

PREDICCIÓN DEL PIB DE COLOMBIA: TÉCNICAS CLÁSICAS VS MACHINE LEARNING.

JONATHAN MONTOYA, VICTOR URIBE, MARIA
CAMILA LOPERA, CRISTIAN CASTRO, KARLA TORRES





ÍNDICE DEL DOCUMENTO

- ENTENDIMIENTO DEL NEGOCIO
Idea del Proyecto
- ENTENDIMIENTO DE LOS DATOS
¿Con que vamos a trabajar?
- PREPARACION DE LOS DATOS
¿Como tratamos los datos para el objetivo?
- MODELADO
¿Qué modelos vamos a usar?
- EVALUACIÓN
¿Como evaluamos el exito de los modelos?

1 — ENTENDIMIENTO DEL NEGOCIO

IDEA DEL PROYECTO

INTRODUCCIÓN

NUESTRA METODOLOGÍA



PIB

El producto interno bruto (PIB) es el valor de mercado de todos los bienes y servicios finales producidos usando los factores de producción disponibles dentro de un país en un periodo determinado.



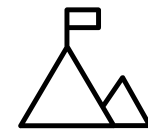
PIB REAL

Al usar los precios de un año base (precios constantes) se conoce como PIB real. Este último es una mejor medida de la actividad económica de un país al medir exclusivamente el cambio en la producción de bienes y servicios en la economía (cantidades), dejando de lado el efecto de las variaciones de los precios.



OBJETIVOS

Comparar Técnicas Clásica y de Machine learning para la predicción del PIB



EJECUCIÓN

Por medio del tratamiento de series de tiempo se van implementar modelos tipo ARIMA, ARIMAX, KNN y VAR para predecir el PIB e identificar cual de estas metodologías posee un mejor desempeño

MARCO TEÓRICO

ARIMA:

Series de tiempo:
Demanda - Fattah, Ezzine,
Aman et al. (2018)

Índice de precios -
Wahyudi (2017)

ARX, SARIMAX

ARX:
Tasa de desempleo a partir
de comentarios - Kwon, K.
Cho, W. and Na, J. (2016)

SARIMAX
Venta de alimentos
perecederos - Arunraj, N.S.,
Ahrens, D., Fernandes, M.
(2016)

KNN/ KNN MULTIVARIANTE

Cross-Validation

Predicción de producción -
Tajmouati, Wahbi, Bedoui
et al. (2021)

Multivariante: Índice de
precios - Ban, , Zhang,
Pang et al. (2013)

REGRESIÓN LINEAL, RIDGE Y LASSO

Fracaso Corporativo -
Pereira, Basto and Silva
(2016)

2 — ENTENDIMIENTO DE LOS DATOS

¿CON QUE VAMOS A TRABAJAR?

