

Smart Maritime Traffic Monitoring

Enrique Ulises Báez Gómez Tagle & Luis Alejandro Guillén Alvarez

Data Engineer

&

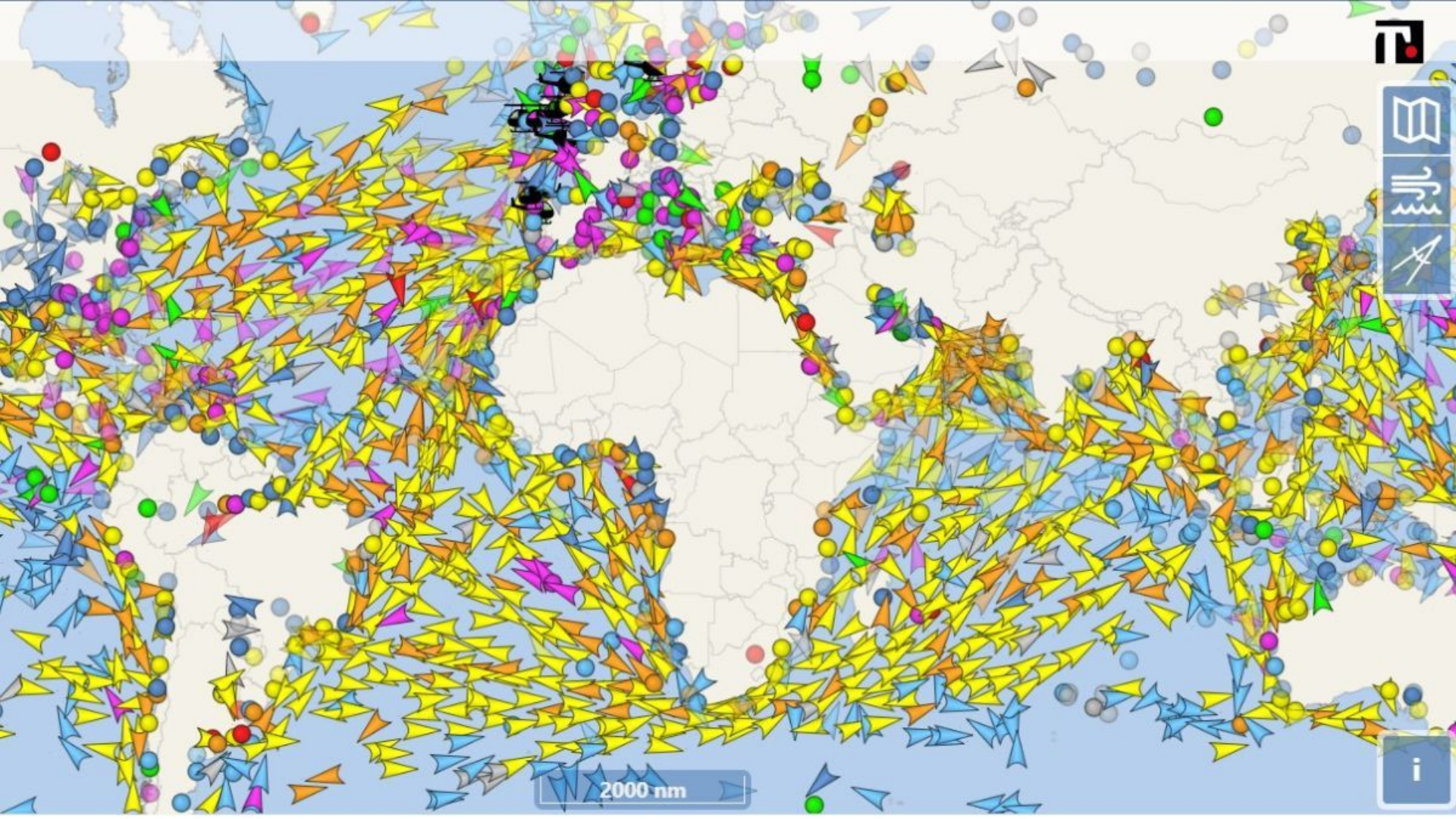
Data Scientist

José María Sarmiento Chávez

Big Data - Master's in Data
Science

UP CDMX





2000 nm

≈ 30
millones

de registros AIS al día



El problema:



- **Riesgos actuales:** colisiones, actividades ilícitas, rutas ineficientes
- **Dificultad:** identificar riesgos, anomalías y patrones en datos masivos.

