

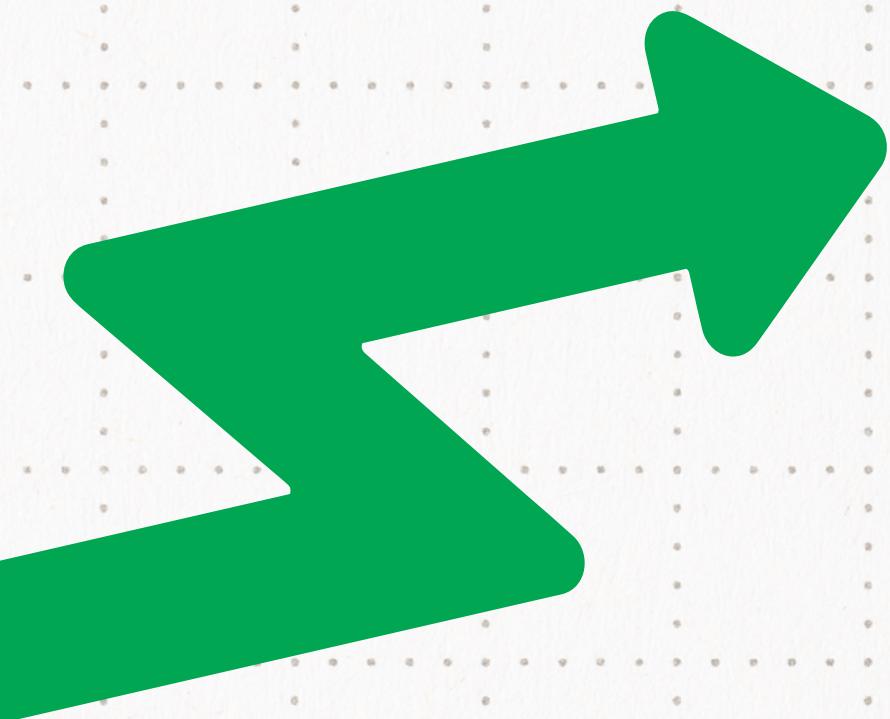
Kernel Ridge Regression para predicir IPC y variables climáticas

Equipo :

- Cázares Trejo Leonardo Damián
- Malváez Flores Axel Daniel
- Peralta Rionda Gabriel Zadquiel



Agenda



- 01 Introducción
- 02 Metodología
- 03 Resultados
- 04 Discusión
- 05 Conclusión

Introducción

Objetivo General

Abordar un **análisis integral** para poder comprender las **relaciones** dinámicas entre diversas **variables** clave en México y poder explorar nuevas variables de impacto que muy posiblemente se están pasando por alto debido a su baja asociación por pertenecer a **diferentes áreas** de interés.

Perspectiva Económico-Ambiental

La investigación se centra en dos enfoques: **primero, determinar** si las condiciones **climáticas**, la **movilidad vehicular** y el **desempeño económico** (IVF) pueden **prever** de manera significativa las **variaciones** en el **IPC** nacional. **Segundo**, evaluar si el **IPC**, el número de **automóviles** y el **IVF** pueden servir como indicadores **predictivos** para las variables **climáticas**.

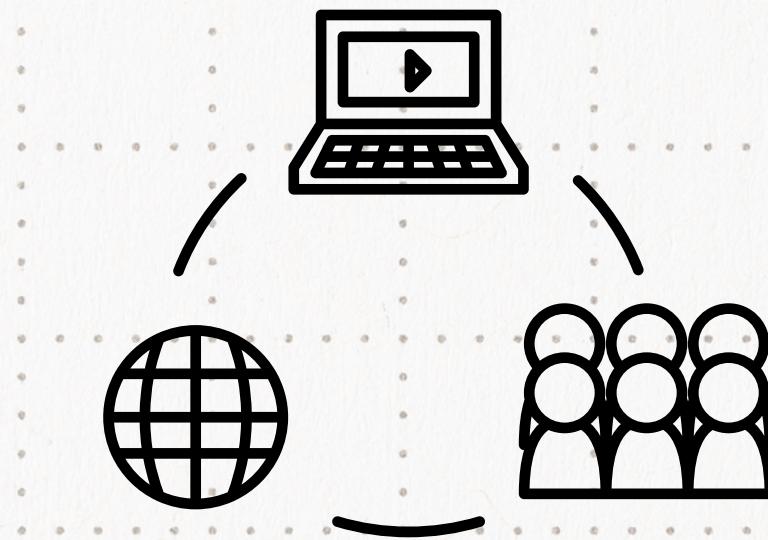
Metodología

$$\hat{y} = f(x_1, x_2, \dots, x_n)$$

Kernel Ridge Regression (KRR) [1]



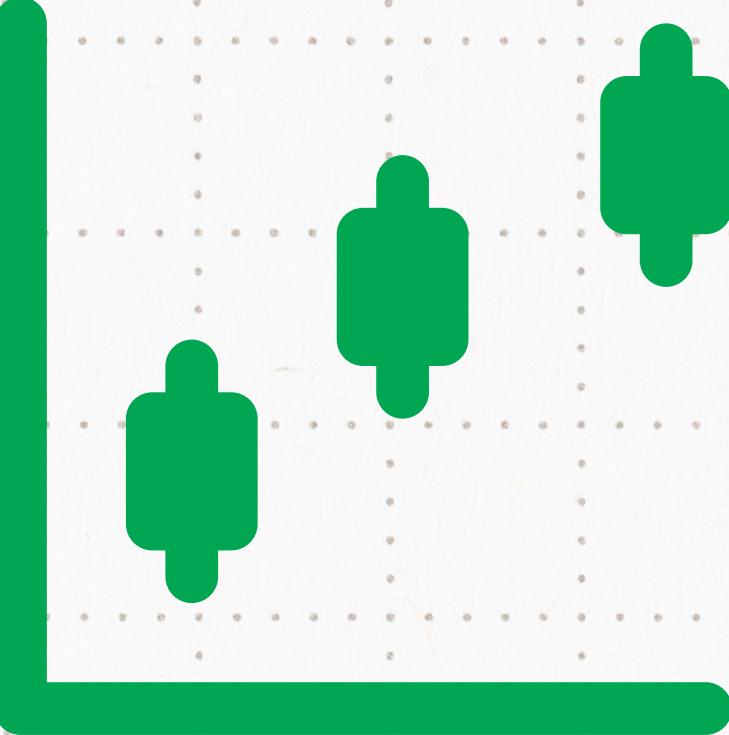
No-lineal



Fácil de
entrenar



Calcula
similitud



Entrenamiento

Kernel

$$\kappa(x_i, x_j) = e^{-\frac{\|x_i - x_j\|^2}{2\sigma^2}}$$

Matriz de similitud

$$[K]_{ij} = \kappa(x_i, x_j)$$

Resolver sistema de ecuaciones = Invertir matriz

$$(K + \lambda I) \vec{a} = (y_1, \dots, y_n)^T$$

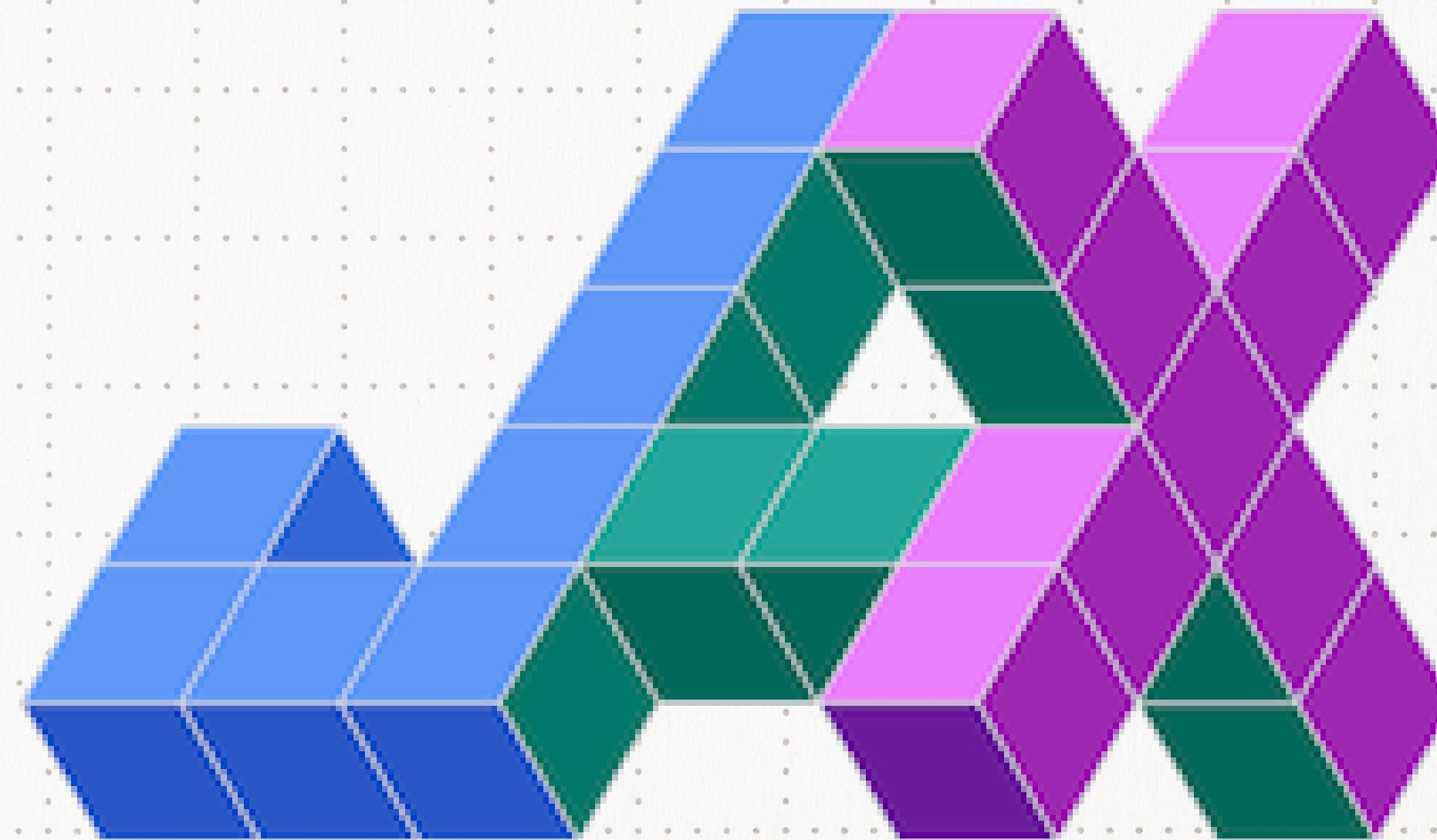
Predictión

$$\hat{y}(x) = \sum_{i=1}^n \alpha_i K(x, x_i)$$



¿Como lo logramos?...

...¡Usando JAX!



01

Conjunto de datos IPC y PIB (Trimestrales)

Datos del PIB e IPC de manera trimestral desde 1980 - cuarto trimestre del año hasta el año 2023 - segundo trimestre del año. Por lo que se emparejaron los datos de PIB e IPC con respecto a sus respectivos años.

