



INGENIERÍA DE SISTEMAS Y COMPUTACIÓN

SISTEMAS INTELIGENTES II

SEGUNDO PARCIAL

Se requiere diseñar y programar un modelo de red neuronal de propagación hacia adelante que permita predecir, dadas las características de un tumor, si éste es benigno o maligno. Para el proceso de aprendizaje se cuenta con un conjunto de datos (DataSet_Cancer.csv) que contiene información sobre uno de los tipos de cáncer más comunes en las mujeres, el cáncer de mama.

A continuación, se proporciona la metadata del dataset:

Nombre de la característica	Descripción
id	Identificador único de cada caso o paciente.
diagnóstico	Diagnóstico del tumor: "B" (benigno) o "M" (maligno).
radio_promedio	Promedio del radio del tumor (distancia media desde el centro hasta el borde).
textura_promedio	Promedio de la desviación estándar de los valores de intensidad en la imagen.
perímetro_promedio	Promedio del perímetro del tumor.
área_promedio	Promedio del área del tumor.
suavidad_promedio	Promedio de las variaciones locales en la longitud de los bordes.
compacidad_promedio	Relación entre el perímetro y el área, comparada con un círculo.
concavidad_promedio	Promedio de la extensión de las concavidades (depresiones) en el contorno del tumor.
puntos_cóncavos_promedio	Promedio de los puntos cóncavos en el contorno del tumor.
simetría_promedio	Promedio de la similitud entre las mitades del tumor.
dimensión_fractal_promedio	Promedio de la "rugosidad" del borde, medido por la dimensión fractal.
error_estándar_radio	Error estándar del radio del tumor.
error_estándar_textura	Error estándar de la textura del tumor.
error_estándar_perímetro	Error estándar del perímetro del tumor.
error_estándar_área	Error estándar del área del tumor.
error_estándar_suavidad	Error estándar de las variaciones locales en la longitud de los bordes.
error_estándar_compacidad	Error estándar de la relación entre el perímetro y el área.
error_estándar_concavidad	Error estándar de la extensión de las concavidades.
error_estándar_puntos_cóncavos	Error estándar de los puntos cóncavos en el contorno del tumor.
error_estándar_simetría	Error estándar de la similitud entre las mitades del tumor.
error_estándar_dimensión_fractal	Error estándar de la rugosidad del borde.



INGENIERÍA DE SISTEMAS Y COMPUTACIÓN

SISTEMAS INTELIGENTES II

SEGUNDO PARCIAL

peor_radio	Valor más alto del radio del tumor.
peor_textura	Valor más alto de la textura del tumor.
peor_perímetro	Valor más alto del perímetro del tumor.
peor_área	Valor más alto del área del tumor.
peor_suavidad	Valor más alto de las variaciones locales en la longitud de los bordes.
peor_compacidad	Valor más alto de la relación entre el perímetro y el área.
peor_concavidad	Valor más alto de la extensión de las concavidades.
peores_puntos_cóncavos	Valor más alto de los puntos cóncavos en el contorno del tumor.
peor_simetría	Valor más alto de la similitud entre las mitades del tumor.
peor_dimensión_fractal	Valor más alto de la rugosidad del borde.

Requerimientos

Se debe hacer el respectivo análisis exploratorio de datos (EDA) para determinar transformaciones del conjunto de datos. Cada decisión deberá estar debidamente justificada.

Luego de haber hecho el EDA, debe generarse un modelo de red neuronal con precisión no inferior al 80% para predecir el tipo de diagnóstico (tumor benigno o tumor maligno).

Para el diseño de la red neuronal se debe considerar la siguiente partición del dataset: Entrenamiento (70%), Validación (20%) y Prueba (10%)

De la arquitectura de la red neuronal definitiva (modelo con precisión $\geq 80\%$) se deben documentar y justificar todos los valores de sus hiperparámetros, además presentar la respectiva curva de aprendizaje. De los modelos previos a la obtención del modelo definitivo se deben mostrar solo sus curvas de aprendizaje.

Del modelo definitivo se deberá dejar registro de las pruebas y su respectiva precisión.

Se debe hacer un comparativo entre el modelo definitivo y cualquier otro modelo previo a este. La comparación se hará basado en metadatos, precisión, curvas de aprendizaje y se deben usar gráficos que apoyen la comparación con su respectiva interpretación.