

**UNIVERSIDAD DEL VALLE DE GUATEMALA**

Data Science

Sección 10

Lynette Garcia



## **Laboratorio 3**

Deep Learning Series

Abner Iván García Alegría - 21285

José Daniel Gomez Cabrera - 21429



<https://github.com/JDgomez2002/data-science-labs.git>

## Lab 3 DS

### 1. Utilice los conjuntos de entrenamiento y prueba de una de las series que utilizó en el Laboratorio.

Se utilizó la serie de ejemplo 1 de consumos del laboratorio anterior para utilizar los conjuntos de entrenamiento y prueba que se realizaron en esa serie.

```
# Dividir en conjuntos de entrenamiento y prueba como el laboratorio anterior
train_data, test_data = train_test_split(data, test_size=0.2, shuffle=False)

print("Datos de entrenamiento:")
print(train_data.head())
print("\nDatos de prueba:")
print(test_data.head())
```

```
Datos de entrenamiento:
      Gasolina regular  Gasolina superior  Diesel  Gas licuado de petróleo
Fecha
NaT          202.64520           308.15682     0.0           194.41048
NaT          205.53096           307.76631     0.0           174.71055
NaT          229.49956           331.91029     0.0           189.23407
NaT          210.68040           315.64808     0.0           174.33061
NaT          208.16434           319.66797     0.0           191.74515

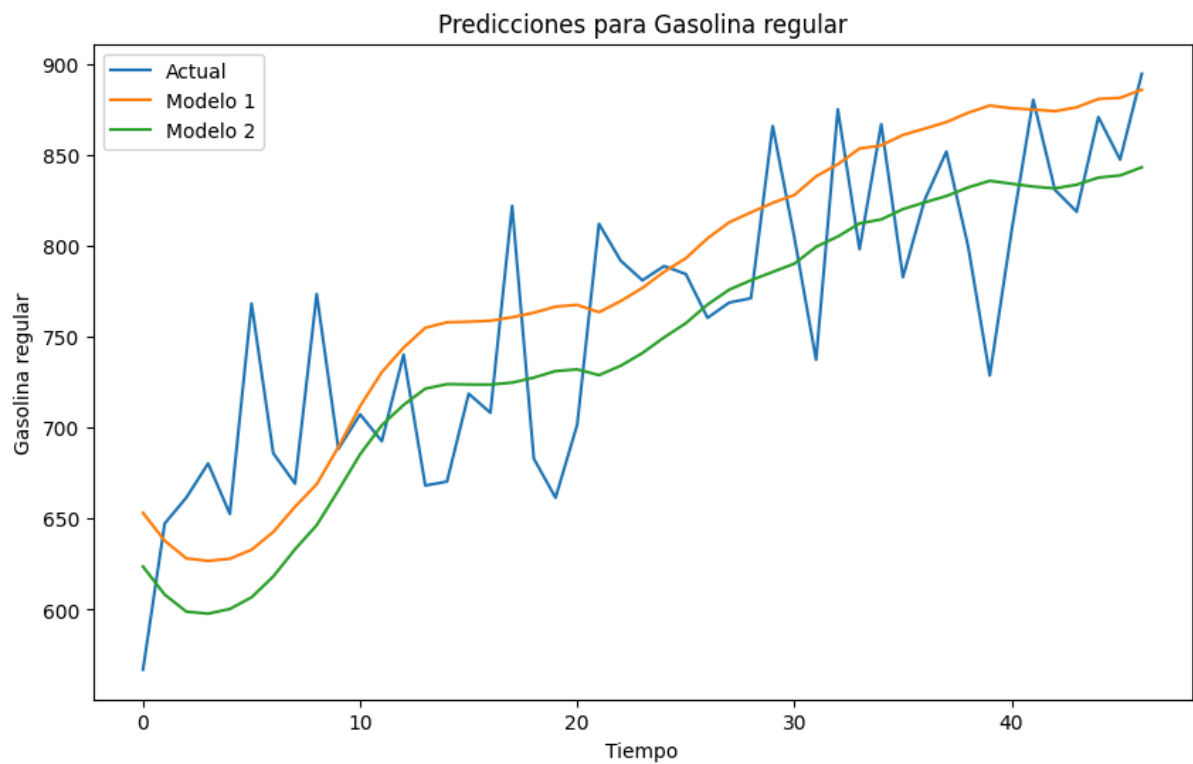
Datos de prueba:
      Gasolina regular  Gasolina superior  Diesel  Gas licuado de petróleo
Fecha
NaT          650.15078           615.55933     0.0           419.95802
NaT          668.78339           632.41570     0.0           419.49382
NaT          616.07930           587.86184     0.0           396.86064
NaT          669.00877           643.04433     0.0           433.91715
NaT          643.07035           620.57221     0.0           430.72589
```

### 2. Haga al menos 2 modelos con configuraciones diferentes usando LSTM por cada serie de las que utilizó. Haga un tuneo de parámetros de los modelos.

## Serie 1 Consumos:

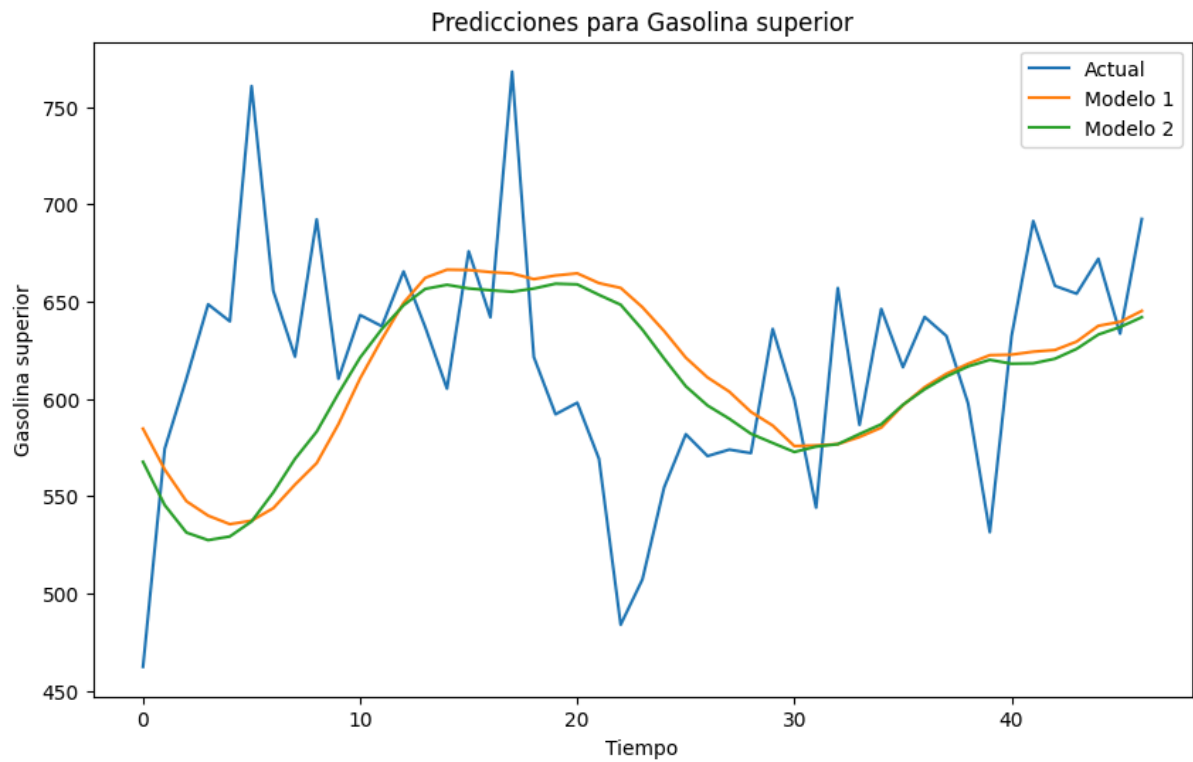
MSE del Modelo 1 para Gasolina regular: 3440.1179512909484

