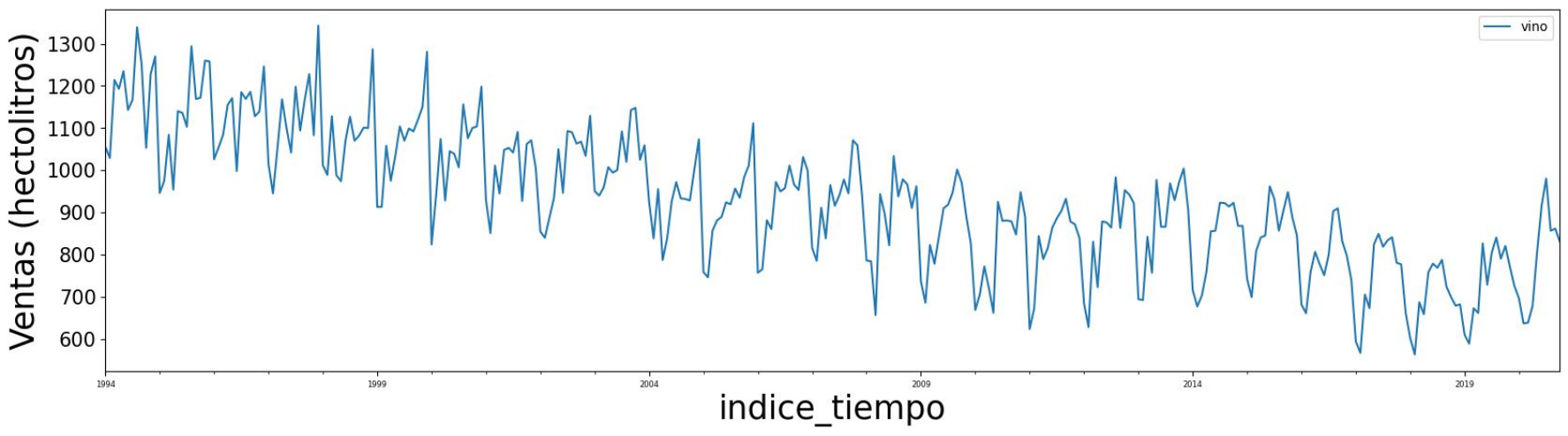
FACULTAD
DE INGENIERIA
Universidad de Buenos Aires

Series Temporales

1B 2024
Sergio Hinojosa

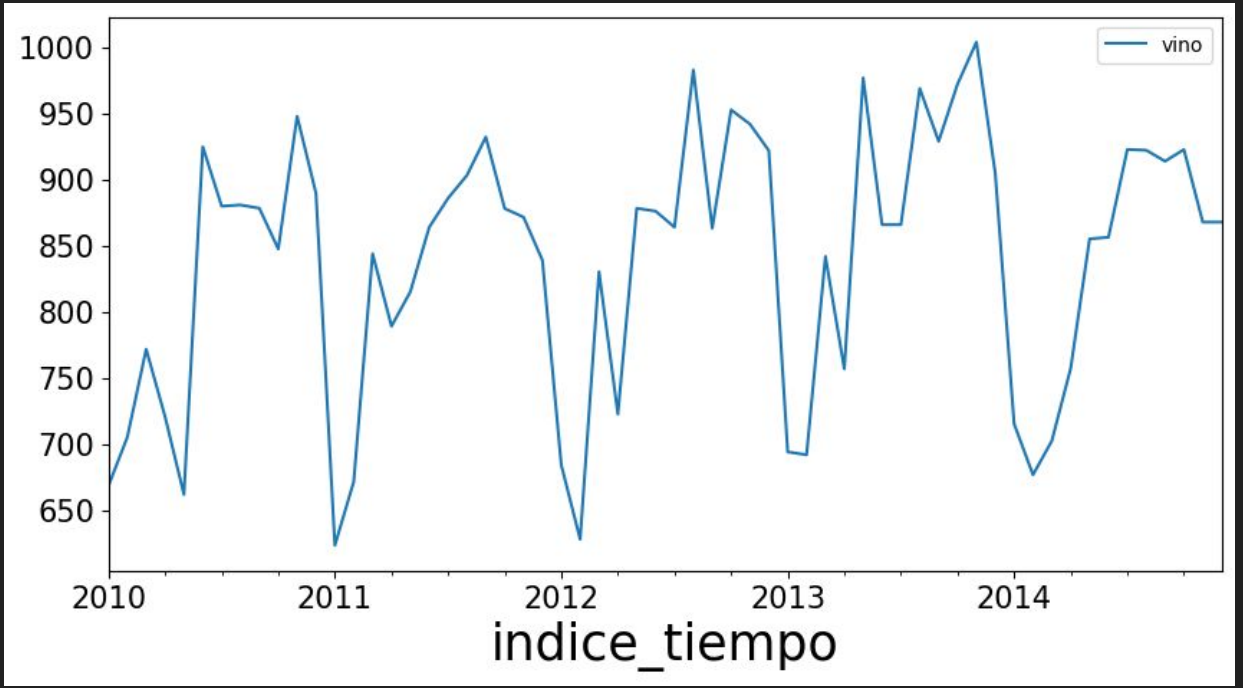
Dataset

Ventas al mercado interno de Producción Nacional de Vino en miles de hectolitros
Frecuencia de actualización: Mensual
Unidades: Miles de hectolitros
322 muestras