“Sin análisis masivo, los datos AIS no generan valor”

Preguntas Clave



01 ¿Qué hacemos?

Analizamos datos AIS para detectar anomalías

02 ¿Para qué?

Mejorar seguridad, eficiencia y toma de decisiones

03 ¿Cómo?



Big Data + Machine Learning
en Google Cloud

04 ¿Quién gana?

Autoridades, empresas,
aseguradoras

Beneficiarios



Autoridades

Seguridad marítima, rescate



Navieras

Optimización de rutas, monitoreo de flotas



Aseguradoras

Gestión de riesgo operativo



Investigadores

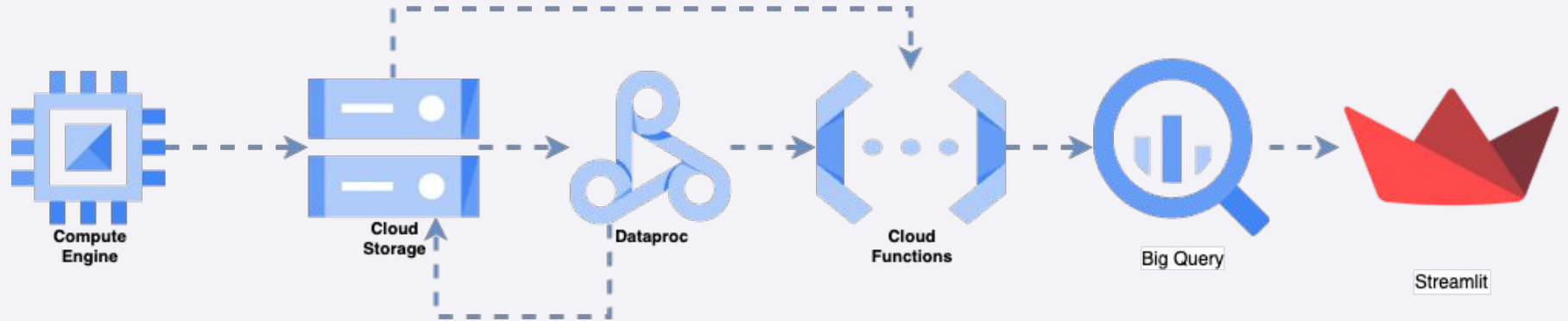
Impacto ambiental



3,119,488,538
registros



Solución propuesta



Componentes técnicos principales



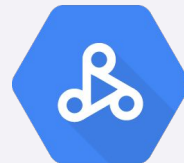
Compute Engine

Scraping de datos AIS desde NOAA



Cloud Storage

Almacenamiento en capas raw & curated



DataProc

Procesamiento batch (jobs raw + curated)



Cloud Functions

Automatiza creación de datasets/tablas y carga a BigQuery



Streamlit

Dashboard interactivo con mapas y visualizaciones



BigQuery

Análisis y consultas optimizadas



Configuración del clúster

DataProc high-memory

1 maestro (4 vCPUs, 32 GB RAM)

4 workers (2 vCPUs, 16 GB RAM)



Datos y Preparación



NOAA - AIS

Datos de 2024
(mensajes de tráfico
marítimo)

Variables clave

Posición (LAT, LON) 📍
Velocidad (SOG) 🕒
Rumbo (COG, Heading) 🧭
Tipo de buque 🚢

Control de calidad

Eliminación de duplicados
(MMSI + fecha/hora)