MSE del Modelo 2 para Gasolina regular: 3093.3984890399893

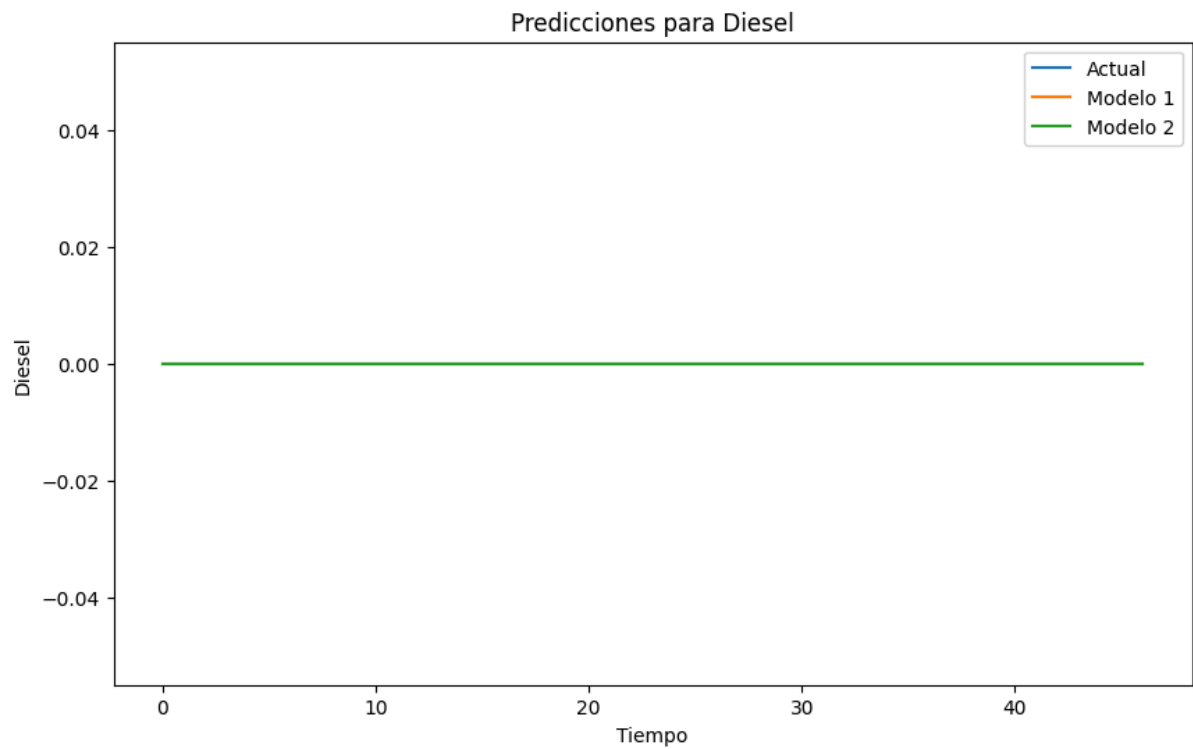


MSE del Modelo 1 para Gasolina superior: 5449.452562880711

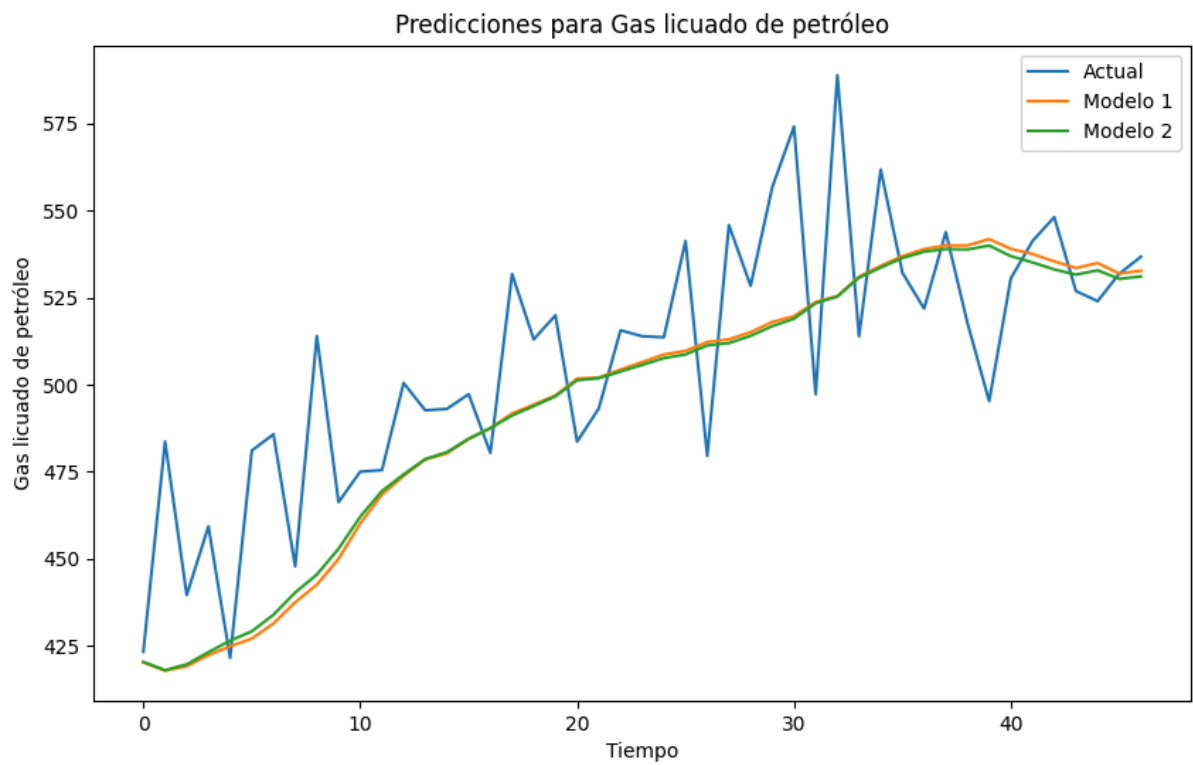
MSE del Modelo 2 para Gasolina superior: 5132.715627185705



MSE del Modelo 1 para Diesel: 0.0  
MSE del Modelo 2 para Diesel: 0.0



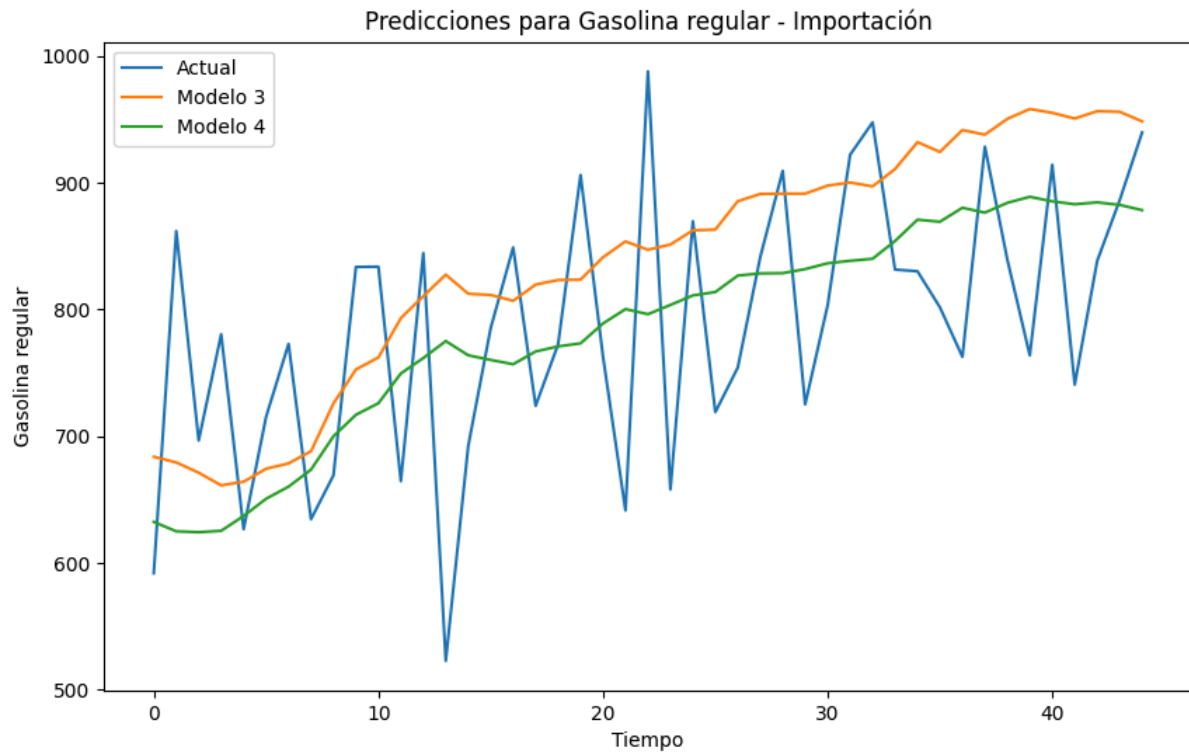
MSE del Modelo 1 para Gas licuado de petróleo: 832.523323018979  
MSE del Modelo 2 para Gas licuado de petróleo: 809.7799852224315



