

## **1er Examen Parcial de Inteligencia Artificial**

Importante: Lea atentamente todo el enunciado antes de comenzar a responder. En todos los casos donde se pide un ejemplo, el mismo debe ser de la vida real.

Como entrega del examen, deberá entregar un cuaderno de Google Colab (formato ipynb), donde estarán incluidos todos los detalles para entender el estudio que ha realizado.

Cuando se pide resolver el ejercicio en un cuaderno de Google Colab, muestra mediante gráficos y explicaciones acordes, las conclusiones a las que ha arribado luego del aprendizaje. Documente cada uno de los pasos del análisis para demostrar su entendimiento sobre el tema.

### **Ejercicio 1**

Se le proporciona un conjunto de datos de marketing <https://archive.ics.uci.edu/dataset/222/bank+marketing> relacionado con campañas de marketing directo de una entidad bancaria. El objetivo es predecir si un cliente se suscribirá o no a un depósito a plazo (variable objetivo "y", donde 1 significa suscripción y 0 significa no suscripción).

La resolución deberá contemplar cada uno de los siguientes pasos:

- 1) **Carga de Datos:** Comience por cargar el conjunto de datos y realizar una exploración inicial para comprender la naturaleza de los datos. ¿Cuántas características hay en el conjunto de datos? ¿Hay valores faltantes? ¿Cómo está distribuida la variable objetivo? ¿Hay desbalance de clases?
- 2) **Preprocesamiento de Datos:** Realice las tareas de preprocesamiento necesarias, como la codificación de variables categóricas y el escalado de características numéricas. Asegúrese de dividir el conjunto de datos en conjuntos de entrenamiento y prueba.
- 3) **Implementación de Modelos:** Implemente tres modelos diferentes para la clasificación de suscripción de depósitos a plazo: un modelo de Random Forest, un modelo XGBoost y un modelo de Regresión Logística. Ajuste los hiperparámetros de cada modelo de manera adecuada.
- 4) **Evaluación de Modelos:** Evalúe el rendimiento de cada modelo en el conjunto de prueba utilizando métricas de clasificación, como AUC, Accuracy y precisión, recall y F1-score. Compare y analice los resultados.
- 5) **Conclusiones:** Proporcione conclusiones finales basadas en el rendimiento de los modelos. ¿Cuál es el modelo que mejor se desempeña en este problema de clasificación? ¿Qué métricas son más importantes en este contexto y por qué?

Preguntas Teóricas Relacionadas:

- a) ¿Por qué es importante dividir el conjunto de datos en: entrenamiento y prueba en este ejercicio?

- b) Describe el concepto de "bagging" en Random Forest y cómo contribuye a la mejora del rendimiento del modelo.
- c) En el contexto de la regresión logística, explique el concepto de regularización. ¿Por qué es importante aplicar la regularización en ciertos casos? Mencione al menos dos técnicas comunes de regularización en regresión logística y sus efectos en el modelo.
- d) ¿Cuáles son algunos de los hiperparámetros clave en un modelo de Random Forest y cómo pueden afectar el rendimiento del modelo?
- e) ¿Por qué podría ser beneficioso ajustar los hiperparámetros de un modelo de Random Forest, incluso en un conjunto de datos relativamente simple como este?

## **Ejercicio 2**

- a. Desarrolle el enunciado de un problema a resolver (con temática de Superhéroes) con un Perceptrón.

Explique por qué puede ser resuelto por un Perceptrón y no usaría otro tipo de Red Neuronal.

Programe el modelo y visualice el resultado mediante uno o más gráficos. Incorpore en el cuaderno de Google Colab todos los comentarios para entender lo que está realizando.

- b. Desarrolle el enunciado de un problema a resolver (con temática de Dibujitos Animados) que deba resolver con una red neuronal distinta a un Perceptrón..

Explique por qué no puede ser resuelto por un Perceptrón pero si con una red neuronal multicapas.

Programe el modelo y visualice el resultado mediante uno o más gráficos. Incorpore en el cuaderno de Google Colab todos los comentarios para entender lo que está realizando. Utilice todos los framework que conoce.

- c. Explique todo lo que sabe sobre el algoritmo mediante el cual una red neuronal aprende.