02

Conjunto de datos clima (lluvia y temp), transporte e IPC

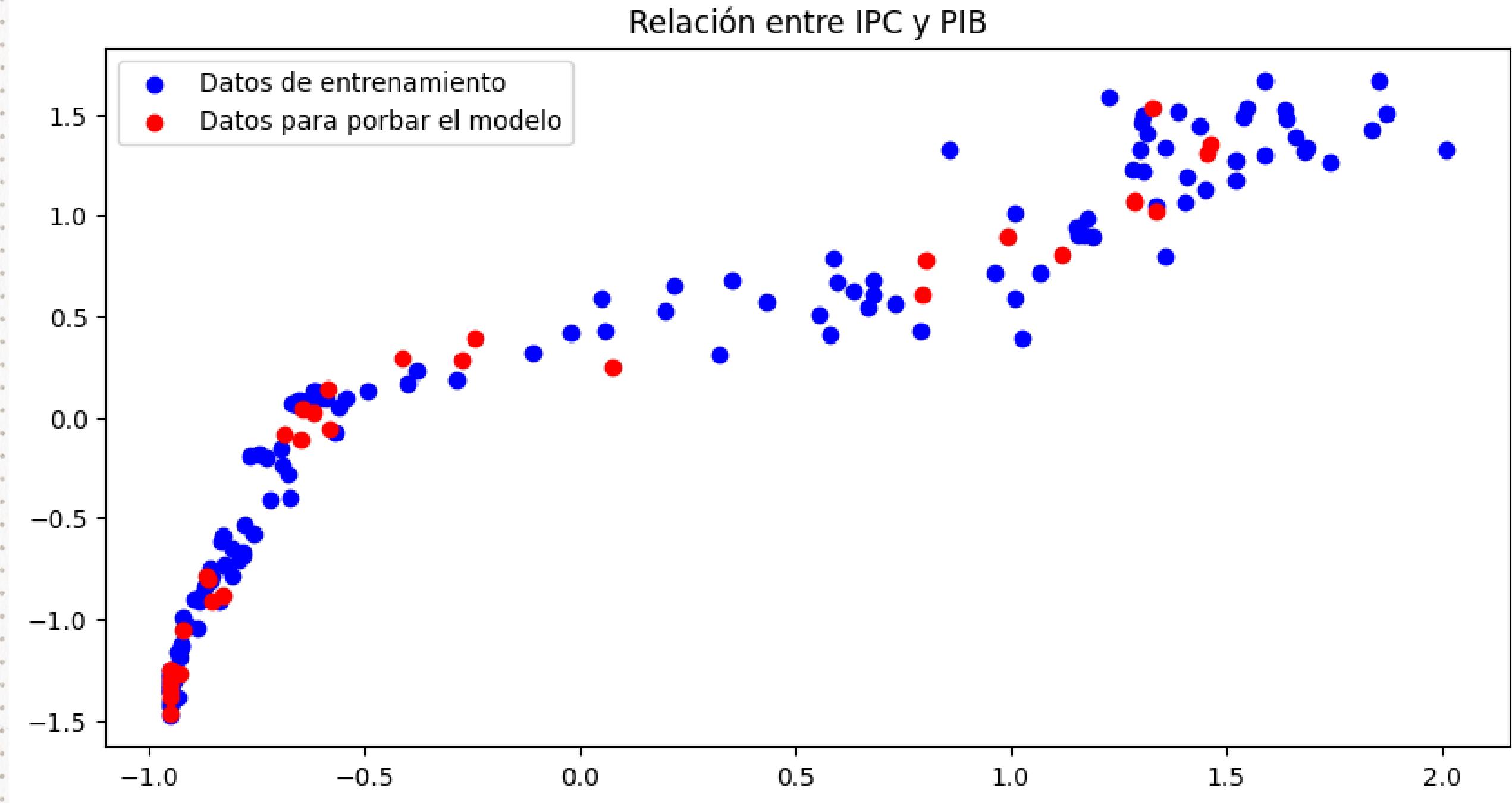
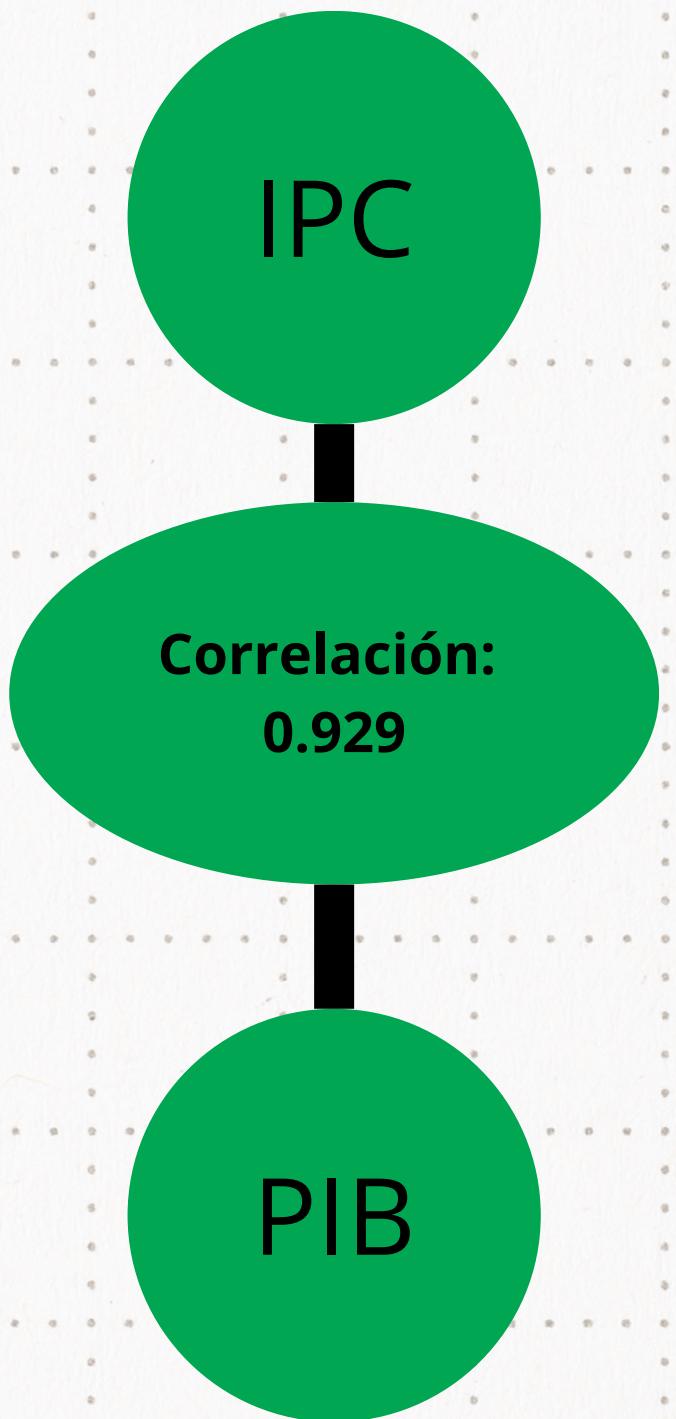
Variables climáticas, de transporte e IPC. Para esto obtuvimos un conjunto de datos comprendidos entre Enero de 1991 a Agosto de 2023, es decir 32 años de información mensual.

03

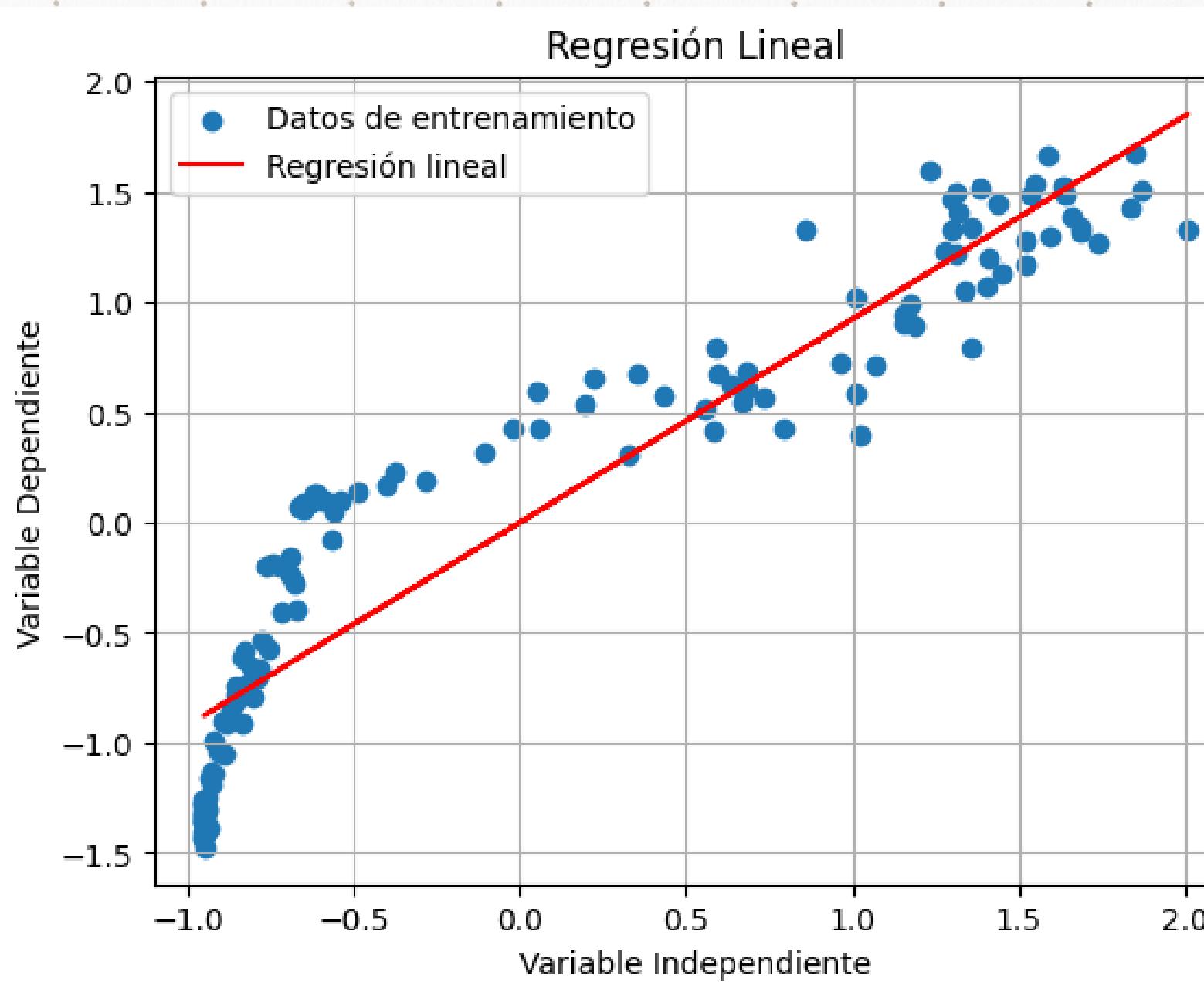
Conjunto de datos clima (lluvia y temp), transporte, IVF e IPC

Conjunto de datos anterior concatenado a las variables del PIB mensual, sin embargo aunque no pudimos conseguir el conjunto de datos de PIB mensual, utilizamos el IVF por estado, ahora el conjunto de datos va desde Enero de 1991 a Junio de 2023.

