

MACHINE LEARNING PARA LA PREDICCIÓN DE SERIES TEMPORALES

Joaquín Amat Rodrigo
Javier Escobar Ortiz



BBVA

AI Factory

CONTADOR DE GENTE QUE CONOCE SKFORECAST



¿QUIÉNES SON ESTOS MATAOS?



Joaquín Amat Rodrigo

Senior Data Scientist @Veeva Systems

LinkedIn 

j.amatrodrigo@gmail.com



Javier Escobar Ortiz

Senior Data Scientist @IKEA

 LinkedIn

javier.escobar.ortiz@gmail.com



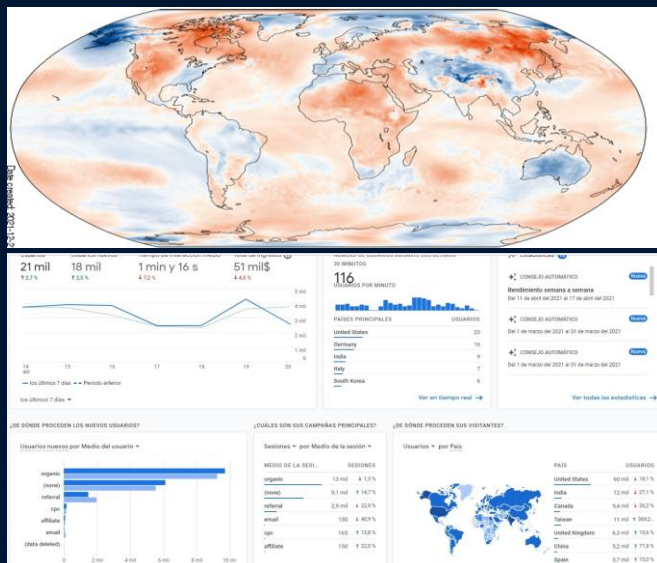
ÍNDICE

- Series temporales y Forecasting
- Estrategias de Forecasting
 - Modelos *multi-step* recursivos
 - Modelos *multi-step* directos
- ¿Por qué skforecast?
- Validación y Optimización de modelos de forecasting
- Feature selection
- Forecasting multi-series
- Forecasting probabilístico
- Explicabilidad
- Material adicional
- Estado del arte
- Queremos escucharos

¿QUÉ ES UNA SERIE TEMPORAL?

SERIES TEMPORALES

Una serie temporal (*time series*) es una sucesión de datos ordenados cronológicamente y espaciados a intervalos iguales o desiguales.



FORECASTING

El proceso de *forecasting* consiste en predecir el valor futuro de una serie temporal, bien **modelando** dicha serie únicamente en función de su **comportamiento pasado** (autorregresivo), o bien empleando otras **variables externas**.



Para crear un modelo de *forecasting*, se utilizan datos históricos con el objetivo de obtener una representación matemática capaz de predecir valores futuros. Esta idea se basa en la premisa de que **el comportamiento futuro de un fenómeno puede explicarse a partir de su comportamiento pasado**.



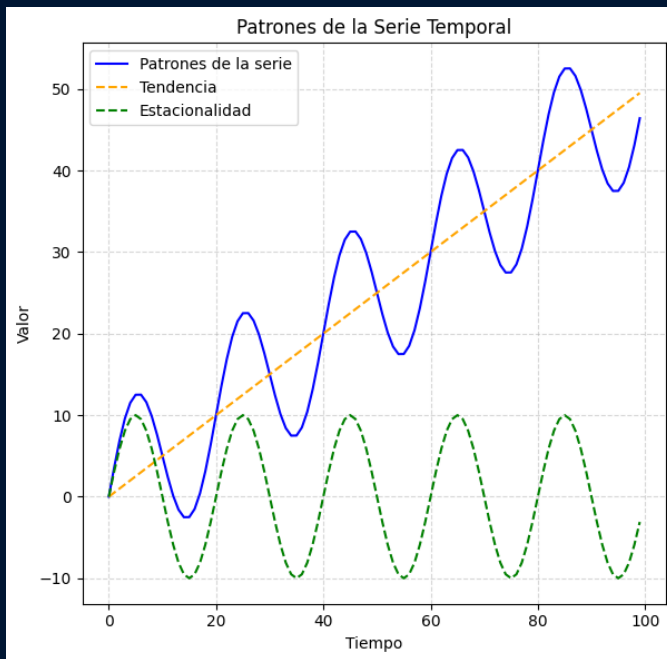
Esto raramente ocurre en la realidad o, al menos, no en su totalidad.

COMPONENTES DE UN FORECAST

Forecast = patrones de la serie + varianza no explicada

COMPONENTES DE UN FORECAST

Forecast = patrones de la serie + **varianza no explicada**

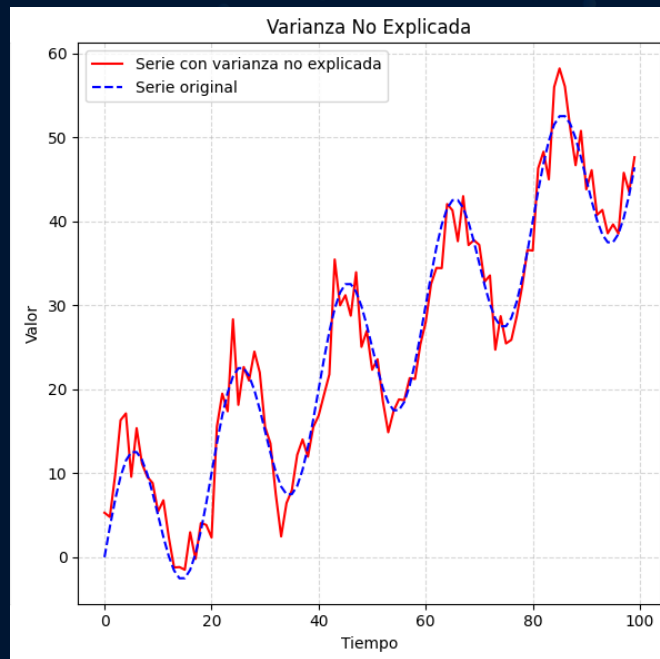


El primer término de la ecuación se refiere a todo aquello que tiene un **carácter repetitivo a lo largo del tiempo** (tendencia, estacionalidad, factores cíclicos, etc.). Se conoce como **componente autoregresivo**.

COMPONENTES DE UN FORECAST

Forecast = patrones de la serie + varianza no explicada

El segundo término representa todo aquello que **influye** en la serie **pero** que **no** está **recogido** (explicado) en sus valores pasados.