ENTENDIMIENTO DE LOS DATOS

Datos trimestrales desde 2007-1 a 2023-1.
65 Observaciones

Variables Macroeconomicas

- PIB REAL
- DESEMPLEO
- IPC
- VARIACIÓN DEL IPC
- CAPACIDAD DE UTILIZACIÓN DE MANUFACTURAS
- TASA DE INTERVENCIÓN BR

Renta Fija

- TES 1 AÑO
- TES 5 AÑOS
- TES 10 AÑOS

Curva Cero Cupon

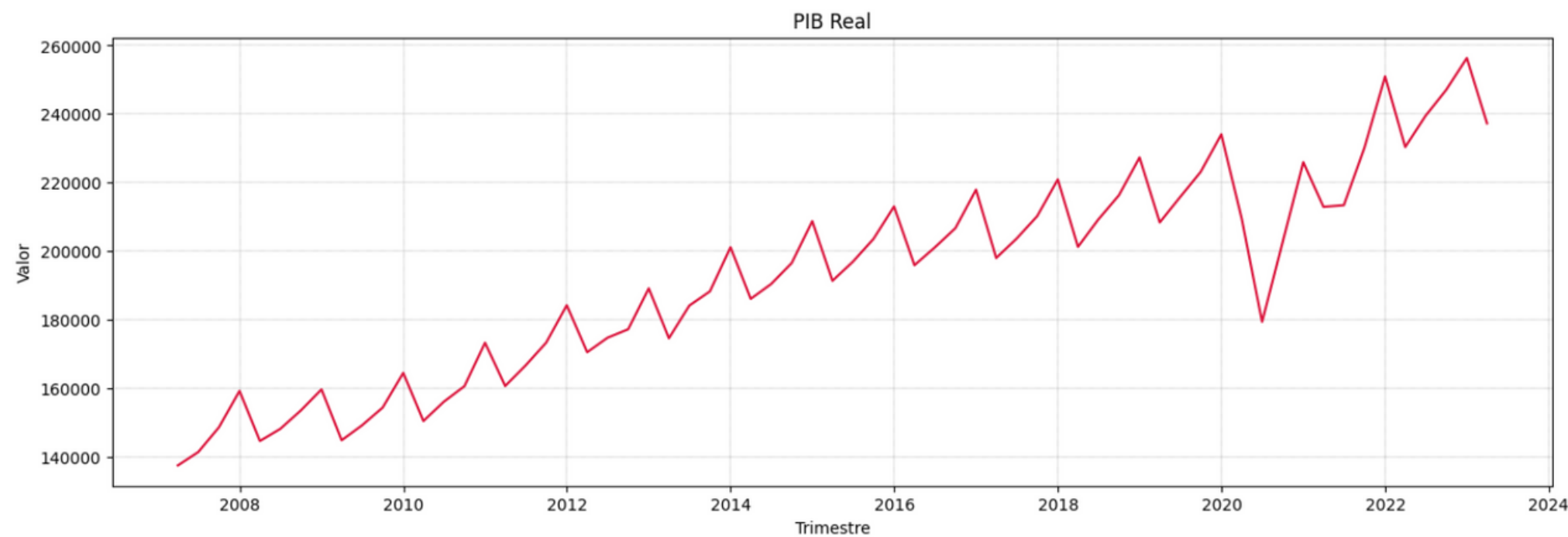
- NIVEL YIELD
- INCLINACIÓN YIELD
- CURVATURA YIELD