## Serie 2 Importación:

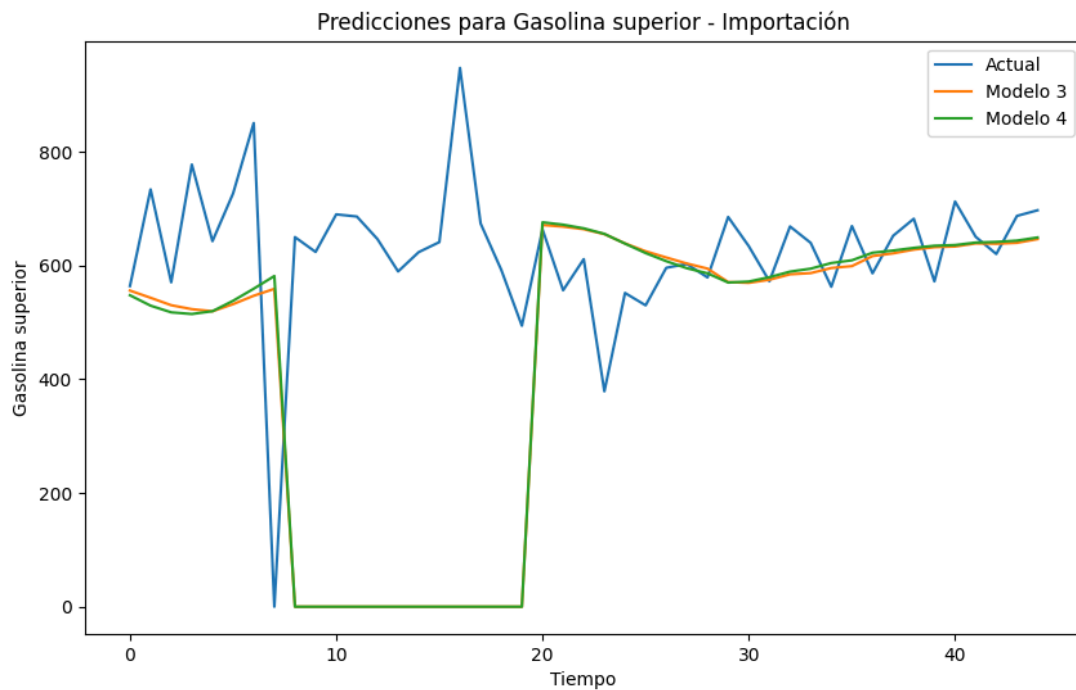
MSE del Modelo 3 para Gasolina regular: 13331.635607973083