COMPONENTES DE UN FORECAST

Forecast = patrones de la serie + varianza no explicada

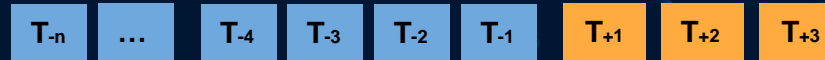
Cuanto mayor sea la importancia del **primer término** respecto al **segundo**, mayor será la probabilidad de **éxito** a la hora de crear modelos de **predicción autorregresivos**.



A medida que el **segundo** término adquiere relevancia, se hace necesario incorporar al modelo **variables adicionales** (si existen), que **ayuden a explicar** el comportamiento observado.

Realizar un buen estudio del fenómeno que se pretende modelar y saber reconocer en qué medida su comportamiento puede explicarse gracias a su pasado, **puede ahorrar muchos esfuerzos innecesarios**.

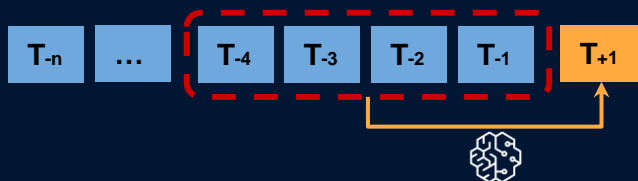
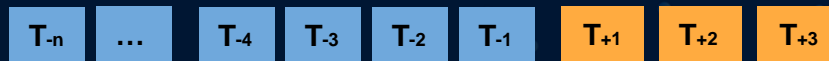
ESTRATEGIAS DE FORECASTING

PREDICCIÓN MULTI-STEP RECURSIVA







 Valor observado
 Valor predicho

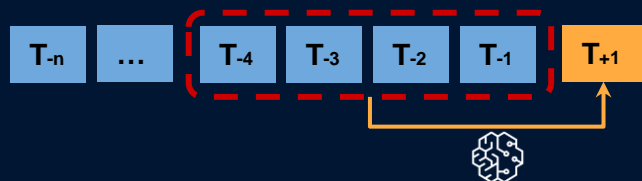
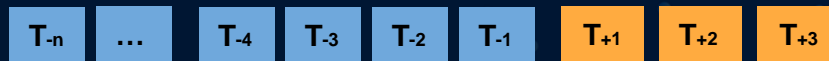
PREDICCIÓN MULTI-STEP RECURSIVA



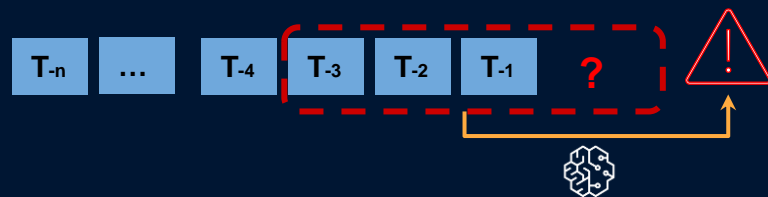
Predicción step 1

-  Valor observado
-  Valor predicho
-  Predictores
-  Modelo para step+1

PREDICCIÓN MULTI-STEP RECURSIVA



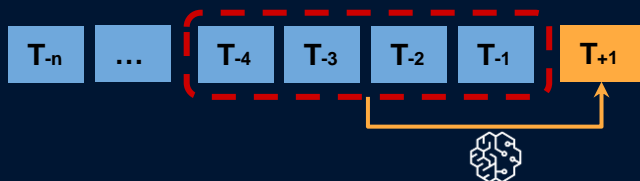
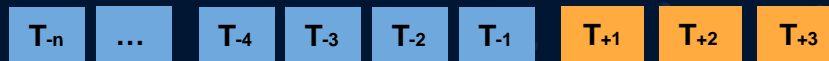
Predicción step 1



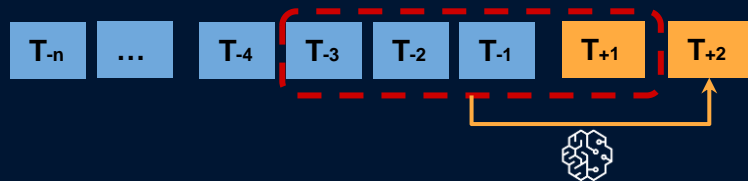
Predicción step 2

- Valor observado
- Valor predicho
- Predictores
- Modelo para step+1

PREDICCIÓN MULTI-STEP RECURSIVA



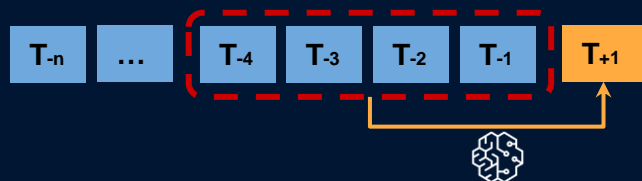
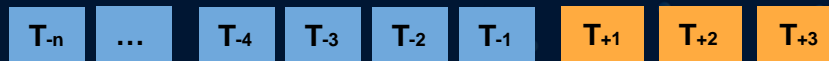
Predicción step 1



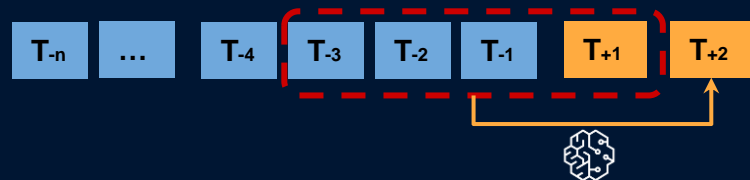
Predicción step 2

- Valor observado
- Valor predicho
- Predictores
- Modelo para step+1

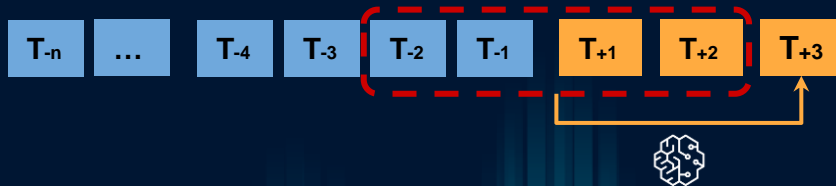
PREDICCIÓN MULTI-STEP RECURSIVA







Predicción step 1



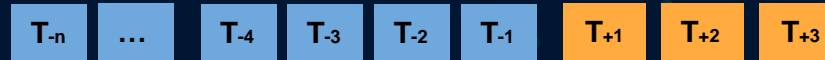
Predicción step 2





Predicción step 3

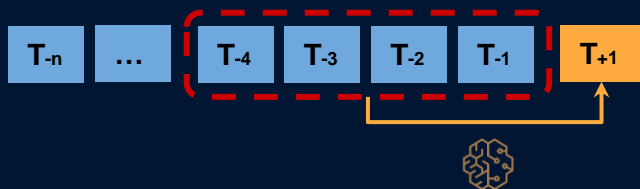
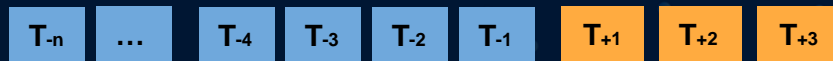
-  Valor observado
-  Valor predicho
-  Predictores
-  Modelo para step+1