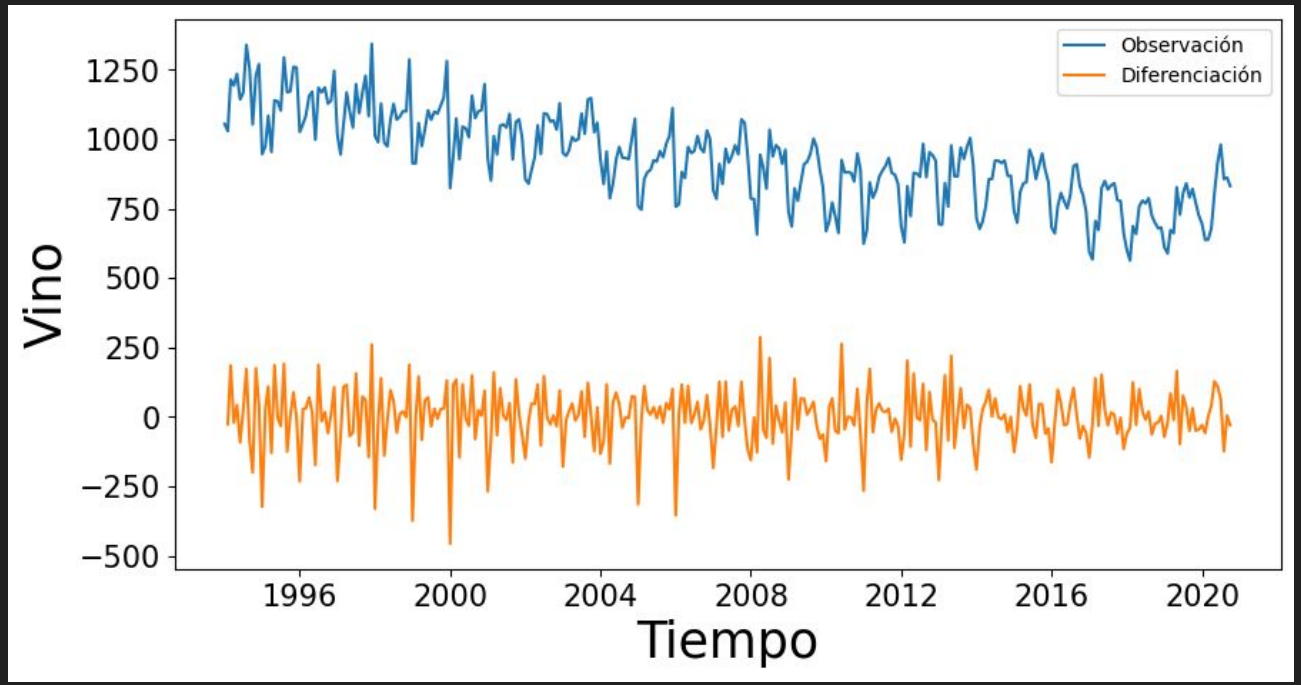
Dataset

Intervalo de 5 años



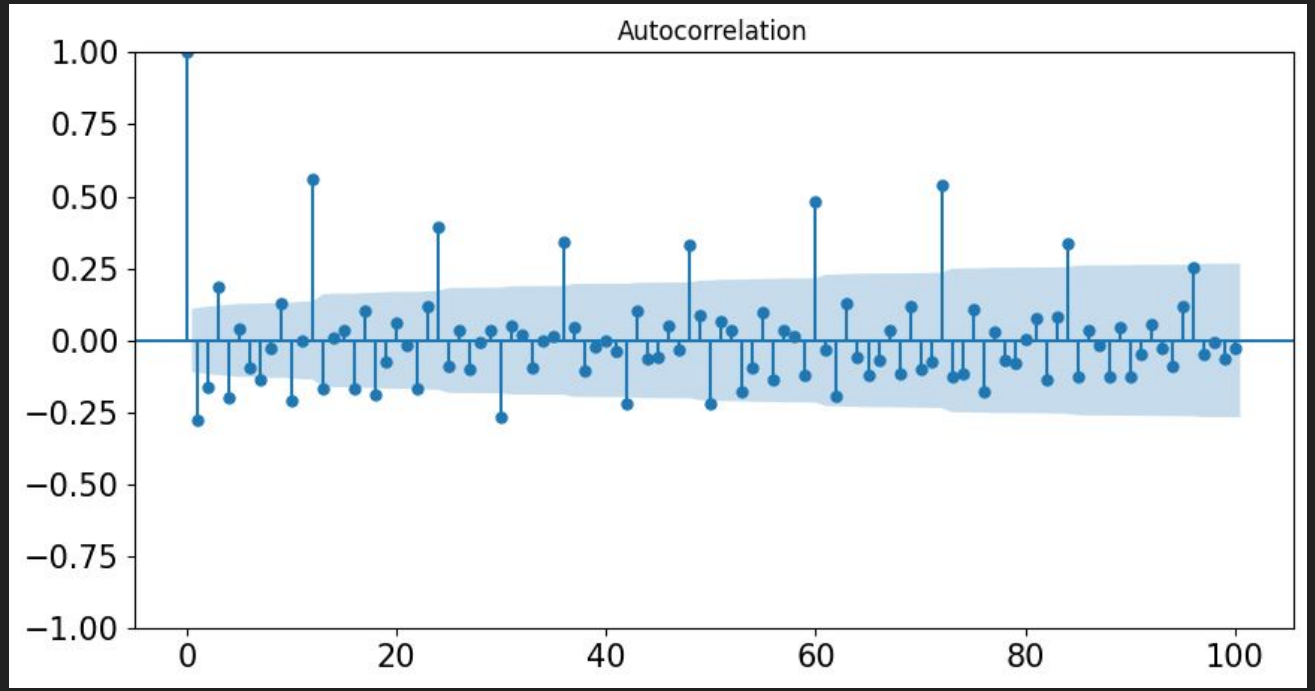
Preprocesamiento

Diferenciación



Preprocesamiento

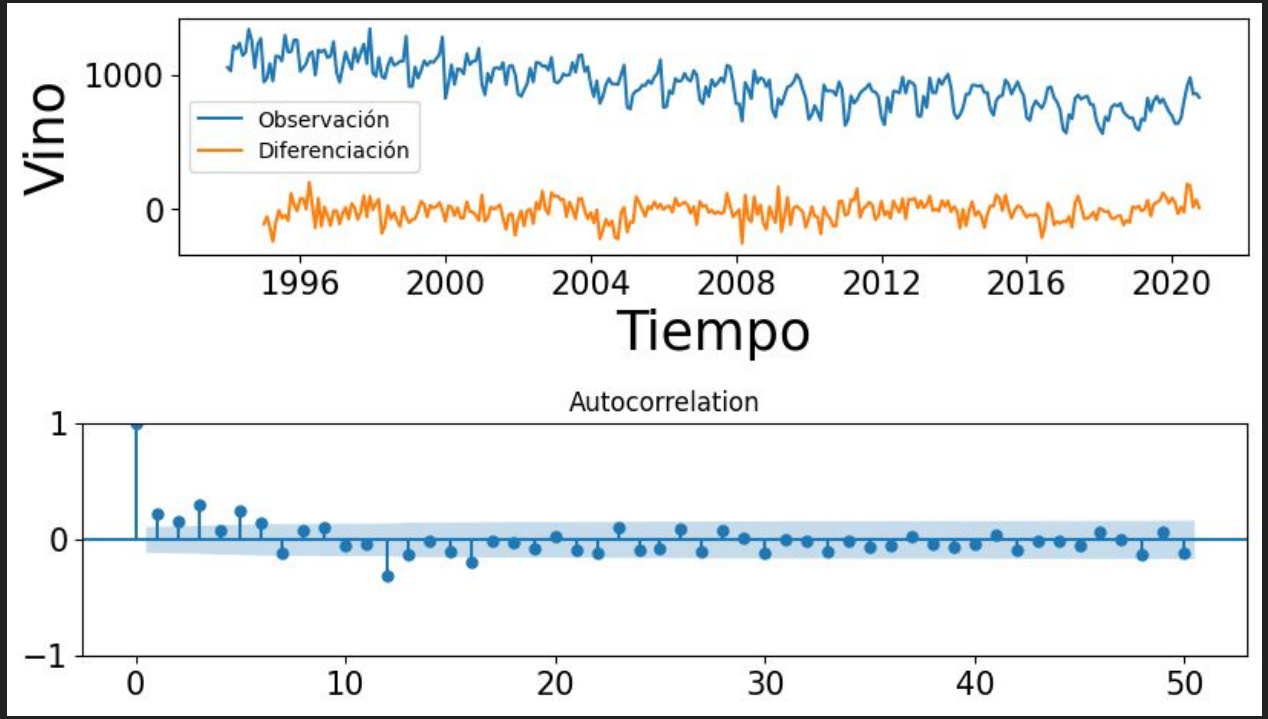
Diferenciación



Preprocesamiento

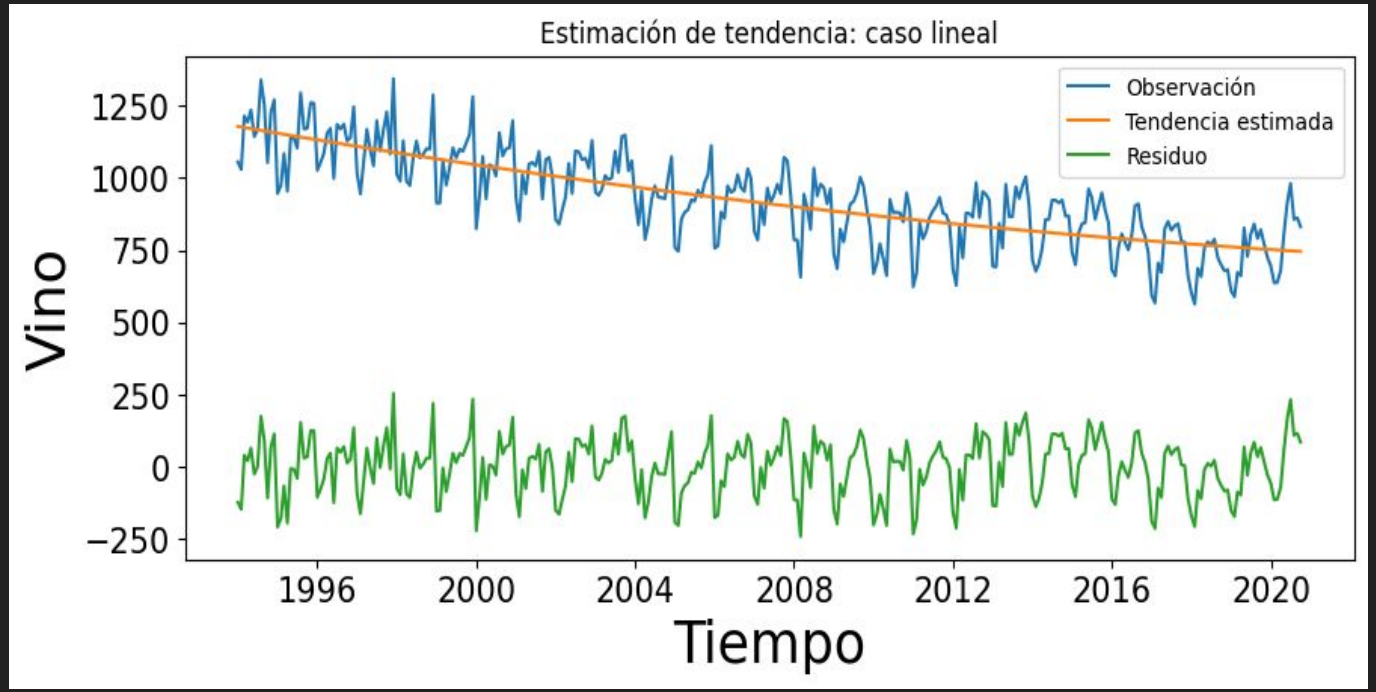
Diferenciación

a 12 pasos