MSE del Modelo 4 para Gasolina regular: 10011.439800713397

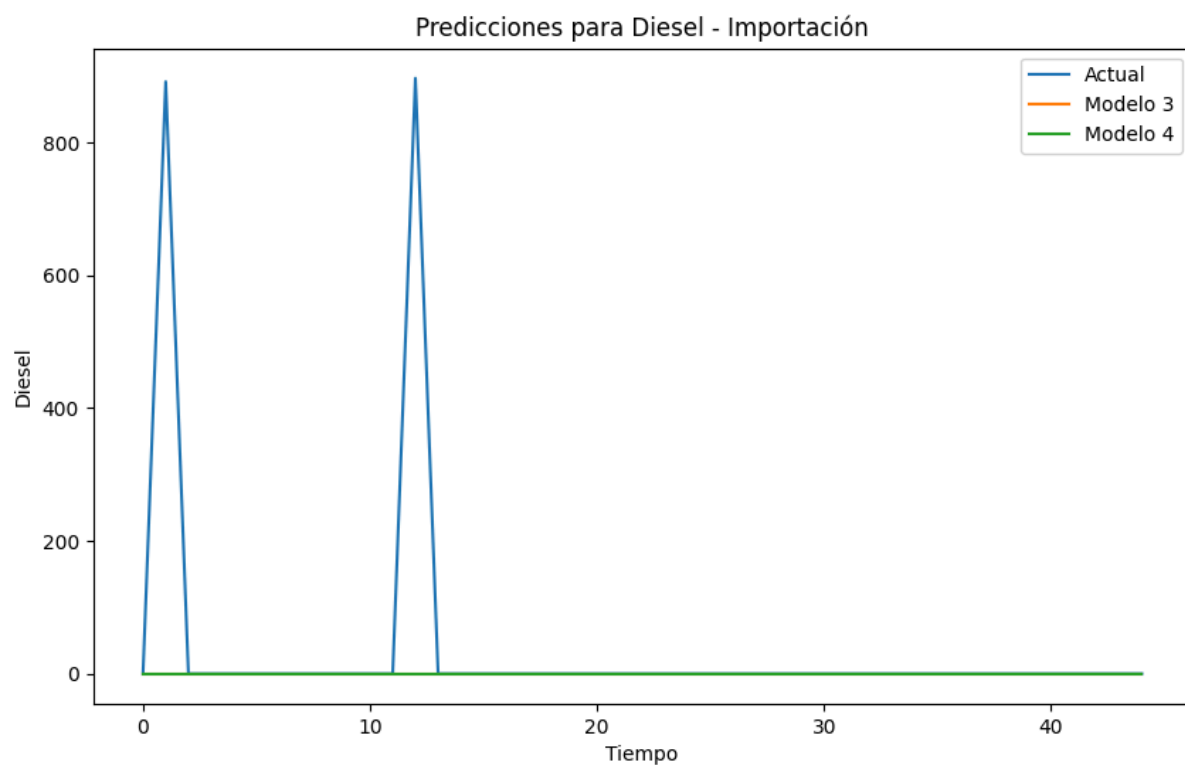


MSE del Modelo 3 para Gasolina superior: 133115.58548004803

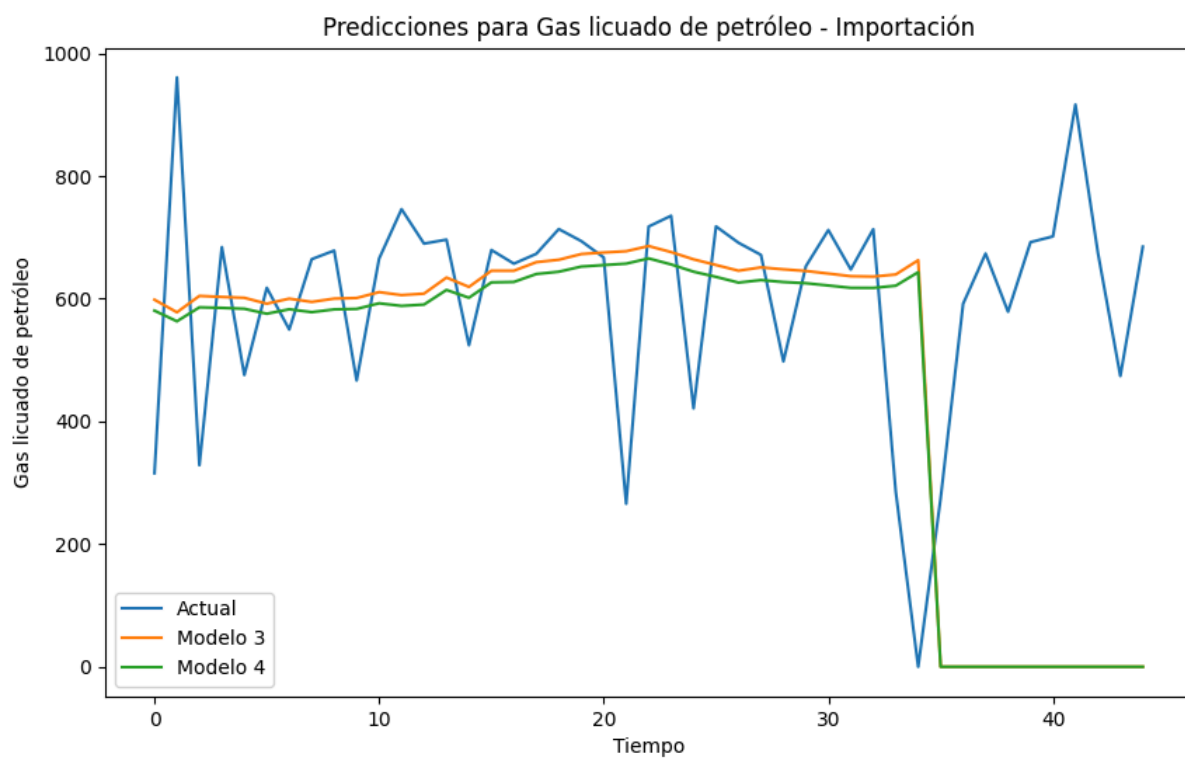
MSE del Modelo 4 para Gasolina superior: 133659.00328645157



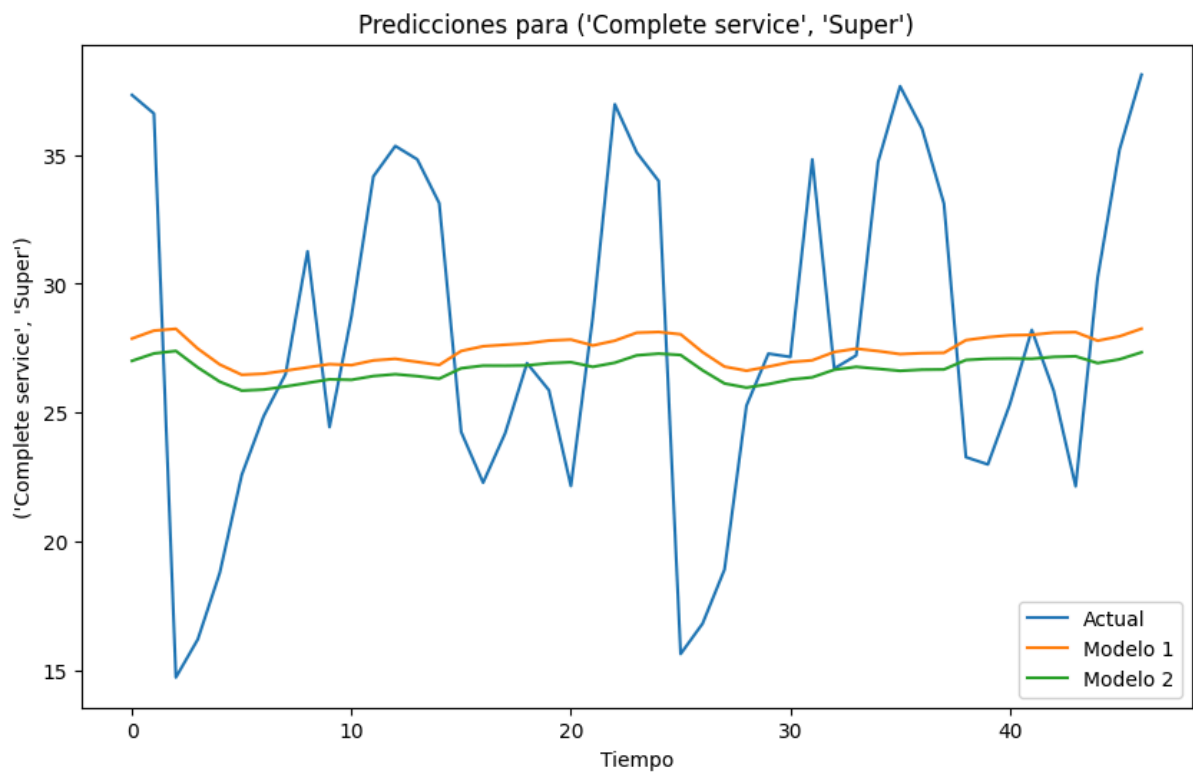
MSE del Modelo 3 para Diesel: 35517.12965576564  
MSE del Modelo 4 para Diesel: 35517.12965576564



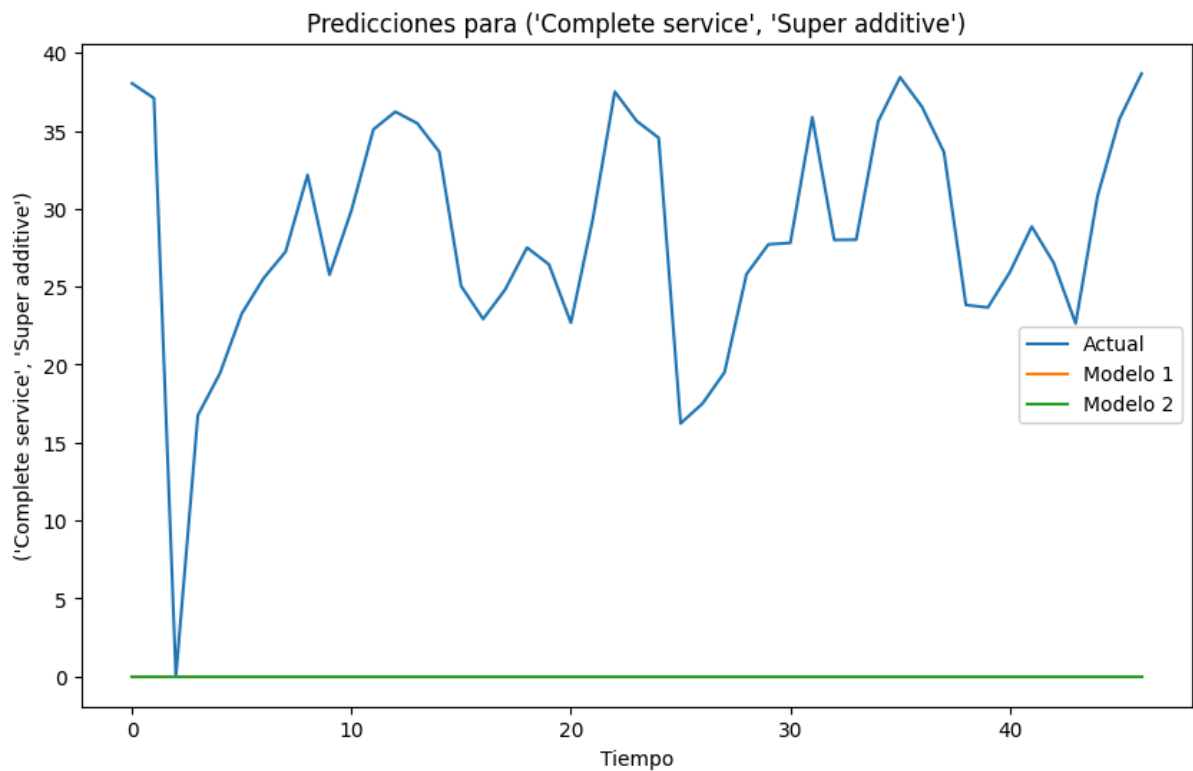
MSE del Modelo 3 para Gas licuado de petróleo: 120378.02141997886  
MSE del Modelo 4 para Gas licuado de petróleo: 119450.83328198142