PREDICCIÓN MULTI-STEP DIRECTA



 Valor observado
 Valor predicho

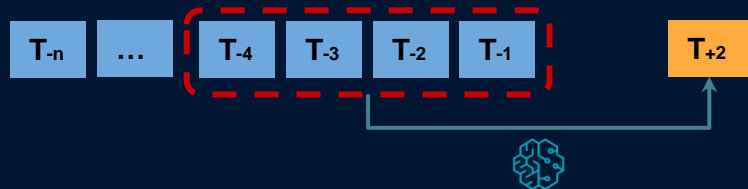
PREDICCIÓN MULTI-STEP DIRECTA



Predicción step 1



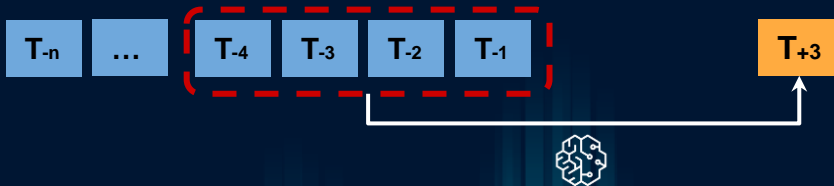
Modelo para step 1



Predicción step 2



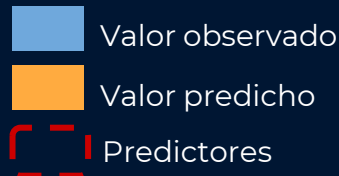
Modelo para step 2



Predicción step 3



Modelo para step 3



¿POR QUÉ SKFORECAST?

¿POR QUÉ SKFORECAST?

Crear modelos de forecasting fácilmente

Entrena modelos con las herramientas que ya conoces:
scikit-learn, **LightGBM** o **XGBoost**.



Responde preguntas clave rápidamente

¿Qué impacto tendría una decisión en tu negocio?



Ahorra tiempo y esfuerzo

*¡Anda! el MVP que empezó hace 4,32 años ha salido bien...
Vaya, no se enfadará mi DS si le digo que hay que poner
los modelos en producción en 3 días, ¿no?*



PUNTOS FUERTES DEL MACHINE LEARNING

Incorporación de variables exógenas

El pasado de una serie temporal (componente autorregresivo) solo explica una parte del comportamiento observado. Las variables exógenas (variables de calendario, temperatura, indicadores económicos, etc.) pueden aportar información muy valiosa para el modelo.

Modelar múltiples series simultáneamente (modelos globales)

Series temporales que comparten patrones comunes pueden modelarse de forma conjunta, lo que permite obtener modelos que generalizan mejor y son más fáciles de mantener en producción. Un modelo capaz de predecir n series vs n modelos individuales.

Aceptar valores ausentes

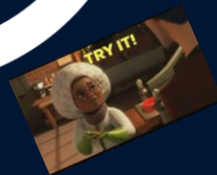
En la práctica, los datos históricos suelen contener valores ausentes.

TODO CONVERGE EN SKFORECAST

MÉTODO
RECURSIVO/
DIRECTO



EXÓGENAS



Pero que me
están contando
estos tíos...



FORECASTERS

Un objeto **Forecaster** es un contenedor completo que proporciona **funciones y métodos esenciales** para entrenar un modelo de previsión y generar predicciones para puntos futuros en el tiempo.

Forecaster	Single series	Multiple series	Recursive strategy	Direct strategy	Probabilistic prediction	Time series differentiation	Exogenous features	Window features
ForecasterRecursive	✓		✓		✓	✓	✓	✓
ForecasterDirect	✓			✓	✓		✓	✓
ForecasterRecursiveMultiSeries		✓	✓		✓	✓	✓	✓
ForecasterDirectMultiVariate		✓		✓	✓		✓	✓
ForecasterRNN		✓		✓				
ForecasterSarimax	✓		✓		✓	✓	✓	

Tabla completa: <https://skforecast.org/latest/#forecasters>

¿CÓMO CREAR UN FORECASTER?

Imaginemos **una serie temporal** de la **demanda eléctrica** de una ciudad y el objetivo es generar un modelo de **forecasting** capaz de predecir la **demanda energética** de toda la ciudad.



DATOS

Demanda eléctrica

Variables exógenas:

- Previsión Temperatura
- Calendario

MODELO



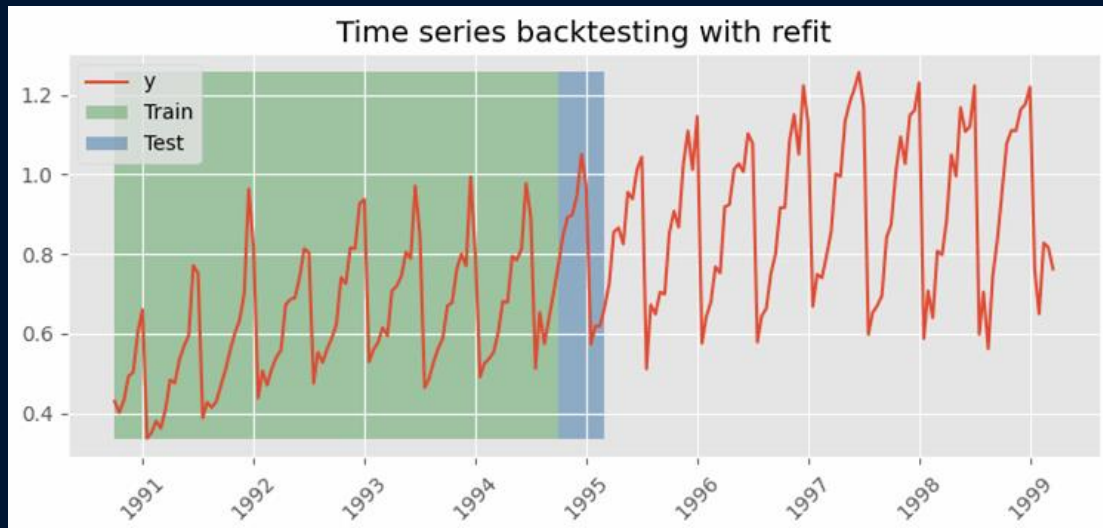
LGBMRegressor

**PERO... ¿CÓMO SÉ QUE
ESTOY MODELANDO
BIEN?**

VALIDACIÓN DE MODELOS (BACKTESTING)

