

Diseño, Desarrollo y Evaluación de Modelos de Predicción de Series Temporales Aplicados al Monitoreo Geotécnico y Estructural

TFG

Juan Marín Fidalgo
Grado de Ingeniería Informática
Área de Inteligencia Artificial

Profesores:

Dr. David Isern Alarcón

Dr. Friman Sánchez Castaño

Contenido

- Objetivos
- Introducción
- Modelos estadísticos
- Modelos de aprendizaje computacional
- Aplicación web
- Comparativa y conclusiones
- Trabajos futuros

Objetivos

Objetivos principales

- Investigar el fundamento teórico de los principales modelos.
- Diseñar su implementación en series temporales de monitoreo.
- Implementar los modelos en distintos tipos de sensores de monitoreo.
- Evaluar el rendimiento de las predicciones de los distintos modelos.

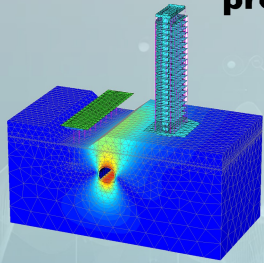
Objetivos

Objetivos secundarios

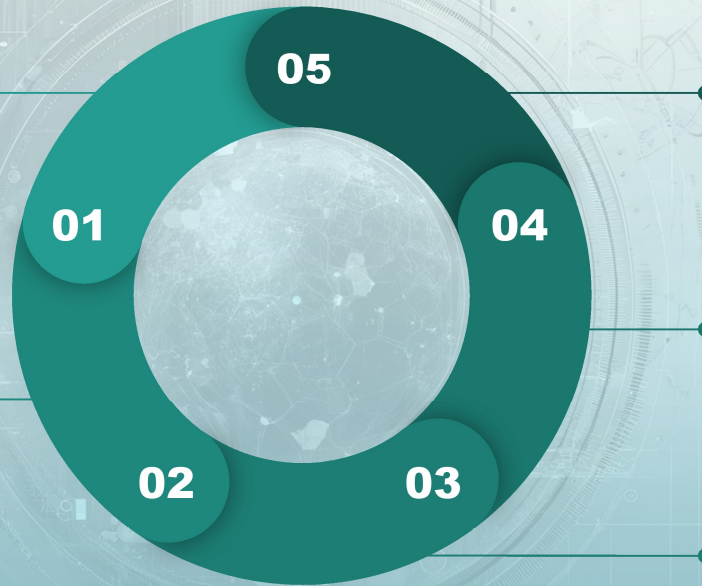
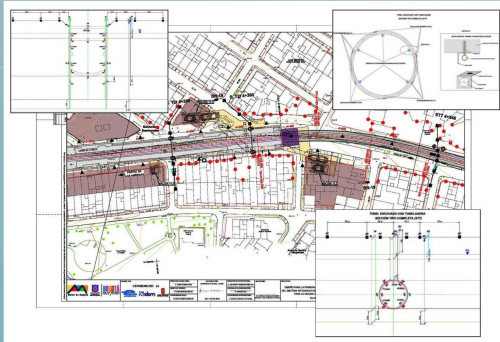
- Implementar una aplicación web que integre distintas tecnologías para visualizar los resultados.
- Comparar el rendimiento de los distintos tipos de modelos.
- Investigar nuevas tendencias.

