

Título de la tarea: Definición de propiedades de una red neuronal para ejercicio de clasificación.

Unidad:4

Ciclo formativo y módulo: Curso especialización en Inteligencia Artificial y Big Data - Casos prácticos de aplicación.

Curso académico: 2022-2023

¿Qué contenidos o resultados de aprendizaje trabajaremos?

Resultados de aprendizaje

- ✓ **RA4.** Aplica técnicas de aprendizaje no supervisado relacionándolas con los tipos de problemas que tratan de resolver.
- ✓ **RA5.** Aplica modelos computacionales de redes neuronales comparándolos con otros métodos de inteligencia artificial.
- ✓ **RA6.** Valora la calidad de los resultados obtenidos en la práctica con sistemas de aprendizaje automático integrando principios fundamentales de la computación.

Contenidos

- 1.- Árbol de decisión.
- 2.- Clustering.
- 3.- Redes neuronales y Deep learning.

[Siguiente »](#)

1.- Descripción de la tarea.



Caso práctico



[Storyset](#) (CC BY-SA)

A Eva le han pedido que entrene un modelo de Inteligencia Artificial que ayude a los médicos a predecir posibles casos de diabetes en pacientes.

Para ello le han facilitado una base de datos con el historial de las personas que han pasado por el hospital de su localidad y se les han hecho análisis de sangre para decidir después si tenían o no Diabetes.

La base de datos se la han dado en un archivo .csv, y no es especialmente grande, por lo que Eva va a hacer una primera aproximación con la herramienta BigML, que le va a permitir hacer un primer prototipo y

detectar posibles aspectos a revisar en la propia base de datos.

Va a utilizar un algoritmo de Árbol de Decisión, que es el que generalmente da mejores resultados en este tipo de primeras aproximaciones con bases de datos pequeñas.

Lo bueno de hacerlo en la plataforma de BigML es que también va a poder enriquecer el informe que entregará a los responsables del hospital un pequeño análisis del caso, pudiendo aportar buenas recomendaciones para el tratamiento previo de los datos que se pretendan utilizar en el desarrollo definitivo.

¿Qué te pedimos que hagas?

Te pedimos que utilices la plataforma BigML, tal como se ha explicado en el caso práctico del primer capítulo de esta unidad. Se da por sentado que sabes acceder, y que tienes usuario para trabajar en ella. Tendrás que entregar un documento con la siguiente información:

- ✓ **Apartado 1: Localiza el dataset "Diabetes completo (spanish)":**
 - ➡ Utiliza el buscador de datasets que tiene la propia plataforma para ello.

- Incorpora una captura de pantalla del dataset mencionado incorporado a tu apartado *Datasets*.
- ✓ **Apartado 2: Observación del dataset:**
 - Incorpora una captura de pantalla del dataset donde se vean al menos 10 categorías con sus tipologías, errores, histogramas, etc.
 - Explica cómo es el dataset: Número de instancias y número de categorías.
 - Explica el tipo de categorías (numéricas, texto, items, categóricas...).
 - Analiza los histogramas de cada categoría y comenta aquellos en los que consideres que hay algún tipo de anomalía.
- ✓ **Apartado 3: Preparación del dataset para entrenamiento y test:**
 - Incorpora una captura de pantalla del proceso en el que defines los porcentajes de datos reservados para entrenamiento y para test.
- ✓ **Apartado 4: Entrenamiento:**
 - Incorpora una captura de pantalla que muestre el árbol de decisión del modelo ya entrenado.
 - Explica los principales resultados: Casos en los que haya resultado positivo o negativo con suficiente confiabilidad.
 - Incorpora capturas de pantalla de los diagramas de confiabilidad (confidence) y predicción (prediction).
- ✓ **Apartado 5: Evaluación:**
 - Incorpora una captura de pantalla en la que se muestre la evaluación del modelo entrenado realizada con el dataset reservado en el apartado 3.
 - Explica el resultado de dicha evaluación, indicando el nivel de confianza obtenido (*Accuracy*) y el nivel de precisión (*Precision*).

NOTA IMPORTANTE

Tendrás que sacar pantallazos de tu ejercicio en BigML.com para incorporarlos al documento que entregarás como tarea. Asegúrate de tener una herramienta que te permita obtenerlos.

2.- Información de interés.



Recursos necesarios y recomendaciones

Recursos necesarios

- ✓ Ordenador personal.
- ✓ Conexión a Internet.
- ✓ Navegador web.
- ✓ Cuenta en BigML (gratuita).
- ✓ Herramienta para obtener capturas de pantalla.