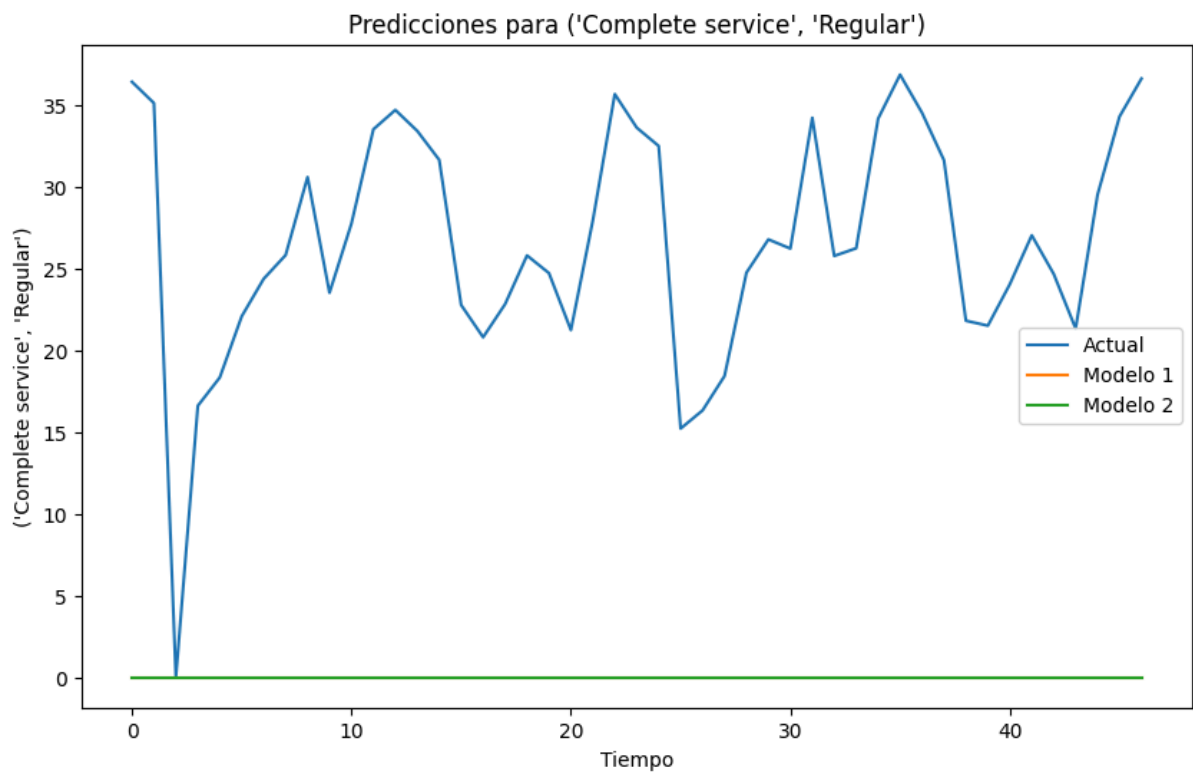
### Serie 3 Precios:



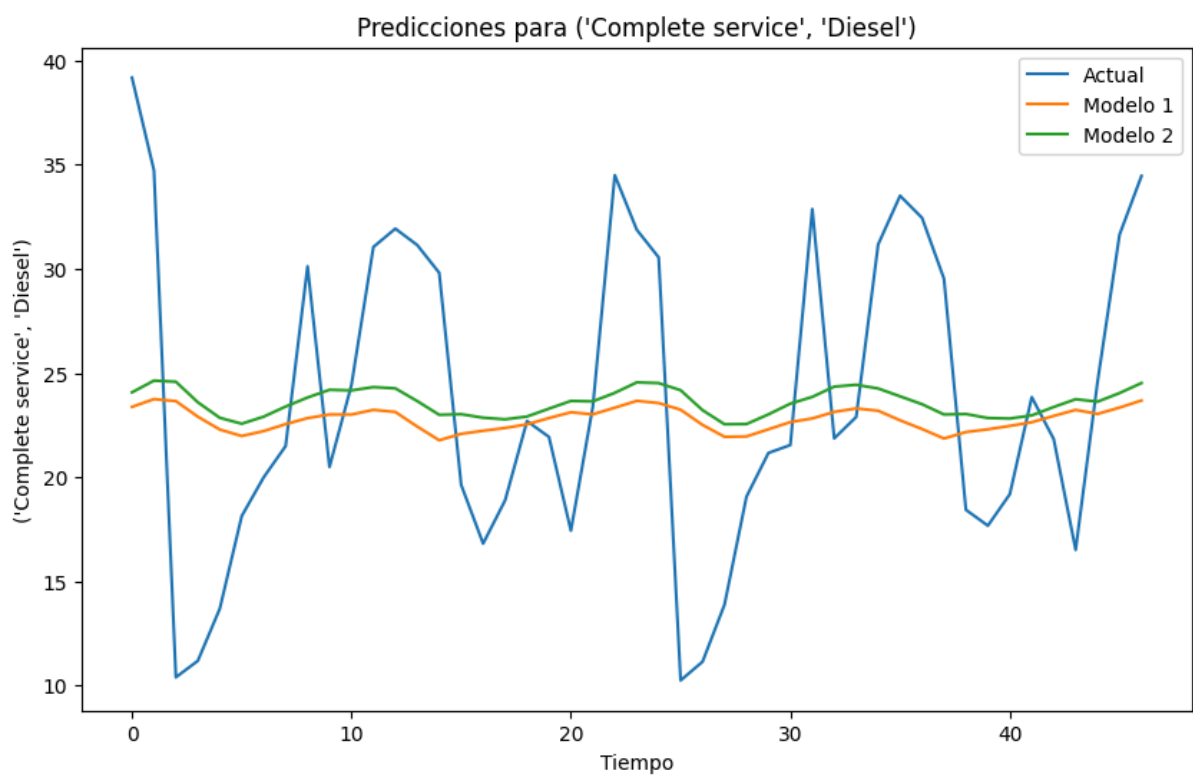
```
MSE del Modelo 1 para ('Complete service', 'Super'): 41.9351953839166
MSE del Modelo 2 para ('Complete service', 'Super'): 43.04010366381156
```



```
MSE del Modelo 1 para ('Complete service', 'Super additive'): 855.7260319148936
MSE del Modelo 2 para ('Complete service', 'Super additive'): 855.7260319148936
```