Introducción

Modelo numérico del proyecto



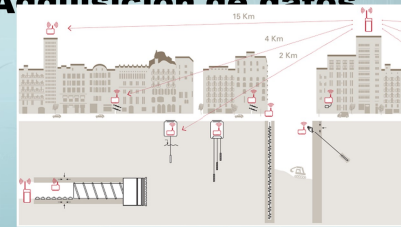
Plan de auscultación



Análisis



Adquisición de datos



Instrumentación



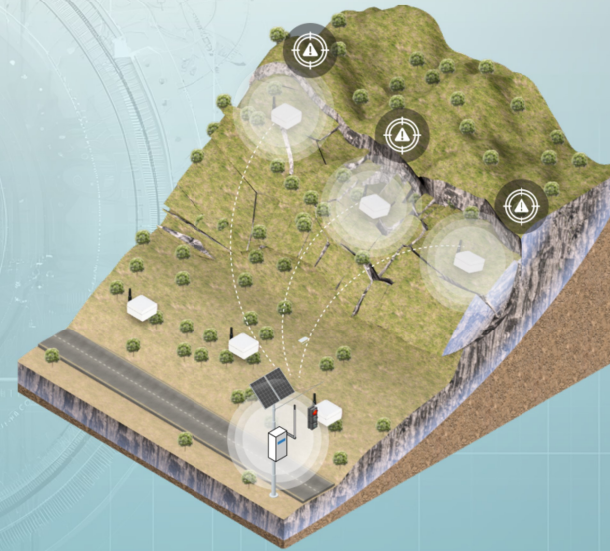
¿Qué es el monitoreo?

Introducción

Monitoreo estructural

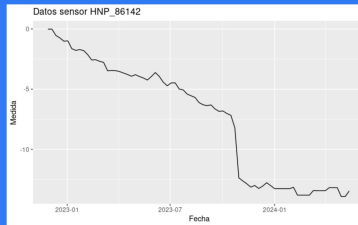


Monitoreo geotécnico

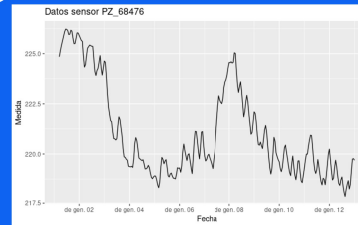


Introducción

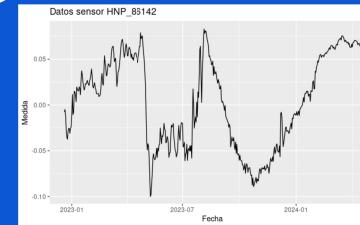
Series no estacionarias



Estacionalidades complejas



Correlaciones no lineales



Características distintivas de las series temporales de monitoreo

Modelos estadísticos

Suavizado exponencial

Componente error

Componente tendencia

Componente estacional

```
ETS(Data ~ error(method = c("A", "M")) +  
      trend(method = c("N", "A", "M")) +  
      season(method = c("N", "A", "M"))  
)
```


