|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| logo berri CIFP  CIFP Ciudad Jardín LHII | | | **Curso de Especialización:** Inteligencia artificial y Big Data | | | |
| **Modulo:**  Sistemas de aprendizaje automático | | | **Cód. Modulo:**  SAPA |
| **Curso:** IABD | | | |
| **1ª Evaluación Final** | | | **Fecha**: 21/03/2023 | | | |
| **Cód. Prof:** PTI3 | | **Profesor:** Ainara Montoya | |  | | |
|  | | | | | | |
| **Datos del Alumno** | | | | | | |
| **Nº** | **Nombre:** | | | | **Firma** | |
| **Apellido 1:** | | | | |
| **Apellido 2:** | | | | |

**Calificación:**

# EXAMEN – PARTE 1

## EJERCICIO 1

Utiliza el modelo de mezcla gaussiana para identificar los clústeres en los datos del fichero blobs\_data.csv.

Utiliza el modelo de mezcla gaussiana para identificar los clústeres. Para elegir el número de grupos representa en una gráfica BIC y AIC para diferente número de grupos.

Representa gráficamente los clústeres (cada uno de un color) y pinta también los centros.

## EJERCICIO 2

En este ejercicio, trabajarás con un conjunto de datos de pruebas creados (regresion\_3variables.csv).

Tu objetivo es construir un modelo de regresión que pueda predecir con precisión la variable dependiente y a partir de las variables independientes x1, x2 y x3.

Deberás:

* Cargar los datos
* Visualizar gráficamente los datos (x1 respecto a y, x2 respecto a y e x3 respecto a y)
* Dividir los datos en conjunto de entrenamiento y pruebas
* Probar diferentes modelos de regresión polinómica, hasta ver cuál es el mejor degree (y explicarlo gráficamente y por medida del error en los conjuntos de entrenamiento y pruebas), no hace falta que uses validación cruzada. ¿Qué medición has usado para evaluar el modelo y por qué? ¿cuál es el mejor degree?
* Probar ahora con Elastic Net, optimizando los hiperparámetros con búsqueda exhaustiva, ¿ha mejorado con esta regularización? ¿cuáles son los mejores datos para los hiperparámetros?

## EJERCICIO 3

Carga el conjunto de datos Breast Cancer desde scikit-learn, este conjunto de datos tiene 30 marcadores a partir de los cuales predecir si estamos ante un caso benigno (1) o maligno (0).

Explora exhaustivamente los datos y divide el conjunto de datos en entrenamiento y pruebas (10% para las pruebas).

Crea un pipeline para preparar los datos y aplicar un algoritmo de máquinas de vector soporte, entrena el modelo (optimiza los hiperparámetros) y evalúa el modelo en el conjunto de pruebas.

Crea un pipeline para preparar los datos y aplicar PCA que explique al menos el 95% de la varianza, aplica esta pipeline al conjunto de entrenamiento y al conjunto de pruebas.

Muestra cuantas dimensiones tenías al principio y cuantas después de aplicar PCA.

Aplica a los datos obtenido de aplicar PCA el algoritmo de máquinas de vector soporte, entrena el modelo (ajusta otra vez los hiperparámetros y guarda el tiempo que tardas en entrenar el modelo) y evalúa el modelo en el conjunto de pruebas.

Compara el error y el tiempo empleado en los dos modelos (sin pca y con pca) e interpreta el resultado.

# EXAMEN – PARTE 2

## EJERCICIO 4

Predice las ventas y los gastos mensuales de una cadena de tiendas minoristas para el próximo mes, teniendo en cuenta los 10 últimos meses anteriores.

Pasos a seguir:

* Preparación de Datos:
  + Cargar los datos históricos de ventas y gastos (retail\_sales.csv).
  + Gráfica de las ventas y los gastos a lo largo del tiempo.
  + Normalizar los datos (normaliza y desnormaliza los datos manualmente)
  + Dividir los datos en conjuntos de entrenamiento, validación y pruebas.
  + Crear secuencias de datos de longitud fija para entrenar la red.
* Construcción del modelo:
  + Utilizar una Red Neuronal Recurrente (RNN) con celdas LSTM (mínimo 2 capas)
  + Añadir una capa densa al final para la predicción de las ventas y los gastos
* Entrena el modelo utilizando la detención temprana
* Evaluación y predicción:
  + Evaluar el modelo en el conjunto de prueba para verificar su rendimiento.
  + Realizar predicciones de las ventas y los gastos para el próximo mes (abril)

Nota: si te fijas en los datos es muy fácil predecir el resultado para abril, comprueba que los resultados que obtienes concuerdan con lo esperado para estar seguro de que está bien hecho.

## EJERCICIO 5

Usa el modelo SSD MobileNet v2 para la detección de objetos en la imagen detección\_objetos.

El resultado que espero es el siguiente:

Imagen que contiene Texto

Descripción generada automáticamente

## EJERCICIO 6

Aplica segmentación de objetos a la imagen segmentacion\_objetos.webp

# CRITERIOS DE EVALUACIÓN

|  |  |
| --- | --- |
| EJERCICIO 1 | PUNTUACIÓN |
| Elegir número de grupos | 0,75 |
| Gráfica | 0,25 |
| (Cara error u omisión resta 0,25) | **1** |

|  |  |
| --- | --- |
| EJERCICIO 2 | PUNTUACIÓN |
| Cargar, explorar y dividir los datos | 0,25 |
| Elegir el mejor degree | 0,75 |
| Aplicar ElasticNet | 0,75 |
| (Cara error u omisión resta 0,25) | **1,75** |

|  |  |
| --- | --- |
| EJERCICIO 3 | PUNTUACIÓN |
| Cargar, explorar y dividir los datos | 0,75 |
| Entrenar y evaluar modelo sin PCA | 0,75 |
| Entrenar y evaluar modelo con PCA | 0,75 |
| (Cara error u omisión resta 0,25) | **2,25** |

|  |  |
| --- | --- |
| EJERCICIO 4 | PUNTUACIÓN |
| Preparación de datos | 0,75 |
| Construcción y entrenamiento del modelo | 0,50 |
| Evaluación y predicción | 0,50 |
| (Cara error u omisión resta 0,25) | **1,75** |

|  |  |
| --- | --- |
| EJERCICIO 5 | PUNTUACIÓN |
| Detectar objetos | 1,75 |
| (Cara error u omisión resta 0,25) | **1,75** |

|  |  |
| --- | --- |
| EJERCICIO 6 | PUNTUACIÓN |
| Aplicar segmentación | 1,50 |
| (Cara error u omisión resta 0,25) | **1,5** |

**PARTE 1:**

|  |  |  |  |
| --- | --- | --- | --- |
| EJERCICIO 1 | EJERCICIO 2 | EJERCICIO 3 | NOTA |
| 1,00 | 1,75 | 2,25 | **5** |

**PARTE 2**

|  |  |  |  |
| --- | --- | --- | --- |
| EJERCICIO 4 | EJERCICIO 5 | EJERCICIO 6 | NOTA |
| 1,75 | 1,75 | 1,50 | **5** |

**TOTAL:**

|  |  |  |
| --- | --- | --- |
| PARTE 1 | PARTE 2 | NOTA |
| 5,00 | 5,00 | **10** |