Predicción económica con inteligencia artificial

¿Qué añade a los métodos econométricos?

Galo Nuño

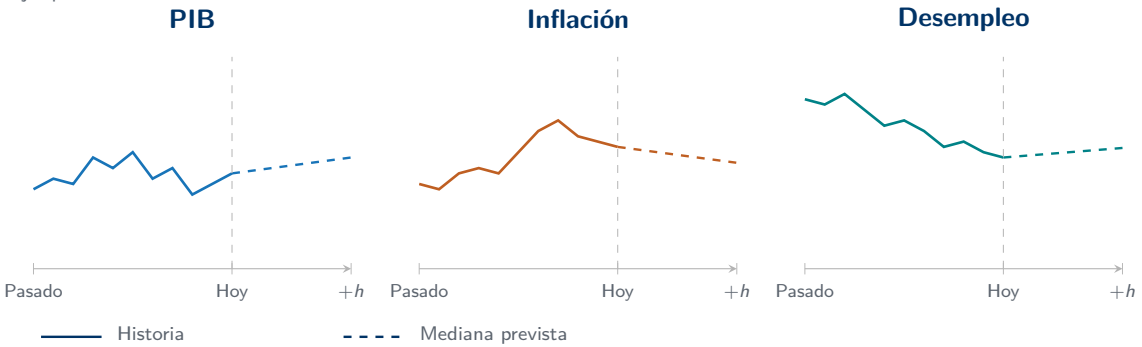
FUNCAS

19 de octubre de 2026

El objetivo: predecir varias series y cuantificar la incertidumbre

PIB, inflación, desempleo: queremos anticipar varias variables a la vez

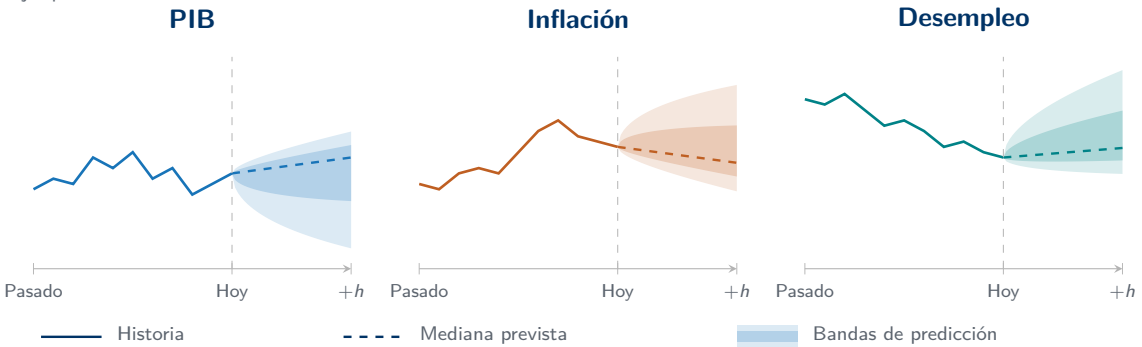
Ejemplo ilustrativo



El objetivo: predecir varias series y cuantificar la incertidumbre

PIB, inflación, desempleo: queremos anticipar varias variables a la vez

Ejemplo ilustrativo

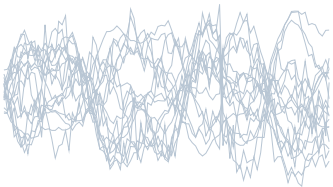


No basta una previsión puntual:
las bandas pueden ser asimétricas

Muchas variables, pocos factores

Stock y Watson: comprimir la información antes de predecir

Panel de series X_t

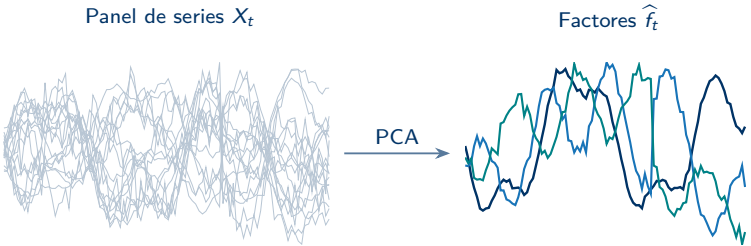


Ejemplo simulado

Stock y Watson (2002), Forecasting Using Principal Components from a Large Number of Predictors.