Modelos estadísticos

ARIMA

AR

Modelo autorregresivo

I

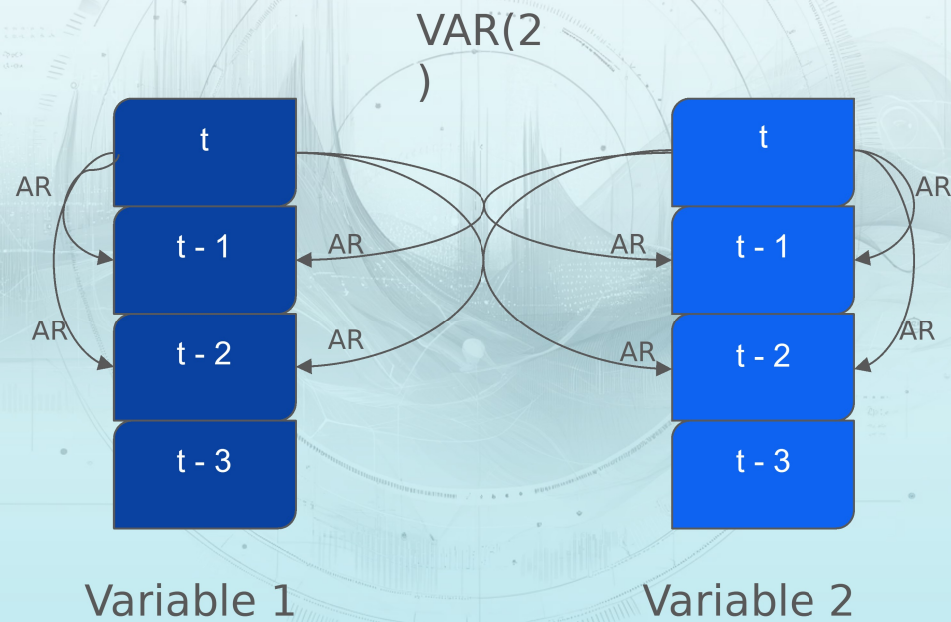
integrado

MA

de medias móviles

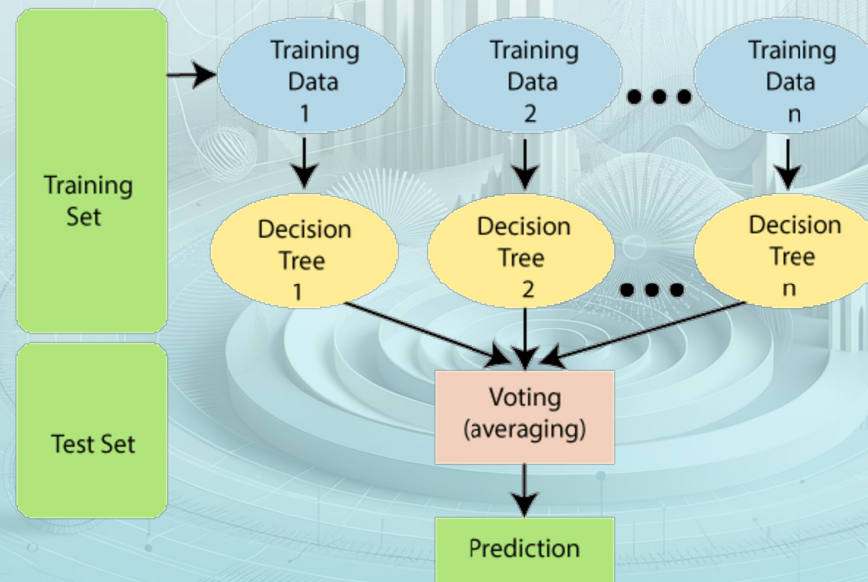
Modelos estadísticos

VAR



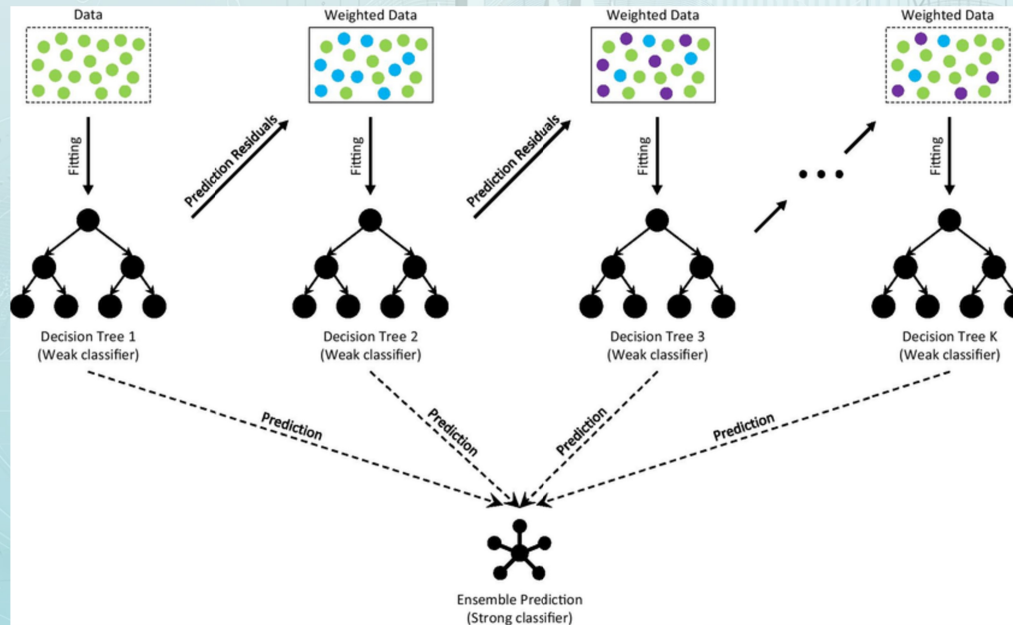
Modelos basados en aprendizaje computacional