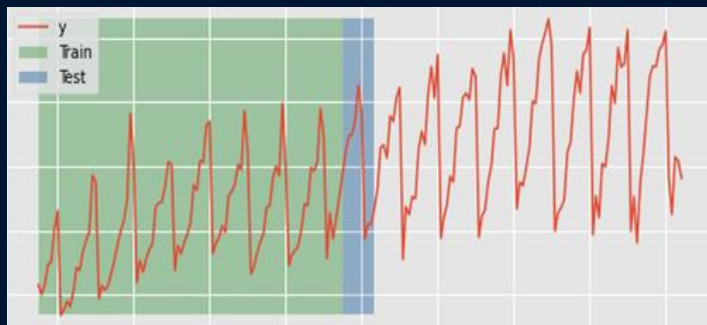
En la previsión de series temporales, el proceso de **backtesting** consiste en **evaluar el rendimiento de un modelo** predictivo aplicándolo retrospectivamente a **datos históricos**.

En otras palabras... **¿Cómo habría funcionado mi modelo en el pasado?**

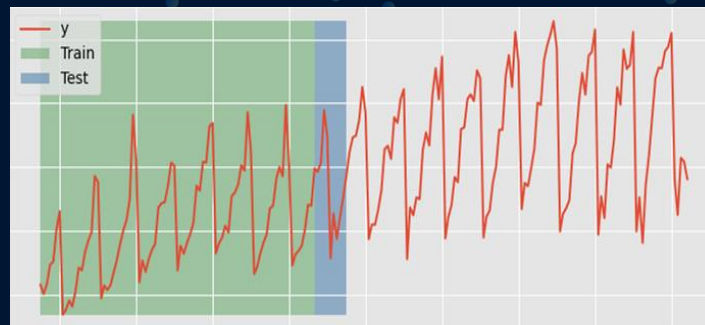


TIPOS DE BACKTESTING

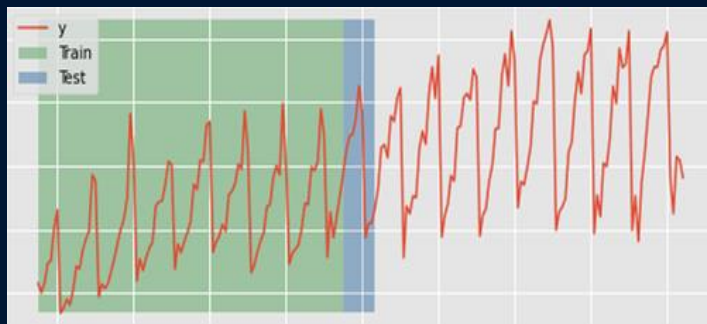
Backtesting sin reentrenamiento



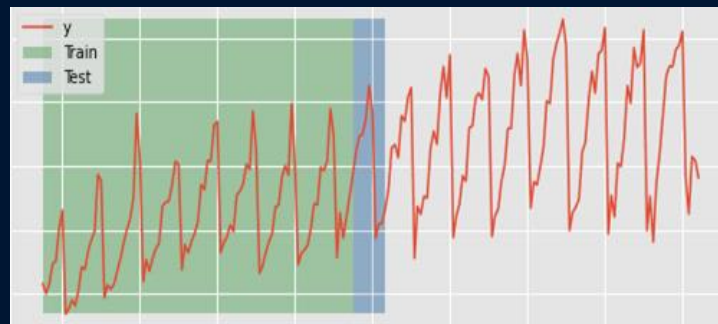
Backtesting con reentrenamiento intermitente



Backtesting con reentrenamiento (time series cross validation, rolling origin)



Backtesting con reentrenamiento (fixed origin)



**ESTO SUENA GENIAL,
PERO... ¿CÓMO OPTIMIZO
MI MODELO?**

BÚSQUEDA DE HIPERPARÁMETROS Y LAGS

Esta búsqueda consiste en probar sistemáticamente **diferentes combinaciones** de hiperparámetros (incluidos los lags) para encontrar la **configuración óptima** que produzca los mejores resultados.

Skforecast combina las diferentes **estrategias habituales** en el mundo del machine learning y usa la técnica de **backtesting** o la de **one step ahead** como método de **validación** para este proceso.

- **Grid search**
- **Random search**
- **Bayesian search**

FEATURE SELECTION

Feature selection es el proceso de selección de un subconjunto de predictores relevantes para su uso en el modelo. Las principales razones para llevar a cabo esta búsqueda son:

- **Simplificar los modelos** para facilitar su interpretación.
- Reducir el **tiempo de entrenamiento**.
- Mejorar la **generalización** reduciendo el sobreajuste (formalmente, reducción de la varianza).

Skforecast es compatible con **los métodos de selección** implementados en **scikit-learn**. Algunos de los más comunes son:

- **Recursive feature elimination (RFE)**
- **Sequential Feature Selection (Forward-SFS, Backward-SFS)**
- **Feature selection based on threshold (SelectFromModel)**

FEATURE SELECTION

Recursive feature elimination (RFECV)

Total number of records available: 8712

Total number of records used for feature selection: 4356

Number of features available: 139

Lags (n=48)

Window features (n=3)

Exog (n=88)

Number of features selected: 52

Lags (n=31) : [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 16, 19, 22, 23, 24, 25, 26, 28, 30, 32, 34, 35, 36, 40, 42, 44, 47, 48]

Window features (n=1) : ['roll_mean_24']

Exog (n=20) : ['hour_day_sin', 'hour_day_cos', 'poly_month_cos_week_of_year_sin', 'poly_week_of_year_sin_week_day_sin', 'poly_week_of_year_sin_week_day_cos', 'poly_week_of_year_sin_hour_day_sin', 'poly_week_of_year_sin_hour_day_cos', 'poly_week_of_year_sin__sunset_hour_cos', 'poly_week_of_year_cos_week_day_sin', 'poly_week_of_year_cos_week_day_cos', 'poly_week_of_year_cos_hour_day_sin', 'poly_week_of_year_cos_hour_day_cos', 'poly_week_day_sin_hour_day_sin', 'poly_week_day_sin_hour_day_cos', 'poly_week_day_sin_sunset_hour_sin', 'poly_week_day_cos_hour_day_sin', 'poly_week_day_cos_hour_day_cos', 'poly_hour_day_sin_hour_day_cos', 'temp_roll_mean_1_day', 'temp']

HANDS ON TIME

TENGO MUCHAS SERIES TEMPORALES...

