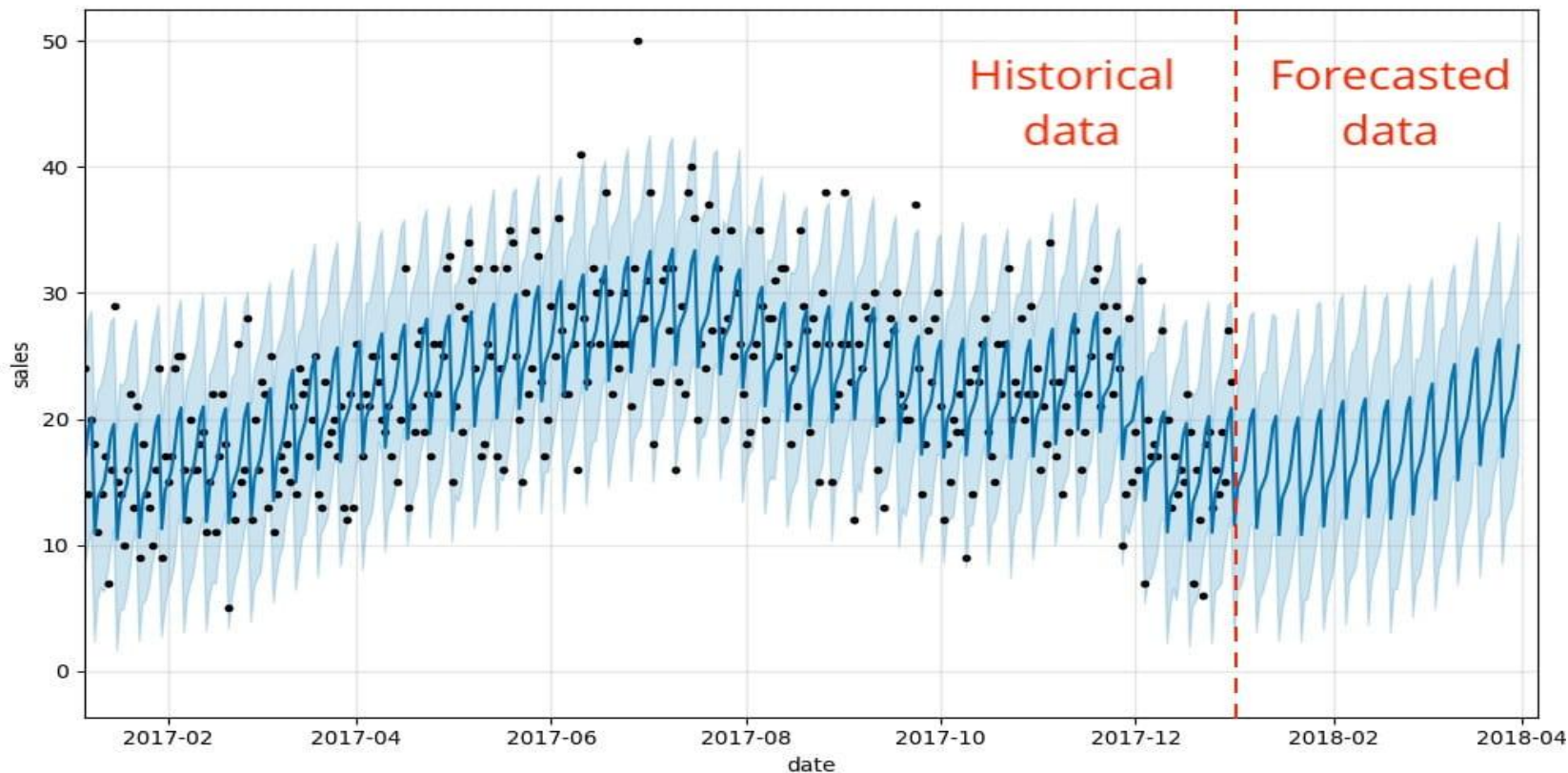


TP Final EEA

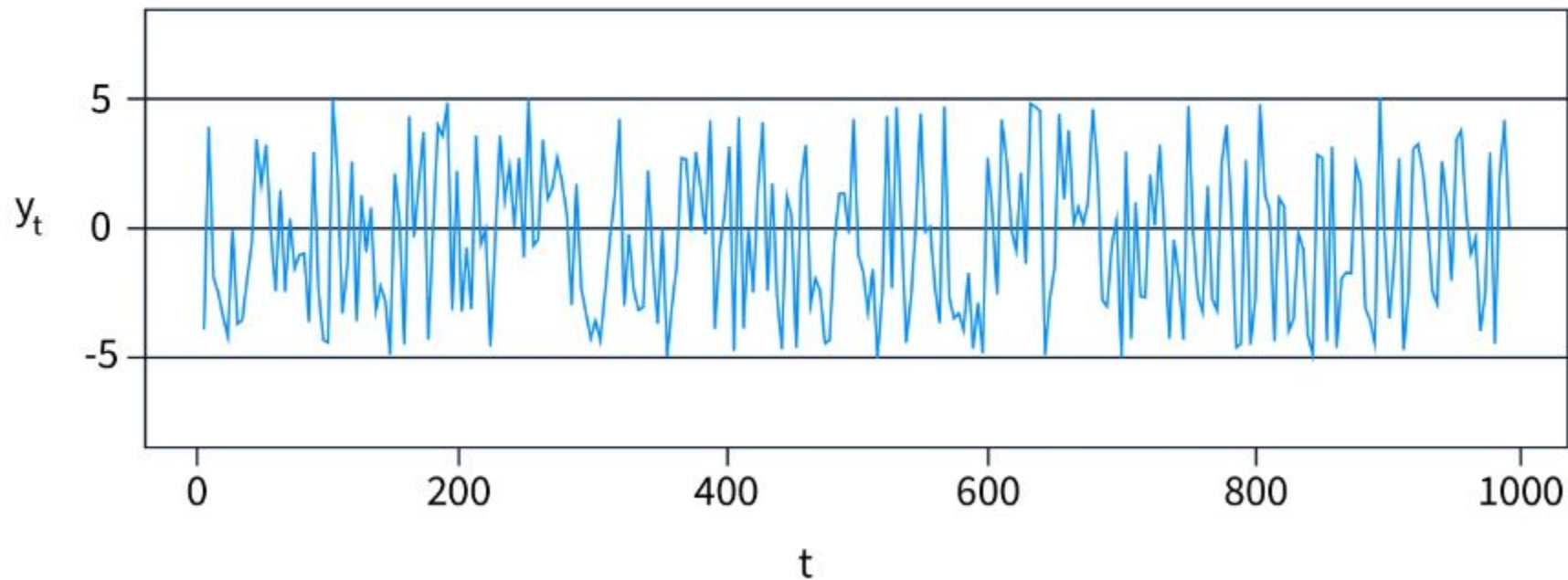
INTRO

Introducción a la Predicción de Series Temporales

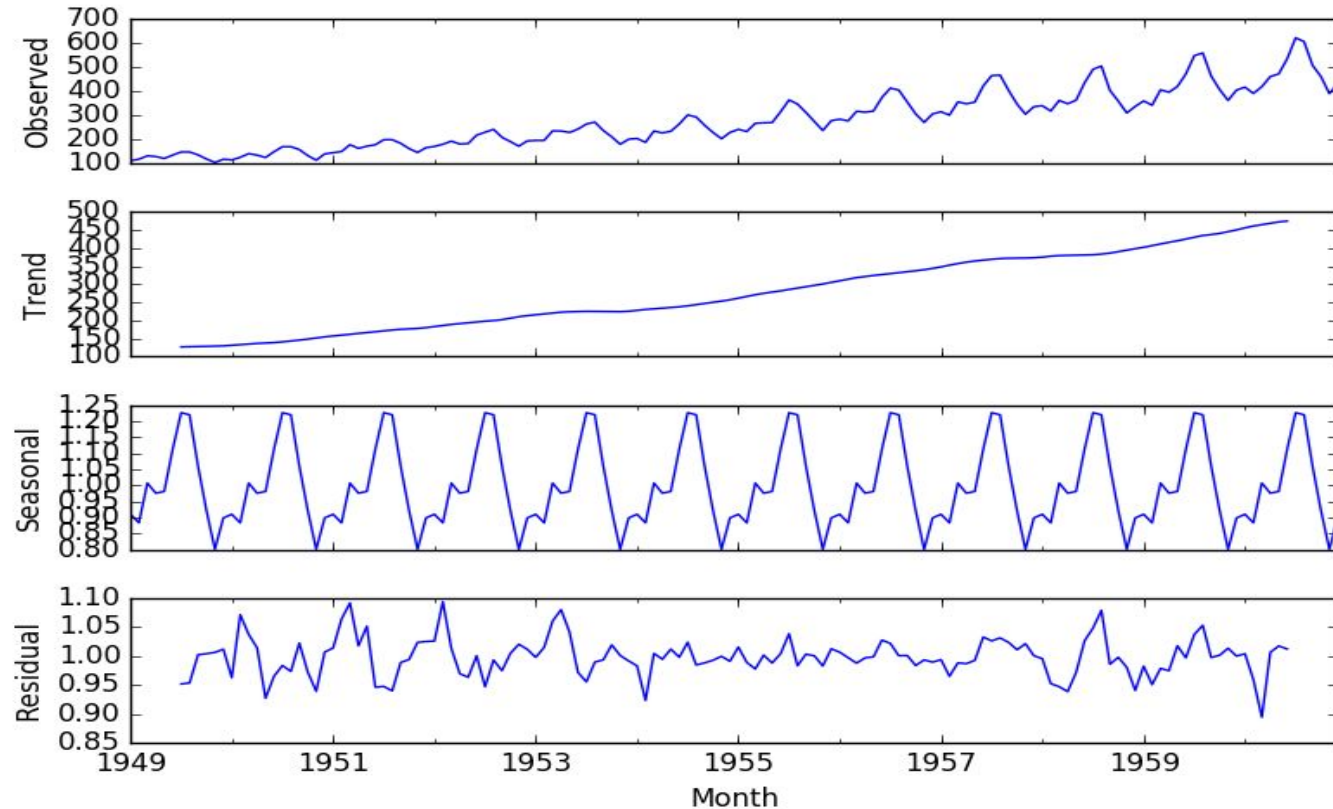


Estacionalidad

Stationary Time Series



Los Componentes de una Serie Temporal



Consideraciones Clave

- Identificación de tendencias, estacionalidad y factores externos.
- Manejo de irregularidades como outliers, valores faltantes y relaciones no lineales.
- El horizonte de pronóstico requerido por el problema

OVERVIEW

Los métodos de la Predicción de Series Temporales

- Modelos estadísticos clásicos
- Enfoques algorítmicos más recientes
- Aprendizaje profundo (Deep Learning)
- Reducción del problema a la regresión tabular



Modelos estadísticos clásicos

- Metodos de Exponential Smoothing
- La familia SARIMAX
- GARCH
- Y muchos otros

SARIMAX (1, 0, 2) (2, 0, 1, 5)

$$y_t = c + \varphi_1 y_{t-1} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \Phi_1 (y_{t-5} + \varphi_1 y_{t-6}) + \Phi_2 (y_{t-10} + \varphi_1 y_{t-11}) + \Theta_1 (\varepsilon_{t-5} + \theta_1 \varepsilon_{t-6} + \theta_2 \varepsilon_{t-7}) + \varepsilon_t$$

Enfoques algorítmicos más recientes (Enterprise)

