|  |  |  |  |
| --- | --- | --- | --- |
| **DAT- 245 Inteligencia Artificial** | | **Fecha:** | **01/10/2024** |
| **CI** | **APELLIDOS Y NOMBRES** | | **NOTA** |
| **12419673** | **Condori Castañeta Diego Eduardo** | |  |

**1. ¿Cuál es el área de su interés que le gustaría investigar multidisciplinariamente con IA? Describa en un párrafo las características del área, métodos y metodologías. (Ej. Cáncer de pulmón, comprende que es un cáncer de pulmón, como se detecta, que parámetros clínicos se toma en cuenta para su detección, factores biológicos, físico y otros)**

**R.** En mi caso se investigará la **hipertensión arterial esencial** utilizando inteligencia artificial de manera multidisciplinaria. La hipertensión arterial esencial es una condición crónica caracterizada por la elevación persistente de la presión arterial sin una causa identificable. Este tipo de hipertensión es común en adultos y puede llevar a complicaciones graves como enfermedades cardiovasculares y renales. Realizando busque de información en sitios web, se tiene la siguiente consideración para su estudio, donde se pueden emplear métodos de análisis de datos masivos (big data) y aprendizaje automático (machine learning) para identificar patrones y factores de riesgo. Las metodologías incluirían el uso de algoritmos de clasificación y regresión para predecir la aparición de hipertensión, así como técnicas de minería de datos para analizar factores biológicos, genéticos, ambientales y de estilo de vida que contribuyen a esta condición. [Además, se podrían integrar datos clínicos y de encuestas de salud para desarrollar modelos predictivos y sistemas de apoyo a la decisión clínica que mejoren la prevención y el manejo de la hipertensión arterial esencial](https://www.kaggle.com/datasets/frederickfelix/hipertensin-arterial-mxico)

**2.   Selección un datatset tabular de al menos 14 columnas, 1000 filas. Si elige imágenes igualmente puede convertir la imagen en datos tabulares de NxM. De esta selección indique cual es la clase o si no tiene.**

**R.** En relación con el área de interés mencionada en la pregunta 1, se ha elegido el conjunto de datos de Hipertensión Arterial México para su análisis. Este dataset incluye la clase “Diagnóstico Hipertensión”, que permite determinar si una persona tiene un diagnóstico positivo o negativo de hipertensión.

Complemente con lo siguiente:

1. Sin el uso de librerías en Python programe el percentil y cuartil de cada columna. Que distribución se puede aplicar en su caso normal, Bernoulli, gaussiana, poisson, otros. Indique la razón de su uso graficando con matplotlib.
2. De al menos tres columnas seleccionadas por usted indique que datos son relevantes de estas, grafique la misma (puede ser dispersión o mapa de calor, otros), indique al menos 4 características por columna seleccionada.
3. Obteniendo la media, mediana, moda con el uso de librerías, grafique un diagrama de cajas-bigote de al menos 3 columnas. Explique el resultado.

3.      Para el preprocesamiento debe usted migrar su dataset de csv u otro formato a arff, una vez realizado ello. Realice un etiqueta onehotencoder, labelencoder, discretización y normalización.

4.      Con el uso de librerías realiza en Python los mismos preprocesamiento del punto 3.

5.      Sin el uso de librerías programe las penalizaciones l1 y l2, aplicando normalización.

6.      Con el uso de EXCEL, realice en el algoritmo genético de f(x)=x2x-1. Al menos tres generaciones. Automatice el cálculo.

7.      Programe el anterior problema en Python con

a.       El uso de DEAP

b.      Sin el uso de DEAP

8.      De las características de altura, peso y talla; realice su propio dataset determinando cuál será su clase. Realice el cálculo de la entropía y ganancia de información.

9.      Usted se escapara al campo luego del examen de IA, deberá llevar varios artículos que no ingresan a su mochila. ¿Cómo optimizaría este problema permitiendo llevar la mayor cantidad de articulos?

10.  En Excel con el uso de formulas convierta un decimal en binario, octal y hexadecimal.