

Análisis y Predicción de Tsunamis: Un Enfoque Educativo

Isabel Bravo Quiñones

Data Scientist

Ingeniero Geólogo



Comprendiendo los Tsunamis

- Los tsunamis son fenómenos naturales de gran escala y poder destructivo.
- Su ocurrencia, a menudo impredecible, ha causado tragedias históricas.
- La ciencia y la tecnología ofrecen nuevas formas de entenderlos.
- El Machine Learning permite detectar **patrones en datos históricos**.
- Este proyecto, con fines **didácticos**, muestra el potencial de la IA en la prevención de desastres.

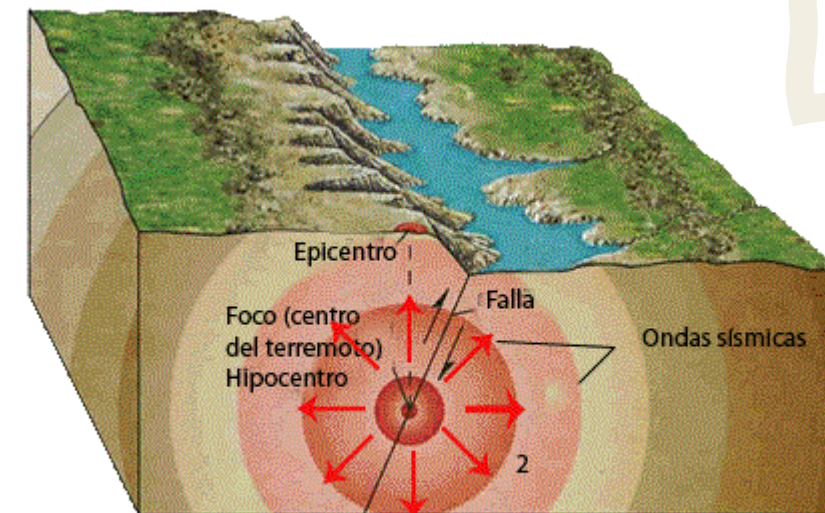
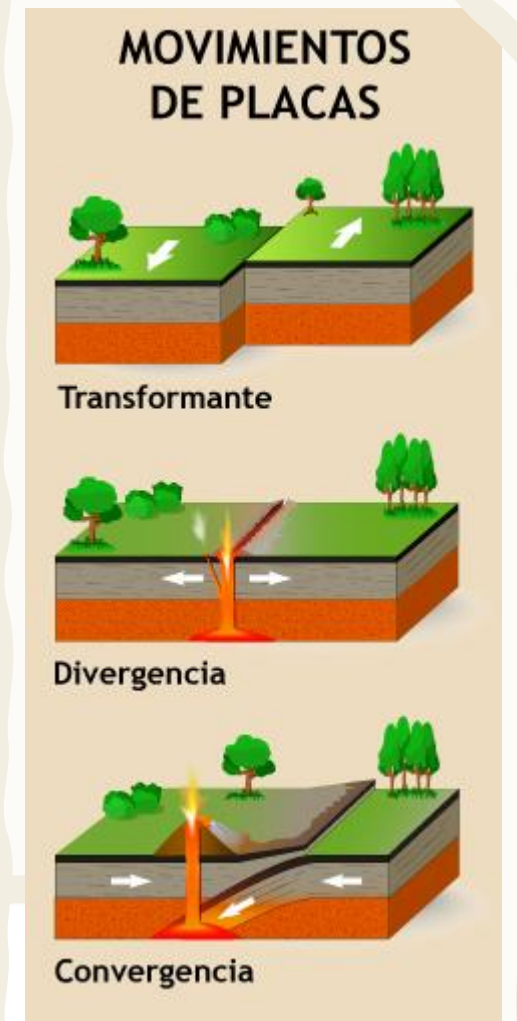
Fecha	Magnitud Mw terremoto	Profundidad del foco (hipocentro)	Tsunami resultante
31 Agosto 2025	6.0 (Terrete)	8 km	No
30 Julio 2025	8.8 (Terrestre)	21 m	Rusia (6 m) Japón (40 cm) Hawaii, Chile
2011	9.1 (Marino)	30 m	Japón (40 m) Fukushima Taiwan, Alaska, Chile
2004	9.0 (Marino)	30 km	Indonesia (30 m) Tailandia (20 m)

El Origen: Terremotos

Un terremoto es la vibración brusca y pasajera de la corteza terrestre, producida por la liberación de energía acumulada en forma de ondas sísmicas. No todos los terremotos son iguales:

- **Tectónicos:** Los más comunes y potentes, causados por el movimiento de las placas tectónicas.
- **Volcánicos:** Asociados a la actividad volcánica.
- **De colapso:** Provocados por el derrumbe de cavernas subterráneas o minas.