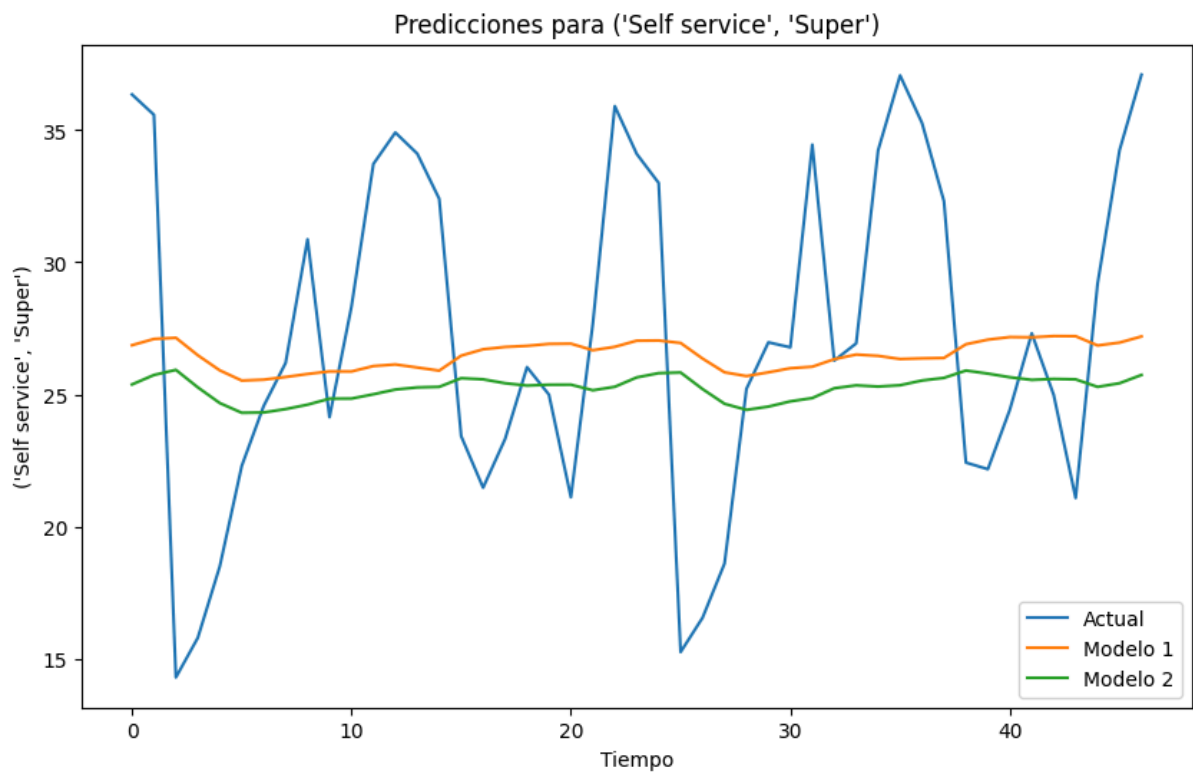


MSE del Modelo 1 para ('Complete service', 'Regular'): 762.2550148936169  
MSE del Modelo 2 para ('Complete service', 'Regular'): 762.2550148936169

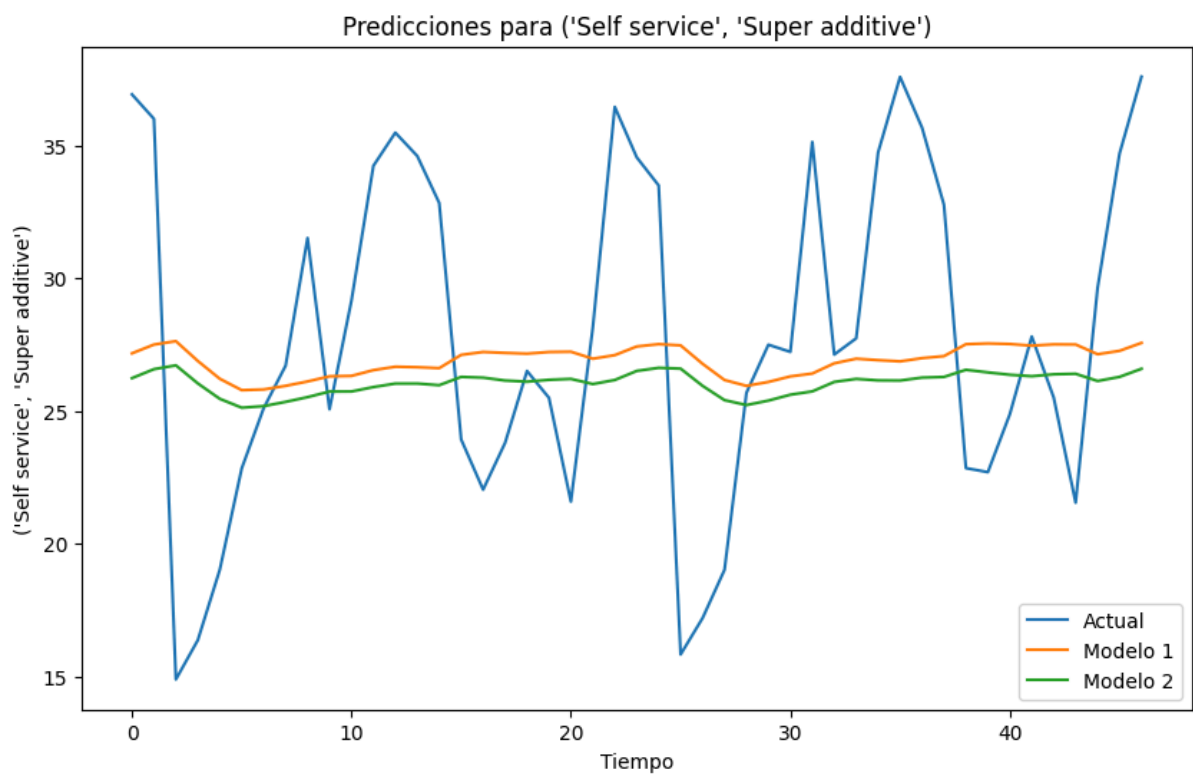


MSE del Modelo 1 para ('Complete service', 'Diesel'): 54.06361567375643  
MSE del Modelo 2 para ('Complete service', 'Diesel'): 51.62784638129211