Muchas variables, pocos factores

Stock y Watson: comprimir la información antes de predecir



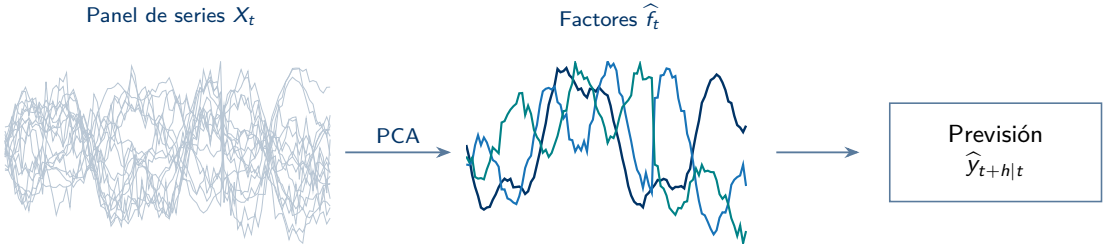
Ejemplo simulado

$$x_{it} = \lambda_i' f_t + e_{it}$$

Stock y Watson (2002), Forecasting Using Principal Components from a Large Number of Predictors.

Muchas variables, pocos factores

Stock y Watson: comprimir la información antes de predecir



$$x_{it} = \lambda_i' f_t + e_{it}$$

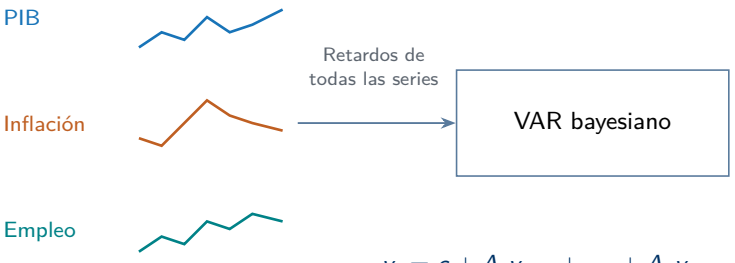
Los factores capturan comovimientos, no toda la variación individual

Stock y Watson (2002), Forecasting Using Principal Components from a Large Number of Predictors.

VAR bayesiano: dinámica conjunta con regularización

Cada variable depende de sus retardos y de los de las demás

Datos observados



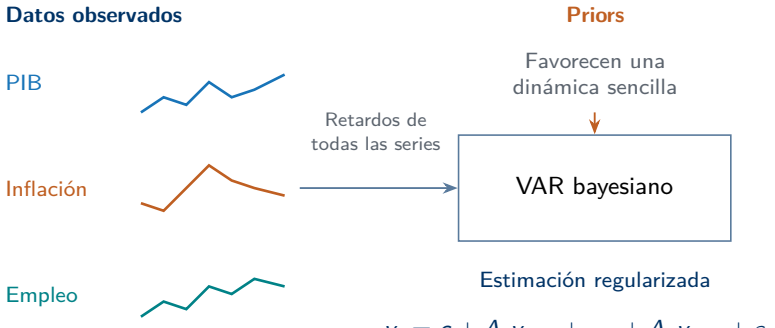
$$y_t = c + A_1 y_{t-1} + \cdots + A_p y_{t-p} + \varepsilon_t$$

Esquema ilustrativo

Giannone, Lenza y Primiceri (2015), Prior Selection for Vector Autoregressions.

VAR bayesiano: dinámica conjunta con regularización

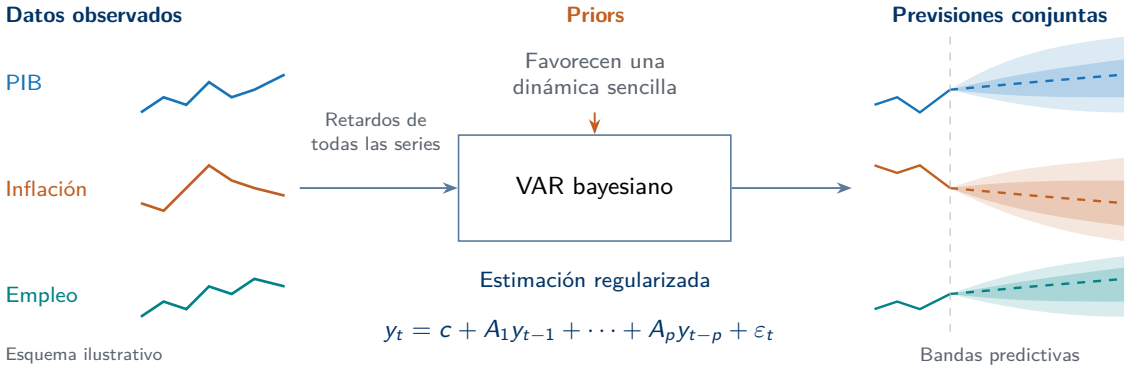
Cada variable depende de sus retardos y de los de las demás



Giannone, Lenza y Primiceri (2015), Prior Selection for Vector Autoregressions.