La energía liberada se mide con la **Escala de Magnitud de Momento Sísmico** (**M_w** (mide la energía liberada), mientras que la **intensidad (Mercalli)** describe cómo se perciben sus efectos en la superficie.



ESQUEMA TERREMOTO

Qué es un Tsunami

Olas Gigantes

Un tsunami no es una ola común; es una serie de **olas oceánicas gigantes** generadas por desplazamientos masivos de agua. La causa más frecuente son los terremotos submarinos, pero también pueden ser provocados por erupciones volcánicas, deslizamientos de tierra o incluso impactos de meteoritos.

Tipos de Tsunamis

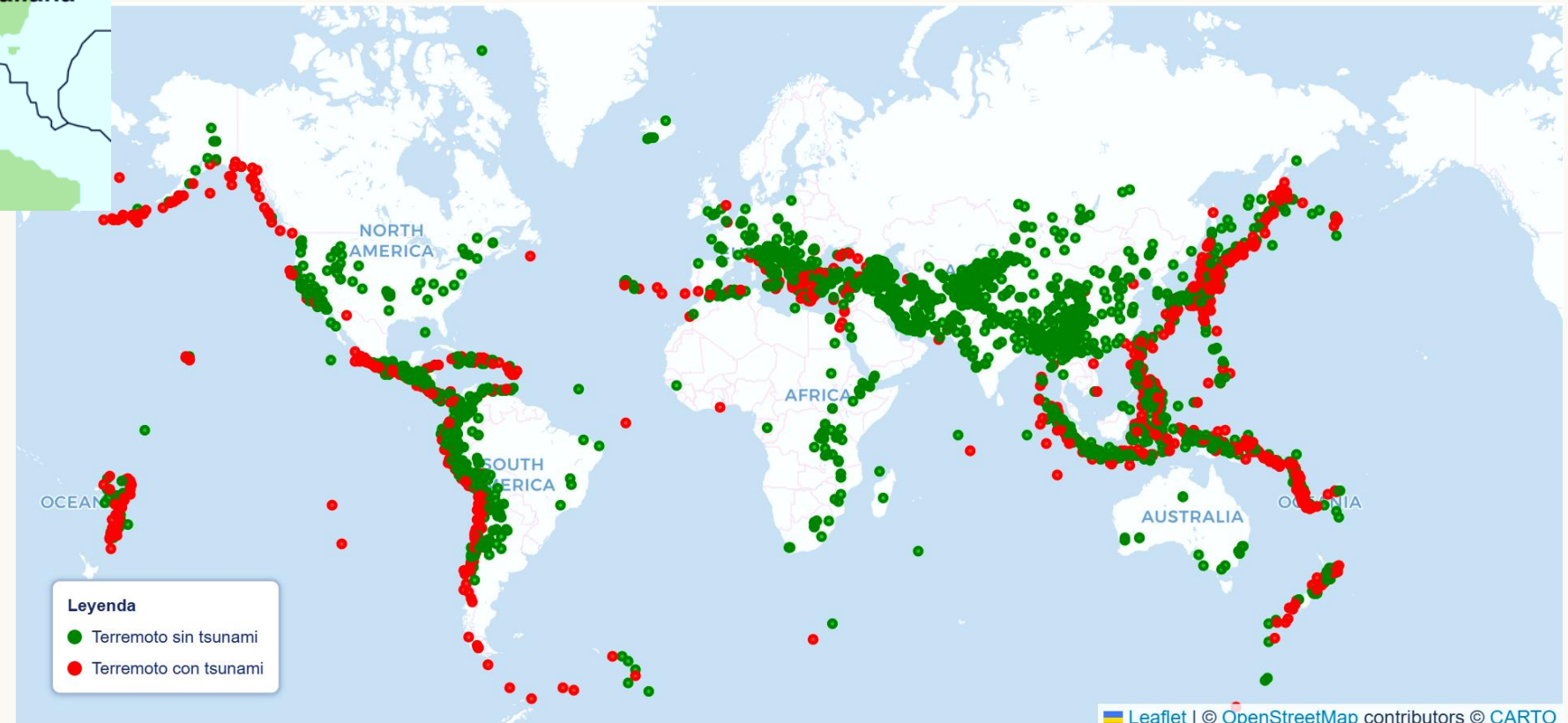
Locales: Impactan zonas costeras cercanas a su origen, con poco tiempo de reacción (minutos).