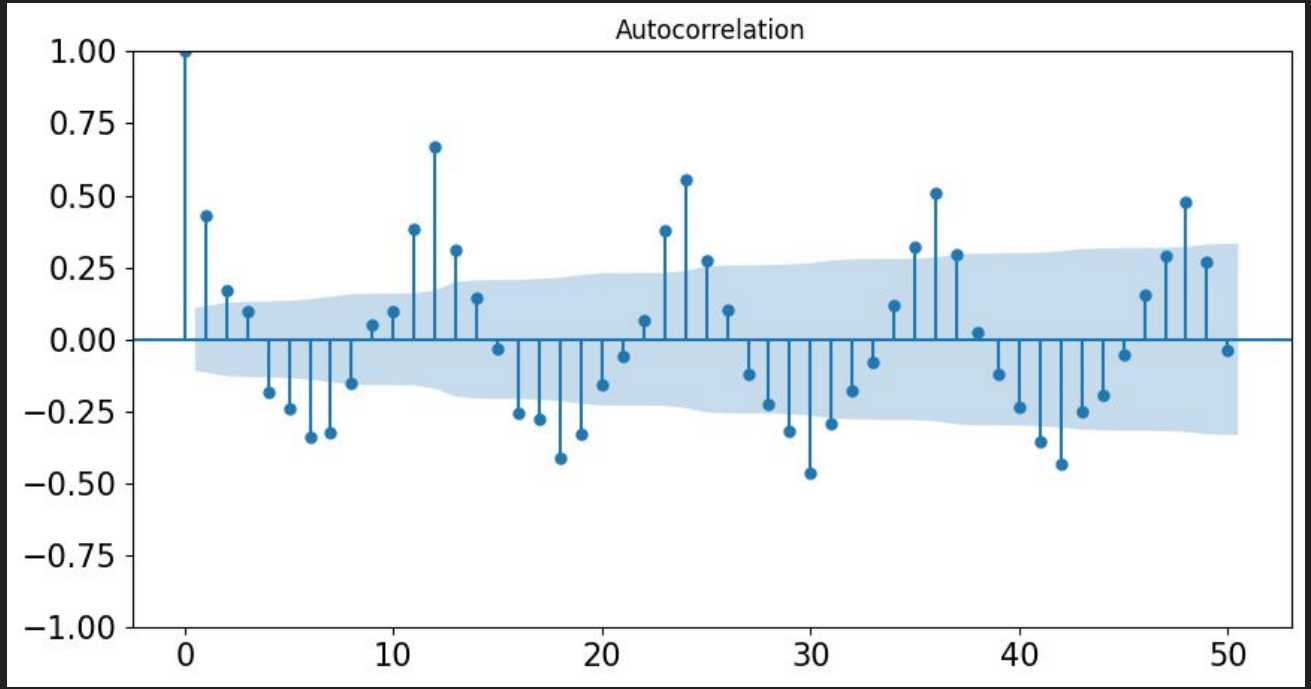
Preprocesamiento

Tendencia lineal



Preprocesamiento

Tendencia lineal



Estacionareidad

Test de Dickey-Fuller

Estadístico	-4.318932
p-valor	0.000412
Lags usados	17.000000
Observaciones usadas	304.000000
Valor del umbral (1%)	-3.452045
Valor del umbral (5%)	-2.871095