Bosques aleatorios



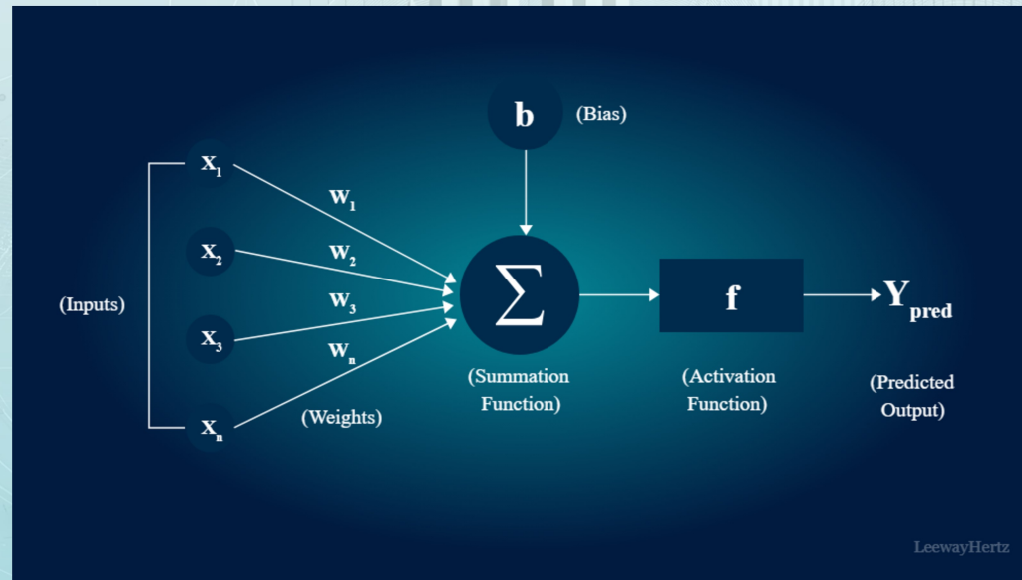
Modelos basados en aprendizaje computacional

Gradient boosting



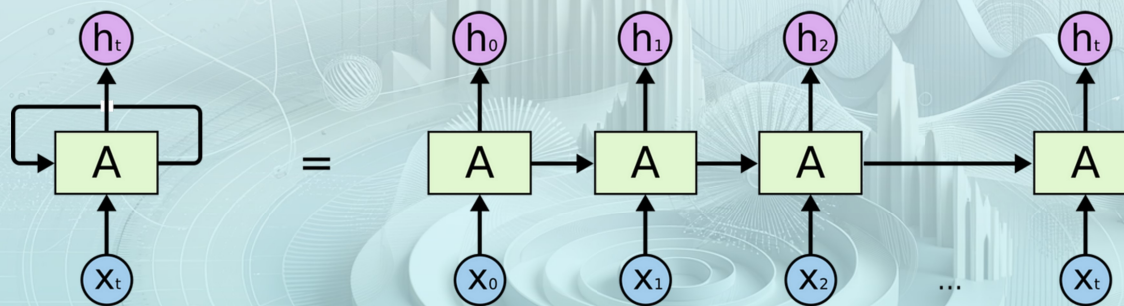
Modelos basados en aprendizaje computacional