Validación de coordenadas
y rangos

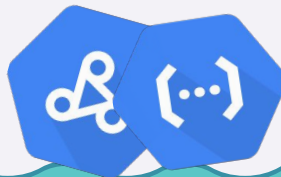
Descarga



RAW



Transformación



Curated layer

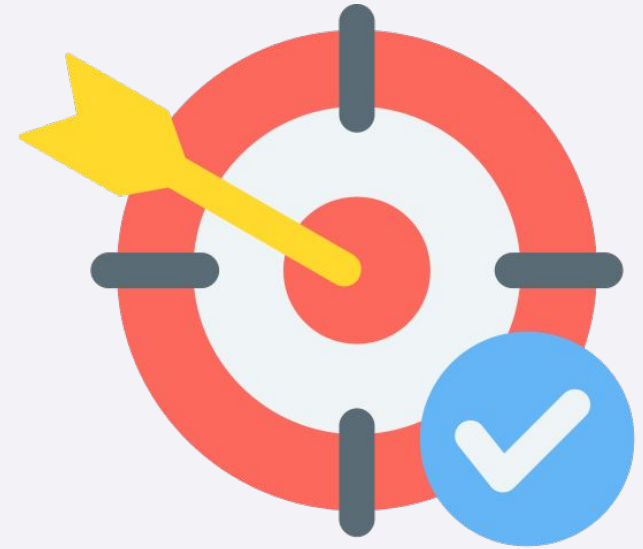


Limitaciones Actuales de la Solución

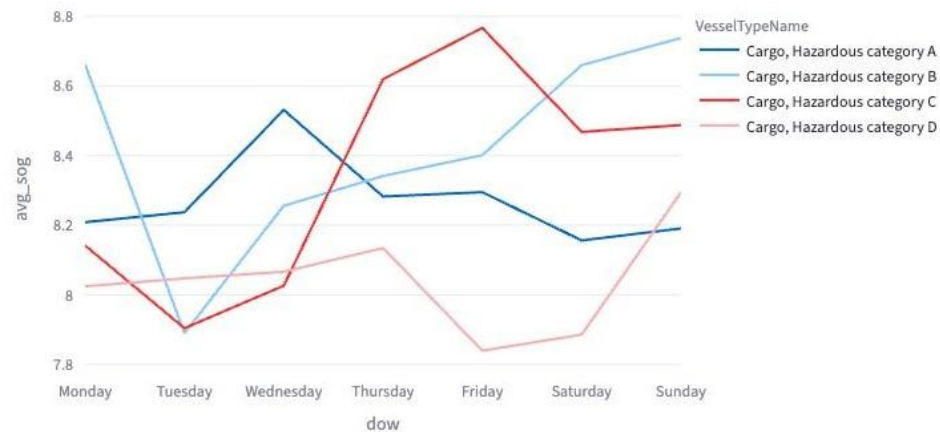


	Descripción	En nuestro proyecto...
Pipeline batch	El procesamiento es por corridas programadas, no en tiempo real	El dashboard refleja datos “al cierre” de cada ejecución, no en vivo
Evolución de esquema básica	Cambios en los datos (nuevos campos, ajustes) requieren pasos manuales	Si NOAA cambia el formato, habría que ajustar código y recargar tablas
Monitoreo limitado	El sistema aún no genera alertas proactivas de fallas o anomalías	Si falla un job de DataProc, nos enteramos revisando logs manualmente

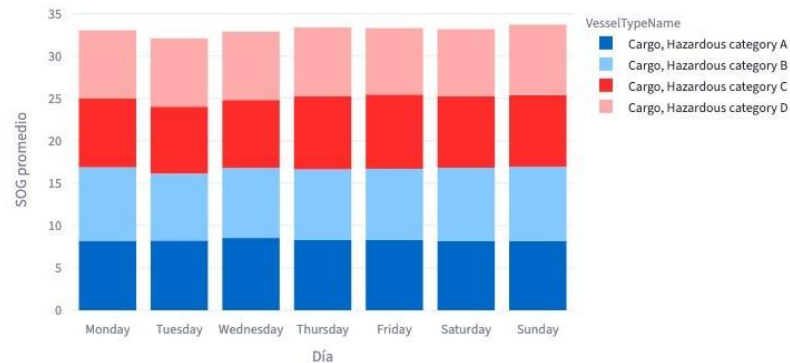
Resultados Y conclusiones



Tendencia semanal de velocidad



Velocidad promedio por día de la semana



Estado más frecuente por día de la semana



Google Cloud

My First Project

bigquery ml

Buscar

Explorador

+ Agregar datos

ML

Mostrar solo los destacados

% de buques con calado > promedio (por tipo)