Lejanos: Pueden viajar miles de kilómetros a través del océano a velocidades de avión, afectando costas distantes horas después del evento original.

El Desafío de la Predicción

La predicción de tsunamis en tiempo real es extremadamente compleja y requiere una red sofisticada de sensores oceánicos y geofísicos. Nuestro proyecto, aunque didáctico, busca simplificar este proceso para entender sus principios fundamentales.

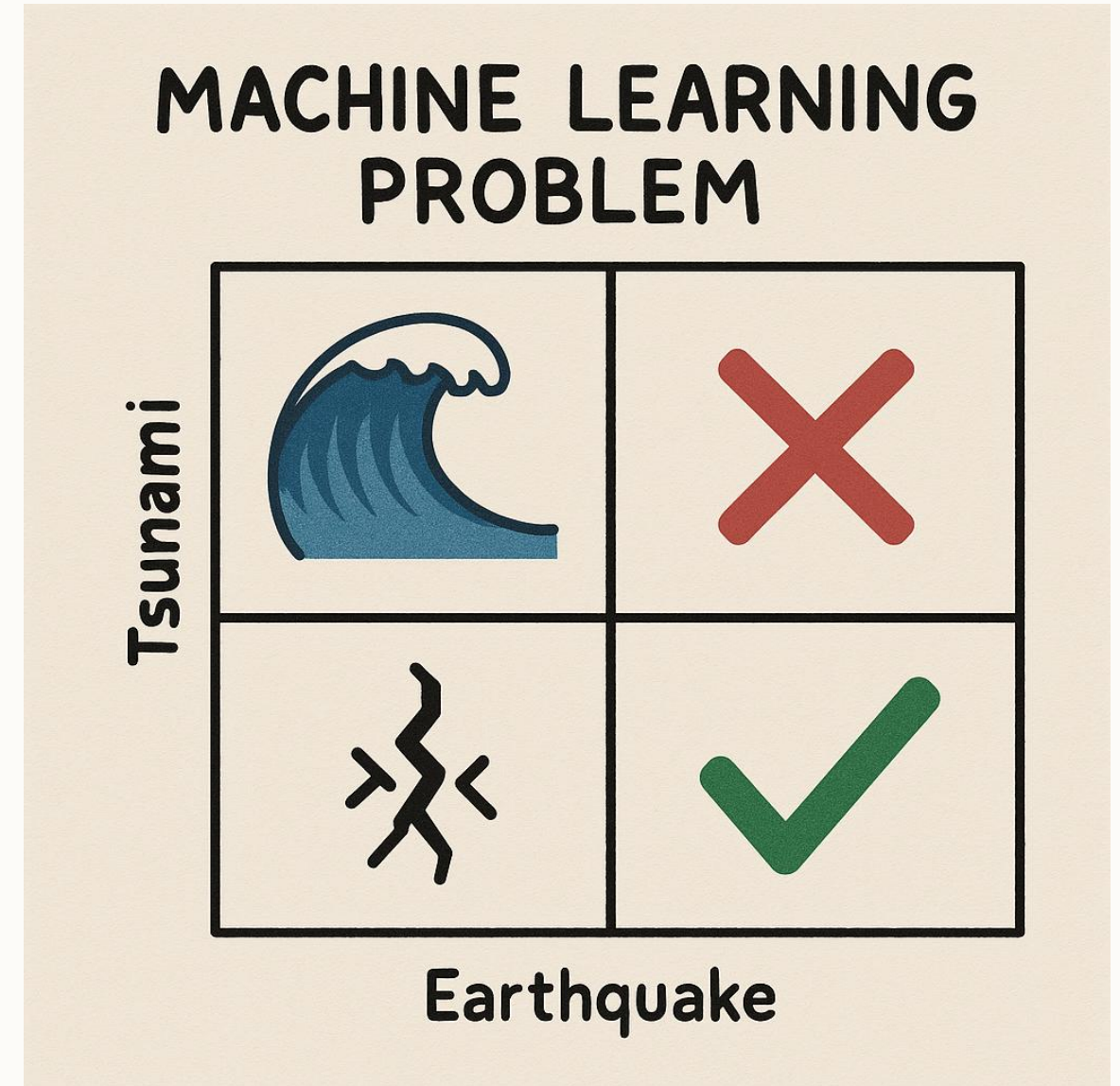
Placas Tectónicas



Nuestro Objetivo: Un Modelo Didáctico de Clasificación

El propósito central de este proyecto es:

- Construir un modelo sencillo de Machine Learning.
- Detectar patrones entre características sísmicas y tsunamis.
- Enfoque **didáctico y divulgativo**.
- No diseñado para predicciones reales.
- La predicción real requiere datos y modelos mucho más complejos.
- Detección de patrones entre **sismos y tsunamis** y la predicción de si un determinado terremoto producirá un tsunami.