Variable Independiente

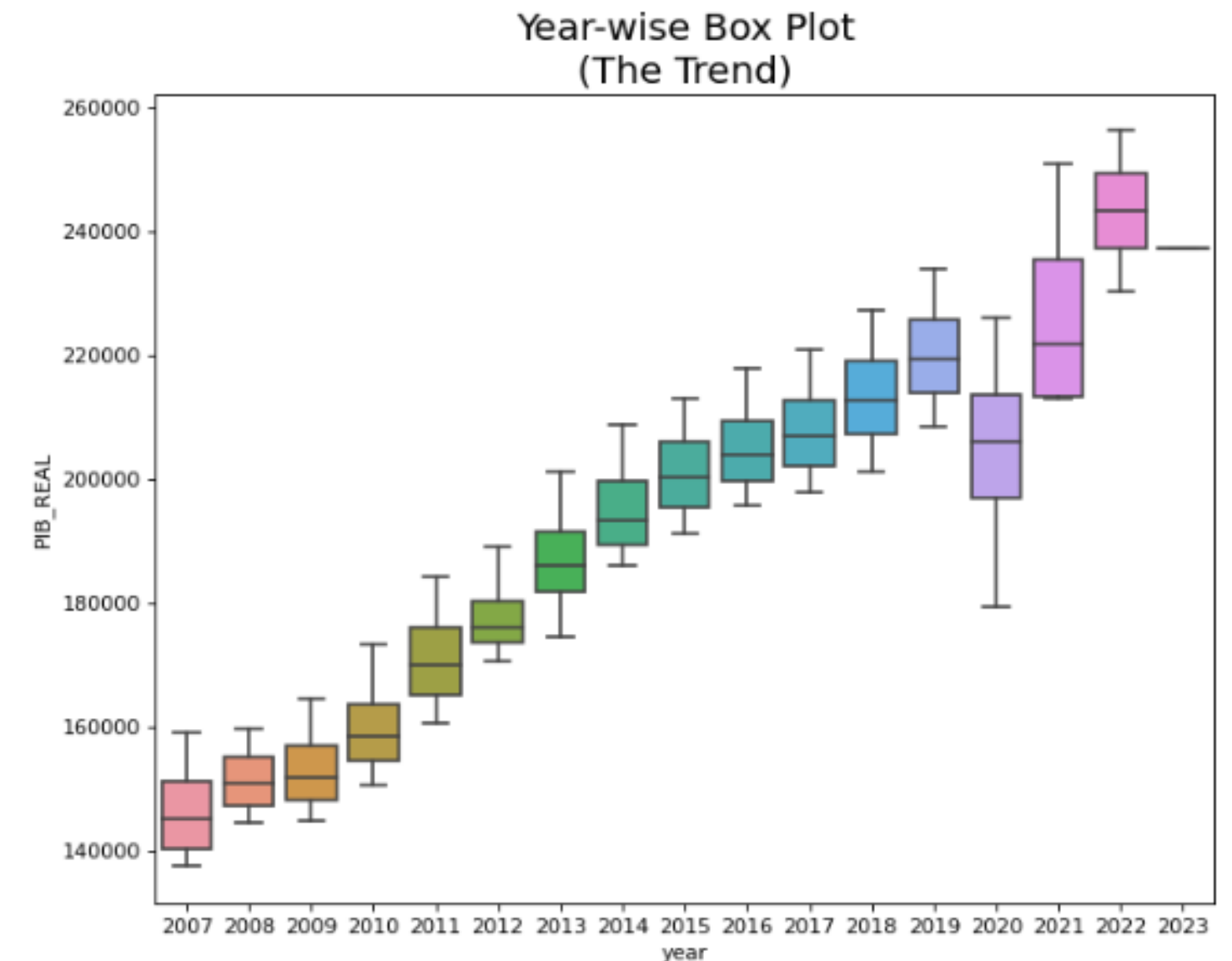
PIB Real

COMPORTAMIENTO DEL PIB REAL

SERIE DE TIEMPO CON TENDENCIA Y ESTACIONAL

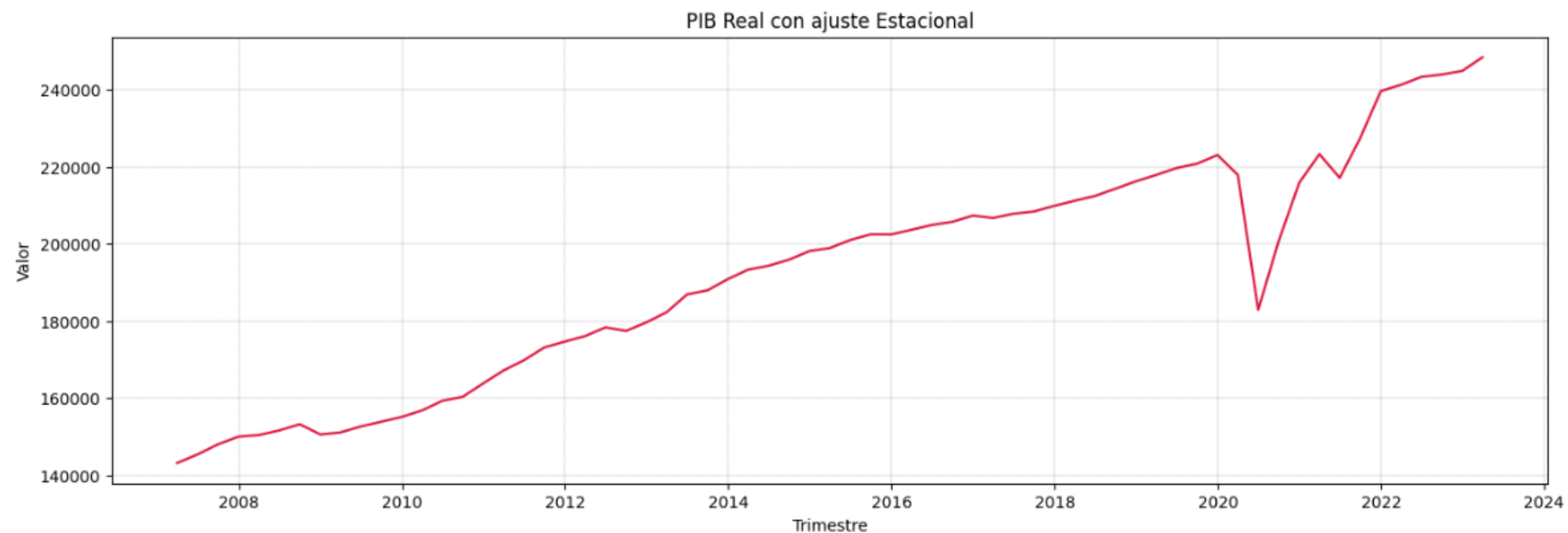


- De la gráfica anual de los Box-Plots se observa de la misma manera que existe una tendencia alcista durante los años teniendo excluyendo el año de 2020 en el cual existió una externalidad, Covid-19, que afectó en gran medida el PIB Real.
- Se observa como existe un componente estacional sobre todo en los meses de Septiembre y Diciembre,
- No hay outliers en la serie debido a que el mismo Box-Plot en ambos casos no los identificó.