CaladoAnomalo

CambiosDireccion

Correlación SOG vs Draft por tipo

Eslera-manga por clase

Estado más frecuente por día de la semana

Resumen por estado (día) + incoherencias simples

Variabilidad de velocidad y rumbo por tipo

Velocidad promedio y desviación (SOG) por día de la sema

Velocidades inusuales

Notebooks

Lienzos de datos

Preparaciones de datos

Preparaciones de datos compartidos

Canalizaciones

Conexiones externas

bg_menos_100

dataset_AIS_2024_silver_refined

Modelos (3)

anomaly_count_geohash9_HOURLY

anomaly_count_geohash9_hourly

anomaly_speed_vesseltypename_daily

table_AIS_2024_silver_refined

anomaly_count_geohash9_HOURLY

Actualizar

Detalles

Inferencia

Entrenamiento

Evaluación

Registro

ARIMA Plus

river-treat-468823-m3_dataset_AIS_2024_silver_refined.anomaly_count_geohash9_HOURLY

26 sept 2025, 12:51:00 a.m. UTC-6

US

Detalles del modelo

Editar

Fecha de modificación26 sept 2025, 12:51:00 a.m. UTC-6

Vencimiento del modeloNunca

Descripción-

Opciones de entrenamiento

Las opciones de entrenamiento son parámetros que se agregaran en la secuencia de comandos para crear este modelo.

Iteraciones reales1

ARIMA automáticoverdadero

Frecuencia de datosPor hora

Orden automático máximo de ARIMA2

Feature Columns

FiltroEscribir el nombre o valor de la propiedad

Nombre del campo	Tipo	Modo	Descripción
No hay resultados para mostrar			

Explorador

+ Agregar datos

Velocid... ame

Ejecutar

Guardar consulta

Descargar

Compartir

Programa

Ab

Buscar recursos de BigQuery

ML

Mostrar solo los destacados

river-treat-468823-m3

Repositorios

Consultas

Consultas compartidas

% de buques con calado > promedio (por tipo)

CaladoAnomalo

CambiosDireccion

Correlación SOG vs Draft por tipo

Eslera-manga por clase

Estado más frecuente por día de la semana

Resumen por estado (día) + incoherencias simples

Variabilidad de velocidad y rumbo por tipo

Velocidad promedio y desviación (SOG) por día de la semana

Velocidades inusuales

Notebooks

Lienzos de datos

Preparaciones de datos

Preparaciones de datos compartidos

Canalizaciones

Conexiones externas

bg_menos_100

dataset_AIS_2024_silver_refined

Velocidad promedio y desviaci...

2 WITH t AS (
3 SELECT
4 VesselTypeName,
5 FORMAT_DATE('%A', DATE(BaseBateLine)) AS dow,
6 EXTRACT(DAYOFWEEK FROM DATE(BaseBateLine)) AS dow_sun1, -- 1=Dom ... 7=Sab
7 SOG
8 FROM `dataset_AIS_2024_silver_refined.table_AIS_2024_silver_refined`
9 WHERE VesselTypeName IS NOT NULL
10)
11 SELECT
12 VesselTypeName,
13 dow,
14 AVG(SOG) AS avg_sog,
15 STDDEV_SAMP(SOG) AS sd_sog,
16 COUNT(*) AS n
17 FROM t
18 GROUP BY VesselTypeName, dow, dow_sun1
19 ORDER BY
20 VesselTypeName,
21 Se completó la consulta