GLOBAL FORECASTING

Independent multi-series forecasting

Existe un **único modelo** para todas las series temporales, pero cada una de ellas es **independiente** de las demás, por lo que significa que los valores pasados de una serie no se utilizan como predictores de las otras. Sin embargo, es útil modelarlas conjuntamente, ya que las series pueden seguir el **mismo patrón intrínseco** en cuanto a sus valores pasados y futuros.

Dependent multi-series forecasting (Multivariate)

Todas las series se modelizan juntas en un **único modelo**, teniendo en cuenta que cada serie temporal **depende** no solo de sus valores pasados, sino también de los valores pasados de las **demás series**. Se espera que el modelo aprenda no solo la información de cada serie por separado, sino que también las **relaciones entre ellas**.

CASOS DE USO TÍPICOS MULTI-SERIES

Independent multi-series forecasting

Las ventas de **1000 productos** en la **misma tienda** pueden no estar relacionadas, pero siguen la **misma dinámica**, la de la tienda

¿Cómo podríamos estimar **el consumo energético de cientos de hogares** en toda una ciudad?

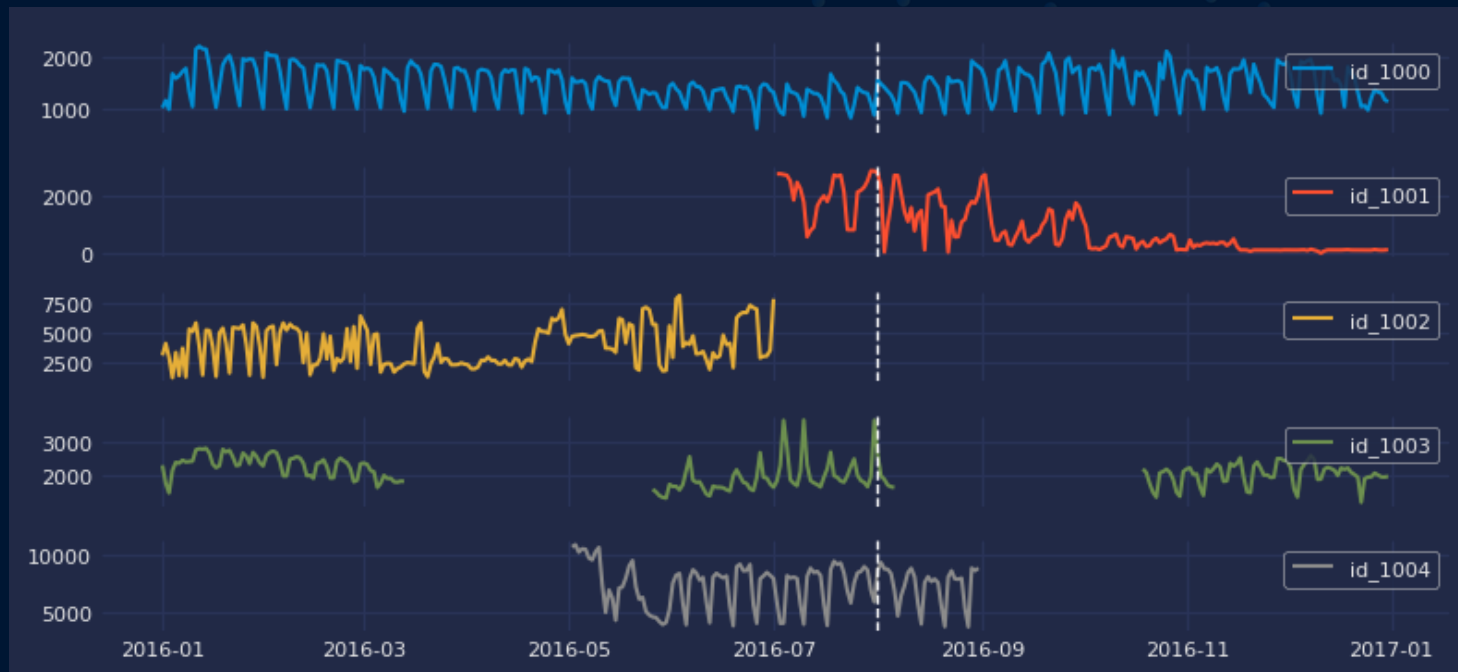
Dependent multi-series forecasting (Multivariate)

¿Cuál será el nivel de **contaminantes** del aire en la próxima semana?

Mediciones efectuadas por todos los **sensores** (caudal, temperatura, presión...) instalados en un compresor de aire industrial

INDEPENDENT MULTI-SERIES FORECASTING

`recursive.ForecasterRecursiveMultiSeries`



INDEPENDENT MULTI-SERIES FORECASTING

Series 1

1	2	3	4	5	6	7
---	---	---	---	---	---	---

Series 2

101	102	103	104	105	106	107
-----	-----	-----	-----	-----	-----	-----

Exogenous variable

a	b	c	d	e	f	g
---	---	---	---	---	---	---

value	id serie	exog
1	series 1	a
2	series 1	b
3	series 1	c
4	series 1	d
5	series 1	e
6	series 1	f
7	series 1	g
101	series 2	a
102	series 2	b
103	series 2	c
104	series 2	d
105	series 2	e
106	series 2	f
107	series 2	g

ForecasterAutoregMultiSeries Training Matrix

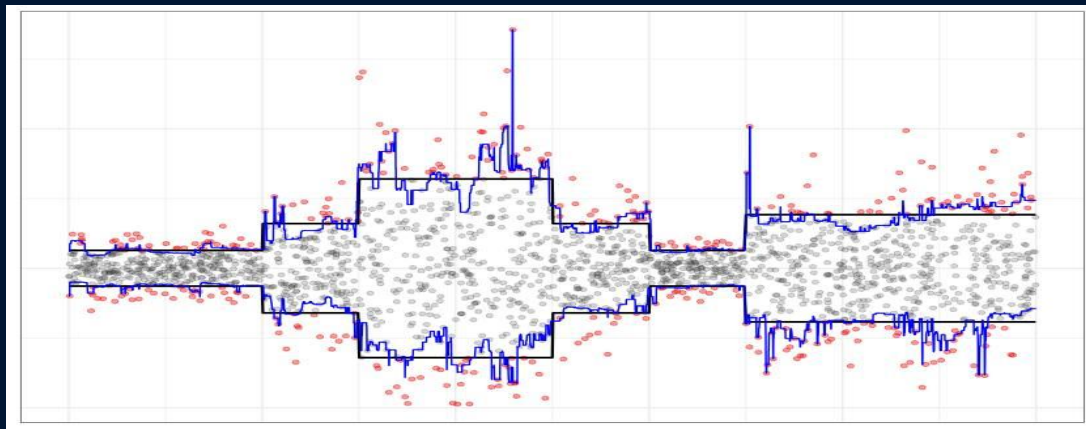
lag 3	lag 2	lag 1	series 1	series 2	exog	Y
1	2	3	1	0	d	4
2	3	4	1	0	e	5
3	4	5	1	0	f	6
4	5	6	1	0	g	7
101	102	103	0	1	d	104
102	103	104	0	1	e	105
103	104	105	0	1	f	106
104	105	106	0	1	g	107

