

Digital NAO- Backlog

Reto 5

**Desarrollo de algoritmos en modelos
con Machine Learning**

Luis Alejandro Azúa Urrutia

NAO ID: 3165

20/02/2026

AI Engineer Core

1	<u>Directora de Comunicación Social</u> Como directora de comunicación quiero un análisis claro de la calidad y funcionalidad de los datos del INEGI 2020 Para proporcionar una base sólida en el entrenamiento de un modelo.
2	<u>Lizbeth Torres, Analista de Datos</u> Como experta en investigación clínica Puedo utilizar métodos de limpieza de datos y transformación a formatos numéricos Para que el modelo de regresión logística pueda procesarlos correctamente.
3	<u>Ingeniería en Machine Learning</u> Como Lizbeth, experta en ML Puedo supervisar el entrenamiento del prototipo base del modelo de regresión logística. Para obtener una aproximación de los resultados de una implementación de esta tecnología en el problema actual.
4	<u>Empleo de técnicas de regularización</u> Como parte del equipo de Machine learning Puedo emplear mis conocimientos en técnicas de regularización y balanceo de clases para evitar overfitting y minimizar los falsos negativos.
5	<u>Presentar resultados</u> Como directora de comunicación quiero organizar los resultados y presentarlos por medio de un reporte final ante la secretaría de salud, de manera clara y comprensible.

- Desarrollar script para hacer un análisis exploratorio bivariado, multivariado de los datos del proyecto para uso posterior.
- Identificar variables clave
- Escribir reporte con hallazgos del análisis exploratorio.

- Script en python para transformar datos, variables numéricas y categóricas de manera específica.

- Desarrollar script de entrenamiento para modelo de regresión logística con sklearn que tome como variable objetivo a HeartDisease y lo demás como predictores.

- Evaluar métricas precision, recall, accuracy y f1-score en conjuntos de entrenamiento y prueba con modelo de sprint 2.
- Evaluar métricas precision, recall, accuracy y f1-score en conjuntos de entrenamiento y prueba con modelo de sprint 3 balancing.

- Presentación ejecutiva de análisis de resultados obtenidos en scripts contrastados con los resultados de sprint 2.

	Requerimientos	Etapas	Estima...	Entregables	Estado
1	- Crear código para hacer un análisis exploratorio bivariado, multivariado de los datos del proyecto para uso posterior.	Etapa 1	4 horas	eda_heart_disease.ipynb.	Done
2	- Escribir reporte con hallazgos del análisis exploratorio	Etapa 1	1 hora	report_eda_heart_disease.pdf	Done
3	- Script en python para transformar datos, variables numéricas y categóricas de manera específica.	Etapa 1	3 hora	processing_heart_disease.py processing_heart_disease.csv	Done
4	- Crear repositorio de Github	Etapa 1	15 minutos	Repositorio GitHub Reto5.pdf	Done
5	- Desarrollar script de entrenamiento para modelo de regresión logística con sklearn que tome como variable objetivo a HeartDisease y lo demás como predictores.	Etapa 2	4 horas	training_model_heart_disease.py	Done
6	- Evaluar métricas precision, recall, accuracy y f1-score en conjuntos de entrenamiento y prueba con modelo de sprint 2.	Etapa 2	2 horas	training_model_heart_disease.py	Done
7	- Script de código para entrenar modelo de regresión logística, además, ejecutar calibración de hyper parámetro C. Considerar validación cruzada con esquema kFold de 5 hojas.	Etapa 3	por definir	regularization_model_heart_disease.py	In-progress
8	- Evaluar métricas precision, recall, accuracy y f1-score en conjuntos de entrenamiento y prueba con modelo de sprint 3.	Etapa 3	2 horas	regularization_model_heart_disease.py	In-progress
9	- Crear guion de python que cargue datos, procese, entrene modelo de regresión logística. Implementar 2 técnicas para desbalance.	Etapa 3	por definir	balancing_model_heart_disease.py	To do
10	- Evaluar métricas precision, recall, accuracy y f1-score en conjuntos de entrenamiento y prueba con modelo de sprint 3 balancing.	Etapa 3	2 horas	regularization_model_heart_disease.py	To do
11	- Presentación ejecutiva de análisis de resultados obtenidos en scripts contrastados con los resultados de sprint 2.	Etapa 3	por definir	results_models_heart_disease.pdf	To do