



Inteligencia Artificial y Big Data

## **Reto del módulo:** **Aprendizaje automático**

**Módulo:** Sistemas de Aprendizaje Automático

## RETO FINAL. MÓDULO 2. APRENDIZAJE AUTOMÁTICO

### Objetivos

Hasta hoy, los equipos de desarrollo de modelos de aprendizaje automático, es decir, los científicos de datos y los equipos de operaciones de IT se han mantenido separados por una barrera invisible formada por intereses particulares y el conocimiento. Mientras que los **científicos de datos** se desenvuelven en un ámbito más matemático/estadístico, los ingenieros de operaciones se mueven en ambientes más computacionales y esto provoca una disyunción. Por tanto, **el objetivo de esta actividad es desarrollar las capacidades**, no solo entrenar modelos de aprendizaje, sino también su despliegue a entornos de producción.

Para lograrlo, nos proponemos con este reto que el estudiante:

1. Entrene uno o varios modelos de aprendizaje automático para la clasificación de imágenes de resonancia magnética.
2. Refine y mejore los modelos.
3. Despliegue los modelos en una plataforma de servicio de modelos de aprendizaje.

### Descripción de la actividad

En la actividad de refuerzo de la unidad formativa 6, desarrollamos un modelo que permitía clasificar imágenes de resonancia magnética procedente de pacientes en los que se había confirmado la presencia de tumores cerebrales y otros que no los poseían. El dataset que usamos solo tenía 98 imágenes de pacientes que no tenían tumor y 155 casos de pacientes que sí lo tenían.

El modelo que usamos fue un VGG16 con los pesos de imagenet, para entrenarlo usamos 30 épocas y aumento de imágenes. Usamos 6 pasos por época, con tamaños de Batch de 32 para entrenamiento y 16 para validación y EarlyStopping usando como métrica para monitorear el entrenamiento accuracy y patience=6. En la parte de clasificación del modelo usamos tres capas densas de 1024, 100 y 64 neuronas, respectivamente, con función de activación relu.

Este modelo nos arrojó una accuracy de 70% en el conjunto de prueba, lo que es bajo. Por tanto, **necesitamos un modelo que nos dé un accuracy** en el conjunto de pruebas no menor que 90%.

Por tanto, tu tarea es **encontrar aquellas combinaciones de parámetros o acciones que permitan obtener un modelo** con una calidad de clasificación mejorada.

Una vez que lo obtengas, deberás desplegarlo a un model server y usar el cliente desarrollado para hacer inferencias sobre el modelo.

Realiza un informe donde se explique claramente la estrategia seguida de las modificaciones en los parámetros y una justificación de cada uno de ellos.

### Criterios de evaluación – Rúbrica

Criterio	Descripción	Puntuación máxima (puntos)	Peso %
Accuracy =>90%	Obtener un modelo cuya accuracy sea igual o mayor que 90%.	80	80%
Despliegue	Desplegar el modelo.	10	10%
Hacer inferencias	Hacer inferencias sobre el modelo desplegado.	10	10%
		<b>100</b>	<b>100 %</b>

## Entrega. Formato y extensión

- Se debe entregar un informe con la descripción detallada de la actividad y las respuestas a todo lo que se solicita en el enunciado. El informe será un PDF con imágenes y texto donde se vean los pasos que se siguieron. Adjunto al informe debe aparecer el notebook con el código Python.
- Se valorará la explicación clara y argumentada, con tablas o estructuras que faciliten esta explicación.

2. Validad que no es una opción de ‘riesgo inaceptable’ y razonadlo aquí (ayudaos de los artículos anteriores y de la propuesta de Reglamento incluida en el Componente 1 al final del Reto).
3. Partiendo de vuestra propuesta, determinad los objetivos del negocio conforme a la primera fase de CRISP-DM (ver Componente 2 incluido al final del reto) respondiendo a estas preguntas:
  - a. cuál es el problema que se desea resolver,
  - b. por qué la necesidad de utilizar Inteligencia Artificial (y la tecnología específica) y
  - c. definir los criterios de éxito (de forma cualitativa o cuantitativa), es decir, lo que presentaremos como medida del éxito del proyecto.
4. Partiendo de vuestra propuesta, determinad una evaluación de la situación conforme a la primera fase de CRISP-DM (ver Componente 2 incluido al final del reto) que os debe ayudar a calificar el estado de la situación antes de iniciar el proceso de IA, considerando aspectos tales como:
  - a. ¿Cuál es el conocimiento previo disponible acerca del problema?,
  - b. ¿Se cuenta con la cantidad de datos requerida para resolver el problema que planteamos?,
  - c. ¿Cuál es la relación coste beneficio de la aplicación de IA incluyendo los beneficios de posicionarnos en el mercado?

Es importante que lo toméis como un primer acercamiento para trabajar vuestras capacidades de unir el aspecto técnico, metodológico y legal en una propuesta de solución tecnológica de