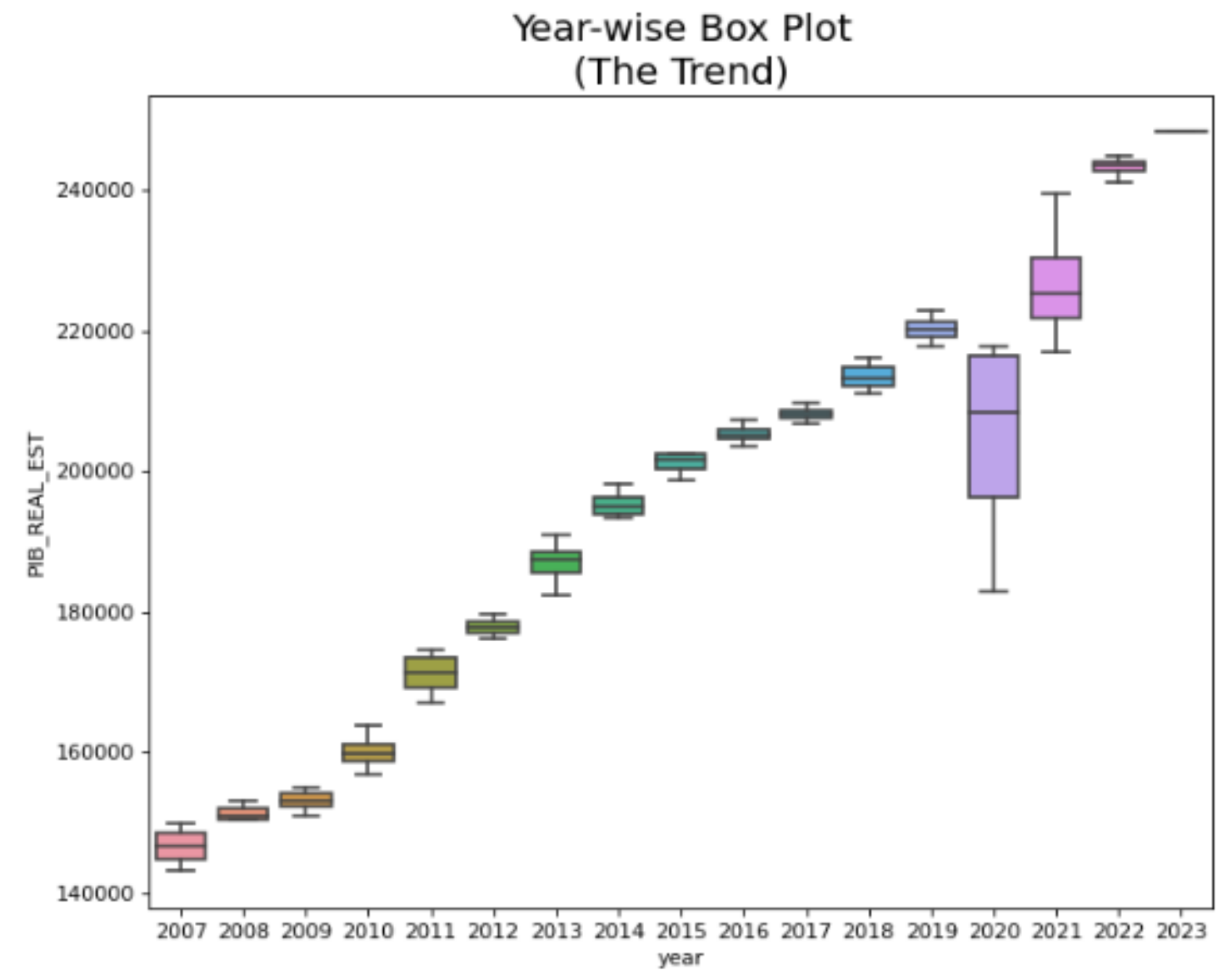


COMPORTAMIENTO DEL PIB REAL

CON AJUSTE ESTACIONAL



- Se observa como el componente tendencial sigue estando latente pero el componente estacional no es tan notorio lo cual se debe al ajuste hecho por el Banco de la República.



3 – TRATAMIENTO DE LOS DATOS

¿CÓMO TRATAMOS LOS DATOS PARA EL OBJETIVO?

TRATAMIENTOS DE LOS DATOS

ESTACIONARIZACIÓN DE LAS VARIABLES PIB REAL Y PIB REAL AJUSTADO
(CON AJUSTE ESTACIONAL)

Una de las partes mas importantes para trabajar con series de tiempo es estacionarizar la serie ya que de esta manera se busca que la media y la varianza de la serie sean constantes en el tiempo.

Por medio de las siguientes técnicas se busca estacionarizar las variables objetivo : PIB Real y PIB Real Ajustado

DIFERENCIAS

Diferencia de entre cada elemento n y $n-1$

DIFERENCIAS LOGARITMICAS

Diferencia de entre cada elemento n y $n-1$ de la serie logarítmica

LOGARITMOS

Logaritmo natural de la serie obojtivo

RAIZ CUADRADA

Raiz cuadrada de la serie objetivo

4 – MODELADO

MODELADO

¿QUÉ MODELOS VAMOS A USAR?

ITécnicas tradicionales y de Machine Learning
propuestas para predecir el PIB

Las métricas seleccionadas para comparar los
modelos son: MSE y MAPE

1

ARIMA

Con 2 y 4 resagos

2

ARX Y SARIMAX

Con 2 y 4 resagos

3

REGRESIONES

Lineal, Lasso y Ridge

4

SVM

.

5

POLINOMIAL

.

6

KNN

uni variante y multivariante

5 – EVALUACIÓN

¿COMO EVALUAMOS EL EXITO DE LOS MODELOS?

RESULTADOS

1

ARIMA

ESCENARIO	MSE	MAPE
Escenario 1	5.85×10^{-3}	99%
Escenario 2	22×10^{-3}	108%
Escenario 3	4.143×10^{-3}	108%
Escenario 4	4.1311×10^{-3}	108%
Escenario 5	5012×10^{-3}	108%
Escenario 6	3.63×10^{-3}	108%

- Escenario 1: Curva cero coupon
- Escenario 2: Curva cero coupon, Macro-variables
- Escenario 3: Macro-Variables
- Escenario 4: Proxies
- Escenario 5: Macrovariables, Proxies
- Escenario 6: Curva cero coupon, Macro-Variables, Proxies

RESULTADOS

2

ARX Y SARIMAX

ESCENARIO	MSE	MAPE
Escenario 1	5.85×10^{-3}	99%
Escenario 2	22×10^{-3}	108%
Escenario 3	4.143×10^{-3}	108%
Escenario 4	4.1311×10^{-3}	108%
Escenario 5	5012×10^{-3}	108%
Escenario 6	3.63×10^{-3}	108%

- Escenario 1: Curva cero coupon
- Escenario 2: Curva cero coupon, Macro-variables
- Escenario 3: Macro-Variables
- Escenario 4: Proxies
- Escenario 5: Macrovariables, Proxies
- Escenario 6: Curva cero coupon, Macro-Variables, Proxies