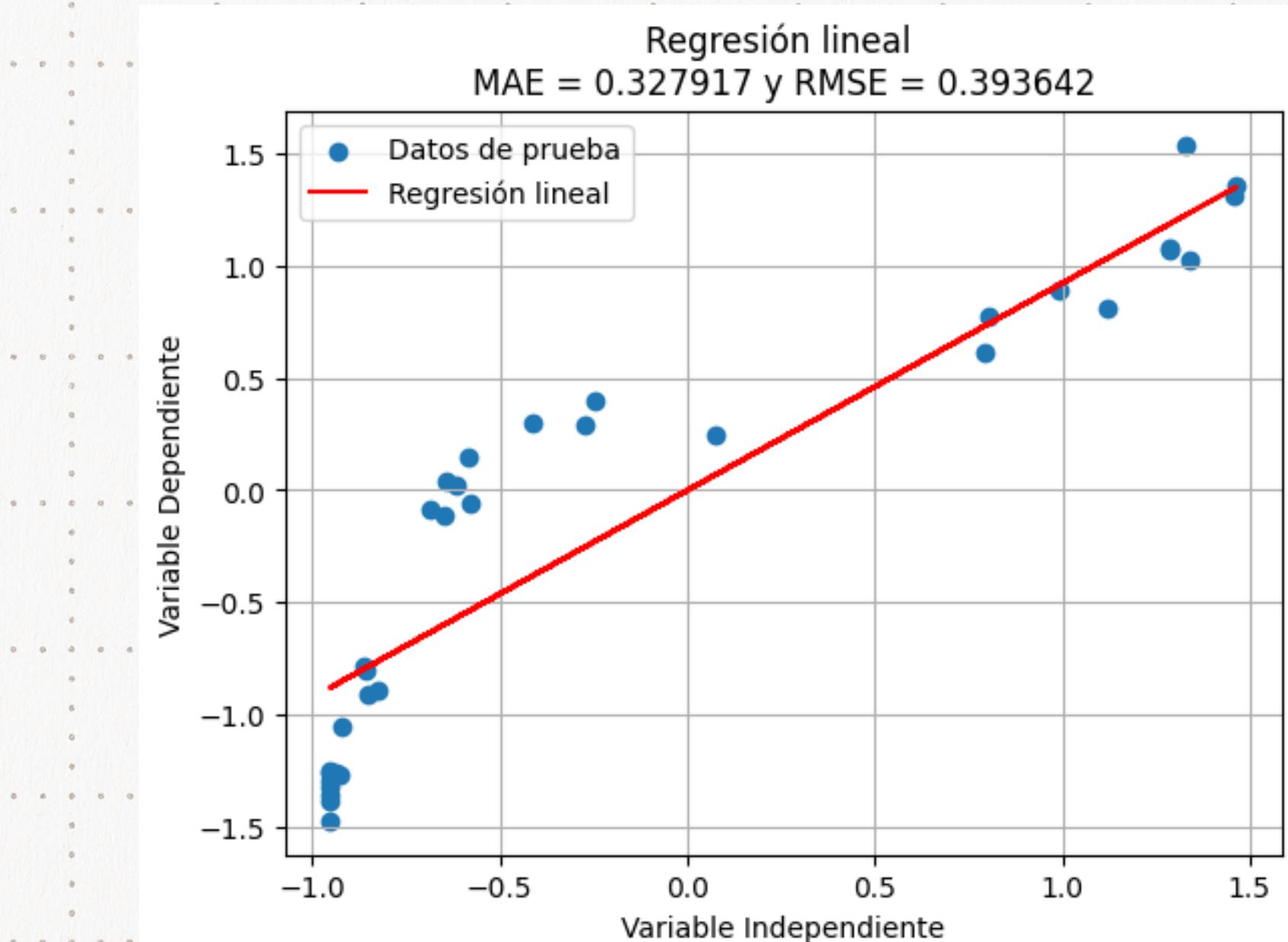
Resultados



PIB vs IPC

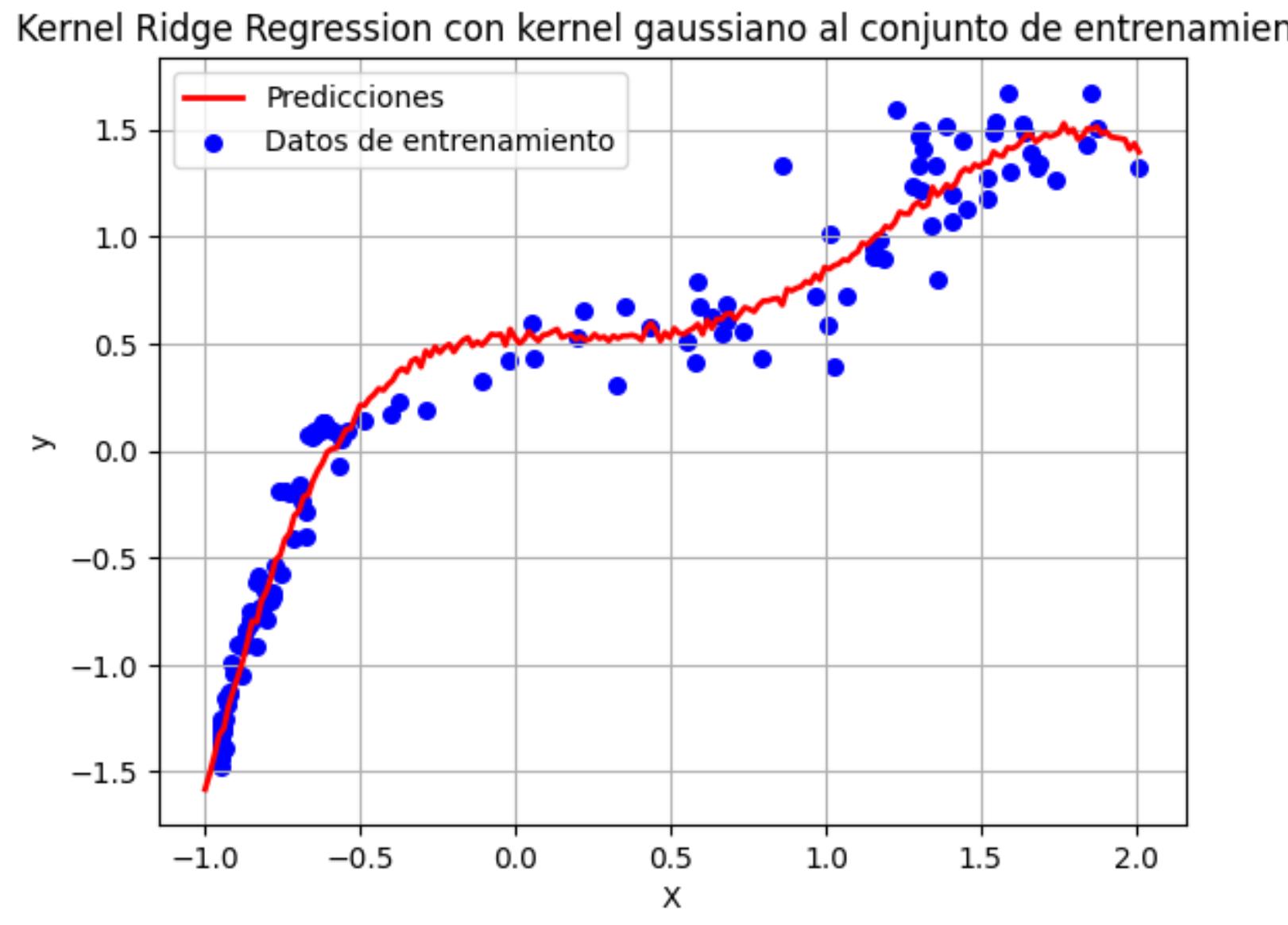


Conjunto de entrenamiento

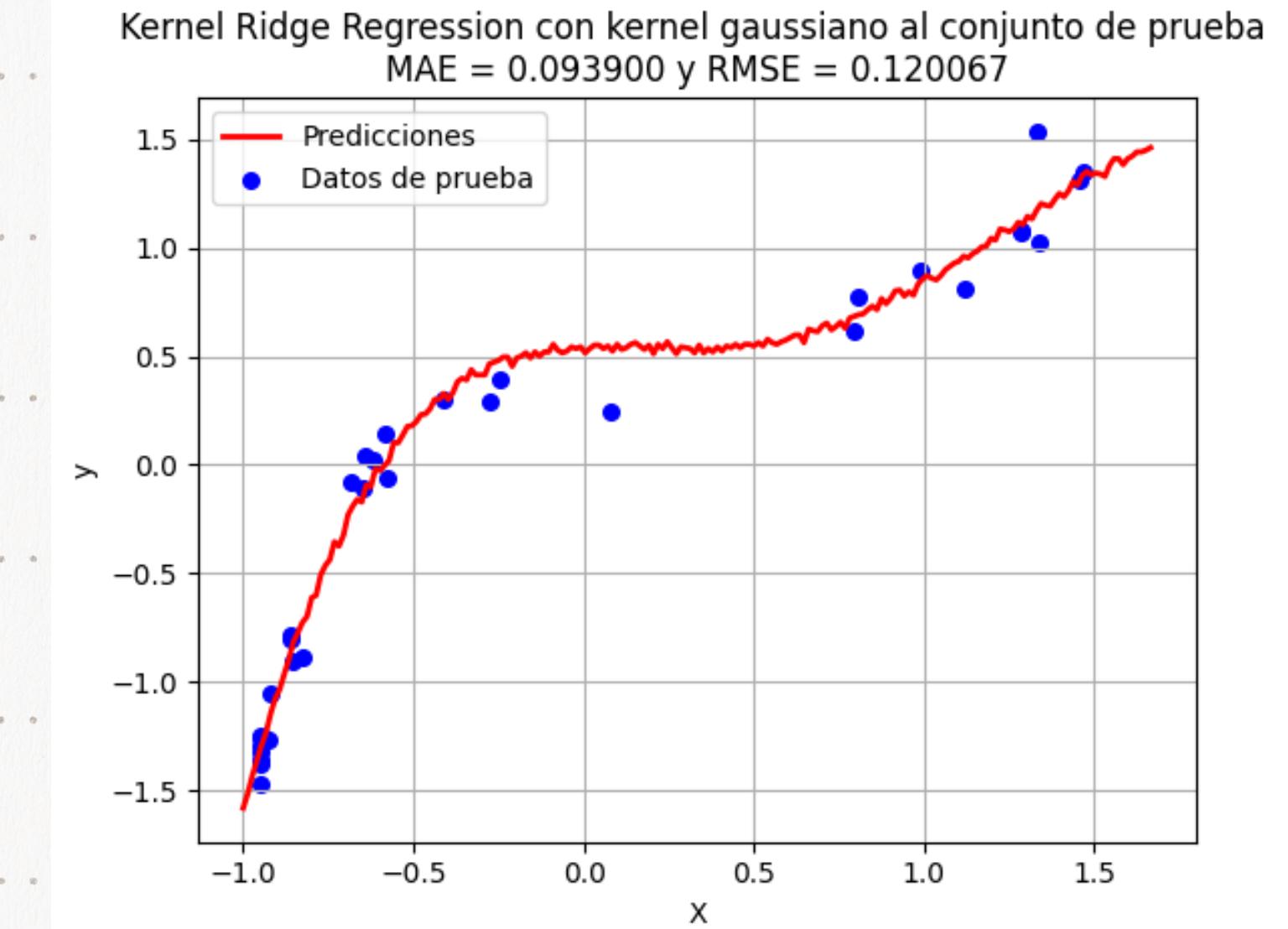


Conjunto de prueba

PIB vs IPC



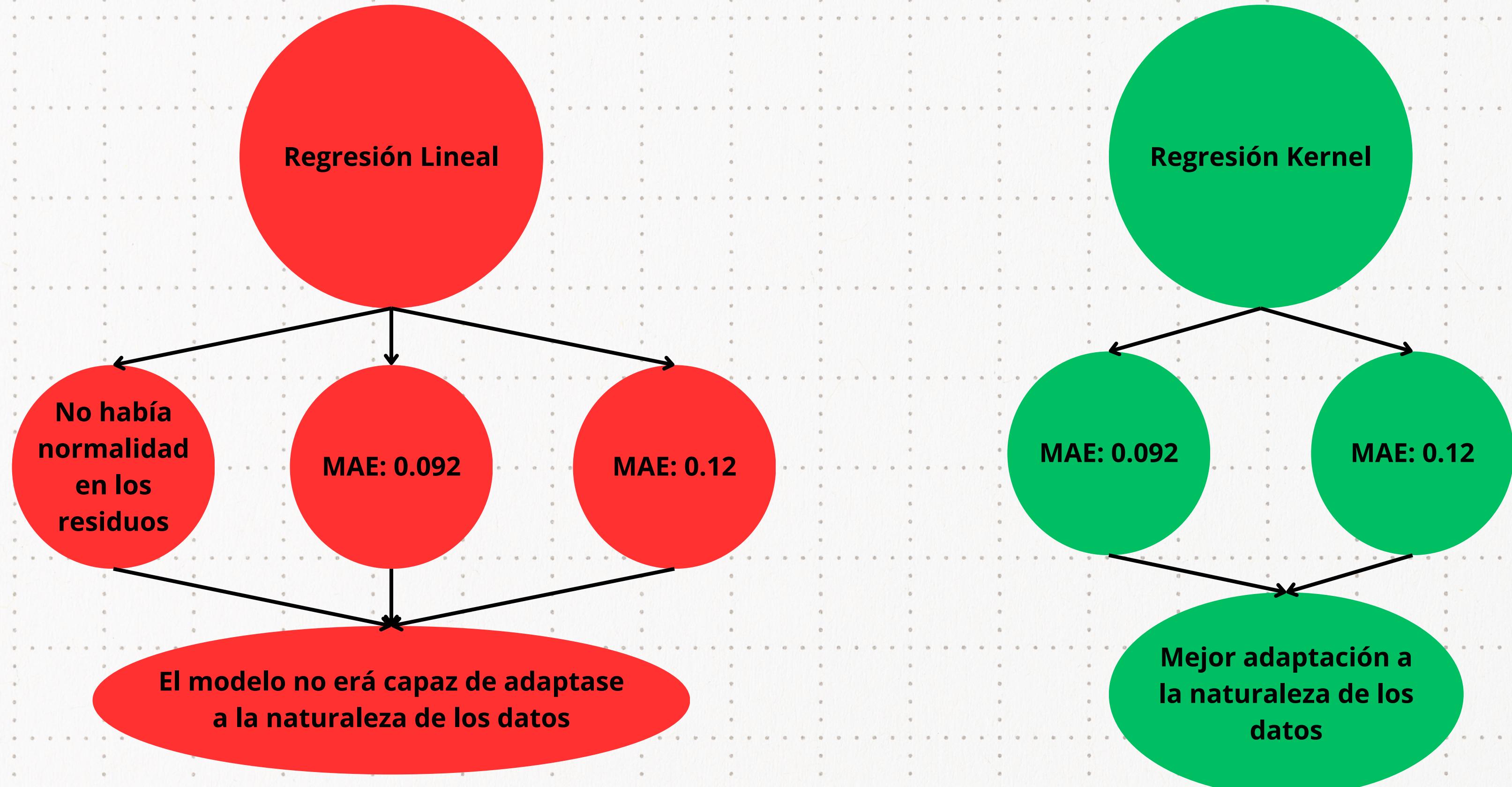
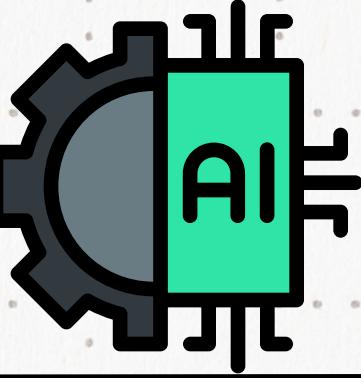
Conjunto de entrenamiento



Conjunto de prueba

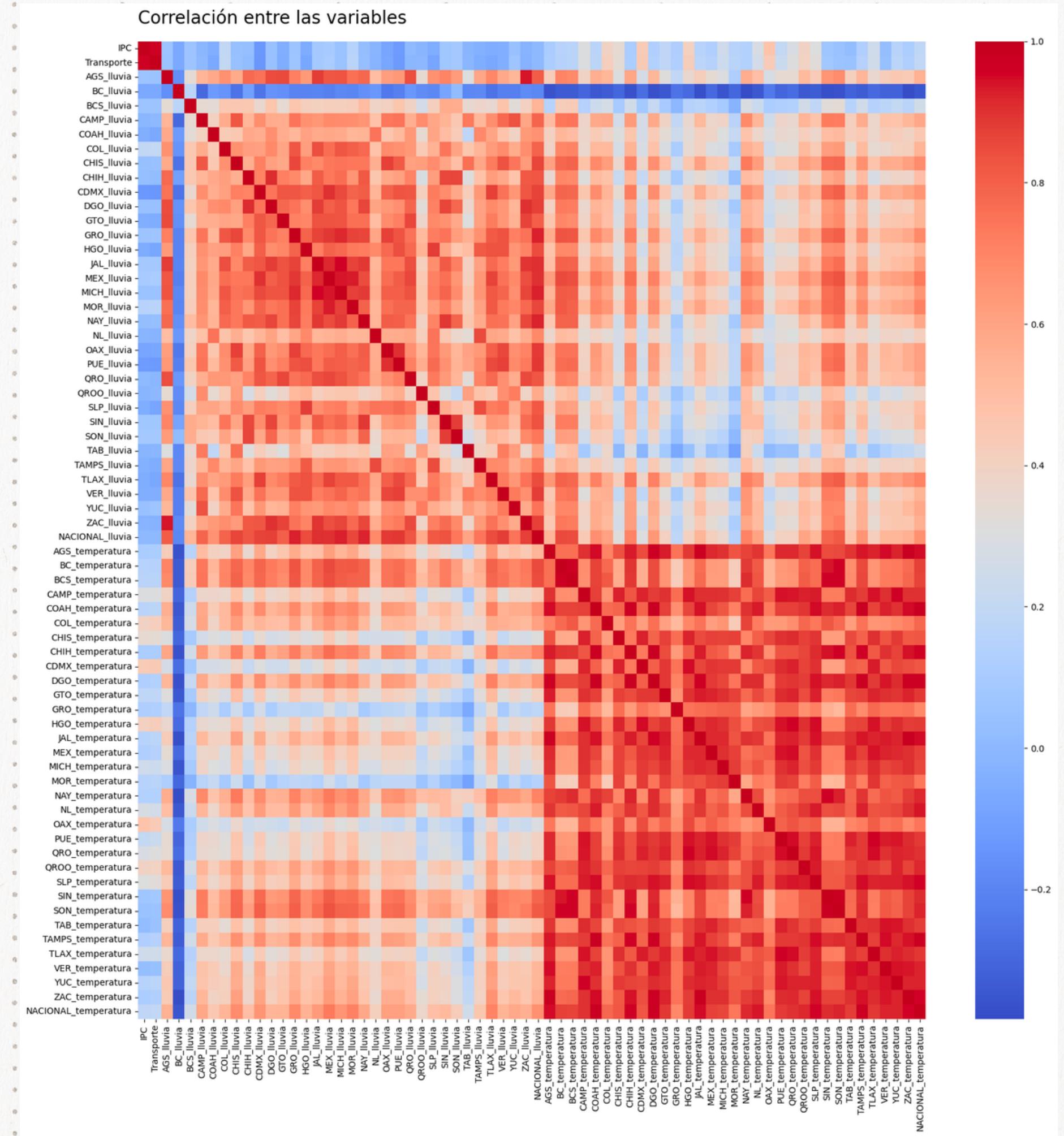
Discusión:

Resultados de la Regresión Lineal vs. Regresión Kernel



Correlaciones

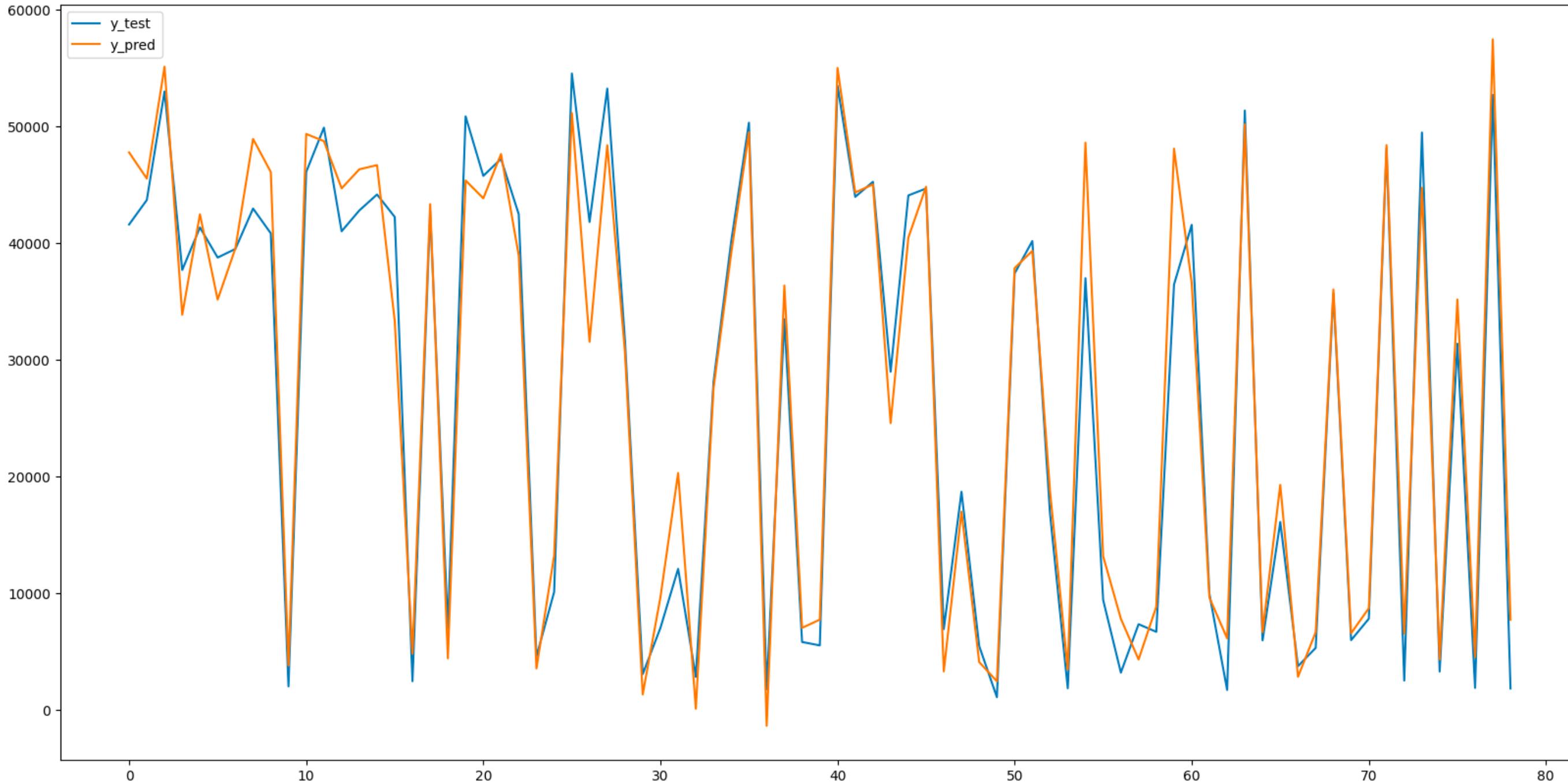
Clima y Transporte vs IPC



Modelos

Clima y Transporte vs IPC

Predicciones sobre el IPC



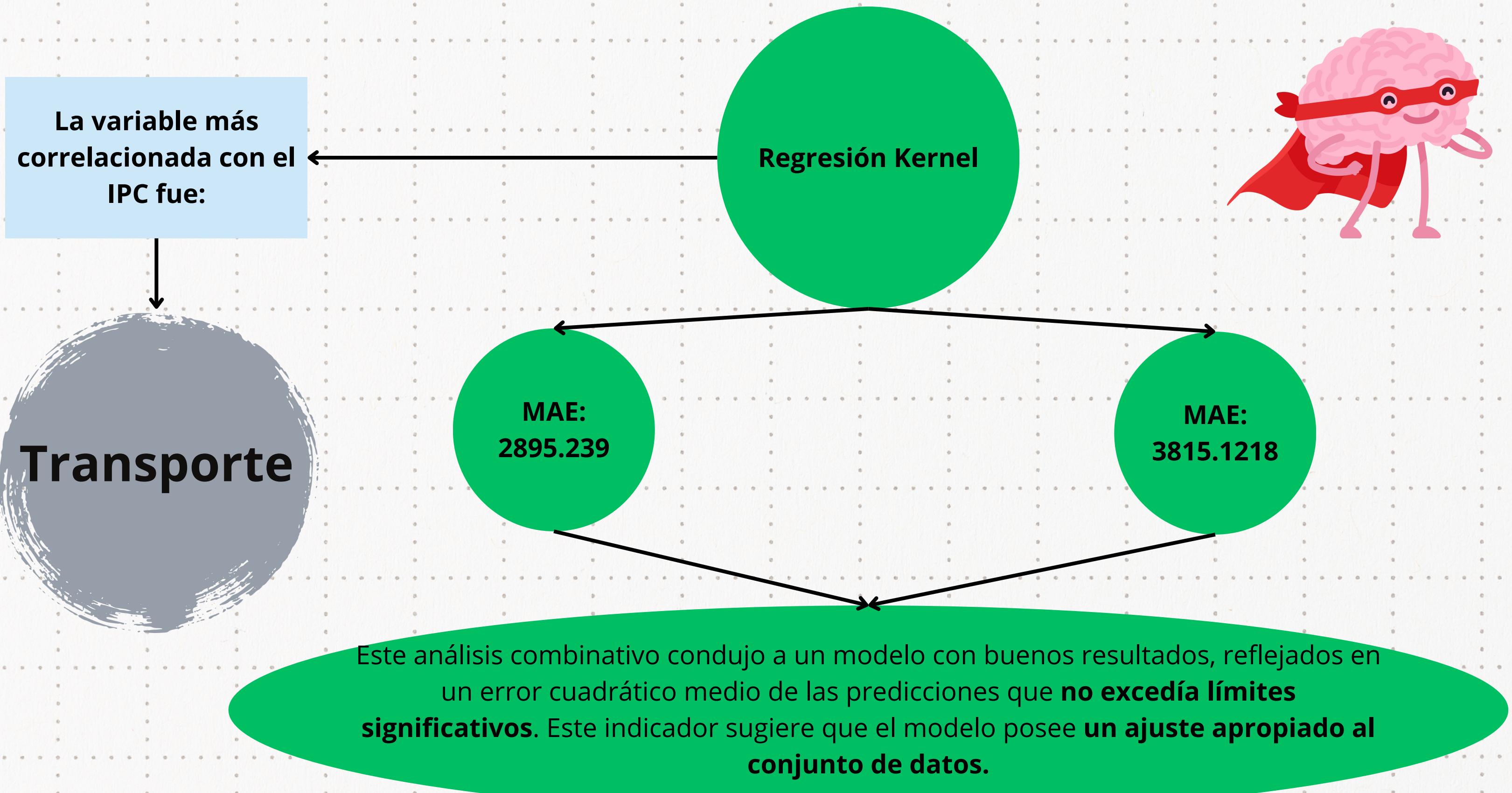
2895.239

MAE

3815.1218

RMSE

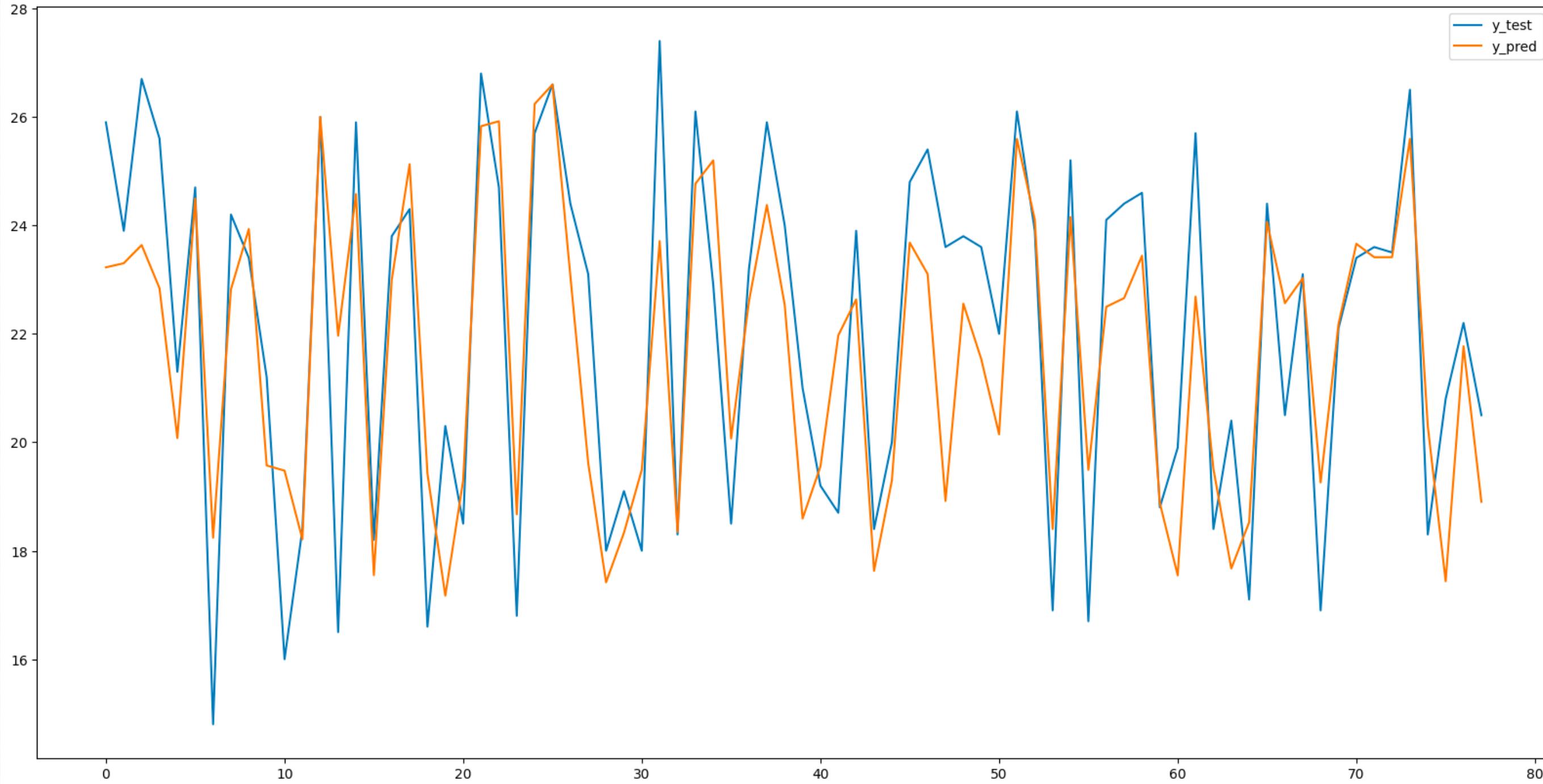
Discusión: Clima y Transporte vs IPC



Modelos

IVF, IPC y coches vs Temperatura Nacional

Predicciones sobre la temperatura nacional



1.5233759

MAE

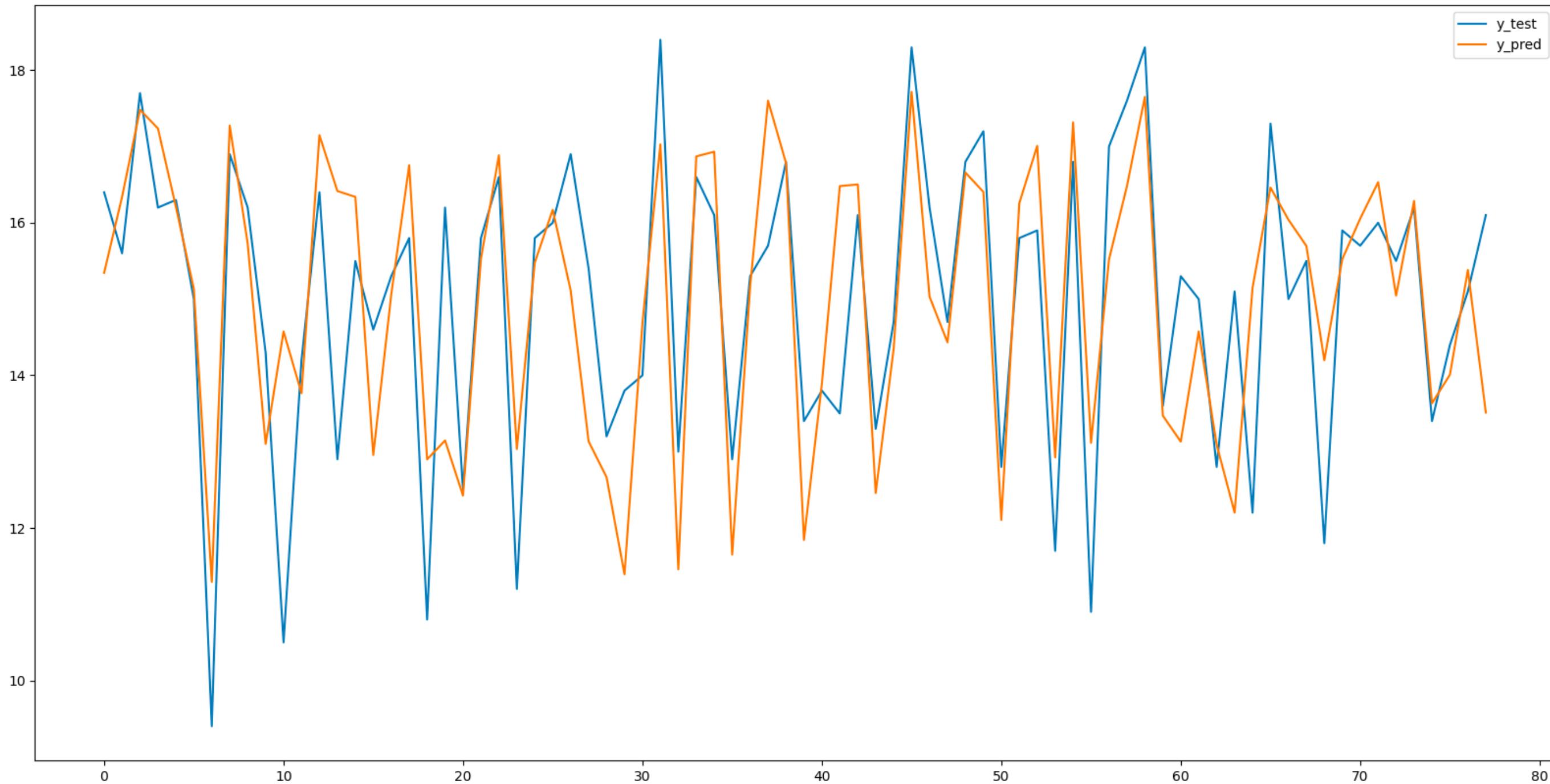
1.9200169

RMSE

Modelos

IVF, IPC y coches vs Temperatura del Estado de México

Predicciones sobre la temperatura del Estado de México



1.0264966

MAE

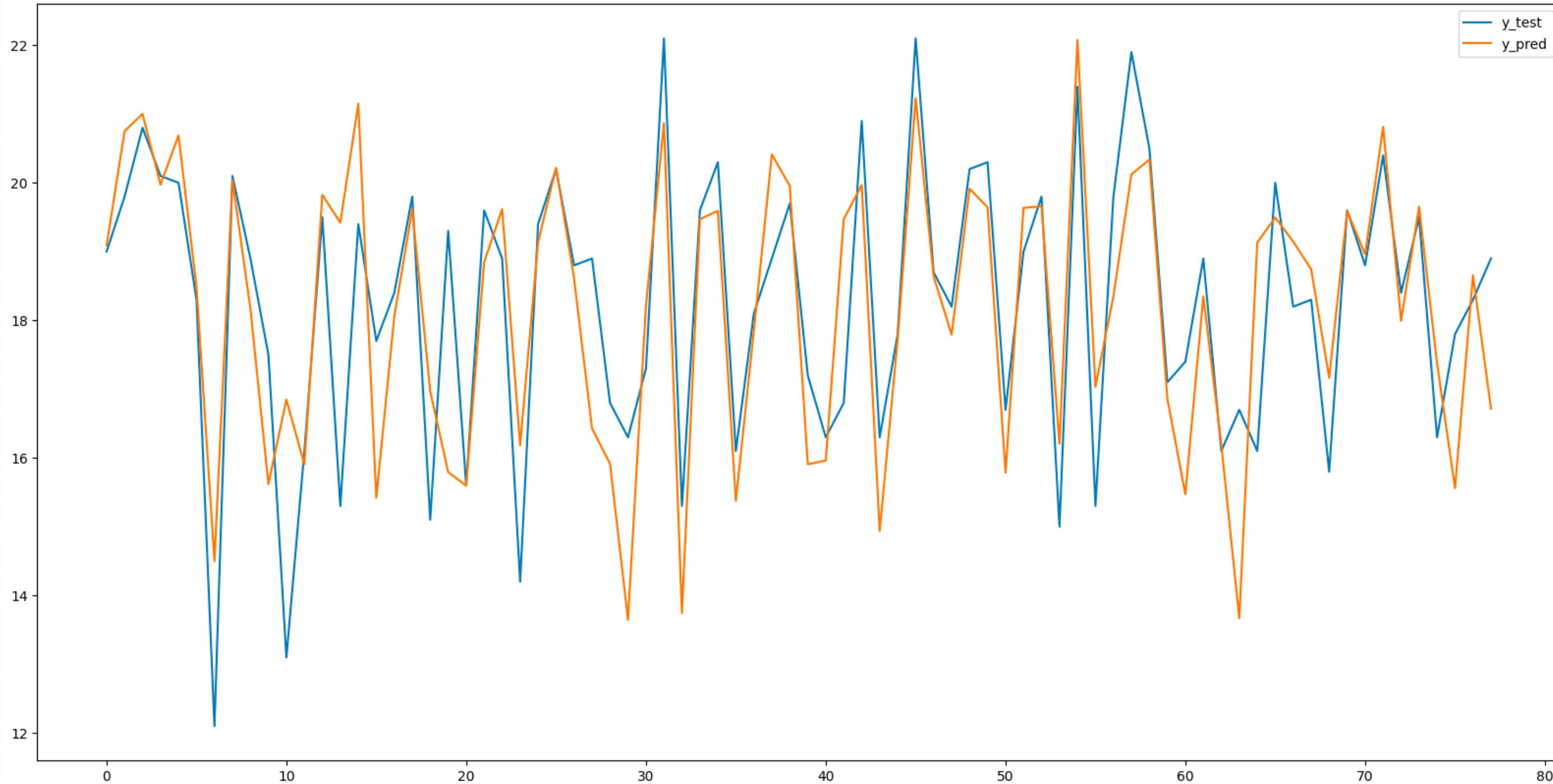
1.3838997

RMSE

Modelos

IVF, IPC y coches vs Temperatura de Puebla

Predicciones sobre la temperatura de Puebla



1.0208962

MAE

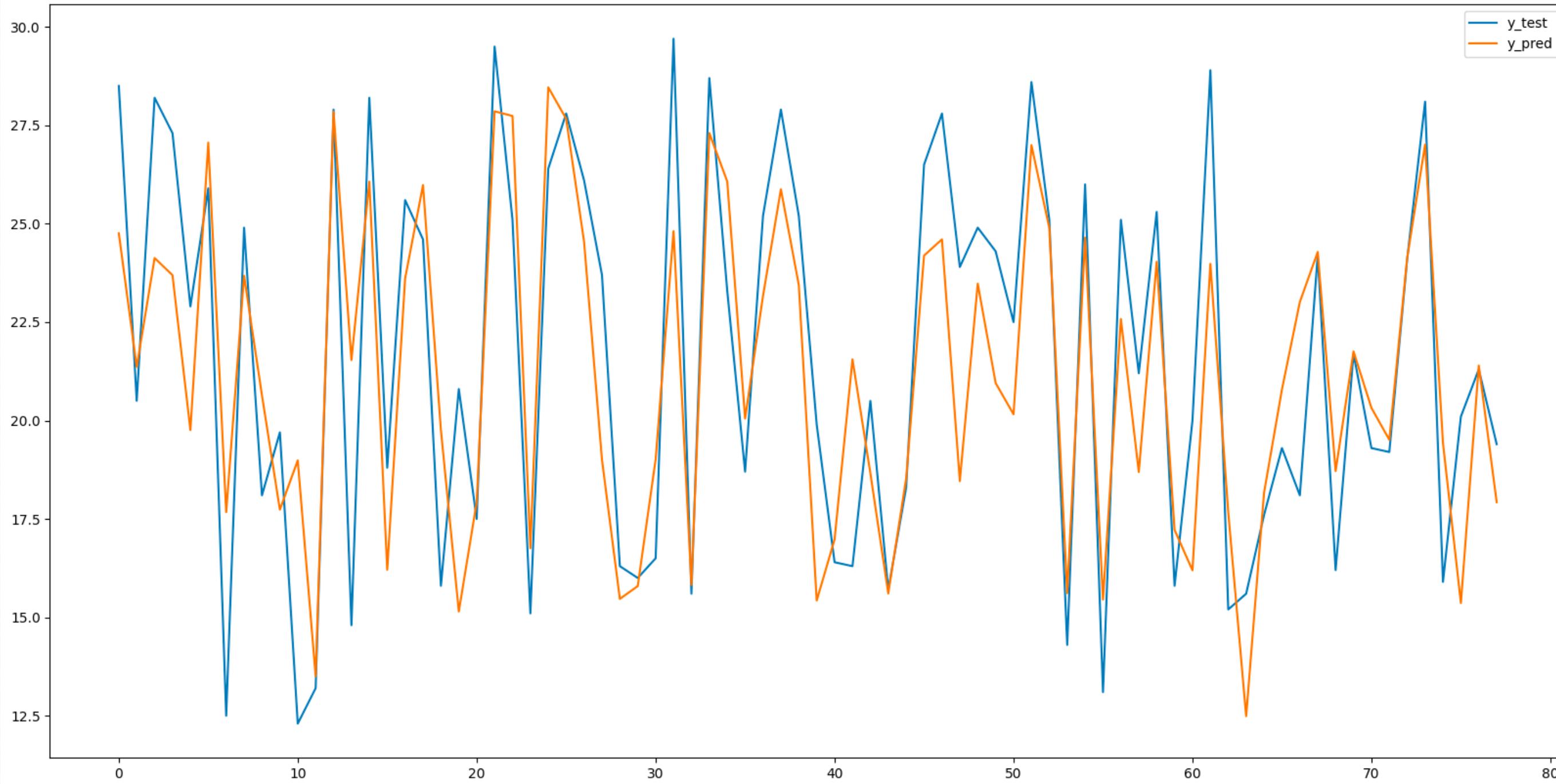
1.4125061

RMSE

Modelos

IVF, IPC y coches vs Temperatura de Nuevo León

Predicciones sobre la temperatura de Nuevo León



2.244553

MAE