VAR bayesiano: dinámica conjunta con regularización

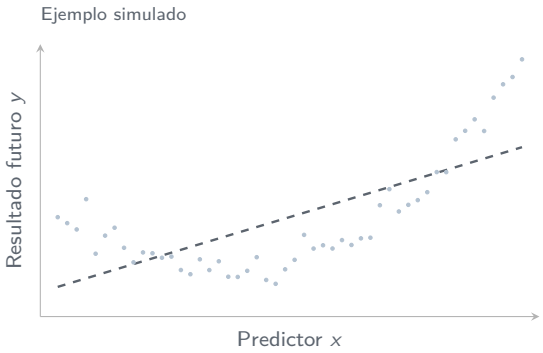
Cada variable depende de sus retardos y de los de las demás



Los priors ayudan a evitar el sobreajuste

Giannone, Lenza y Primiceri (2015), Prior Selection for Vector Autoregressions.

Inteligencia artificial: relaciones más flexibles

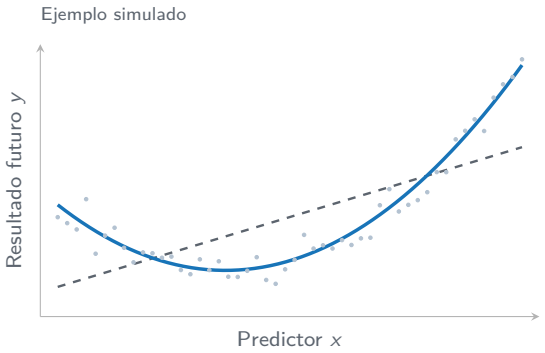


Relación lineal

$$y_{t+h} = \beta' X_t + \varepsilon_{t+h}$$

Goulet Coulombe et al. (2022), How Is Machine Learning Useful for Macroeconomic Forecasting?
Medeiros et al. (2021), Forecasting Inflation in a Data-Rich Environment: The Benefits of Machine Learning Methods.

Inteligencia artificial: relaciones más flexibles



Relación lineal

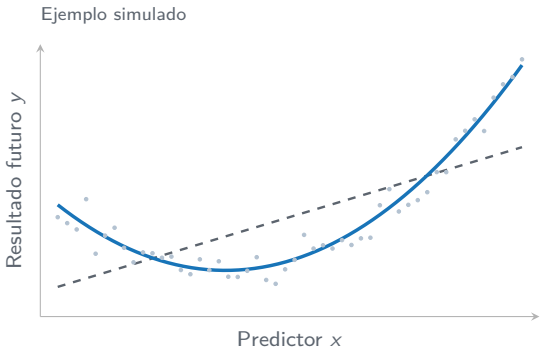
$$y_{t+h} = \beta' X_t + \varepsilon_{t+h}$$

Relación no lineal

$$y_{t+h} = f(X_t) + \varepsilon_{t+h}$$

Goulet Coulombe et al. (2022), How Is Machine Learning Useful for Macroeconomic Forecasting?
Medeiros et al. (2021), Forecasting Inflation in a Data-Rich Environment: The Benefits of Machine Learning Methods.

Inteligencia artificial: relaciones más flexibles



Relación lineal

$$y_{t+h} = \beta' X_t + \varepsilon_{t+h}$$

Relación no lineal

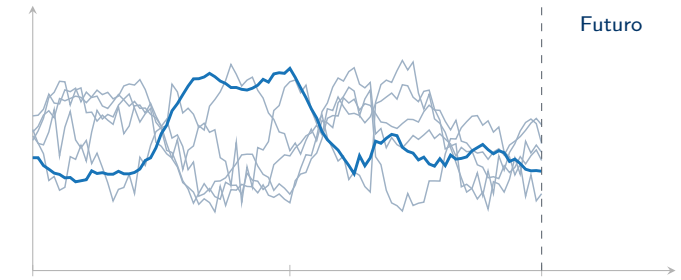
$$y_{t+h} = f(X_t) + \varepsilon_{t+h}$$

Importancia de la regularización y evaluación fuera de muestra

Goulet Coulombe et al. (2022), How Is Machine Learning Useful for Macroeconomic Forecasting?
Medeiros et al. (2021), Forecasting Inflation in a Data-Rich Environment: The Benefits of Machine Learning Methods.

