

Para el dataset **diabetes.csv** realizar desarrolla las siguientes tareas:

EXPERIMENTO DE PRUEBA

1. Entrenamiento de un modelo
2. Registrar dicho modelo
3. Exponerlo a través de una API, y comprobar que hace las predicciones correctamente

EXPERIMENTOS COMPLETOS

1. Dividir el dataset en training (70%), validación (20%) y test (10%)
2. Entrenar el modelo con los datasets de training y validación
3. Ajustar los parámetros, utilizando diferentes combinaciones (por ejemplo con técnicas de Grid Search)
4. Hacer una selección de los mejores y registrarlos en MLflow. Para elegir lo mejores considera las diferentes métricas: *accuracy*, *precision*, *recall* y *F1 score*. Utiliza *criterios de carácter médico* para decidir qué métrica(s) son más relevantes
5. Exponer el modelo finalmente seleccionado a través de la API de MLflow
6. Evaluar las métricas reales del modelo sobre el dataset de test (20%). Utiliza para ello un script que obtenga las predicciones de todo el dataset a través de la API
7. Elabora tus conclusiones

ENTREGABLES:

- Tabla de parámetros y métricas de los modelos que hayas decidido incluir en el registro, especificando la razón de por qué es útil registrar estos y no los de otros runs que hayas realizado
- Script de obtención de las métricas de la API (aplicado al dataset de test)
- Conclusiones

NOTA: *Se adjunta EDA del dataset para facilitar la comprensión de los datos*