



CURSO: CMP 5002 - DATA MINING
COLEGIO: POLITÉCNICO
Semestre: 1er Semestre 2024/2025

Tarea 6: Ejercicio usando las técnicas de autoencoders sparse and denoising.

Objetivo:

- Implementar una arquitectura autoencoder (AE) eficiente para analizar datos de imágenes y maximizar la clasificación de los individuos.

Modalidad: Trabajo en equipo. Por tanto, cada integrante debe preparar su parte (asignada por el líder) y dominar las técnicas evaluadas completamente.

Problema:

1. Dado el conjunto de datos **fase_datasets.zip** (se encuentra en la misma carpeta de orientación del proyecto en el D2L). Se desea:
 - Implementar dos arquitecturas AE (sparse y denoising) que permitan generar un espacio latente capaz de maximizar el desempeño de clasificación de individuos. Se debe variar los parámetros de cada AE al menos 3 veces para facilitar la búsqueda del mejor espacio reducido (ej: encoding_dim, regularization, etc). **(3 puntos)**
 - Mostrar los gráficos de los espacios reducidos para cada combinación de parámetros (Un subplot tipo grid) y plot de la función loss en el entrenamiento y validación. Tener en cuenta que, con cada variación de parámetros del punto anterior, se produce un nuevo modelo (ej: AE con dim=5, es diferente del AE con dim=50). **(2 puntos)**
 - Para cada escenario, mostrar la reconstrucción visual y el error de reconstrucción basado en alguna métrica estándar (ej: MSE). **(2 puntos)**
 - Aplicar un modelo supervisado o no supervisado (libre elección) que corrobore la calidad de los espacios generados por los AEs. Mostrar los resultados de las métricas de clasificación escogidas. **(2 puntos)**.
 - Determinar cuál es el mejor esquema de clasificación (AE + método de clasificación) entre todos los evaluados y aplicarlo al conjunto test externo para ver los resultados de las métricas de clasificación. **(1 punto)**.
 - Cargar al D2L los códigos implementados (archivo compactado) dentro del plazo de entrega.