Valor del umbral (10%)	-2.571861

El estadístico es menor que los 3 umbrales, por lo tanto puedo rechazar la hipótesis nula. Por otro lado, el p-valor es muy chico.
Según este test la serie diferenciada es ESTACIONARIA.

Estacionareidad

Test KPSS

Estadístico	0.040966
p-valor	0.100000
Lags usados	7.000000
Valor del umbral (10%)	0.347000
Valor del umbral (5%)	0.463000
Valor del umbral (2.5%)	0.574000
Valor del umbral (1%)	0.739000

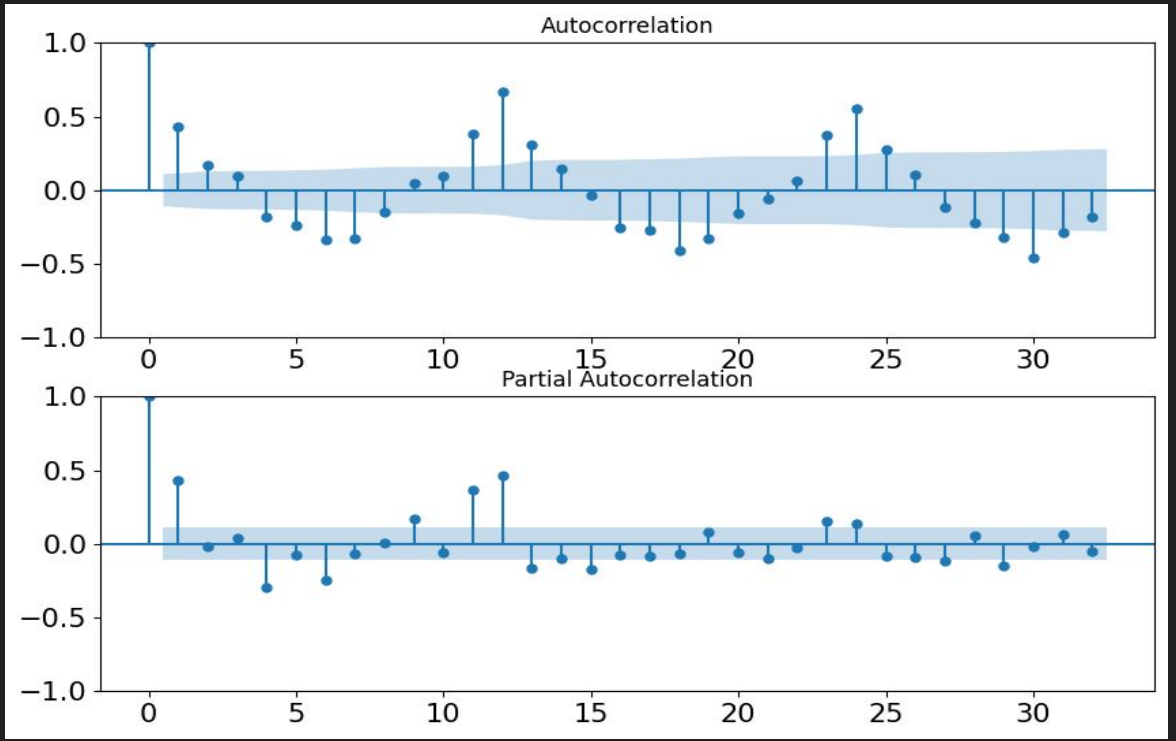
En este caso el p-valor no es significativamente chico por lo que no puedo rechazar el test, puedo asumir entonces que la serie es ESTACIONARIA.