Resultados de la consulta

Guardar los resultados

Información del trabajo

Resultados

Visualización

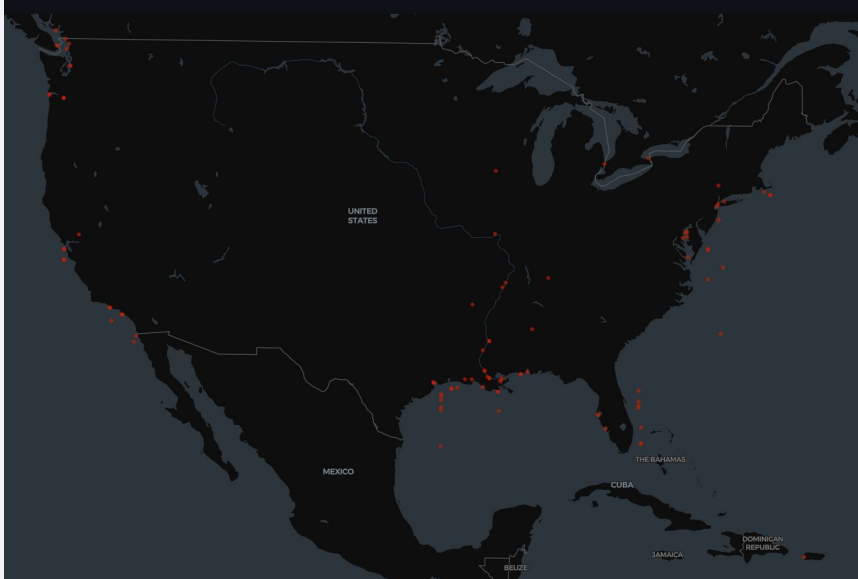
JSON

Detalles de la ejecución

Gráfico de ejecución

File	VesselTypeName	dow	avg_sog	sd_sog	n
1	Anti-pollution equipment	Monday	1.19949058473...	4.262006914665...	524716
2	Anti-pollution equipment	Tuesday	1.181959916231...	4.096016030672...	544937
3	Anti-pollution equipment	Wednesday	1.557231715034...	4.560903738736...	553118
4	Anti-pollution equipment	Thursday	1.236306457387...	4.111486055254...	530747
5	Anti-pollution equipment	Friday	0.899460874041...	3.504203500876...	490030
6	Anti-pollution equipment	Saturday	0.597358047363...	2.883965316987...	468293
7	Anti-pollution equipment	Sunday	0.509906794432...	2.580444498144...	465858
8	Cargo, Hazardous category A	Monday	8.4133384199555...	6.968534646006...	3008526
9	Cargo, Hazardous category A	Tuesday	8.375714579532...	6.948625154785...	2990797

Resultados por página: 50 1 - 50 de 31



app

Calado anómalo

Cambios de dirección

Correlación SOG vs Draft

Edora manga por clase

Resumen estado + Incohere...

Varibilidad vel y rumbo

Velocidades inusuales

Velocidad por día semana

Estado más frecuente sema...

Detección de anomalías AIS

Render the anomaly detection page.

Deploy

Detección de anomalías en tráfico marítimo

Rango de fechas (entrenamiento y predicción)