Predicción económica con inteligencia artificial

Desafíos



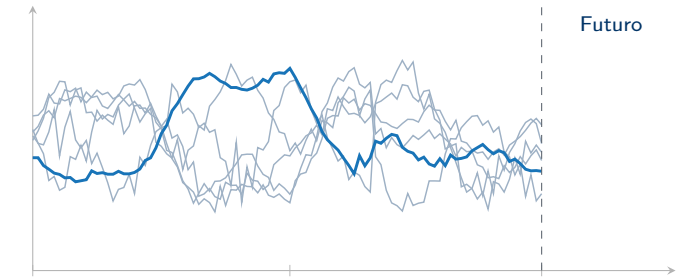
Ejemplo simulado

Muchas series

Actividad, precios, empleo, finanzas

Predicción económica con inteligencia artificial

Desafíos



Ejemplo simulado

Muchas series

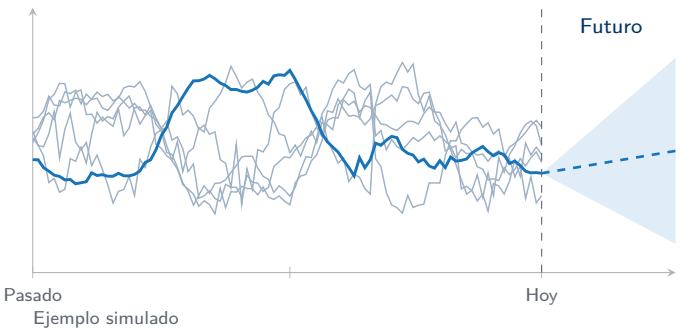
Actividad, precios, empleo, finanzas

Historia limitada

Pocas observaciones y relaciones que cambian

Predicción económica con inteligencia artificial

Desafíos



Muchas series

Actividad, precios, empleo, finanzas

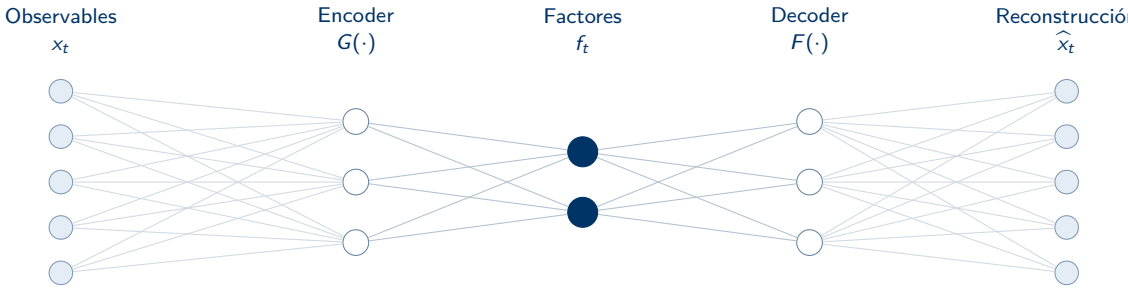
Historia limitada

Pocas observaciones y relaciones que cambian

El reto es extraer una señal que siga siendo útil mañana

Deep Dynamic Factor Models

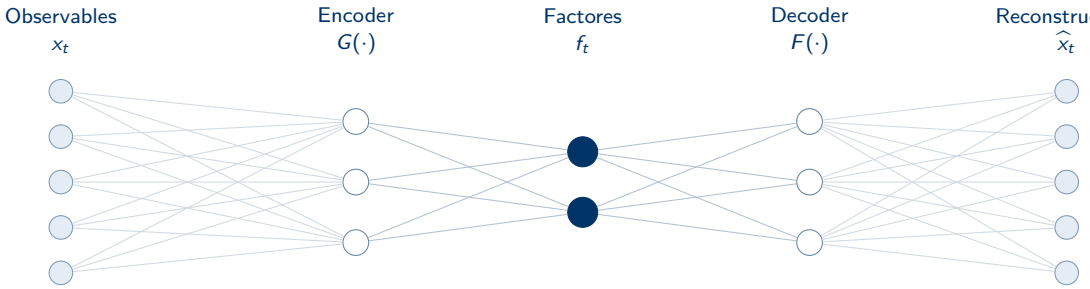
Andreini, Izzo y Ricco: una representación de pocos factores, ahora no lineal



Andreini, Izzo y Ricco (2023), Deep Dynamic Factor Models, CREST Working Paper 2023-08.