Ambos test dan indicio que la serie es ESTACIONARIA.

Modelo ARIMA

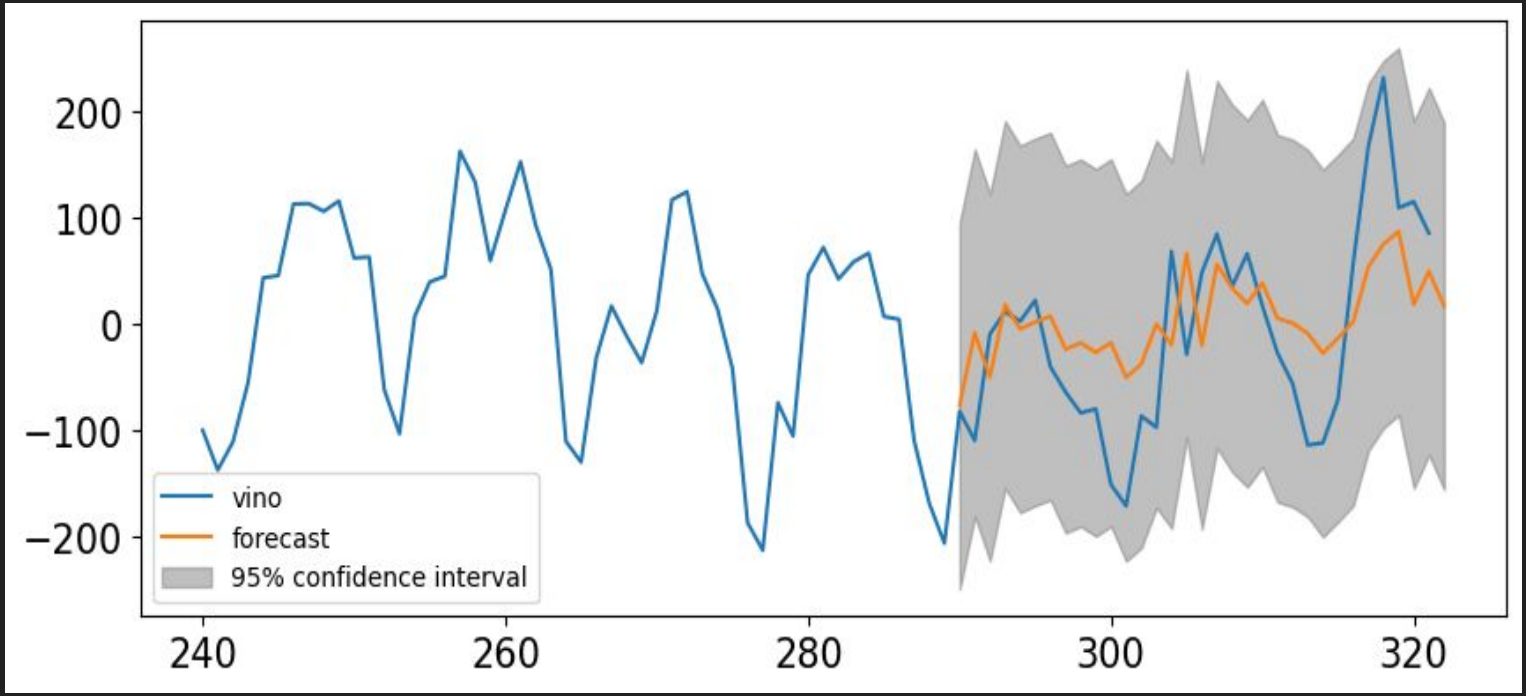
p = 1, q = 2, d = 0

Log Likelihood	-1900.638
AIC	3811.276
BIC	3830.148
Ljung-Box (L1) (Q)	0.03
Prob(Q)	0.86