Redes neuronales autorregresivas NNAR



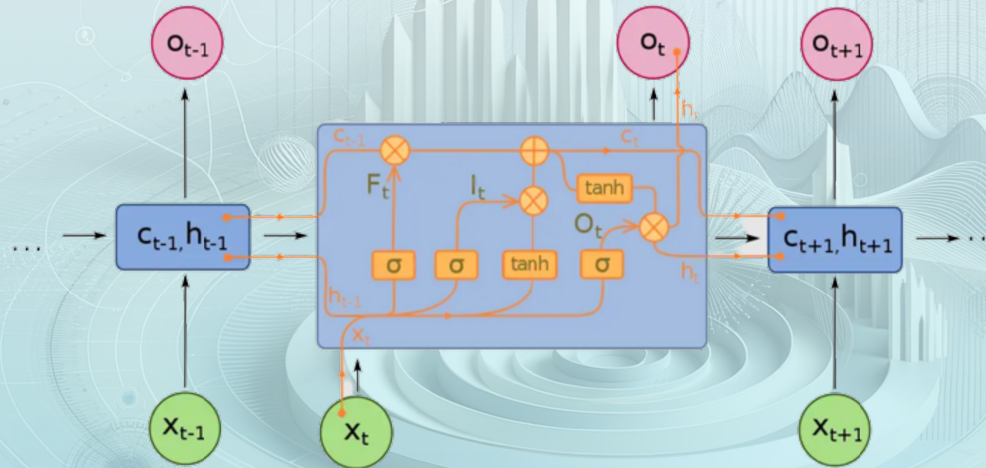
Modelos basados en aprendizaje computacional

Redes neuronales recurrentes RNN



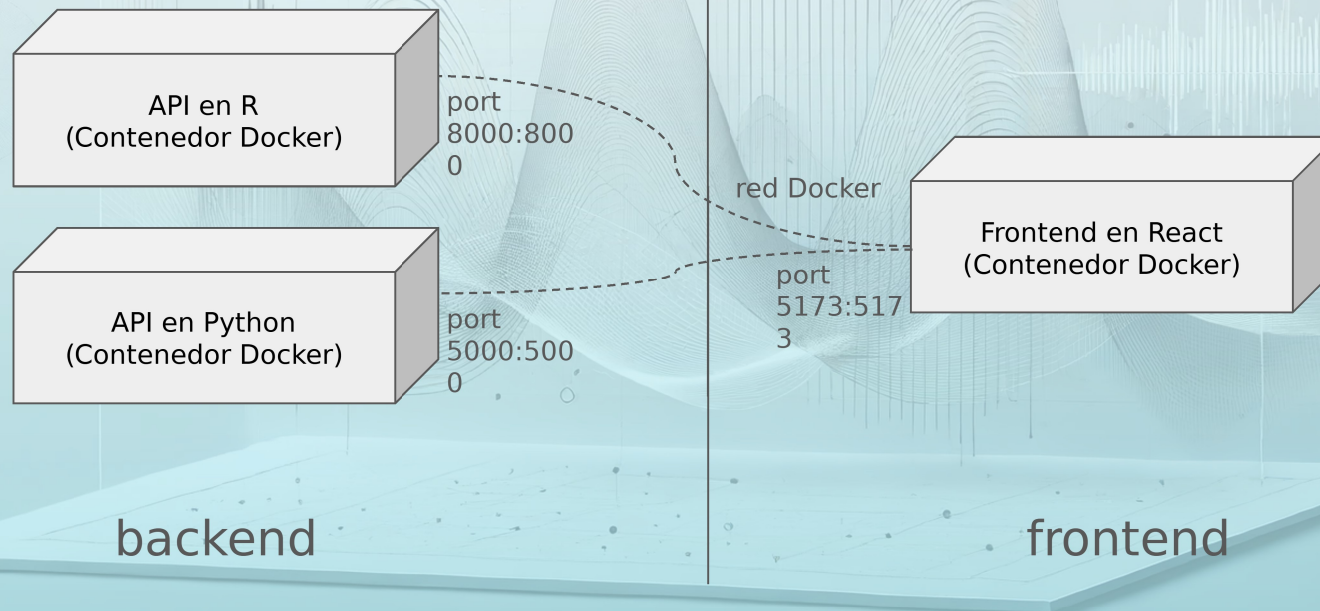
Modelos basados en aprendizaje computacional