2024/09/24 - 2024/09/28

Tipos de buque

Choose options

Métrica a modelar

Número de mensajes

Frecuencia de agregación

HOURLY

Identificador de serie

geohash0

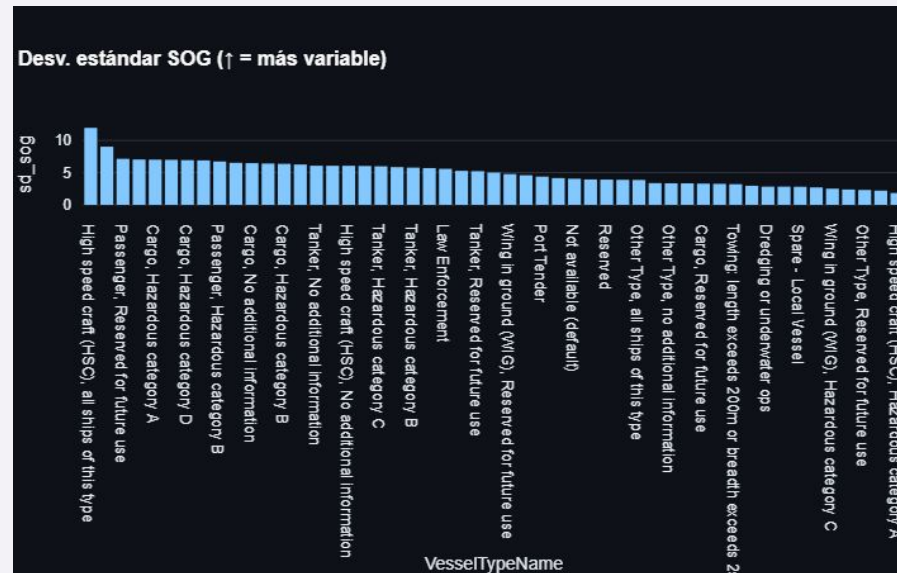
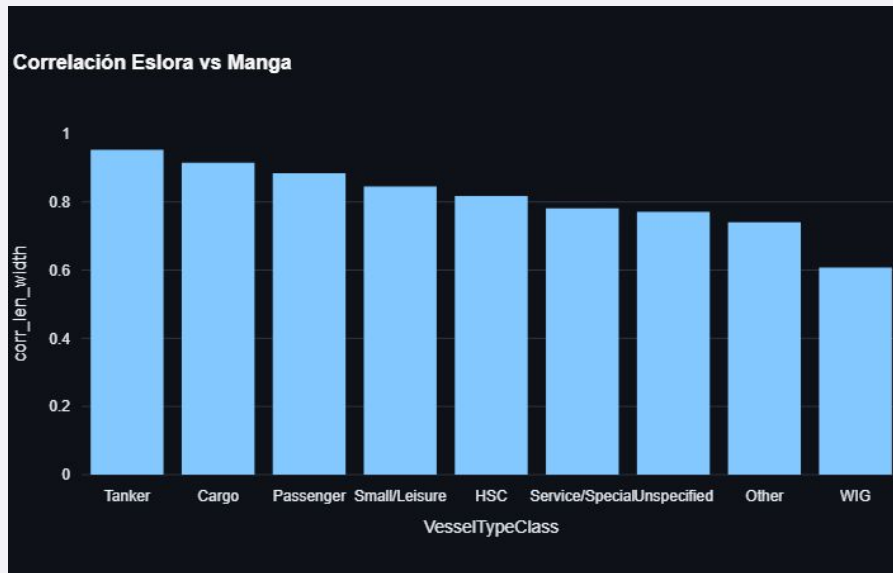
Umbral de probabilidad de anomalía

0.95

Entrenar modelo

Modelo entrenado correctamente.

Detectar anomalías





<https://github.com/enriquegomeztagle/MCD-BigData-SmartMaritimeTrafficMonitoring-FinalProject/tree/main>

```
apps/  
├── .streamlit/  
├── secrets.toml  
├── lib/  
│   ├── __pycache__/  
│   ├── __init__.py  
│   ├── bq.py  
│   ├── queries.py  
│   ├── query_utils.py  
│   └── ui.py  
├── pages/  
│   ├── 1_🗺️_Calado_anómalo.py  
│   ├── 2_🔄_Cambios_de_dirección.py  
│   ├── 3_🔗_Correlación_SOG_vs_Draft.py  
│   ├── 4_📄_Eslora_manga_por_clase.py  
│   ├── 5_📄_Resumen_estado+_incoherencias.py  
│   ├── 6_📊_Variabilidad_vel_y_rumbo.py  
│   ├── 7_📈_Velocidades_inusuales.py  
│   ├── 8_📅_Velocidad_por_día_semana.py  
│   ├── 9_📊_Estado_más_frecuente_semanal.py  
│   └── 10_👤_Detección_de_anomalías_AIS.py  
├── sql/  
│   ├── avg_speed_deviation_day_week.sql  
│   ├── calado_anomalo.sql  
│   ├── cambios_direccion.sql  
│   ├── corr_sog_draft.sql  
│   ├── eslora_manga.sql  
│   ├── incoherencias_estado.sql  
│   ├── most_frequent_state_by_week.sql  
│   ├── resumen_estado.sql  
│   ├── variabilidad_vel_rumbo.sql  
│   └── velocidades_inusuales.sql  
├── app.py  
├── requirements.txt  
└── credentials.json
```

¡Gracias!

¿Dudas?

"Eso es lo que es un barco, ¿sabes? No es sólo una quilla y un casco y una cubierta y velas, eso es lo que necesita un barco. Pero lo que un barco es, en realidad, es libertad"

- Jack Sparrow