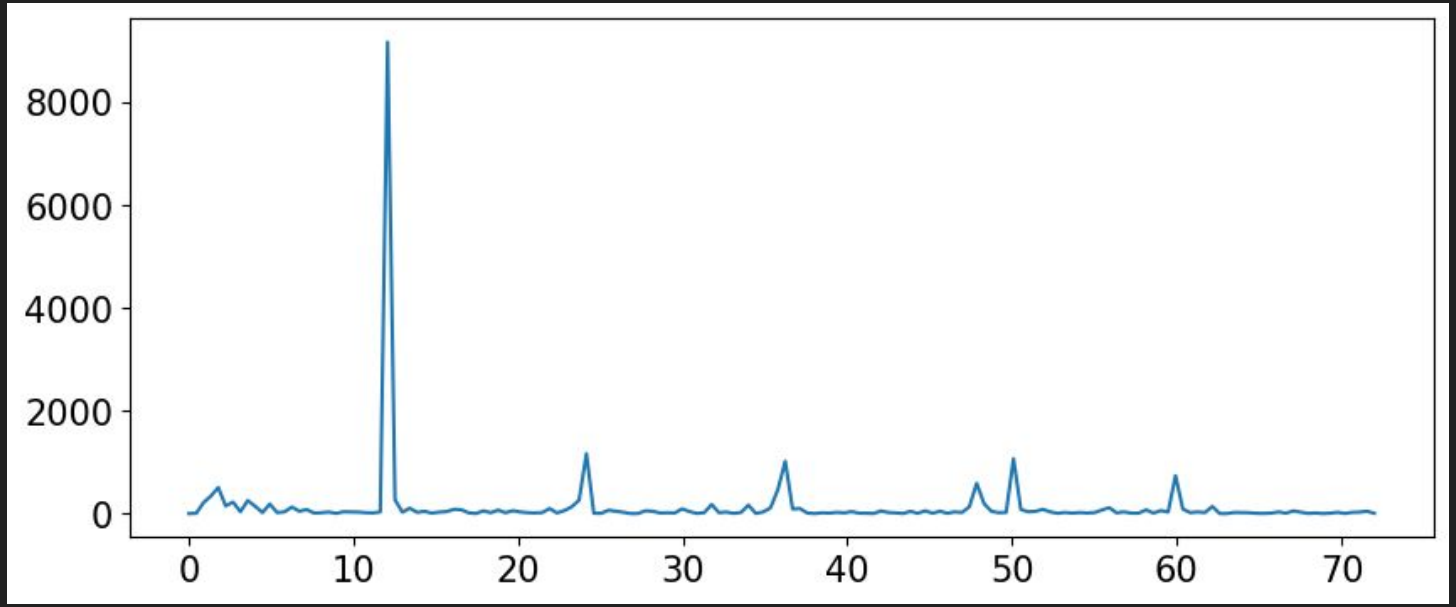


Modelo ARIMA

Predicción



Analisís espectral

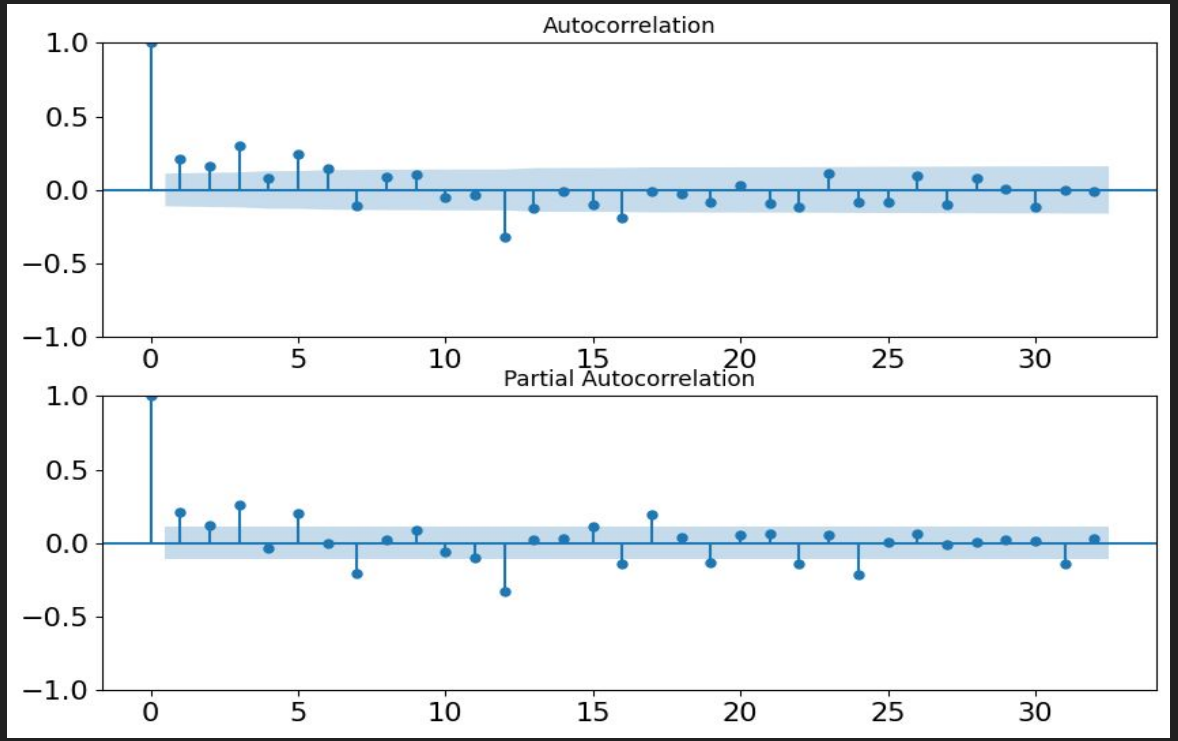


Modelo SARIMA

p = 1, d = 0, q = 2