Redes neuronales con memoria de largo y corto plazo LSTM



La aplicación web

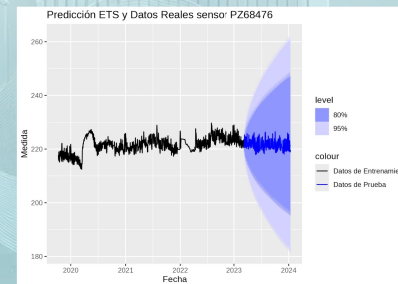
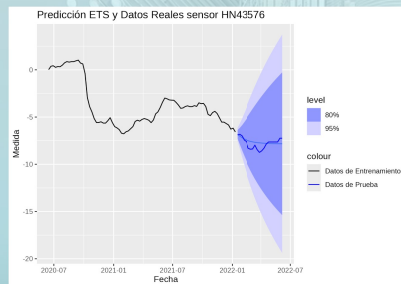
<https://github.com/jmarinfi/tfg-ts-jmf>



Conclusiones

Mejores resultados ETS

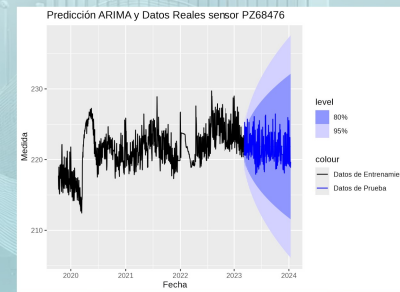
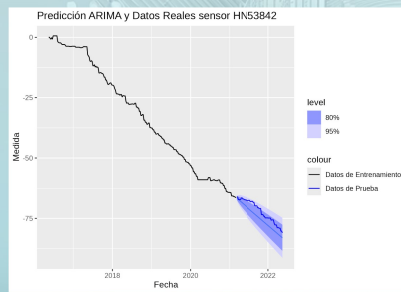
Sensor	Modelo	ME	RMSE	MAE	MPE	MAPE
HN43576	ETS	-0.182	0.511	0.412	1.98	5.12
PZ68476	ETS	-0.368	1.72	1.43	-0.172	0.643



Conclusiones

Mejores resultados ARIMA

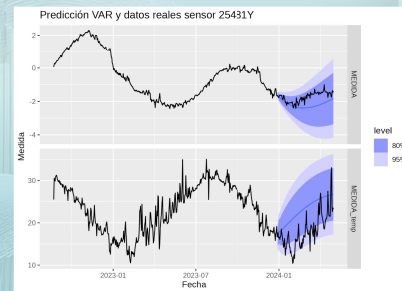
Sensor	Modelo	ME	RMSE	MAE	MPE	MAPE
HN53842	ARIMA	2.88	3.08	2.88	-4.00	4.00
PZ68476	ARIMA	-0.460	1.89	1.60	-0.215	0.723



Conclusiones

Mejores resultados VAR

Sensor	Modelo	ME	RMSE	MAE	MPE	MAPE
25481Y	VAR	0.427	0.525	0.440	-27.4	28.0



Conclusiones

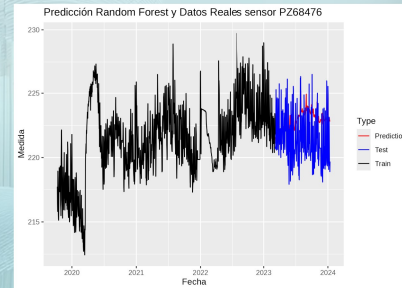
Modelos estadísticos

- ❖ Buen rendimiento con estacionalidades diarias o semanales, y con tendencias lineales. Estos casos son poco comunes en el monitoreo.
- ❖ Mal rendimiento con estacionalidades anuales y relaciones no lineales. Estos casos son los más comunes en el monitoreo.
- ❖ Mejor rendimiento en sensores geotécnicos que estructurales.
- ❖ Pueden resultar útiles en caso de predicciones a corto plazo en series con estacionalidades y tendencias regulares.