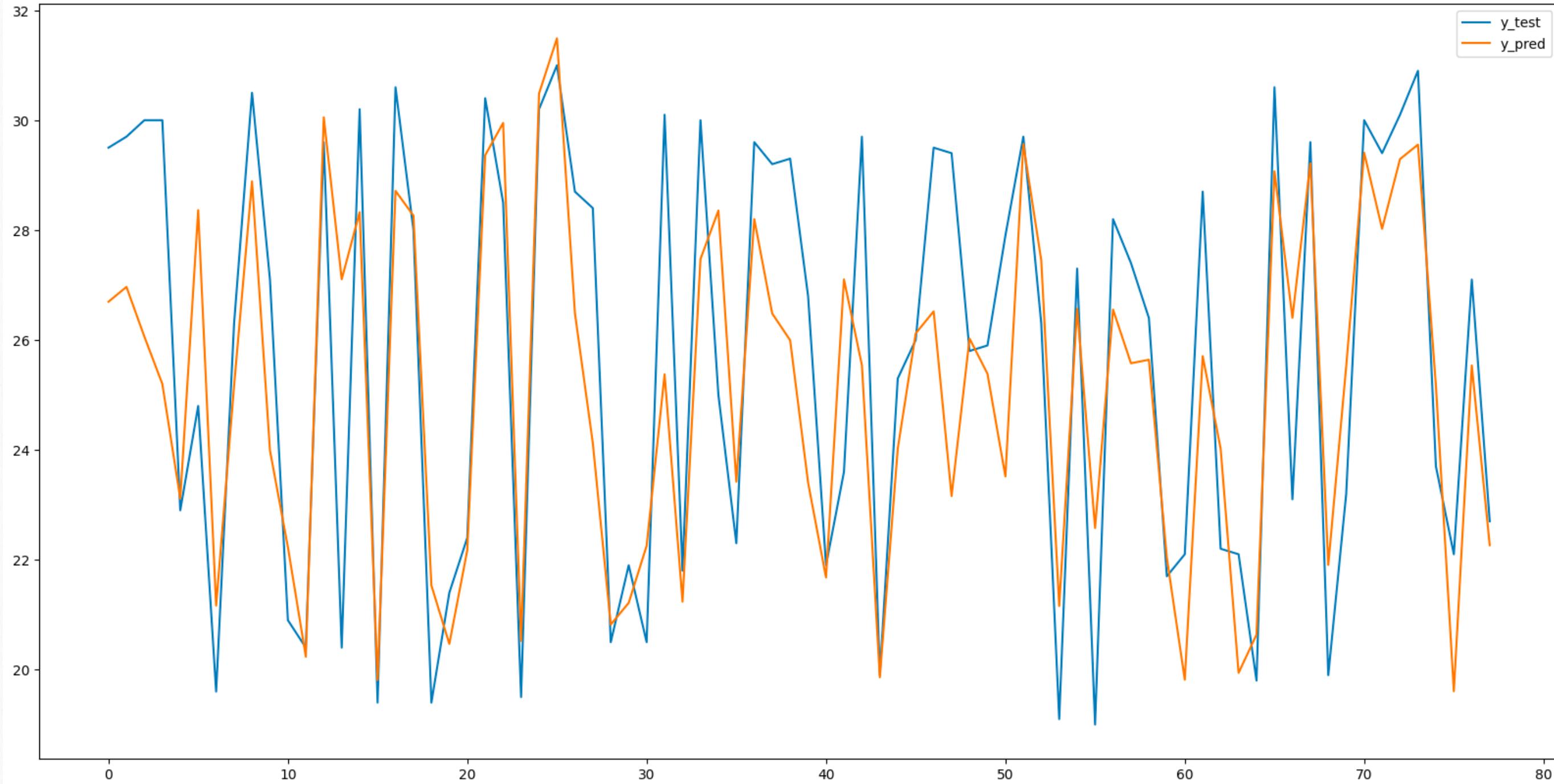
2.8061838

RMSE

Modelos

IVF, IPC y coches vs Temperatura de Sinaloa

Predicciones sobre la temperatura de Sinaloa



1.8475219

MAE

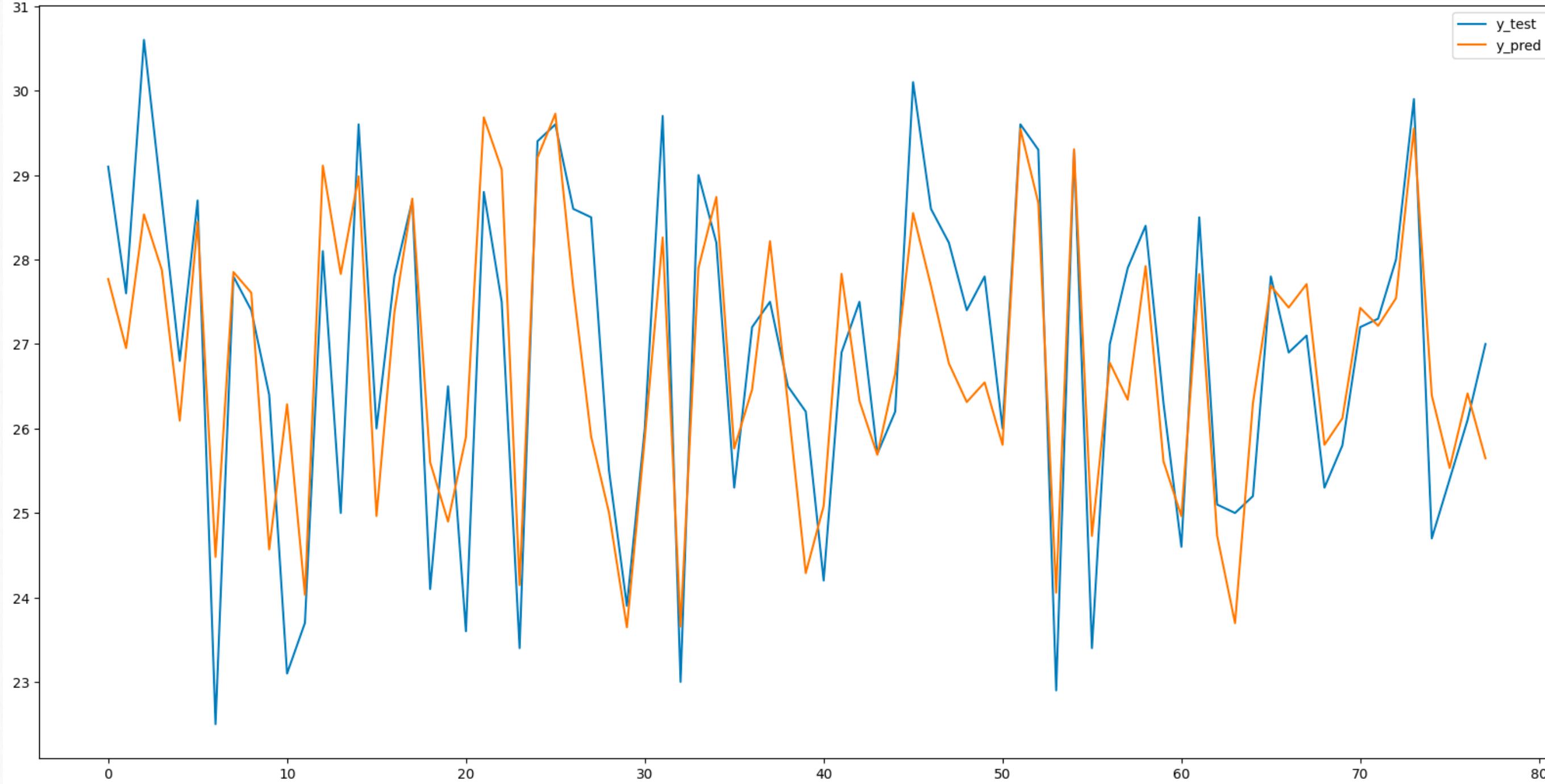
2.3554316

RMSE

Modelos

IVF, IPC y coches vs Temperatura de Quintana Roo

Predicciones sobre la temperatura de Quintana Roo



0.85811955

MAE

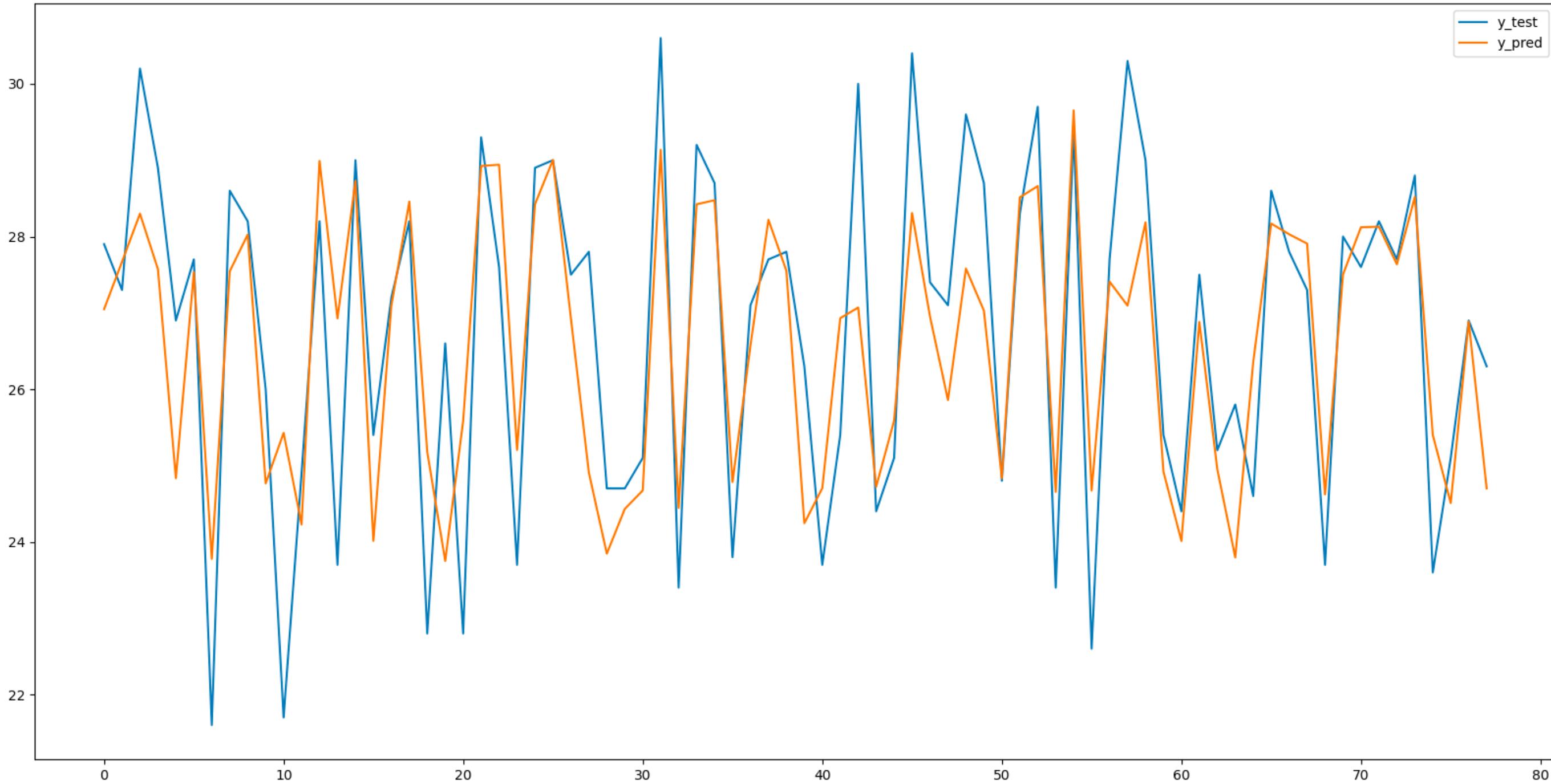
1.1013886

RMSE

Modelos

IVF, IPC y coches vs Temperatura de Yucatán

Predicciones sobre la temperatura de Yucatán



1.0562049

MAE

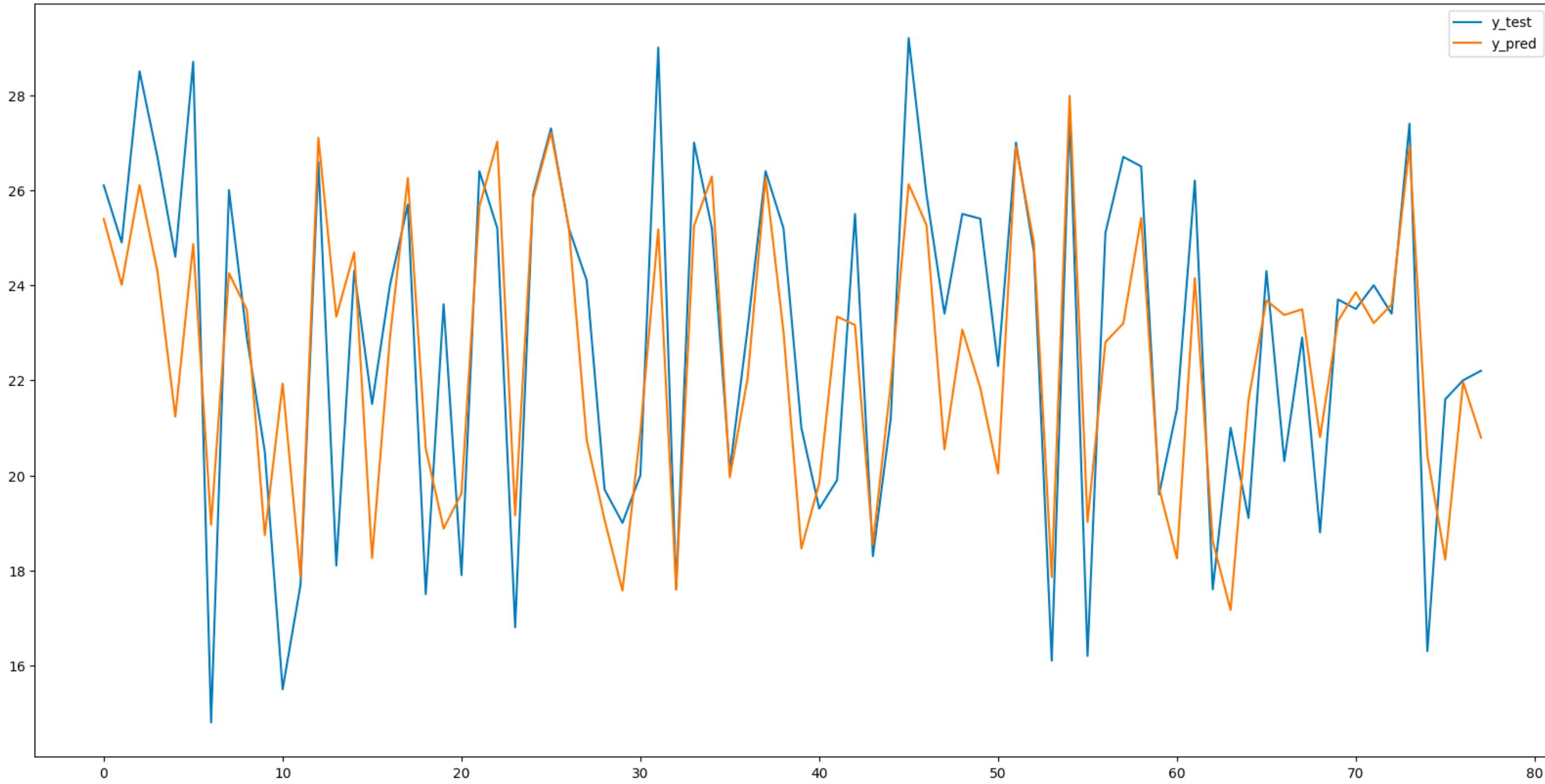
1.3877538

RMSE

Modelos

IVF, IPC y coches vs Temperatura de San Luis Potosí

Predicciones sobre la temperatura de San Luis Potosí



1.7629331

MAE

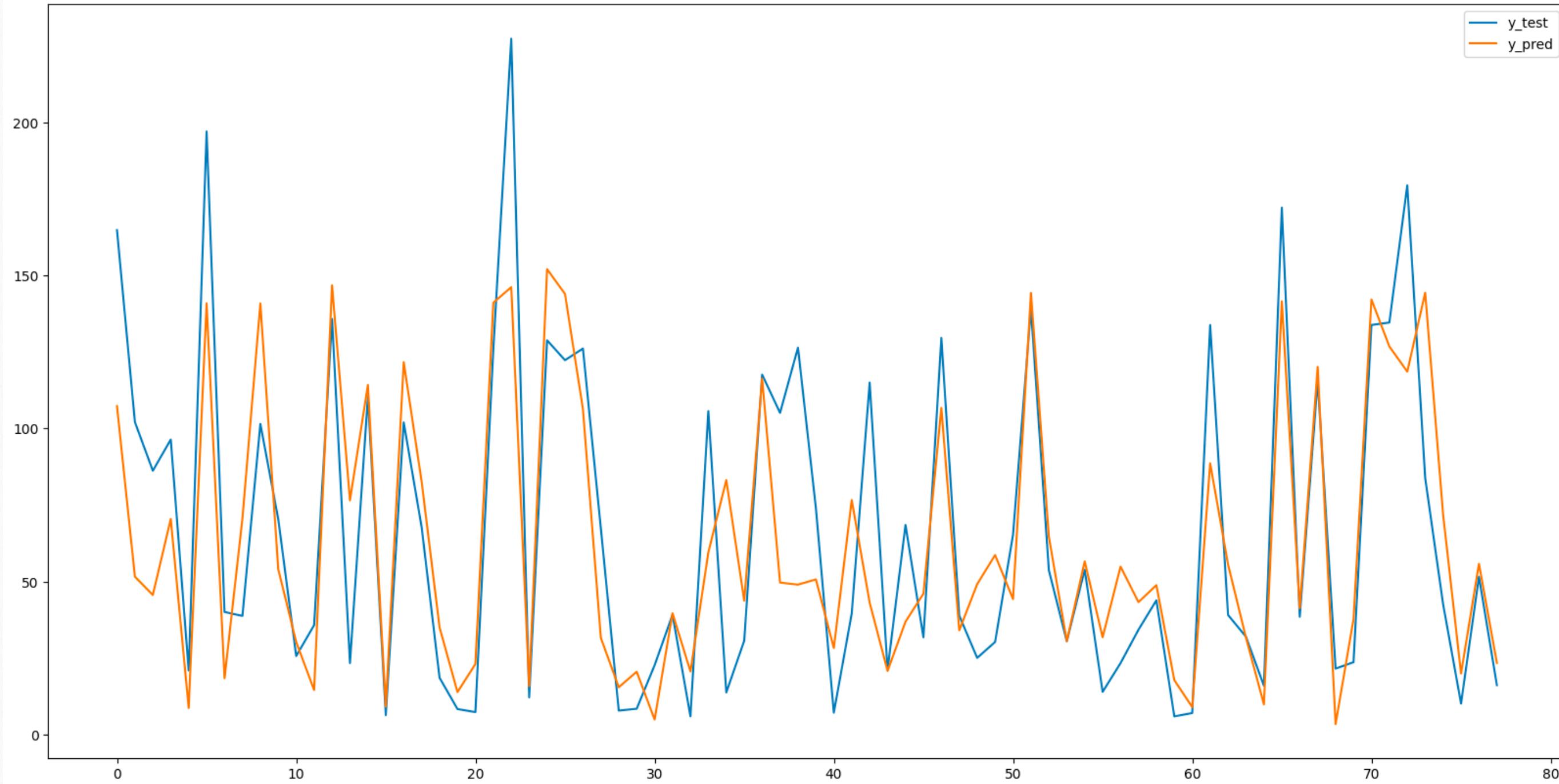
2.2793207

RMSE

Modelos

IVF, IPC y coches vs Lluvia Nacional

Predicciones sobre la lluvia nacional



22.554369

MAE

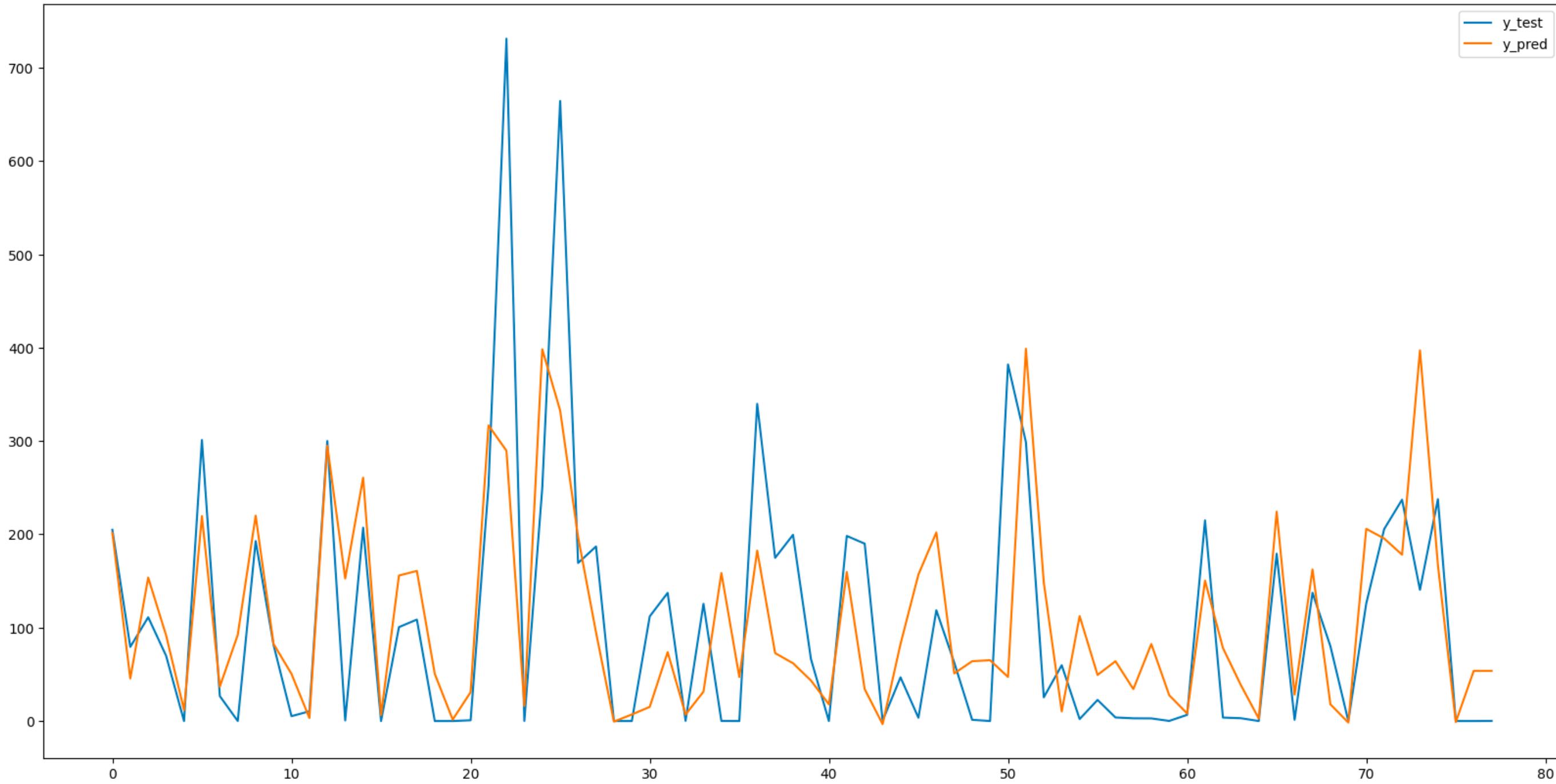
30.15665

RMSE

Modelos

IVF, IPC y coches vs Lluvia Colima

Predicciones sobre la lluvia de Colima