HANDS ON TIME

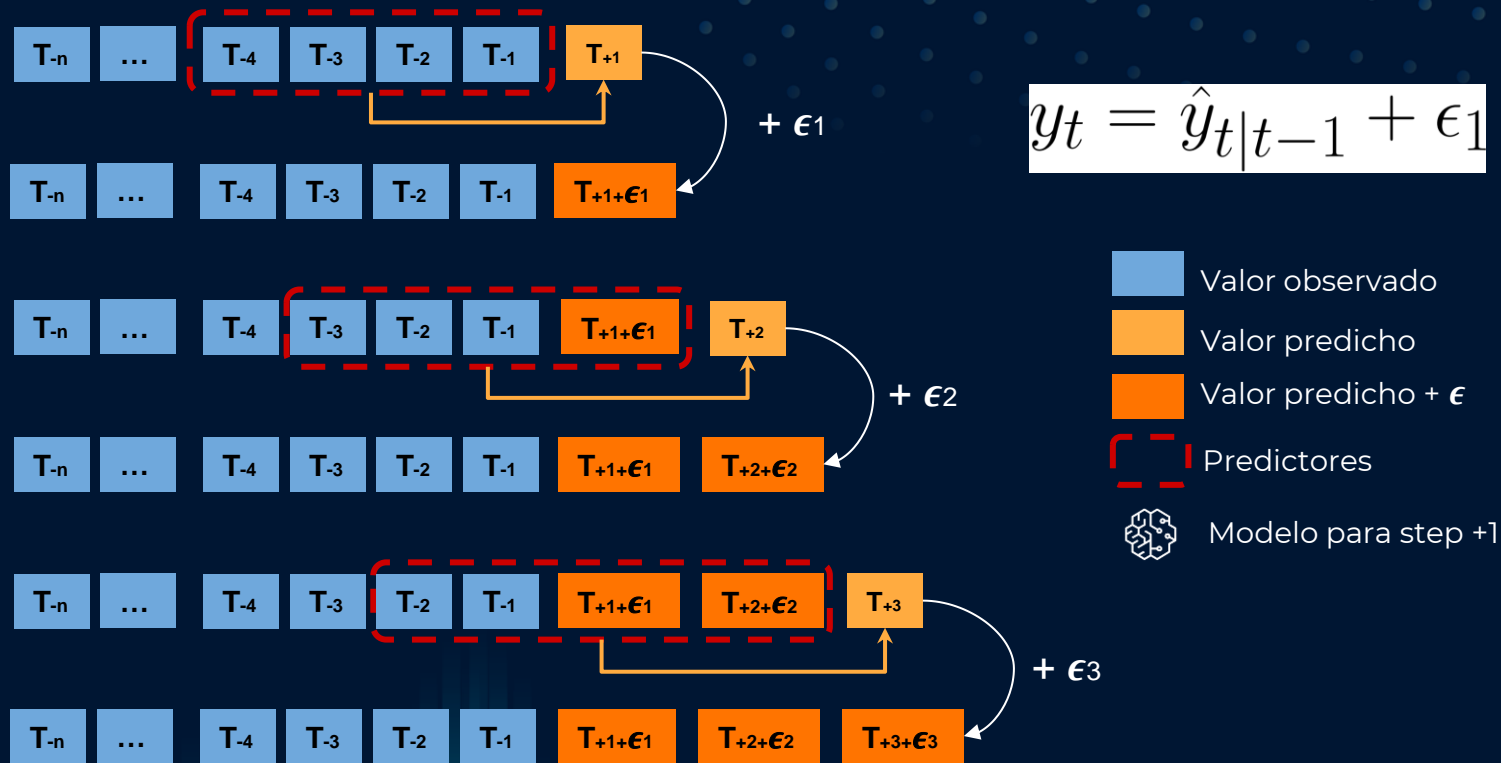
VALE, PERO ME HAN PEDIDO INTERVALOS DE PREDICCIÓN

FORECASTING PROBABILÍSTICO: INTERVALOS DE PREDICCIÓN

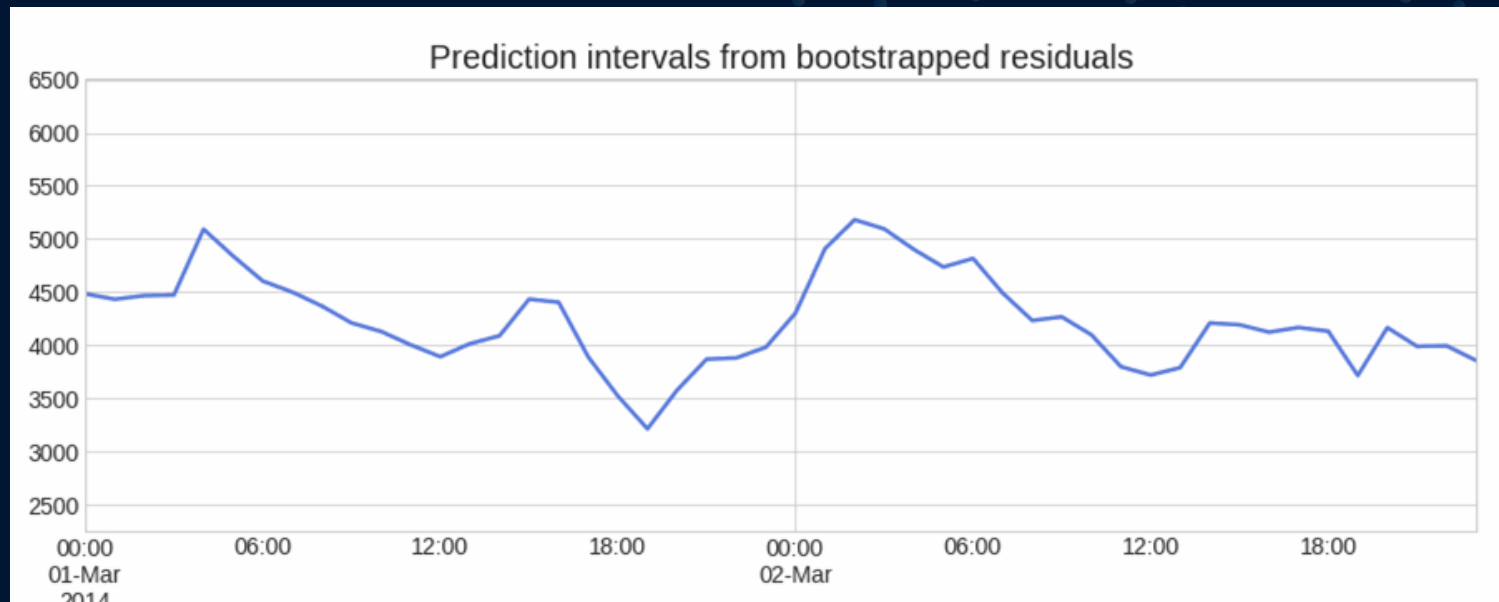
Un intervalo de predicción define el intervalo dentro del cual es de esperar que se encuentre el verdadero valor de y con una probabilidad determinada. Por ejemplo, se espera que el intervalo de predicción del 98% contenga el verdadero valor de la predicción con un 98% de probabilidad.



