

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 | <p><u>Directora de Comunicación Social</u></p> <p>Como directora de comunicacion
 quiero un analisis claro de la calidad y funcionalidad de los datos del INEGI 2020
 Para proporcionar una base solida en el entrenamiento de un modelo.</p> | <ul style="list-style-type: none"> - 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 analisis exploratorio. |
| 2 | <p><u>Lizbeth Torres, Analista de Datos</u></p> <p>Como experta en investigación clínica
 Puedo utilizar metodos de limpieza de datos y transformacion a formatos numericos
 Para que el modelo de regresión logistica pueda procesarlos correctamente.</p> | <ul style="list-style-type: none"> - Script en python para transformar datos, variables numericas y categoricas de manera especifica. |
| 3 | <p><u>Ingenieria en Machine Learning</u></p> <p>Como Lizbeth, experta en ML
 Puedo supervisar el entrenamiento del prototipo base del modelo de regresión logisitica.
 Para obtener una aproximación de los resultados de una implementación de esta tecnologia en el problema actual.</p> | <ul style="list-style-type: none"> - Desarrollar script de entrenamiento para modelo de regresion logistica con sklearn que tome como variable objetivo a HeartDisease y lo demas como predictores. |
| 4 | <p><u>Empleo de tecnicas de regularización</u></p> <p>Como parte del equipo de Machine learning
 Puedo emplear mis conocimientos en tecnicas de regularización y balanceo de clases para evitar overfitting y minimizar los falsos negativos.</p> | <ul style="list-style-type: none"> - Evaluar metricas precision, recall, accuracy y f1-score en conjuntos de entrenamiento y prueba con modelo de sprint 2. - Evaluar metricas precision, recall, accuracy y f1-score en conjuntos de entrenamiento y prueba con modelo de sprint 3 balancing. |
| 5 | <p><u>Presentar resultados</u></p> <p>Como directora de comunicación
 quiero organizar los resultados y presentarlos por medio de un reporte final ante la secretaria de salud, de manera clara y comprensible.</p> | <ul style="list-style-type: none"> - Presentación ejecutiva de analisis 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 analisis exploratorio	Etapa 1	1 hora	report_eda_heart_disease.pdf	Done	
3	- Script en python para transformar datos, variables numericas y categoricas de manera especifica.	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 regresion logistica con sklearn que tome como variable objetivo a HeartDisease y lo demas como predictores.	Etapa 2	4 horas	training_model_heart_disease.py	Done	
6	- Evaluar metricas 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 codigo para entrenar modelo de regresion logistica, ademas, ejecutar calibracion de hyper parametro C. Considerar validacion cruzada con esquema kFold de 5 hojas.	Etapa 3	por definir	regularization_model_heart_disease.py	In-progress	
8	- Evaluar metricas 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 regresion logistica. Implementar 2 tecnicas para desbalance.	Etapa 3	por definir	balacing_model_heart_disease.py	To do	
10	- Evaluar metricas precision, recall, accuracy y f1-score en conjuntos de entrenamiento y prueba con modelo de sprint 3 balacing.	Etapa 3	2 horas	regularization_model_heart_disease.py	To do	
11	- Presentación ejecutiva de analisis de resultados obtenidos en scripts contrastados con los resultados de sprint 2.	Etapa 3	por definir	results_models_heart_disease.pdf	To do	