Deep Dynamic Factor Models

Andreini, Izzo y Ricco: una representación de pocos factores, ahora no lineal

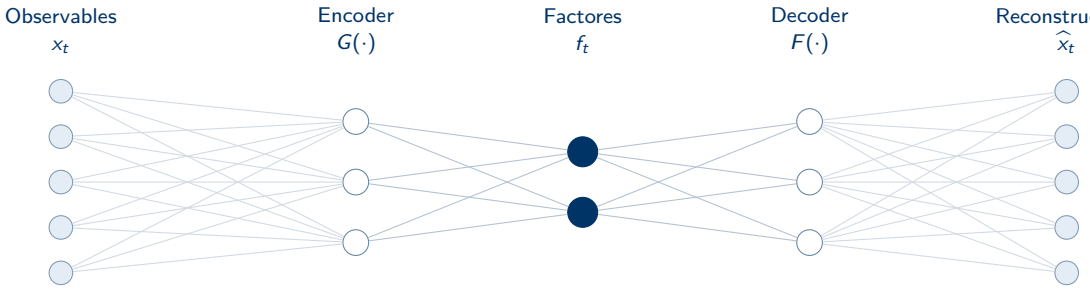


$$f_t = G(x_t) \qquad x_t = F(f_t) + e_t$$

Andreini, Izzo y Ricco (2023), Deep Dynamic Factor Models, CREST Working Paper 2023-08.

Deep Dynamic Factor Models

Andreini, Izzo y Ricco: una representación de pocos factores, ahora no lineal



$$f_t = G(x_t) \qquad x_t = F(f_t) + e_t$$

La estructura factorial permanece y la relación con los datos gana flexibilidad

Andreini, Izzo y Ricco (2023), Deep Dynamic Factor Models, CREST Working Paper 2023-08.

Transformers: combinar información del pasado

Ejemplo de PatchTST: una ventana de observaciones se convierte en un token

1. Dividir la historia

2. Relacionar ventanas

3. Predecir

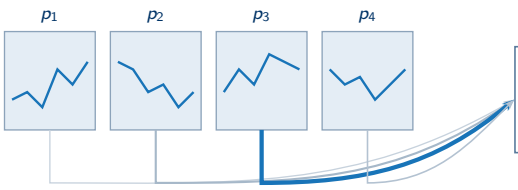
Esquema ilustrativo. Flecha más gruesa: mayor peso de atención.

Nie, Nguyen, Sinthong y Kalagnanam (2023), A Time Series Is Worth 64 Words, ICLR.

Transformers: combinar información del pasado

Ejemplo de PatchTST: una ventana de observaciones se convierte en un token

1. Dividir la historia



Esquema ilustrativo. Flecha más gruesa: mayor peso de atención.

2. Relacionar ventanas

Atención + red no lineal

La atención asigna pesos según el contenido de cada ventana.

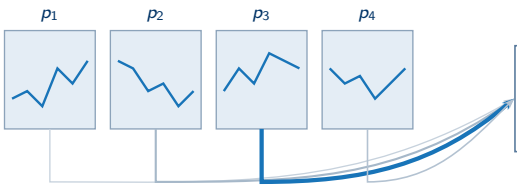
3. Predecir

Nie, Nguyen, Sinthong y Kalagnanam (2023), A Time Series Is Worth 64 Words, ICLR.

Transformers: combinar información del pasado

Ejemplo de PatchTST: una ventana de observaciones se convierte en un token

1. Dividir la historia



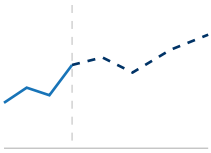
Esquema ilustrativo. Flecha más gruesa: mayor peso de atención.

2. Relacionar ventanas

Atención + red no lineal

La atención asigna pesos según el contenido de cada ventana.

3. Predecir



La salida produce los valores futuros.

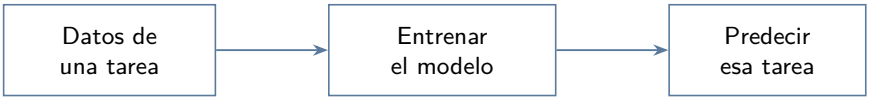
La atención aprende qué partes de la historia combinar para predecir

Nie, Nguyen, Sinthong y Kalagnanam (2023), A Time Series Is Worth 64 Words, ICLR.