Conclusiones

Mejores resultados Random Forest

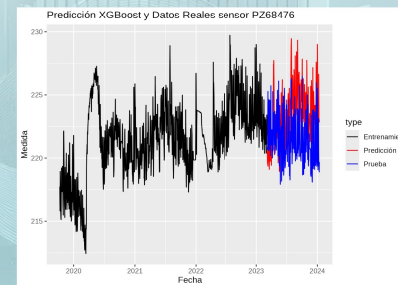
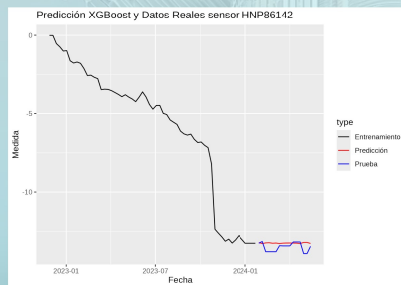
Sensor	Modelo	ME	RMSE	MAE	MPE	MAPE
PZ68476	RF	-1.472649	2.436916	2.096979	-0.671946	0.949856



Conclusiones

Mejores resultados Gradient Boosting

Sensor	Modelo	ME	RMSE	MAE	MPE	MAPE
HNP86142	GB	-0.26072	0.39015	0.30810	1.887008	2.246678
PZ68476	GB	-2.04306	3.25785	2.77410	-0.928171	1.254884



Conclusiones

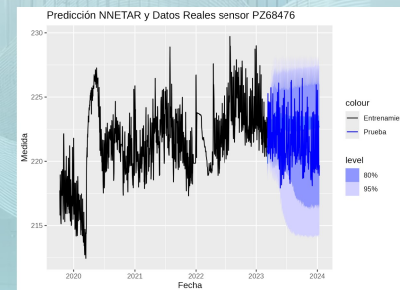
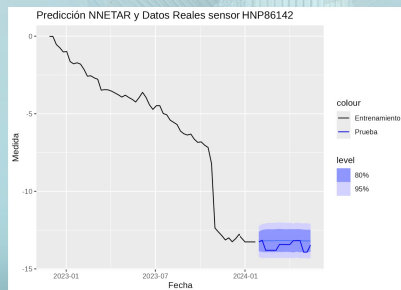
Modelos de ML basados en árboles de decisión

- ❖ Gradient Boosting supera ligeramente en rendimiento a Random Forest.
- ❖ Los únicos casos que muestran buen rendimiento son con disponibilidad de gran cantidad de datos (caso poco común en monitoreo) o con ausencia de tendencia (caso poco relevante en monitoreo).
- ❖ Modelos sencillos de implementar pero con poca aplicación práctica en monitoreo.
- ❖ Pueden resultar útiles en caso de sensores automatizados con alta frecuencia de muestreo y, por lo tanto, en series con gran cantidad de datos.

Conclusiones

Mejores resultados NNAR

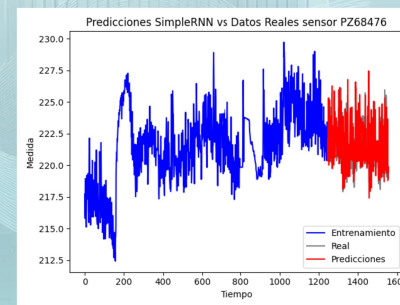
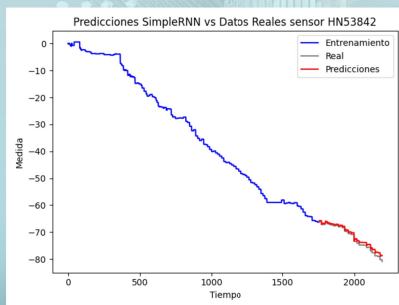
Sensor	Modelo	ME	RMSE	MAE	MPE	MAPE
HNP86142	NNAR	-0.324	0.427	0.332	2.36	2.42
PZ68476	NNAR	-1.36	2.27	1.99	-0.622	0.900