INTERVALOS DE PREDICCIÓN: BOOTSTRAPPED RESIDUALS



INTERVALOS DE PREDICCIÓN: BOOTSTRAPPED RESIDUALS



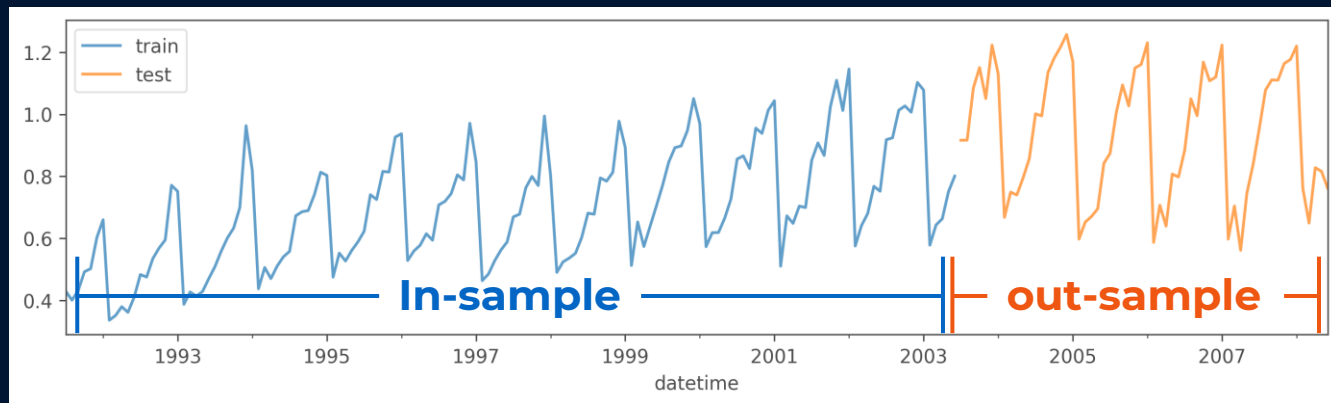
INTERVALOS DE PREDICCIÓN: TIPOS DE RESIDUOS

In-sample residuals

Residuos calculados usando el **conjunto de entrenamiento**, ideales para evaluar el ajuste del modelo a los **datos históricos**.

Out-sample residuals

Residuos generados en el **conjunto de validación**, útiles para entender cómo el modelo generaliza a **datos nuevos**.



INTERVALOS DE PREDICCIÓN: RESIDUOS CONDICIONADOS

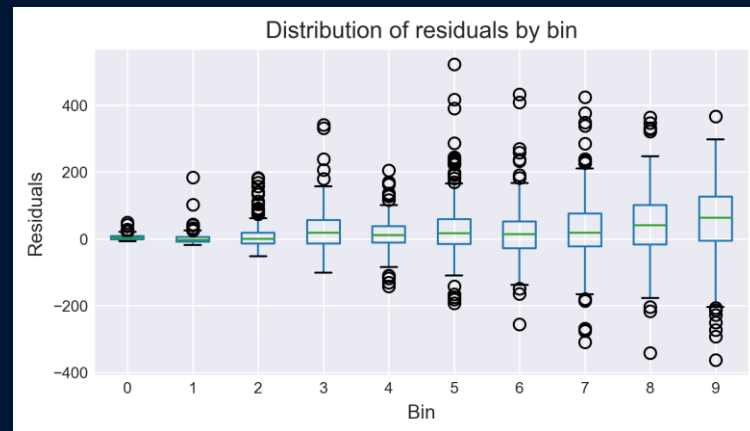
Residuos condicionados con las predicciones (In or Out)

Residuos organizados en **intervalos o bins** según el valor de las **predicciones**. Esto permite ajustar la incertidumbre del bootstrapping **mejorando la cobertura (coverage)** de los intervalos.

Binner intervals

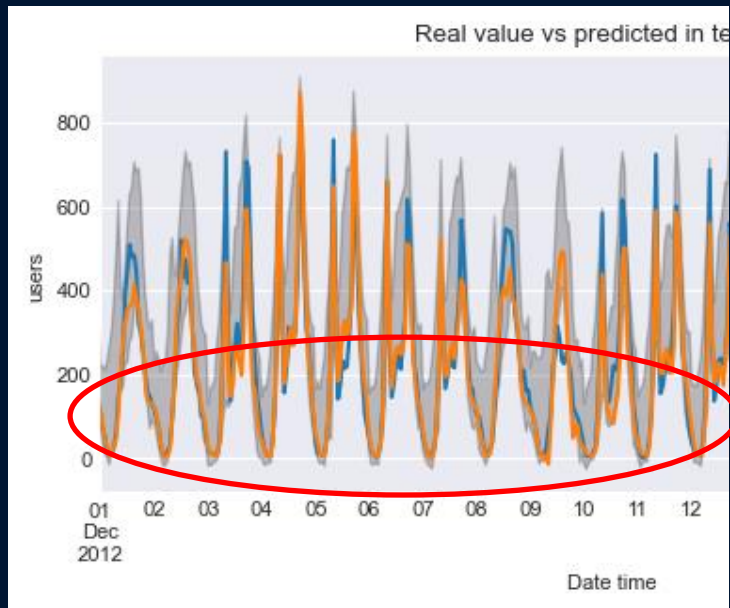
{bin: predictions Interval}

```
{0.0: (-8.229467171553717, 11.116037535200665),  
1.0: (11.116037535200665, 31.879155847370434),  
2.0: (31.879155847370434, 75.9019071402224),  
3.0: (75.9019071402224, 124.5691653220086),  
4.0: (124.5691653220086, 170.35484312260417),  
5.0: (170.35484312260417, 218.96823239624555),  
6.0: (218.96823239624555, 278.6496576655771),  
7.0: (278.6496576655771, 355.13229168292287),  
8.0: (355.13229168292287, 486.1660497574729),  
9.0: (486.1660497574729, 970.517259284916)}
```