- Prophet
- Orbit
- Greykite

PROPHET



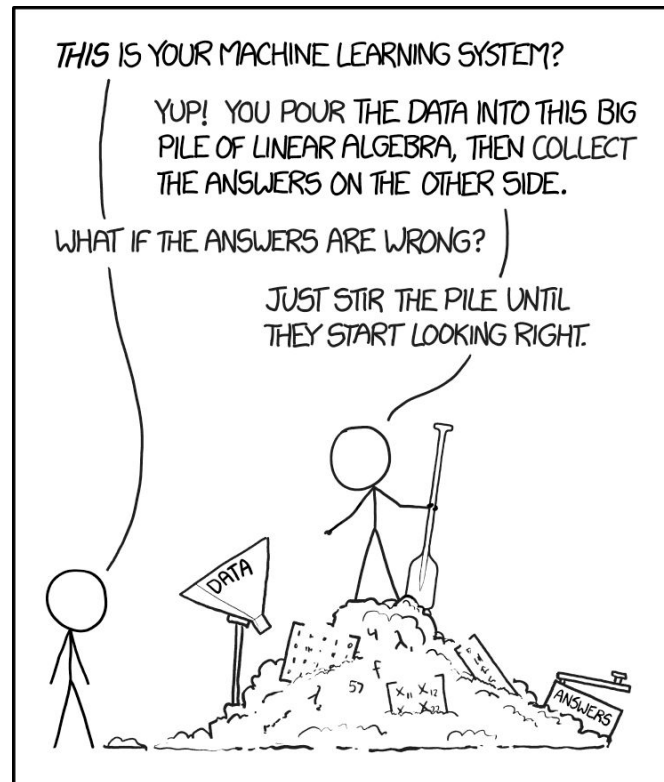
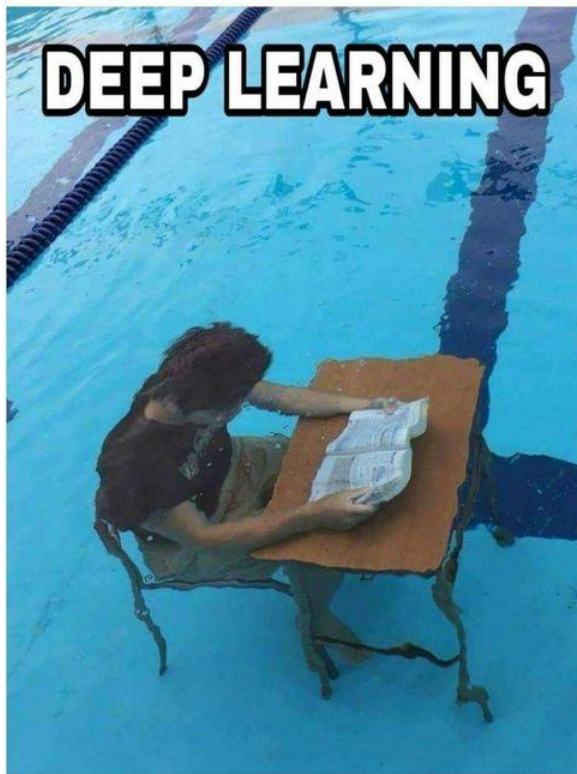
Orbit



GREYKITE

Deep Learning

- CNN
- LSTM
- GRU
- RNN
- Transformers



Regresión Tabular

¿Podemos usar lo que ya sabemos?



Diseño del experimento

- Dataset: US Births, 7305 obsevaciones, diario
- La métrica: MAPE (Error Porcentual Absoluto Medio)
- Pronóstico de predicción: todo el año 1988 (366 días)
- Predicción recursiva para modelos con horizonte de pronóstico igual a 1
- Baseline: copiar los valores de 1987