Datos y Preparación

Fuente de Datos	La base de nuestro análisis son los datos históricos abiertos de la NOAA (NCEI) , una fuente fiable de información sobre eventos sísmicos y tsunamis a nivel mundial. (API tsunamis y API terremotos)
Variables Clave	Hemos considerado variables como la magnitud (Mw) , profundidad , ubicación (latitud/longitud) , año y región del evento sísmico. La variable objetivo es binaria: oceanicTsunami (0 = no tsunami, 1 = tsunami).
Limpieza y Unificación	Se realizó una limpieza rigurosa, eliminando registros incompletos y unificando las magnitudes a la escala Mw para garantizar la consistencia de los datos.

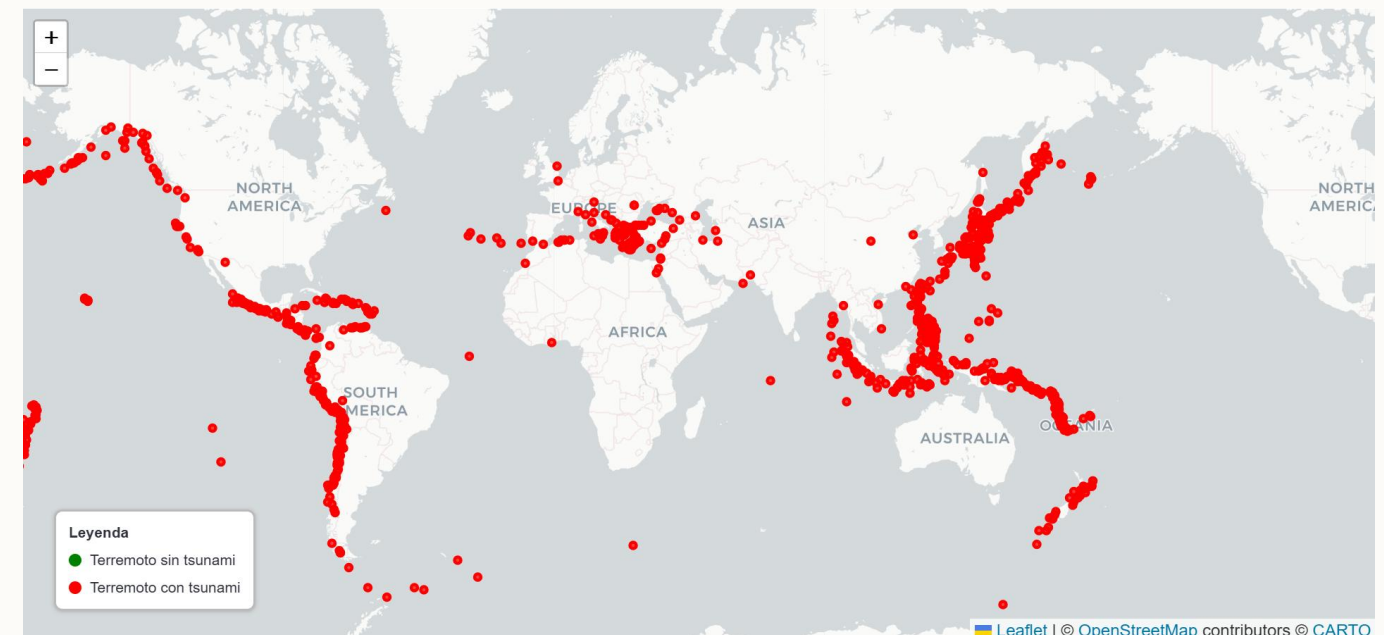
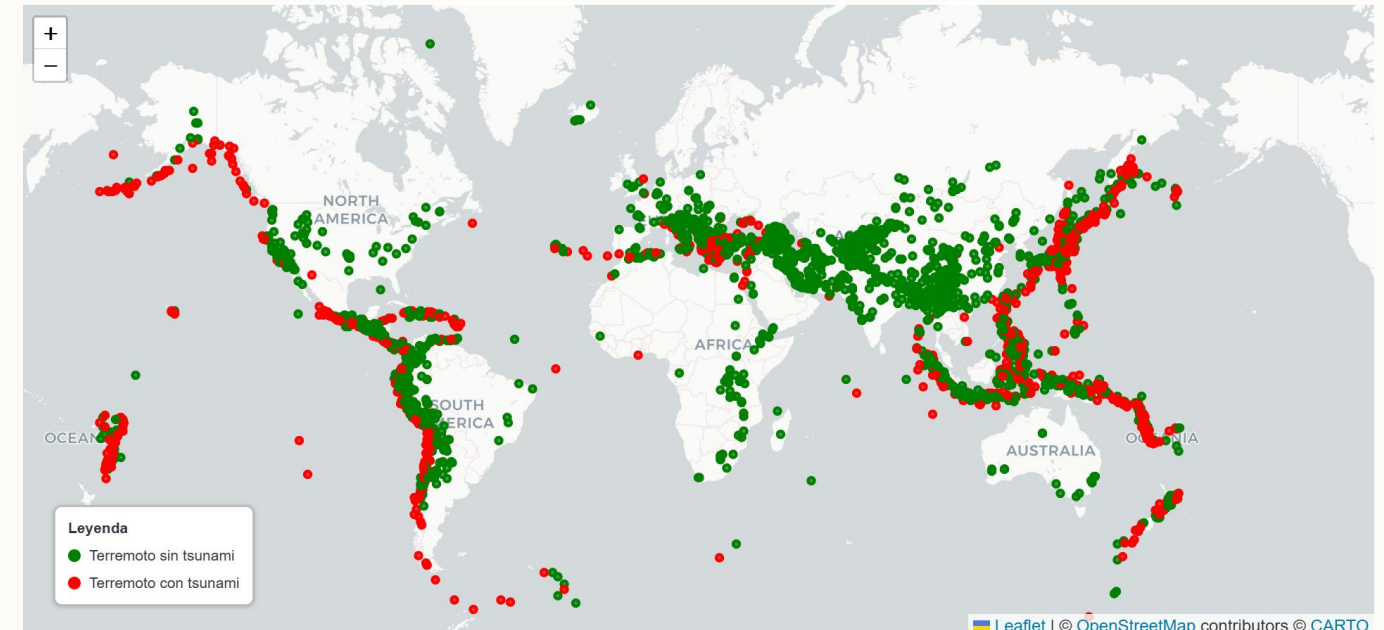
Exploración de Datos: Un Desequilibrio Crucial