67.91536

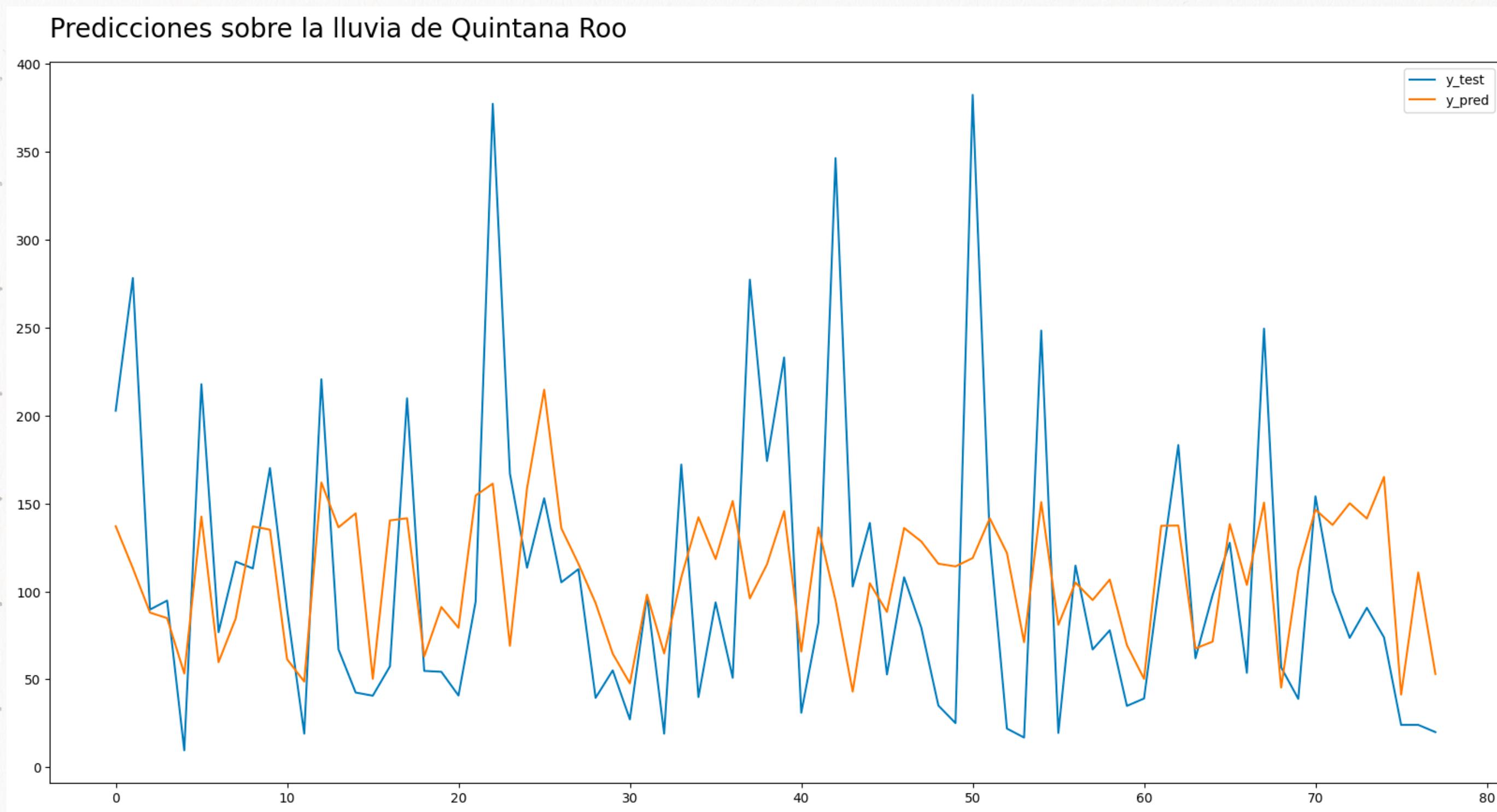
MAE

103.615074

RMSE

Modelos

IVF, IPC y coches vs Lluvia Quintana Roo



56.831257

MAE

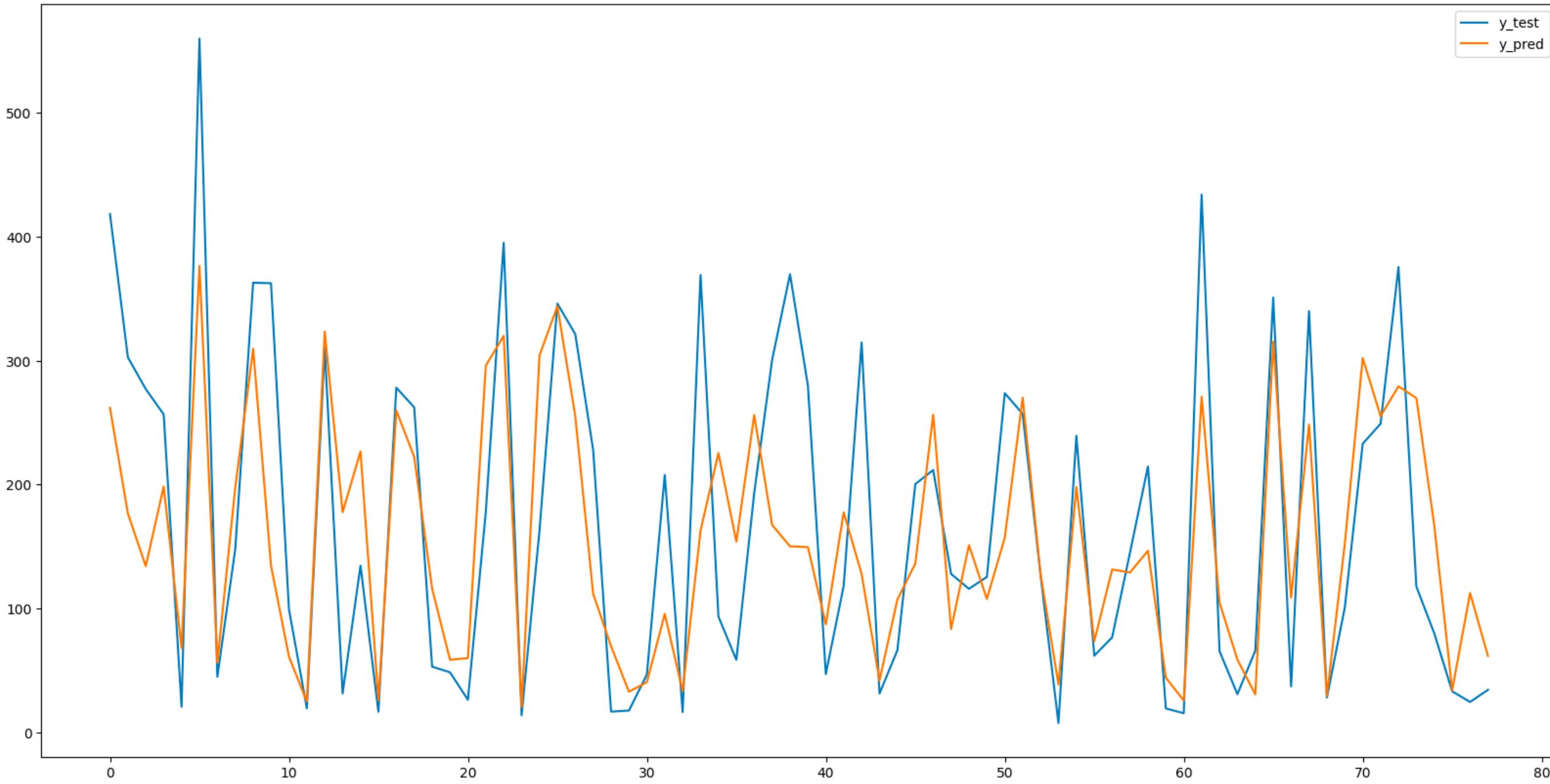
76.56804

RMSE

Modelos

IVF, IPC y coches vs Lluvia Chiapas

Predicciones sobre la lluvia de Chiapas



66.732635

MAE

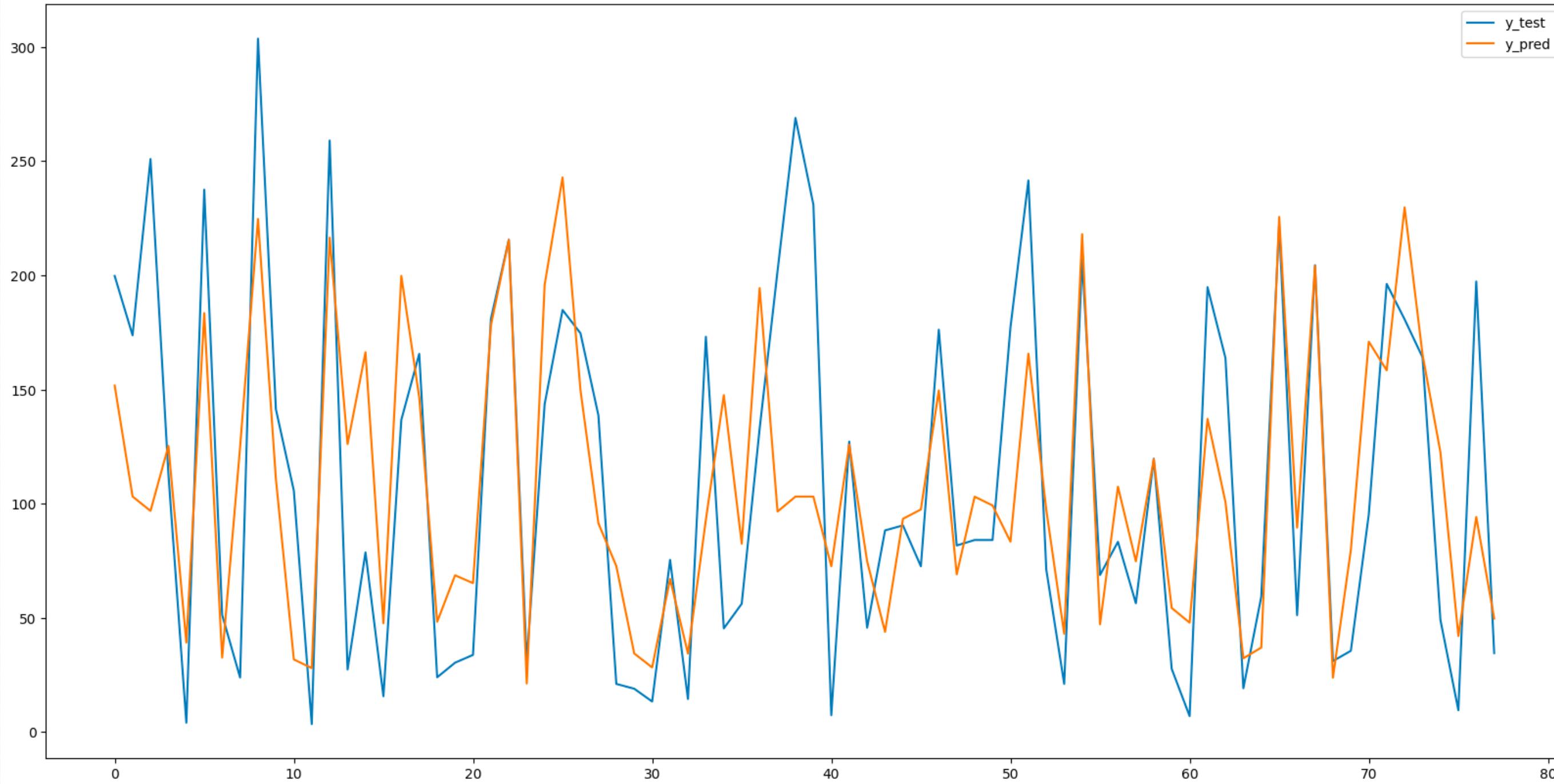
87.82864

RMSE

Modelos

IVF, IPC y coches vs Lluvia Campeche

Predicciones sobre la lluvia de Campeche



42.962452

MAE

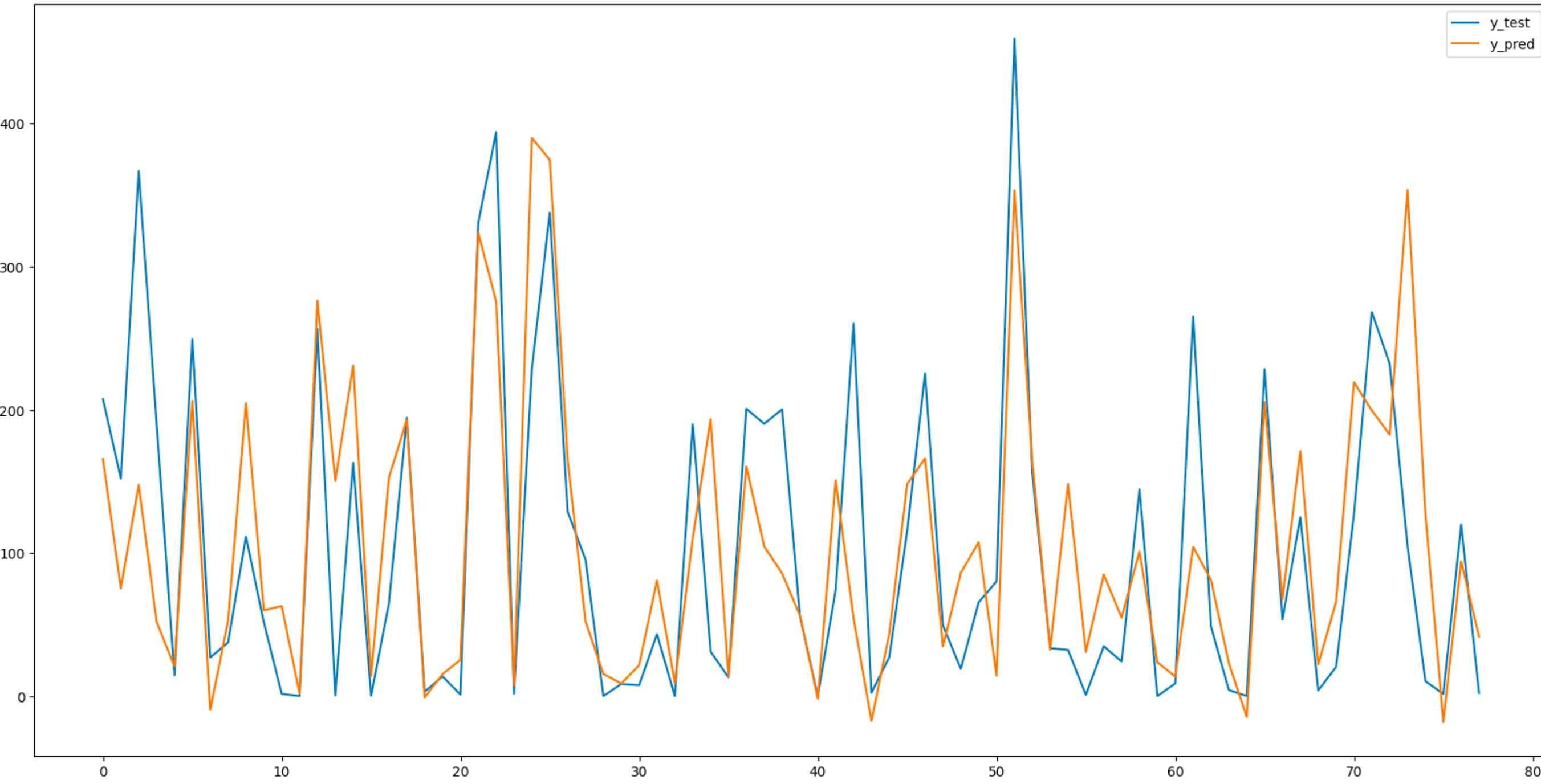
55.77219

RMSE

Modelos

IVF, IPC y coches vs Lluvia Morelos

Predicciones sobre la lluvia de Morelos



52.953743

MAE

75.98079

RMSE

Discusión:

IVF, IPC y transporte vs Variables climáticas

No existía una correlación significativa entre variables económicas y de transporte con las variables climáticas

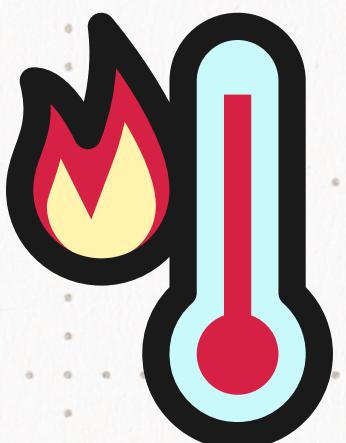
Regresión Kernel

Temp. Nacional

Error cuadrático medio por debajo de 1.5 grados indica precisión en predicciones climáticas, económicas y demográficas.