Recomendaciones

- ✓ Antes de abordar la tarea:
 - Lee con detenimiento la unidad, consulta los enlaces para saber más, examina el material proporcionado por el profesor y aclara las dudas que te surjan con él.
 - Realiza el examen online de la unidad, y consulta nuevamente las dudas que te surjan. Solo cuando lo tengas todo claro, debes abordar la realización de la tarea.
- ✓ Tendrás que elaborar un documento indicando explicando el proceso de entrenamiento con la técnica de Machine Learning (Aprendizaje Automático) de un caso real. En la medida de lo posible te recomendamos que investigues un poco acerca de esa casuística (detección de casos de diabetes).



Indicaciones de entrega

Una vez realizada la tarea, el envío se realizará a través de la plataforma. El archivo se nombrará siguiendo las siguientes pautas:

Apellido1_Apellido2_Nombre_SAA_Tarea04

[« Anterior](#) [Siguiente »](#)

3.- Evaluación de la tarea.



Criterios de evaluación implicados

Criterios de evaluación RA4

- ✓ a) Se han caracterizado los tipos de problemas que el aprendizaje no supervisado trata de resolver.
- ✓ b) Se han caracterizado las técnicas de aprendizaje no supervisado utilizadas para la resolución de dichos tipos de problemas.
- ✓ c) Se han aplicado algoritmos utilizados en el aprendizaje no supervisado.
- ✓ d) Se ha optimizado el modelo de aprendizaje no supervisado validando datos de prueba.

Criterios de evaluación RA5

- ✓ a) Se han evaluado los modelos neuronales para elegir el más adecuado para cada clase de problema.
- ✓ b) Se han aplicado técnicas de aprendizaje profundo (deep learning) para entrenar redes de neuronas.
- ✓ c) Se han comparado las redes de neuronas artificiales con otros métodos de inteligencia artificial.
- ✓ d) Se ha reconocido una red de neuronas entrenada a partir de un conjunto de datos.

Criterios de evaluación RA6

- ✓ a) Se ha valorado la conveniencia de los algoritmos propuestos para dar solución a los problemas planteados.
- ✓ b) Se ha evaluado la aplicación práctica de los principios y técnicas básicas de los sistemas inteligentes.
- ✓ c) Se han integrado los principios fundamentales de la computación en la práctica para seleccionar, valorar y crear nuevos desarrollos tecnológicos.
- ✓ d) Se han desarrollado sistemas y aplicaciones informáticas que utilizan técnicas de los sistemas inteligentes.
- ✓ e) Se han desarrollado técnicas de aprendizaje computacional dedicadas a la extracción automática de información a partir de grandes volúmenes de datos.

¿Cómo valoramos y puntuamos tu tarea?

Rúbrica de la tarea	
Apartado 1: Se aporta captura de pantalla del dataset "Diabetes completo (spanish) incorporado al apartado Datasets del usuario.	0,5 puntos
Apartado 2: Se aportan capturas de pantalla del dataset donde se ven	0,5 puntos

al menos 10 categorías.	
Apartado 2: Se redacta una explicación del dataset indicando número de instancias y número de categorías.	1 punto
Apartado 2: Se redacta una explicación de la tipología de las categorías.	1 punto
Apartado 2: Se redacta una explicación sobre las categorías que podrían contener errores o alteraciones a la vista de sus histogramas.	1 punto
Apartado 3: Se aporta captura de pantalla del proceso en el que se divide el dataset original en dos dataset para entrenamiento y test.	1 punto
Apartado 4: Se aporta captura de pantalla que muestre el diagrama de árbol del modelo ya entrenado.	1 punto
Apartado 4: Se redacta una explicación de los principales resultados del entrenamiento a partir de la información del diagrama de árbol	1 punto
Apartado 4: Se aporta captura de pantalla de los diagramas de confiabilidad y predicción del modelo entrenado	1 punto
Apartado 5: Se aporta captura de pantalla que muestre el resultado de la evaluación del modelo hecha a partir del dataset test	1 punto
Apartado 5: Se redacta una explicación con las principales conclusiones de la evaluación del modelo, indicando al menos los niveles de confianza y precisión	1 punto
Redacción clara y correcta, sin errores ortográficos	Se resta 0,1 puntos por cada error ortográfico o expresiones incorrectas.

« Anterior