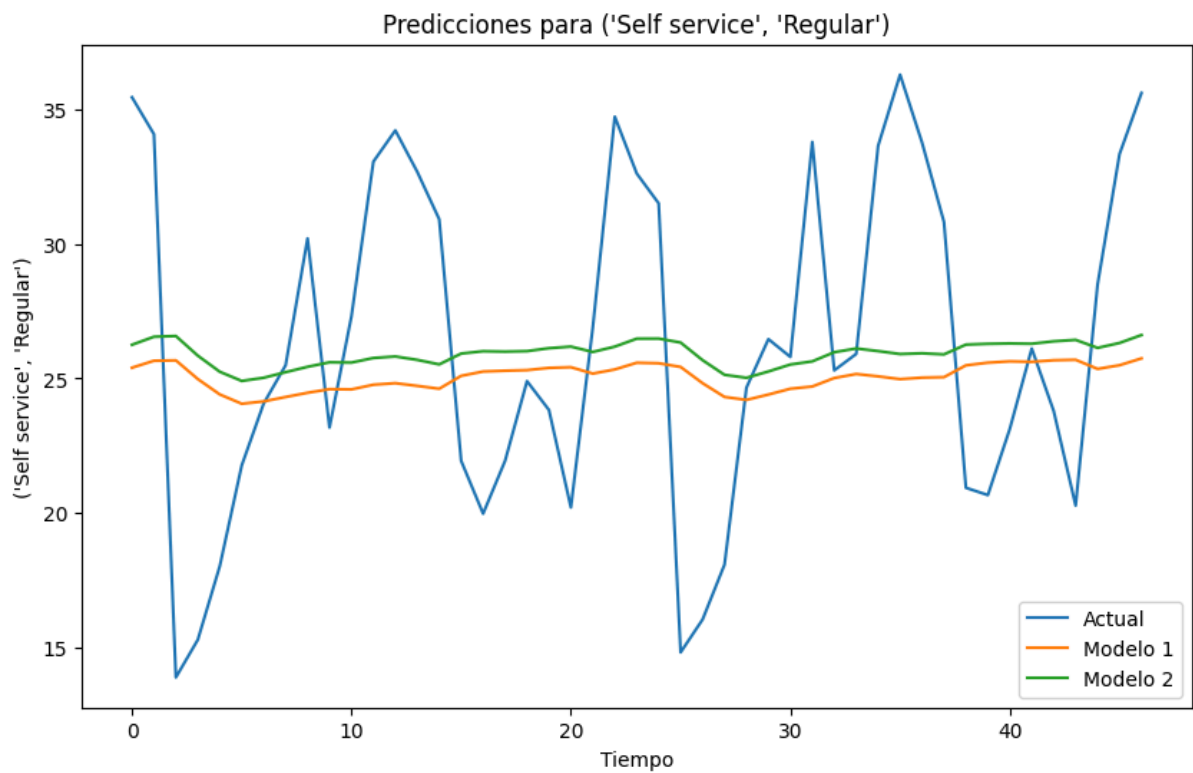




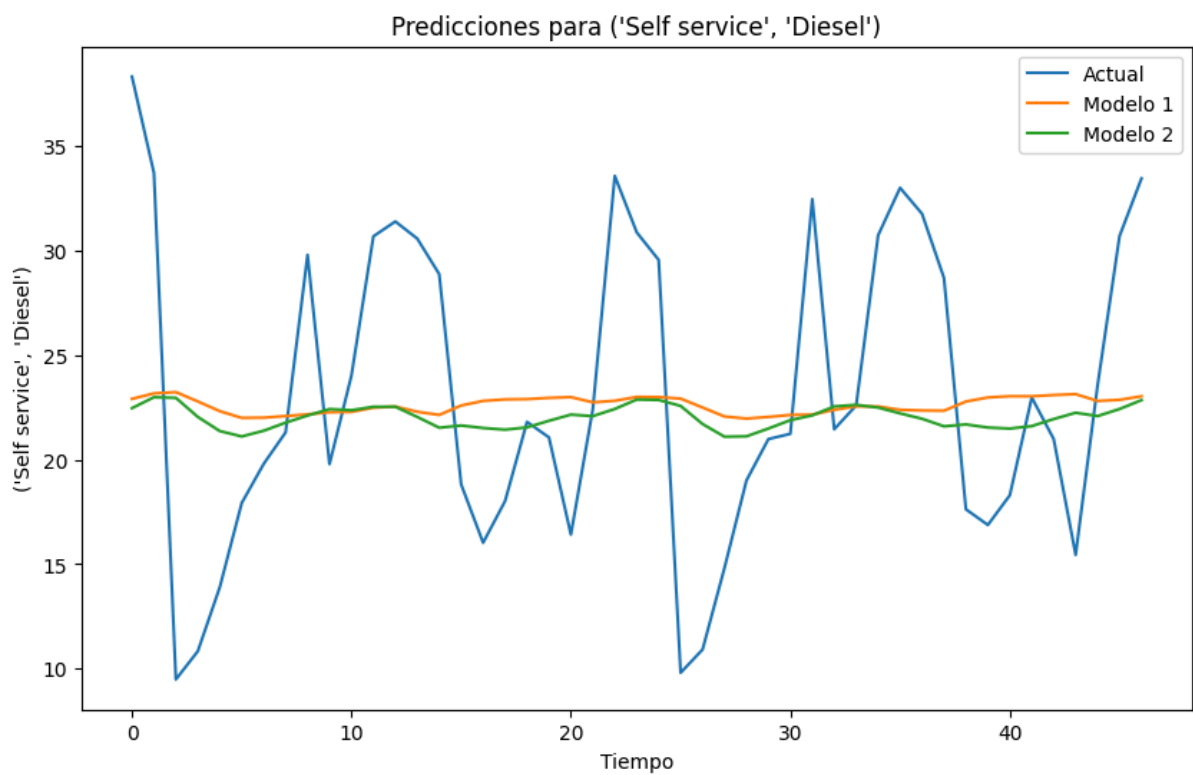
```
MSE del Modelo 1 para ('Self service', 'Super'): 41.32162203639675
MSE del Modelo 2 para ('Self service', 'Super'): 44.37650691695577
```



```
MSE del Modelo 1 para ('Self service', 'Super additive'): 41.183281064743134
MSE del Modelo 2 para ('Self service', 'Super additive'): 42.871430879783894
```



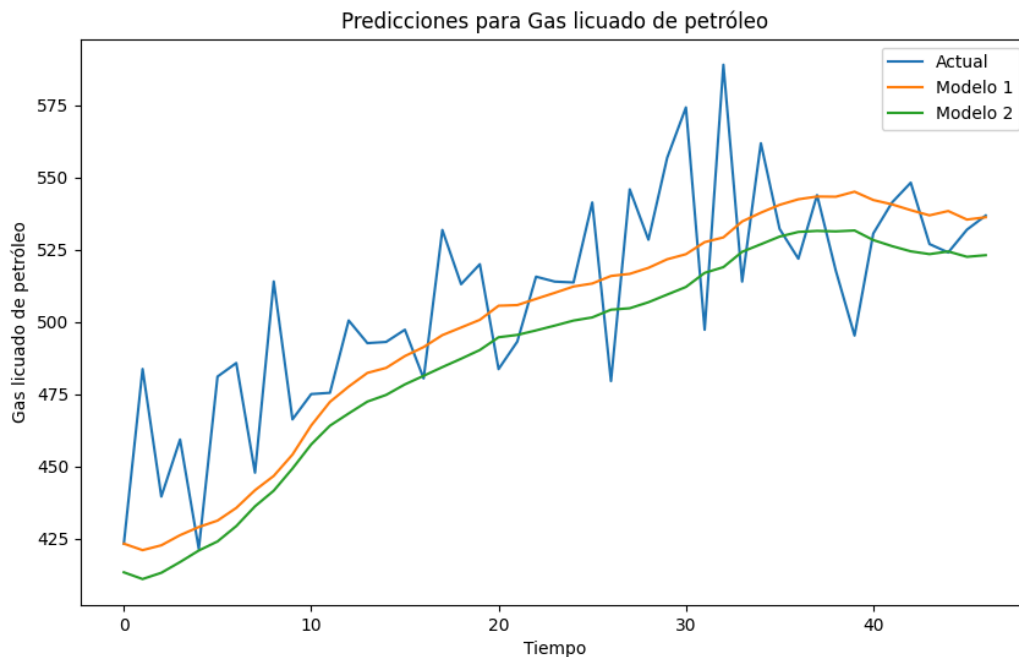
MSE del Modelo 1 para ('Self service', 'Regular'): 40.156121077449306  
MSE del Modelo 2 para ('Self service', 'Regular'): 38.30763339582786



MSE del Modelo 1 para ('Self service', 'Diesel'): 54.45211301970529  
MSE del Modelo 2 para ('Self service', 'Diesel'): 51.991709639742616

### 3. Use el mejor modelo para predecir.

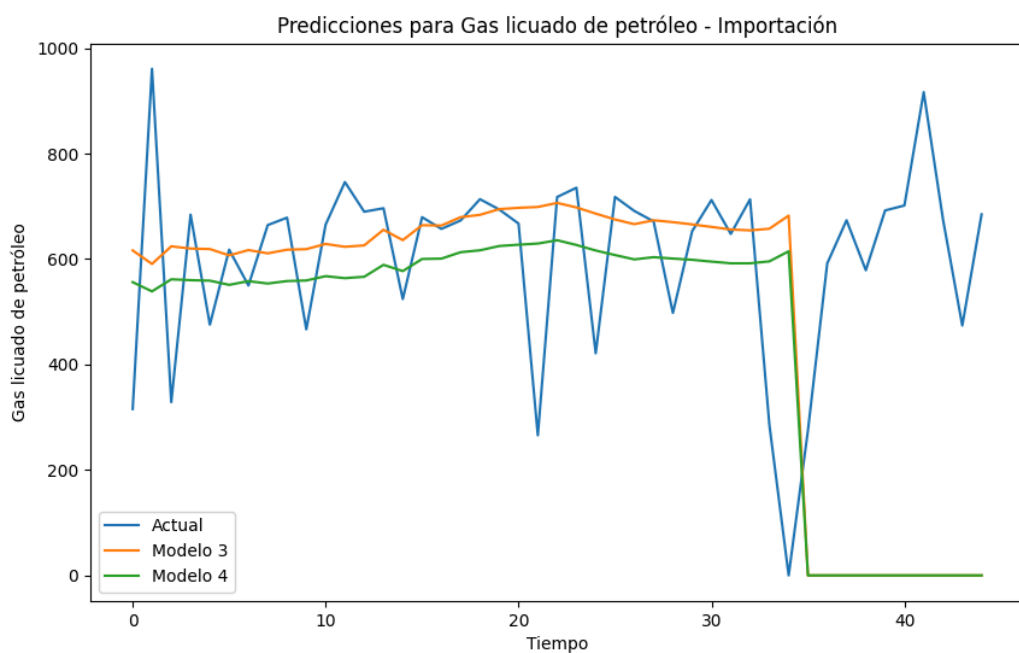
#### Serie 1 Consumos: mejor modelo



MSE del Modelo 1 para Gas licuado de petróleo: 749.7884263909983  
MSE del Modelo 2 para Gas licuado de petróleo: 1001.669483891946

**R//** Al comparar los modelos vemos que el modelo 1 de gas licuado es el mejor debido a su MSE muy bajo a comparación con los otros modelos.