Un hallazgo fundamental en la exploración de datos es el **desequilibrio significativo** en el dataset:

- Aproximadamente el **70%** de los terremotos **no generan tsunamis**.
- Solo el **30%** sí los generan.

Este desequilibrio es vital, ya que un modelo que simplemente prediga "no tsunami" en todos los casos obtendría una alta precisión pero sería inútil. Por ello, nos centramos en métricas más adecuadas:

- **F1 Score:** Equilibra precisión y recall.
- **Recall:** Esencial para **evitar falsos negativos** (detectar la mayor cantidad posible de tsunamis reales).



Modelado:



Algoritmos Probados

- Regresión Logística
- Random Forest
- XGBoost
- **LightGBM**
- CatBoost
- SVM
- kNN

*DBSCAN (no útil).



Balanceo de Clases

Para abordar el desequilibrio entre clases ("tsunami" vs. "no tsunami"), aplicamos la técnica **SMOTE (Synthetic Minority Over-sampling Technique)**, que crea ejemplos sintéticos de la clase minoritaria.



Optimización y Selección

Utilizamos **GridSearch** para optimizar los hiperparámetros de cada modelo. La selección final se basó en el modelo con el mejor **F1 Score** (capacidad para identificar tsunamis reales) y un **recall** competitivo.

Resultados

Resultado de los diferentes modelos evaluados

Métricas Obtenidas

Tras el entrenamiento y validación, los resultados fueron :

- Un **Recall alto**, indicando una buena capacidad para detectar los casos positivos (tsunamis reales).
- Un **F1 Score competitivo**, mostrando un buen equilibrio general.

Esto significa que nuestro modelo puede **identificar una gran proporción de tsunamis reales** basándose en datos sísmicos históricos.