Algunos estados mostraron errores en predicciones climáticas menores a 1.1 grados (Estado de México, Puebla, Quintana Roo y Yucatán).

Aunque las predicciones de lluvia tuvieron errores más altos debido a las unidades mayores, se considera un margen de error tolerable.



Adaptación de no lineal del algoritmo reveló asociaciones significativas entre variables económicas, demográficas y climáticas.



Conclusión

La exploración de la **interacción** entre **variables económicas** y **ambientales** es un **desafío** que requiere un enfoque minucioso y algoritmos especializados. La **regresión Kernel Gaussiano** emerge como una **herramienta prometedora** para modelar relaciones complejas en conjuntos de datos de naturaleza diversa. A pesar de la inicial baja correlación, esta herramienta permite la identificación y modelado preciso de relaciones no lineales significativas. La convergencia entre la ciencia de datos y la comprensión de las relaciones entre variables económicas y ambientales puede brindar un marco sólido para tomar decisiones informadas que equilibren el crecimiento económico con la sostenibilidad ambiental en el horizonte del futuro.

Agradecimientos



Dr. Christian

Agradecemos al Dr. Christian Pelagio por habernos acompañado durante un semestre y ofrecernos su inmenso apoyo y conocimiento. Aunado a su colaboración en ofrecernos los datos utilizados para este estudio.

Referencias



Murphy, K. P. (2013). Machine learning : a probabilistic perspective. Cambridge, Mass. [u.a.]: MIT Press. ISBN: 9780262018029 0262018020

Bishop, C. M. (2007). Pattern Recognition and Machine Learning (Information Science and Statistics). Springer. ISBN: 0387310738
<https://jax.readthedocs.io/en/latest/> (05/12/2023. 20:00hrs)

Pedro Domingos. 2020. Every Model Learned by Gradient Descent Is Approximately a Kernel Machine.

Chmiela, S., Tkatchenko, A., Sauceda, H. E., Poltavsky, I., Schütt, K. T., Müller, K.-R., *Science Advances*, 3(5), 2017, e1603015.

Chmiela, S., Sauceda, H. E., Müller, K.-R., Tkatchenko, A., *Nature Communications*, 9(1), 2018, 3887.