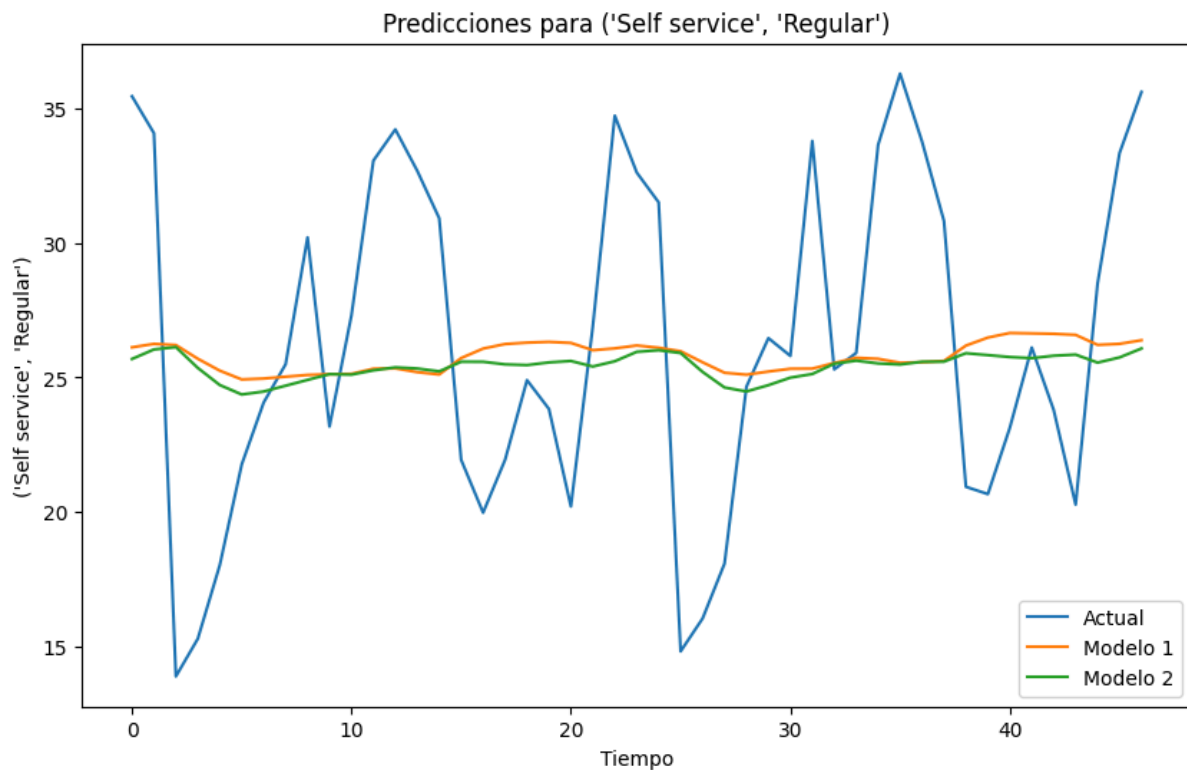
#### Serie 2 Importaciones: mejor modelo



MSE del Modelo 3 para Gas licuado de petróleo: 121939.40761854194  
MSE del Modelo 4 para Gas licuado de petróleo: 119156.11951533101

**R//** El mejor modelo para la serie 2 fue el modelo 4 del gas licuado debido a que obtuvo el MSE más bajo que los otros modelos.

### **Serie 3 precios: mejor modelo**



```
MSE del Modelo 1 para ('Self service', 'Regular'): 39.718506470405075
MSE del Modelo 2 para ('Self service', 'Regular'): 38.90436681053843
```

**R//** Al comparar con todos nuestros modelos realizados vimos que estos dos modelos tuvieron el mejor resultado pero nuestro mejor modelo es el modelo 2. Esto se debe a que tiene un valor de MSE (Error Cuadrático Medio) más bajo que el Modelo 1. Un MSE más bajo indica que las predicciones del Modelo 2 están, en promedio, más cerca de los valores reales ("Actual") que las predicciones del Modelo 1. Por lo tanto, el Modelo 2 es más preciso al igual que al realizar la predicción para el Modelo 2 es el mejor que nos resultó de los demás modelos por esta razón el modelo 2 de la serie 3 de los precios en la gasolina regular se lleva la corona debido a su MSE que es bajo a comparación de los demás.

### **4. ¿Cuál predijo mejor? ¿Son mejores que los modelos creados en el laboratorio pasado? ¿Cómo lo determinaron?**

A nuestro criterio, los mejores modelos de cada Serie fueron: En consumos, las predicciones para gasolina regular. En las series de importación, también las predicciones de para gasolina regular. Finalmente para la serie de precios, las predicciones para gasolina regular en autoservicio, fueron las mejores. Sin embargo, debido a que el modelo de la serie de

consumos fue el que tuvo un menor MSE, se puede suponer que fue el que tuvo un mejor desempeño.

Vemos que los mejores modelos para predecir, son los de este laboratorio debido a que los índices de MSE y también las gráficas reflejan que los modelos tuvieron mejores resultados. Sin embargo, aunque son mejores que los resultados del laboratorio anterior, tampoco se ve que sean super exactos a los datos reales. Esto debido a que la curvatura y variabilidad de los datos es tan grande e irregular que no es muy difícil poder entrenar un modelo para responder a picos y valles tan irregulares. No obstante, los resultados fueron más fieles que los del laboratorio pasado.

Tanto el MSE como las gráficas fueron importantes para poder determinar que los resultados mejoraron, pero el factor que más ayudó a poder cuantificar la mejoría de los modelos fue el MSE debido a que la gráfica fue muy similar, sin embargo solo visualmente es imposible poder determinar que tan fiel o no es el modelo para predecir respecto a los datos reales, pero con el MSE se puede percibir dicha variabilidad respecto a los datos reales.

## Conclusiones

1. Efectividad de los Modelos LSTM: Los modelos LSTM implementados en este laboratorio demostraron ser más efectivos que los modelos utilizados en el laboratorio anterior, especialmente en términos de precisión de predicciones, como lo evidencia el menor MSE (Error Cuadrático Medio). Esto sugiere que, a pesar de la complejidad de los datos y la variabilidad, el uso de técnicas de deep learning como LSTM puede mejorar significativamente el rendimiento de los modelos predictivos.
2. Importancia del MSE como Métrica de Evaluación: A lo largo del análisis, el MSE fue una métrica clave para evaluar la precisión de los modelos. A pesar de que las gráficas ofrecieron una visión visual del rendimiento, el MSE proporcionó una cuantificación más precisa de la exactitud de los modelos. Este enfoque permitió identificar de manera objetiva cuál de los modelos tuvo un mejor desempeño.
3. Desafíos en la Predicción de Series Temporales Complejas: Aunque los modelos actuales superaron a los anteriores, se observó que la gran variabilidad y los patrones irregulares de las series temporales utilizadas presentan desafíos significativos para el entrenamiento de modelos precisos. Esto subraya la necesidad de continuar afinando los modelos y posiblemente explorar enfoques adicionales para manejar datos altamente irregulares.
4. Selección de los Mejores Modelos por Serie: En cada una de las series analizadas (consumos, importaciones, y precios), se logró identificar un modelo que sobresalió en términos de rendimiento. Sin embargo, es importante notar que, aunque estos modelos mostraron mejoras, ninguno logró una exactitud perfecta, lo que indica que aún hay espacio para la optimización y refinamiento en futuros experimentos.