--- Resumen de resultados (media CV) ---

Modelo	f1_mean	recall_mean
LightGBM	0.765322	0.798553
CatBoost	0.759346	0.820003
XGBoost	0.756302	0.774743
RandomForest	0.739660	0.771143
SVM	0.708525	0.839072
LogisticRegression	0.707135	0.840298
KNN	0.685428	0.785458

Resultado de los scores en test del mejor modelo (LightGBM):

F1 Score: 0,73%

Recall: 0,75%

Demo App

1) Datos

Sube tu CSV procesado

Drag and drop file here

Limit 200MB per file • CSV

Browse files



demo.csv

162.1KB



...o indica una ruta local (opcional)

CSV cargado correctamente (subida).

2) Filtros

Vista

Todos



Rango de año

-250

202

3) (Opcional) Modelo

Carga tu modelo .pkl (pipeline)

Drag and drop file here

Limit 200MB per file • PKL

Browse files



final_mod...

430.3KB



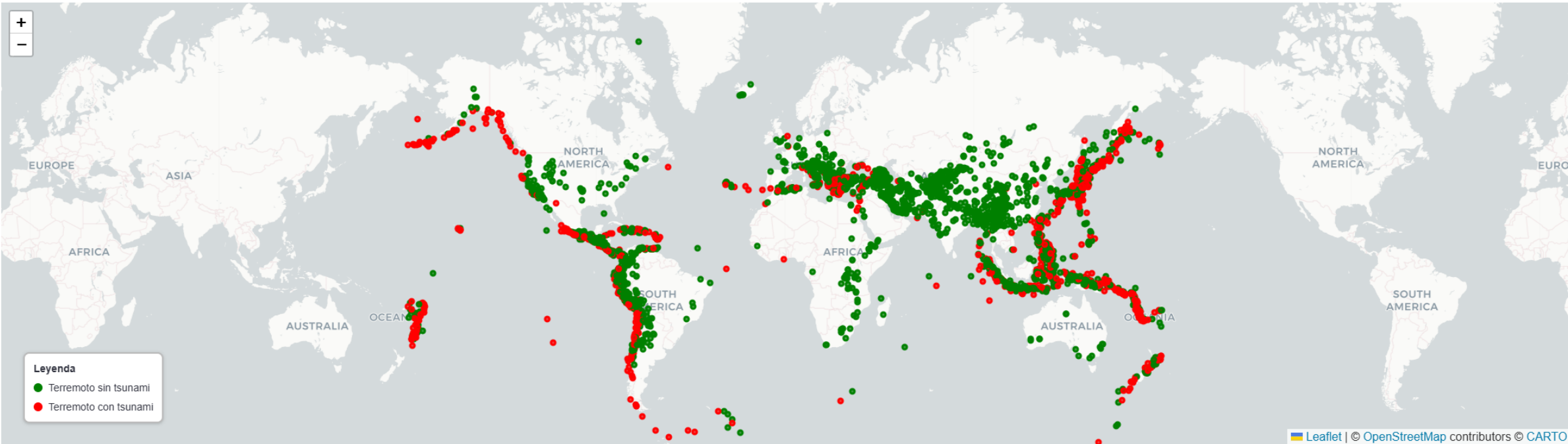
...o ruta local al .pkl

Predicción y visualización de tsunamis a partir de terremotos

Carga tu dataset para visualizar terremotos y tsunamis. Opcionalmente, carga tu modelo para predicciones puntuales.

Faltaban columnas **opcionales** y se han creado por defecto: ['locationName']

Mapa histórico



Predicción puntual (opcional)

Magnitud (Mw)

6,50

-

+

Profundidad (km)

20,00

-

+

Latitud

0,00

Longitud

0,00

-

+

Intensidad (opcional)

País

AFGHANISTAN

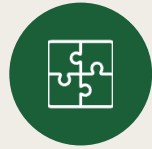
regionCode_eq

0

Año (opcional)

Prededir tsunami

Impacto y Conclusiones



ML: Una Pieza del Puzzle

El Machine Learning es una herramienta poderosa que puede encontrar patrones complejos en los datos. Sin embargo, para la predicción de tsunamis reales, **no es suficiente**. Requiere integrar datos geológicos y oceanográficos entre otros.



Valor Educativo y Divulgativo

El verdadero valor de este proyecto reside en su función **educativa y divulgativa**. Demuestra cómo la ciencia de datos puede aplicarse a problemas del mundo real y la importancia del análisis crítico de los datos y resultados.



El Camino Hacia Adelante

El futuro pasa por la **integración de más fuentes de datos** y el desarrollo de modelos más complejos, incluyendo el aprendizaje profundo, para acercarnos a sistemas de alerta temprana de tsunamis más fiables.

¡GRACIAS!