Del Transformer al modelo fundacional

La arquitectura procesa los datos; el preentrenamiento permite transferir lo aprendido

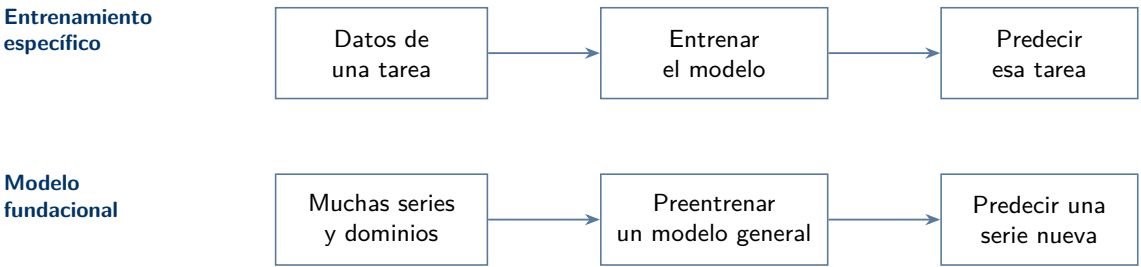
Entrenamiento específico



Ansari et al. (2024), Chronos. Das et al. (2024), TimesFM. Rasul et al. (2024), Lag-Llama.

Del Transformer al modelo fundacional

La arquitectura procesa los datos; el preentrenamiento permite transferir lo aprendido

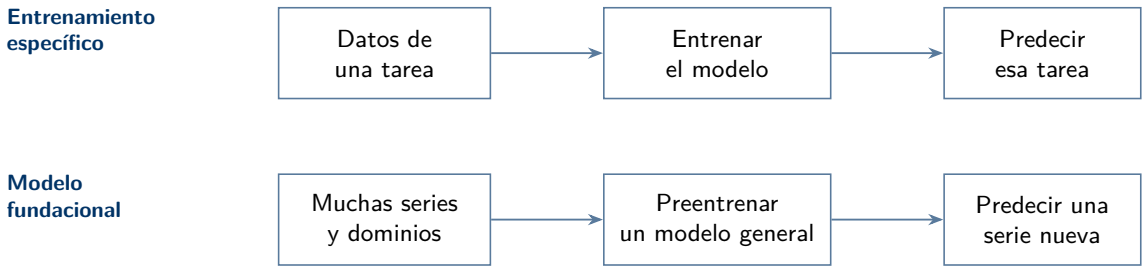


Zero-shot: se aporta su historia, sin volver a ajustar los parámetros.

Ansari et al. (2024), Chronos. Das et al. (2024), TimesFM. Rasul et al. (2024), Lag-Llama.

Del Transformer al modelo fundacional

La arquitectura procesa los datos; el preentrenamiento permite transferir lo aprendido



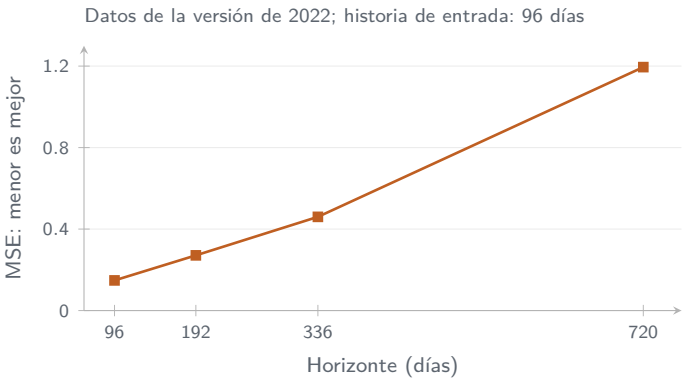
Zero-shot: se aporta su historia, sin volver a ajustar los parámetros.

Ejemplos: Chronos (2024) TimesFM (2024) Lag-Llama (2024)

Ansari et al. (2024), Chronos. Das et al. (2024), TimesFM. Rasul et al. (2024), Lag-Llama.