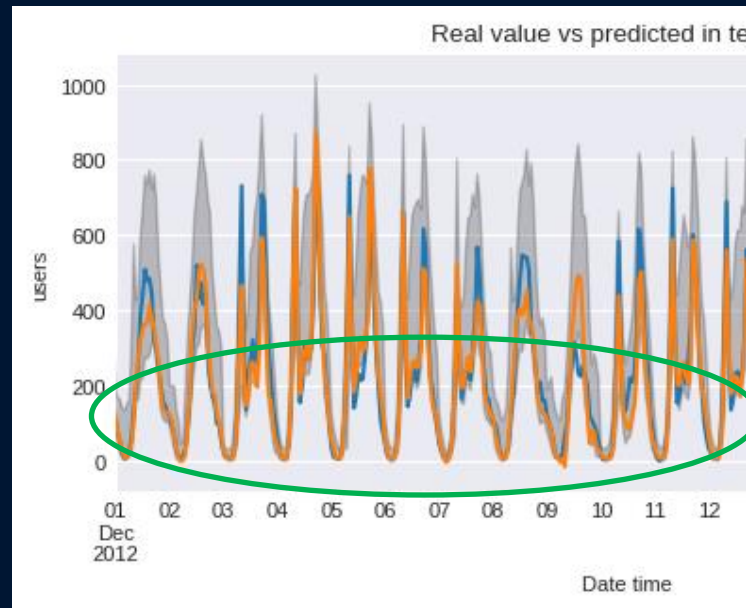


INTERVALOS DE PREDICCIÓN: RESIDUOS CONDICIONADOS

Out-sample residuals no binned
interval coverage: 75.4 %



Out-sample residuals binned
interval coverage: 84.5 %



**Y... ¿SI QUIERO
EXPLICAR TODO ESTO?**

FORECASTER EXPLAINABILITY: FEATURE IMPORTANCES

Calcular la importancia de las variables (*feature importances*) es una técnica utilizada en *machine learning* para determinar la relevancia o importancia de cada *feature* (o variable) en la predicción de un modelo. En otras palabras, mide **en qué medida contribuye cada *feature* al resultado** del modelo.

- **Regresores lineales:** coeficientes de regresión.
- **Modelos de árboles:** mean decrease impurity or permutation feature importance methods.

Example LGBMRegressor

	feature	importance
7	Temperature	570
0	lag_1	470
2	lag_3	387
1	lag_2	362
6	lag_7	325
5	lag_6	313
4	lag_5	298
3	lag_4	275

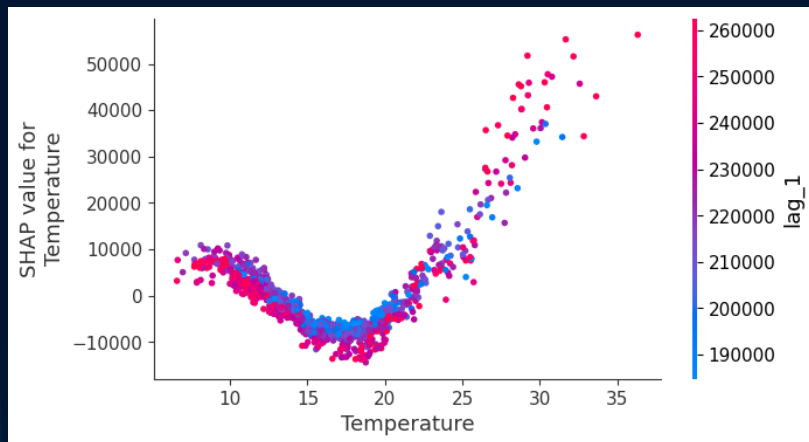
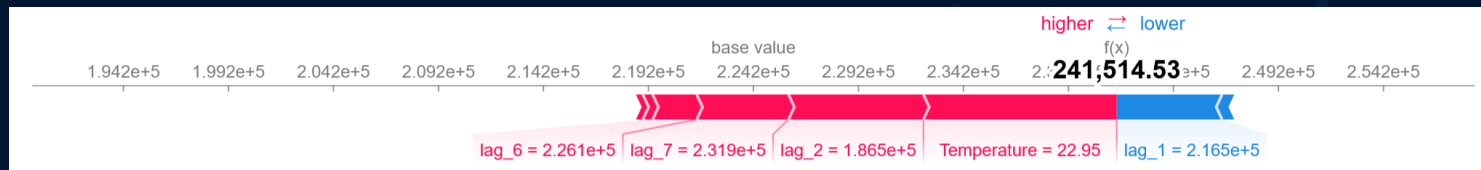
FORECASTER EXPLAINABILITY: SHAP VALUES

Los valores **SHAP** (SHapley Additive exPlanations) proporcionan una clara comprensión de cómo las variables y sus valores impactan en las predicciones, tanto visual como numéricamente.

- El **regresor interno** utilizado por el Forecaster.
- Las **matrices** derivadas de las series temporales. Estas matrices pueden ser las utilizadas para **entrenar** el Forecaster o las necesarias durante la fase de **predicción**.


FORECASTER EXPLAINABILITY: SHAP VALUES

Más información: https://skforecast.org/latest/user_guides/explainability




¿TODAVÍA HAY MÁS?

MATERIAL ADICIONAL

 **Introducción al forecasting** - Basics of forecasting concepts and methodologies

 **Quick start** - Get started quickly with skforecast

 **User guides** - Detailed guides on skforecast features and functionalities

 **Examples and tutorials** - Learn through practical examples and tutorials to master skforecast


 **FAQ and tips** - Find answers and tips about forecasting

 **API Reference** - Comprehensive reference for functions and classes

ESTADO DEL ARTE



QUEREMOS ESCUCHAROS 😊

- Estrellita en GitHub ★
- Compartir tu opinión en LinkedIn 
- Reportar bugs y/o sugerir mejoras
- Contribuir con nuevo código o casos de uso



¿PREGUNTAS?

Joaquín Amat Rodrigo
Javier Escobar Ortiz