TABULAR

Preparación de datos

y
0
1
2
3
4
5

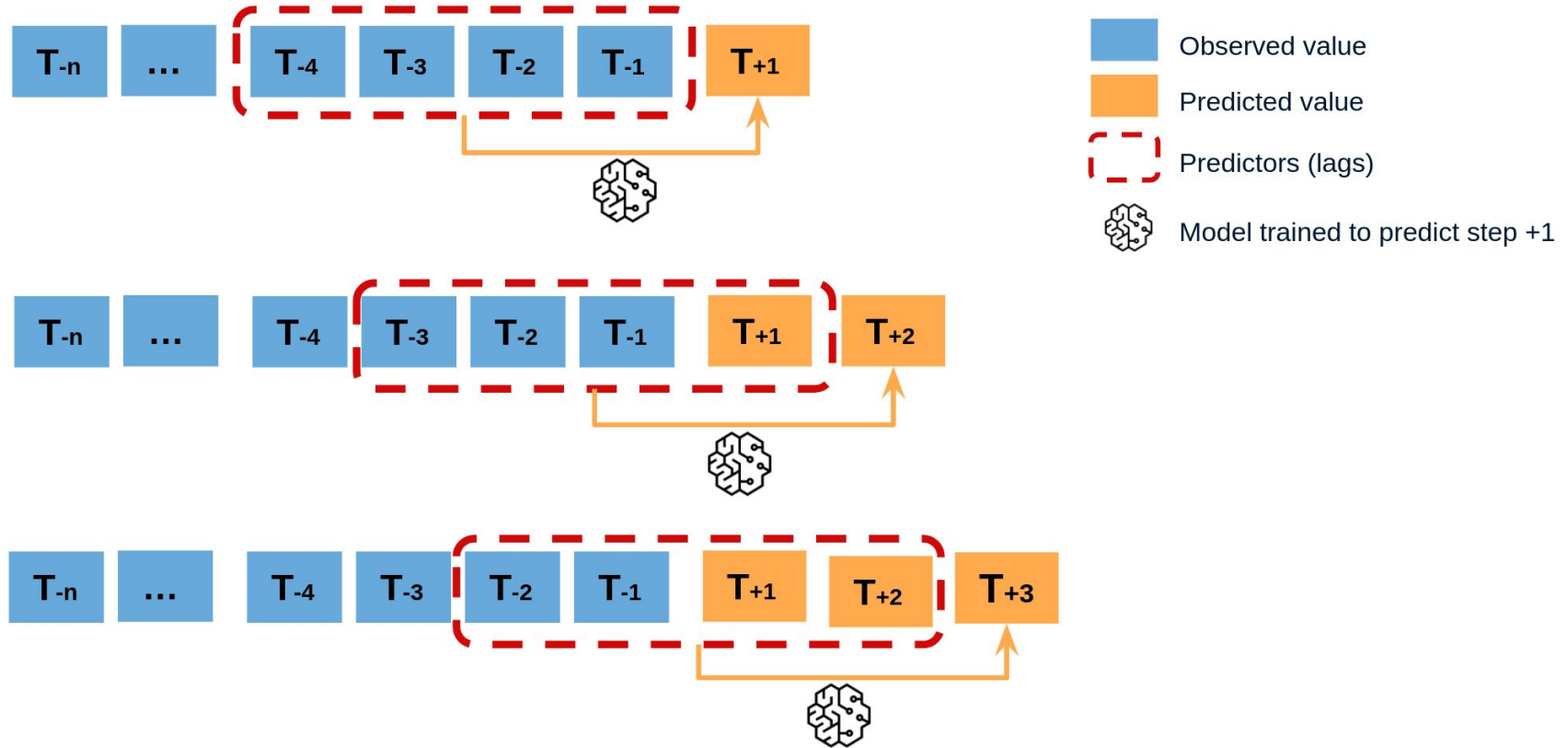
make_reduction



y	lag_1	lag_2	lag_3
0	-	-	-
1	0	-	-
2	1	0	-
3	2	1	0
4	3	2	1
5	4	3	2

x

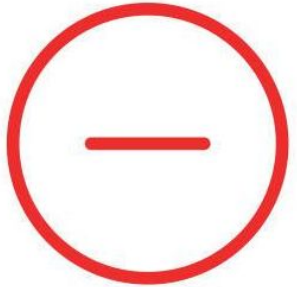
Predicción con un Horizonte mayor a uno



Pros y Contras



Fácil de utilizar variables exógenas para la predicción.
Utiliza información de múltiples series temporales
Utilizamos herramientas ya conocidas



La pérdida de datos durante la tabularización puede ser fatal.
El horizonte de pronóstico es Uno.
Inherente multicolinealidad fuerte.

CONCLUSION

Particularidades del conjunto de datos:

- Pequeña cantidad de datos
- No hay variables exógenas
- Período de tiempo diario
- Pronóstico a largo plazo
- Solo una serie temporal

Datos del performance de los modelos

Modelo	MAPE
Baseline	0.138
MA	0.135
AR	0.109
ARMA	0.065
ARIMA	0.058
SARIMA	0.048

Modelo	MAPE
LinReg	0.083
XGBoost	0.061
Transformer	0.092
Prophet	0.039
Orbit	0.109
Kite	0.033

Las preguntas principales

- ¿Tengo variables exógenas?
- ¿Tengo múltiples series temporales con una naturaleza común?
- ¿Hay datos faltantes en mi conjunto de datos?
- ¿Cuán grande es mi conjunto de datos?
- ¿Cuál es mi horizonte de pronóstico?