Más complejidad no garantiza una mejor previsión

Zeng et al. (2023): tipos de cambio, predicción multivariante



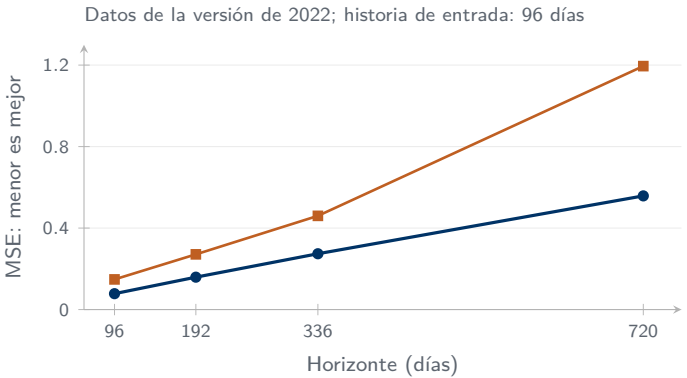
FEDformer

Transformer con descomposición y análisis en frecuencias.

Zeng, Chen, Zhang y Xu (2023), Are Transformers Effective for Time Series Forecasting? arXiv v2 (2022), tabla 2.

Más complejidad no garantiza una mejor previsión

Zeng et al. (2023): tipos de cambio, predicción multivariante



FEDformer

Transformer con descomposición y análisis en frecuencias.

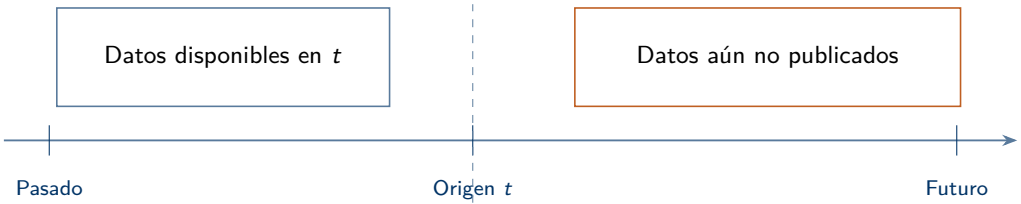
DLinear-S

Modelo lineal sobre tendencia y residuo. Comparte pesos entre las series.

En este experimento, la referencia lineal obtiene menor error

Zeng, Chen, Zhang y Xu (2023), Are Transformers Effective for Time Series Forecasting? arXiv v2 (2022), tabla 2.

La evaluación debe reconstruir qué sabía el analista en cada fecha

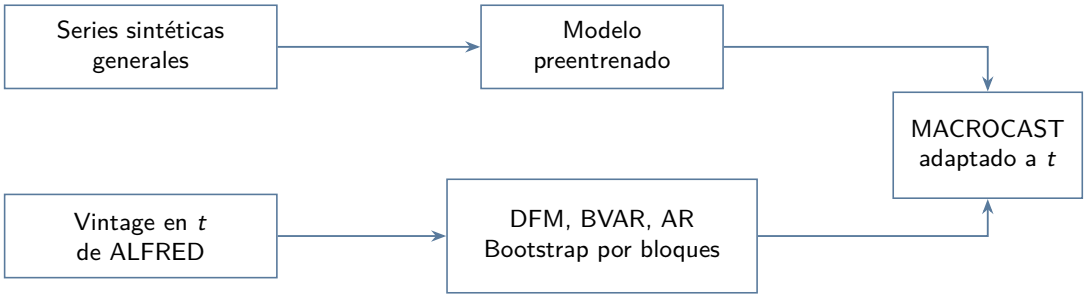


Contaminación temporal
El preentrenamiento puede contener observaciones futuras.

Sesgo de revisión
El dato revisado puede diferir del dato conocido en t.

Preentrenamiento sintético y adaptación a cada vintage histórico

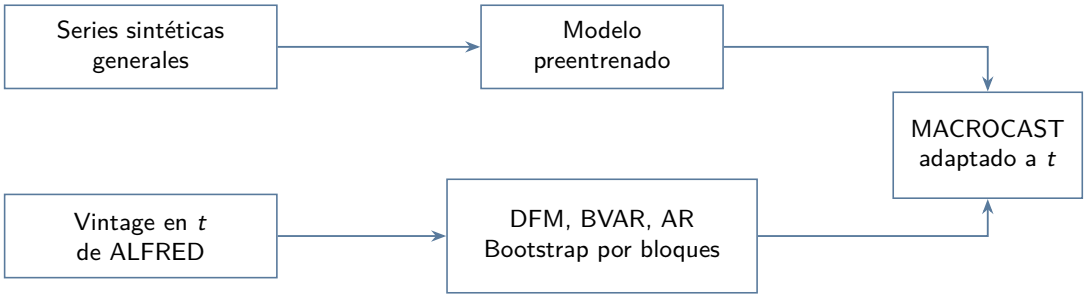
MACROCAST: la econometría enseña a la red



Carriero, Pettenuzzo y Shekhar (2026), MacroCast, arXiv:2606.28670.

