - Compara los accuracies obtenidos y explica qué revela esto sobre la idoneidad de la accuracy como métrica en contextos de clasificación con clases desbalanceadas
6. Explique de manera intuitiva, utilizando el gráfico de una parábola con el eje horizontal dado por el parámetro  $\theta$  y eje vertical dado por la función de pérdida  $\mathcal{L}$ , la técnica de optimización llamada Gradiente Descendiente. Mencione cuál es uno de los problemas inherentes del algoritmo.
7. Dadas las siguientes ecuaciones normales:

$$\Theta = (X^T X)^{-1} X^T \mathbf{y} \quad (1)$$

Si  $\mathbf{x} \in \mathbb{R}^{1 \times N}$  es una fila de la matriz  $X$ , es decir  $\mathbf{x} = [x_1, x_2, \dots, x_N]$  nuestros datos, para los cuales tenemos 250 ejemplos. Además sea  $\theta \in \mathbb{R}^{6 \times 1}$  e  $y_i \in \mathbb{R}$ . ¿Cuáles son las dimensiones de las siguientes matrices-vectores?

- $X^T$
- $X^T X$
- $(X^T X)^{-1}$
- $X^T \mathbf{y}$
- $\mathbf{x}$
- $(X^T X)^{-1} X^T$

(g) **y**

8. Explique detalladamente porque el modelo de regresión lineal en 2D (utilizando una recta) no es eficiente a la hora de resolver problemas de clasificación binaria. Utilice gráficos y una regla de clasificación dada la recta para desarrollar sus explicaciones.
9. Dadas las clases  $\text{Clases} = \{\text{miedo}, \text{felicidad}, \text{amor}, \text{tristeza}, \text{violencia}\}$ , un modelo entrenado para análisis de sentimiento, generó la siguiente matriz de confusión:

Tabela 3: Matriz de Confusión

		Clases Predichas				
		mied	felici	amor	triste	violen
Clases Reales	mied	1620	30	60	0	0
	felici	0	1390	110	80	100
	amor	0	40	1090	40	0
	triste	0	350	30	1020	90
	violen	0	50	30	300	800

- (a) ¿Cuál fue la clase donde el modelo clasifico de manera más correcta?
  - (b) ¿Cuál fue la clase donde el modelo tuvo más problemas?, ¿Qué emociones está confundiendo en este caso? (HINT: normalice las filas)
10. Explique cada parte involucrada en una clasificación binaria dentro de la función Binary Cross Entropy Loss:

$$\mathcal{L}(y, \hat{y}) = - \sum_{k=1}^n y \cdot \log(\sigma(x_i)) + (1 - y) \cdot \log(1 - \sigma(x_i)) \quad (2)$$

recuerde que en general, bajo nuestra notación  $\hat{y} = \sigma(x)$ . Utilice las clases  $c \in \{0, 1\}$  para ejemplificar que ocurre en un problema de clasificación binaria y cómo la expresión anterior nos ayuda a "castigar" el modelo cuando comete errores.

11. En relación al escalamiento de características:
  - (a) Explique porque es necesario escalar los datos al momento de entrenar utilizando gradiente descendiente. ¿Cómo se ve afectado el modelo de Gradiente Descendiente si no escalamos en un escenario de características con notables diferencias en sus escalas?

(b) Explique al menos dos formas diferentes de escalar datos mencionando el rango de salida de cada método.

12. ¿Qué es el learning rate o tasa de aprendizaje en el contexto de Gradiente Descendiente?

13. Dado el modelo lineal:

$$g_{\theta}(x) = \theta_1 x + \theta_0, \quad (3)$$

y la función de pérdida:

$$\mathcal{L}(\theta, x_i, y_i) = \frac{1}{n} \sum_{k=1}^n (g_{\theta}(x_i) - y_i)^2 \quad (4)$$

y dados los siguientes parámetros de inicialización:

- $\alpha = 0.1$
- $\theta_0 = 0$
- $\theta_1 = 0$

realice dos iteraciones del algoritmo de gradiente descendiente de forma estocástica para el siguiente dataset:

x	y
1	1
2	0.5
3	2

Tabela 4: Datos de entrenamiento

14. Dado el modelo lineal:

$$g_{\theta}(x) = \theta_1 x + \theta_0, \quad (5)$$

y la función de pérdida:

$$\mathcal{L}(\theta, x_i, y_i) = \frac{1}{n} \sum_{k=1}^n (g_{\theta}(x_i) - y_i)^2 \quad (6)$$

y dados los siguientes parámetros de inicialización:

- $\alpha = 0.1$
- $\theta_0 = 0$
- $\theta_1 = 0$

escriba como se representaría una iteración del modelo en **batch enviando todos los ejemplos de entrenamiento en una sola iteración**:

x	y
1	1
2	0.5
3	2

Tabela 5: Datos de entrenamiento

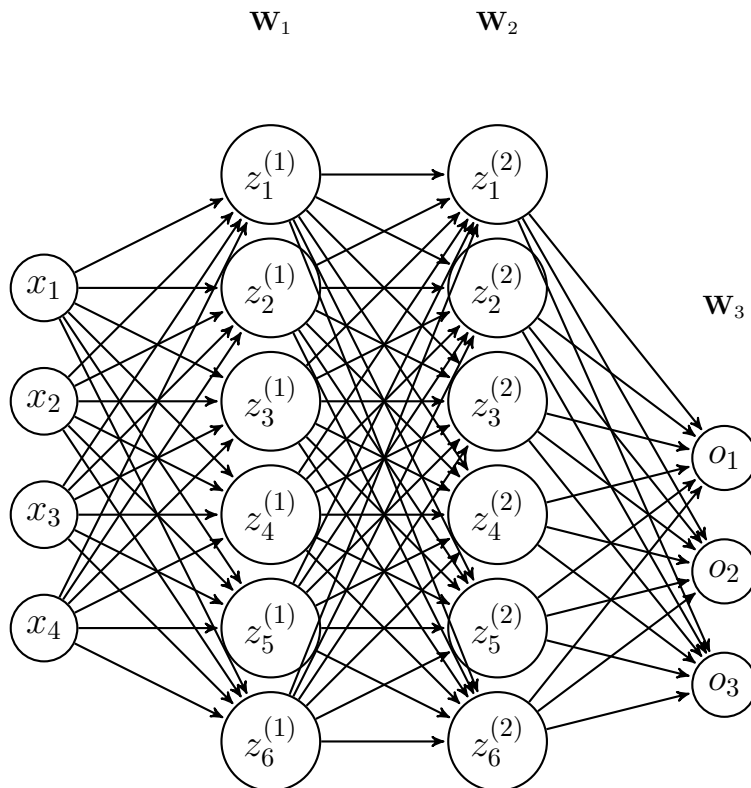
15. En relación a la transformada de Fourier:

(a) Cómo nos puede ayudar en el contexto de Machine Learning (justifique su respuesta de manera detallada y con ejemplos).

- (b) ¿Por qué el cálculo de los coeficientes de una serie de Fourier puede ser pensado como una regresión?
- (c) ¿Qué representan los coeficientes en el contexto de audio digital?
16. Explique cada parte que conforma una neurona artificial. Expresé el modelo tanto como un diagrama como con su fórmula matemática. ¿Existe una única manera de entrenar una red neuronal?
17. ¿Qué es una red neuronal?, ejemplifique utilizando una red de tres neuronas.
18. Dado un vector de entrada  $\mathbf{x} \in \mathbb{R}^d$  con  $d = 10$ . Queremos construir un modelo de red neuronal con dimensión de entrada 10, dimensión interna 60 y dimensión de salida 2. Escribe la ecuación de forma anidada utilizando la notación siguiente:
- (a)  $W_i$  matriz de pesos de la capa  $i$
- (b)  $b_i$  vector de sesgo de la capa  $i$
- (c)  $\sigma$  función sigmoide

Recuerda que el modelo tiene dos capas de parámetros o dos modelos lineales involucrados, además recuerda que la operación sigmoide se realiza componente a componente. Puedes escribir tu respuesta en una sola línea o utilizando dos ecuaciones.

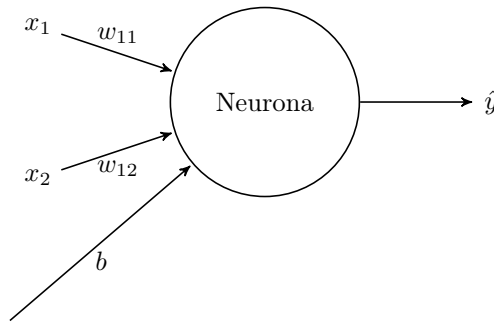
19. Dado el siguiente diagrama de una red neuronal donde  $x_j$  con  $j = 1, \dots, 4$  siendo las componentes de entrada o información de entrada y además, donde el sesgo o bias es igual a cero para cada capa, es decir  $\mathbf{b}_i = 0, \forall i$ :



Responda las siguientes preguntas:

- (a) Cuántas capas tiene este modelo.
- (b) Cuántos parámetros tiene este modelo.

- (c) Sea  $z_i^{(j)}$  la salida de la neurona  $i$  en la capa  $j$ . Cuál es la dimensión de las matrices involucradas si los modelos dentro de cada potencial de activación son modelos lineales de la forma  $W_i \mathbf{x} + \mathbf{b}_i$ .
- (d) Escriba las ecuaciones que representan a este modelo (ya sea de forma anidada o como un sistema de ecuaciones).
20. Suponiendo que usted utilizará audio en una de sus aplicaciones utilizando machine learning y asumiendo que no tiene el poder de computo suficiente (o los datos suficientes) para poder entrenar un modelo con muchos parámetros.
- (a) Nombre 2 de las características de audio que se pueden utilizar para poder extraer información de las fuentes de audio.
- (b) Explique en que escenarios utilizaría cada una de ellas, en particular, bajo el escenario de clasificar archivos de audio.
21. Dada la neurona artificial con 3 parámetros y función de activación  $\sigma(x) = \frac{x}{|x|}$  para  $|x| \neq 0$ ; y dado el dataset de la tabla 6



- (a) Encuentra los valores de los parámetros  $w_{11}$ ,  $w_{12}$ ,  $b$  para que el modelo pueda imitar correctamente la tabla de datos.
- (b) Para la siguiente tabla de datos no es posible hallar una combinación de números que pueda imitar la tabla:

x	y	clase
-	-	-
-	+	-
+	-	-
+	+	+

Tabela 6: Tabla de clases

¿Cuál fue la solución que encontraron los investigadores a este problema?

- (c) Escriba la ecuación matemática que representa a esta neurona artificial.