RESULTADOS

3

REGRESIONES

Lineal, Lasso y Ridge

ESCENARIO	REGRESION	MSE	MAPE	R^2
Escenario 1	Lineal	3.45×10^{-3}	76%	0.056
Escenario 2	Lasso	4.178×10^{-3}	88%	-0.0017
Escenario 3	Lasso	4.189×10^{-3}	72%	-0.0017
Escenario 4	Ridge	3.9×10^{-3}	72%	0.052
Escenario 5	Ridge	3.84×10^{-3}	224%	0.081
Escenario 6	Ridge	3.908×10^{-3}	88%	0.065

- Escenario 1: Curva cero coupon
- Escenario 2: Curva cero coupon, Macro-variables
- Escenario 3: Macro-Variables
- Escenario 4: Proxies
- Escenario 5: Macrovariables, Proxies
- Escenario 6: Curva cero coupon, Macro-Variables, Proxies

RESULTADOS

3

REGRESIONES

Lineal, Lasso y Ridge

ESCENARIO	REGRESION	MSE	MAPE	R^2
Escenario 1	Lineal	3.45×10^{-3}	76%	0.056
Escenario 2	Lasso	4.178×10^{-3}	88%	-0.0017
Escenario 3	Lasso	4.189×10^{-3}	72%	-0.0017
Escenario 4	Ridge	3.9×10^{-3}	72%	0.052
Escenario 5	Ridge	3.84×10^{-3}	224%	0.081
Escenario 6	Ridge	3.908×10^{-3}	88%	0.065

- Escenario 1: Curva cero coupon
- Escenario 2: Curva cero coupon, Macro-variables
- Escenario 3: Macro-Variables
- Escenario 4: Proxies
- Escenario 5: Macrovariables, Proxies
- Escenario 6: Curva cero coupon, Macro-Variables, Proxies

RESULTADOS

4

SVM

ESCENARIO	MSE	MAPE
Escenario 1	4.2×10^{-3}	98%
Escenario 2	4.294×10^{-3}	97%
Escenario 3	4.286×10^{-3}	97%
Escenario 4	4.286×10^{-3}	108%
Escenario 5	4.286×10^{-3}	108%
Escenario 6	4.286×10^{-3}	108%

- Escenario 1: Curva cero coupon
- Escenario 2: Curva cero coupon, Macro-variables
- Escenario 3: Macro-Variables
- Escenario 4: Proxies
- Escenario 5: Macrovariables, Proxies
- Escenario 6: Curva cero coupon, Macro-Variables, Proxies

RESULTADOS

5

POLINOMIAL

ESCENARIO	MSE	MAPE
Escenario 1	5.85×10^{-3}	99%
Escenario 2	22×10^{-3}	108%
Escenario 3	4.143×10^{-3}	108%
Escenario 4	4.1311×10^{-3}	108%
Escenario 5	5012×10^{-3}	108%
Escenario 6	3.63×10^{-3}	108%

- Escenario 1: Curva cero coupon
- Escenario 2: Curva cero coupon, Macro-variables
- Escenario 3: Macro-Variables
- Escenario 4: Proxies
- Escenario 5: Macrovariables, Proxies
- Escenario 6: Curva cero coupon, Macro-Variables, Proxies

RESULTADOS

6

KNN

ESCENARIO	MSE	MAPE
Escenario 1	5.85×10^{-3}	99%
Escenario 2	22×10^{-3}	108%
Escenario 3	4.143×10^{-3}	108%
Escenario 4	4.1311×10^{-3}	108%
Escenario 5	5012×10^{-3}	108%
Escenario 6	3.63×10^{-3}	108%

- Escenario 1: Curva cero coupon
- Escenario 2: Curva cero coupon, Macro-variables
- Escenario 3: Macro-Variables
- Escenario 4: Proxies
- Escenario 5: Macrovariables, Proxies
- Escenario 6: Curva cero coupon, Macro-Variables, Proxies

**MUCHAS
GRACIAS POR
SU ATENCIÓN**