Preentrenamiento sintético y adaptación a cada vintage histórico

MACROCAST: la econometría enseña a la red



Información disponible en t tanto al entrenar como al evaluar

Carriero, Pettenuzzo y Shekhar (2026), MacroCast, arXiv:2606.28670.

Una caja de herramientas cada vez más integrada

Econometría

Factores y dinámica conjunta
Regularización y datos en tiempo real

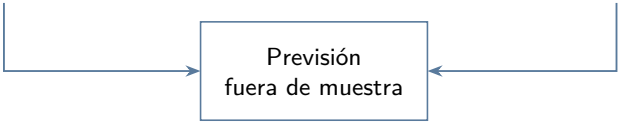
Una caja de herramientas cada vez más integrada

Econometría

Factores y dinámica conjunta
Regularización y datos en tiempo real

Inteligencia artificial

Representaciones no lineales
Aprendizaje transferible entre series



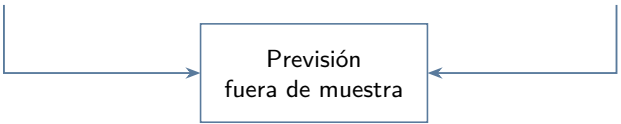
Una caja de herramientas cada vez más integrada

Econometría

Factores y dinámica conjunta
Regularización y datos en tiempo real

Inteligencia artificial

Representaciones no lineales
Aprendizaje transferible entre series



¿Qué supuestos imponemos al modelo
y qué relaciones dejamos que aprenda de los datos?

Muchas gracias

Referencias (1/2)

Stock, J. H. y Watson, M. W. (2002).
Forecasting Using Principal Components from a Large Number of Predictors.
Journal of the American Statistical Association, 97(460), 1167–1179. doi:10.1198/016214502388618960

Giannone, D., Lenza, M. y Primiceri, G. E. (2015).
Prior Selection for Vector Autoregressions.
The Review of Economics and Statistics, 97(2), 436–451. Artículo y versiones

Goulet Coulombe, P., Leroux, M., Stevanovic, D. y Surprenant, S. (2022).
How Is Machine Learning Useful for Macroeconomic Forecasting?
Journal of Applied Econometrics, 37(5), 920–964. doi:10.1002/jae.2910

Medeiros, M. C., Vasconcelos, G. F. R., Veiga, Á. y Zilberman, E. (2021).
Forecasting Inflation in a Data-Rich Environment: The Benefits of Machine Learning Methods.
Journal of Business & Economic Statistics, 39(1), 98–119. doi:10.1080/07350015.2019.1637745

Andreini, P., Izzo, C. y Ricco, G. (2023).
Deep Dynamic Factor Models.
CREST Working Paper 2023-08, versión de 20 de mayo de 2023.

Referencias (2/2)

Nie, Y., Nguyen, N. H., Sinthong, P. y Kalagnanam, J. (2023).
A Time Series Is Worth 64 Words: Long-term Forecasting with Transformers.
International Conference on Learning Representations (ICLR). arXiv:2211.14730

Ansari, A. F. et al. (2024).
Chronos: Learning the Language of Time Series.
Transactions on Machine Learning Research. OpenReview: gerNCVqqtR; arXiv:2403.07815

Das, A., Kong, W., Sen, R. y Zhou, Y. (2024).
A Decoder-Only Foundation Model for Time-Series Forecasting.
Proceedings of the 41st International Conference on Machine Learning, PMLR 235, 10148–10167. Enlace

Rasul, K. et al. (2024).
Lag-Llama: Towards Foundation Models for Probabilistic Time Series Forecasting.
arXiv:2310.08278v3, versión de 8 de febrero de 2024.

Zeng, A., Chen, M., Zhang, L. y Xu, Q. (2023).
Are Transformers Effective for Time Series Forecasting?
Proceedings of the AAAI Conference on Artificial Intelligence, 37(9), 11121–11128. DOI. Datos: arXiv v2 (2022), tabla 2.

Carriero, A., Pettenuzzo, D. y Shekhar, S. (2026).
MacroCast: A Vintage-Consistent Time Series Foundation Model for Real-Time Macroeconomic Forecasting.
arXiv:2606.28670.