Conclusiones

Mejores resultados RNN

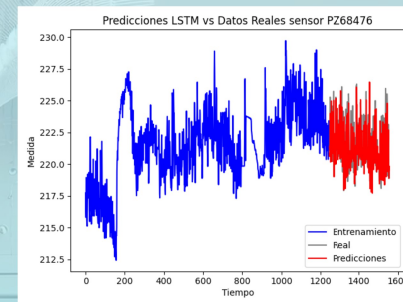
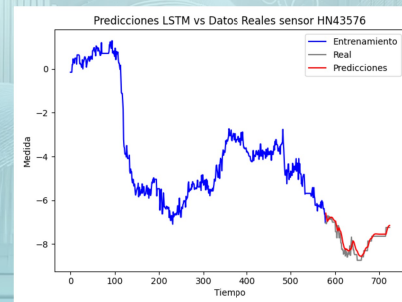
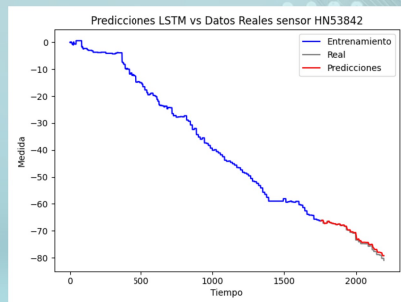
Sensor	Modelo	ME	RMSE	MAE	MPE	MAPE
HN53842	RNN	-0.94644	1.0404	0.94714	1.29143	1.29249
PZ68476	RNN	-0.08906	1.49458	1.14509	-0.04255	0.51599



Conclusiones

Mejores resultados LSTM

Sensor	Modelo	ME	RMSE	MAE	MPE	MAPE
HN53842	LSTM	-0.35303	0.59076	0.40854	0.46156	0.5445
HN43576	LSTM	-0.12443	0.22939	0.18188	1.52506	2.29599
PZ68476	LSTM	0.26396	1.50068	1.11164	0.1162	0.50044



Conclusiones

Modelos de ML basados en redes neuronales

- ❖ La implementación de NNAR con una sola capa, cuyos inputs son un número de rezagos de la serie resulta en un rendimiento muy similar a los modelos estadísticos
- ❖ El rendimiento de las predicciones con redes neuronales RNN y LSTM es muy superior al resto de modelos, destacando también los sensores geotécnicos, en especial los que tienen gran cantidad de datos.
- ❖ La implementación con la librería TensorFlow de Python resulta sencilla para cualquier escenario de monitoreo.
- ❖ Pueden resultar útiles en cualquier caso, ya que logran capturar dependencias no lineales y estacionalidades complejas, aunque haya relativamente poca disponibilidad de datos históricos.

Conclusiones generales

- ❖ Los modelos que han mostrado mejor rendimiento son los basados en redes neuronales RNN y LSTM.
- ❖ La tipología de sensor que mejores resultados ha ofrecido son los geotécnicos, los cuales presentan estacionalidades y tendencias más regulares.
- ❖ Los modelos estadísticos sólo resultan útiles en predicciones a corto plazo de series con estacionalidades cortas (diarias o semanales), o con tendencias lineales.
- ❖ Los modelos basados en árboles de decisión sólo resultan útiles cuando se dispone de gran cantidad de datos históricos o en series sin tendencias.

Trabajos futuros

- ❖ Explorar el rendimiento de los modelos de aprendizaje computacional entrenando los modelos con otras variables predictoras.
- ❖ Explorar el rendimiento de nuevos modelos de redes neuronales tales como redes convolucionales o transformers.
- ❖ Experimentar la implementación de modelos híbridos y evaluar su rendimiento.
- ❖ Evolucionar la implementación de la aplicación web para permitir que se pueda comunicar con una base de datos, o que se pueda alimentar con datos externos.

