



Unidades Didácticas 1, 2 y 3: Actividad de aprendizaje 1

Comportamiento Humano e Integración de la Inteligencia Artificial

Para aplicar los conocimientos adquiridos durante las tres primeras unidades didácticas, se propone esta actividad de aprendizaje. Deberá realizarse de forma **individual**.

1. Objetivo

Utiliza lo practicado en este laboratorio para resolver un caso práctico “real” usando la librería Keras y una arquitectura MLP para predecir el carácter del tumor cancerígeno detectado en 569 pacientes. Practicaremos el uso de arquitecturas con varias capas ocultas y las ventajas/desventajas que esto implica frente a arquitecturas menos profundas y con más neuronas por capa. Usaremos Adam como función de modificación de matriz de pesos (optimizer) de la forma que se indica en el apartado de “Implementación”. Responde a las preguntas que se plantean en “Cuestiones”.

2. Enunciado

2.1. Implementación

Crea el notebook L2P3-Cancer.ipynb. El programa en Python deberá responder a los siguientes puntos:

1. El fichero cancer_dataset.csv contiene 32 columnas. La primera columna contiene un identificador anonimizado del paciente. La segunda (diagnosis) contiene el diagnóstico (Benigno o Maligno). El resto son 30 características morfológicas. Divide el fichero en dos datasets disjuntos, uno para entrenar y otro para predecir.
2. Crea un MLP que has de entrenar y validar con la información del dataset de entrenamiento. Usa como optimizer ‘Adam’, funciones de activación ‘ReLU’ y ‘sigmoide y función de error ‘Mean Squared Error’.

Prueba distintas arquitecturas (en número de capas y número de neuronas por capa).

Notas:

- o Deberás categorizar el diagnóstico usando el método de keras to_categorical.
3. Una vez creada la red neuronal y viendo que puede predecir correctamente, usar los valores del archivo de predicción y predice si los tumores de esos pacientes son benignos o malignos.

2.2. Cuestiones

Continúa en la memoria respondiendo a las siguientes cuestiones

1. Explica cómo has llevado a cabo la normalización de los datos de entrada

2. ¿Cuál es la mejor arquitectura? Dibújala y justifícala con una tabla que recoja los valores de *loss* y *accuracy* para el conjunto de entrenamiento y el de validación para las distintas pruebas que has llevado a cabo.
3. Con esa arquitectura, determina el error que se obtiene para el dataset de predicción usando los valores de diagnosis de ese fichero.
4. Ve eliminando características (de menos relevantes a más) para ver cómo influyen en la capacidad predictiva de la red. Para ello debes entrenar/validar cada vez con una neurona/dato de entrada menos, hasta que la **accuracy** baje más de un 10% respecto a la recogida en la cuestión. Analiza estos resultados.

3. Instrucciones de entrega

- **Extensión:** El fichero en formato .zip, a adjuntar, deberá contener la memoria .pdf y el código Python.
- **Nombre del fichero:** Actividad1_Nombre_Apellido1_Apellido2
- **Formato de entrega:** Seguir las instrucciones y fechas especificadas en el aula virtual.
- **El plagio de fuentes no referenciadas o entre compañeros significará el suspenso en la actividad de aprendizaje.**

4. Evaluación

- Resolución correcta de los supuestos planteados (desarrollo y uso adecuado de los conceptos básicos de las unidades didácticas) en la actividad de aprendizaje: 10 puntos.
- Formato/presentación/extensión/corrección ortográfica y gramatical: 1 punto.

No se evaluarán las actividades que se entreguen después de la fecha límite establecida en el aula virtual.

WELCOME
TO
UAX

OPENUAX
UNIVERSIDAD ALFONSO X EL SABIO

GRACIAS

UAX.COM