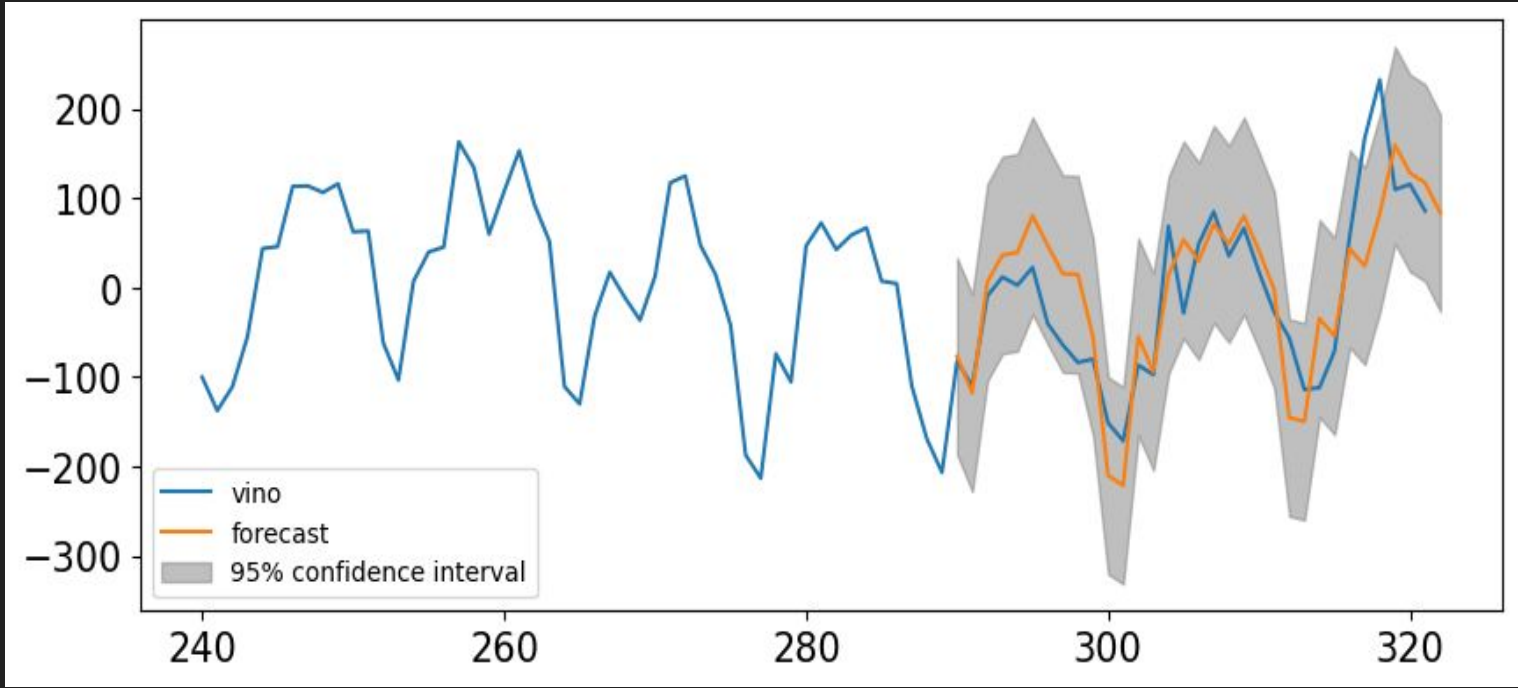
P = 3, D = 1, Q = 3, S = 12

Log Likelihood	-1700.961
AIC	3421.922
BIC	3459.288
Ljung-Box (L1) (Q)	0.02
Prob(Q)	0.89



Modelo ARIMA

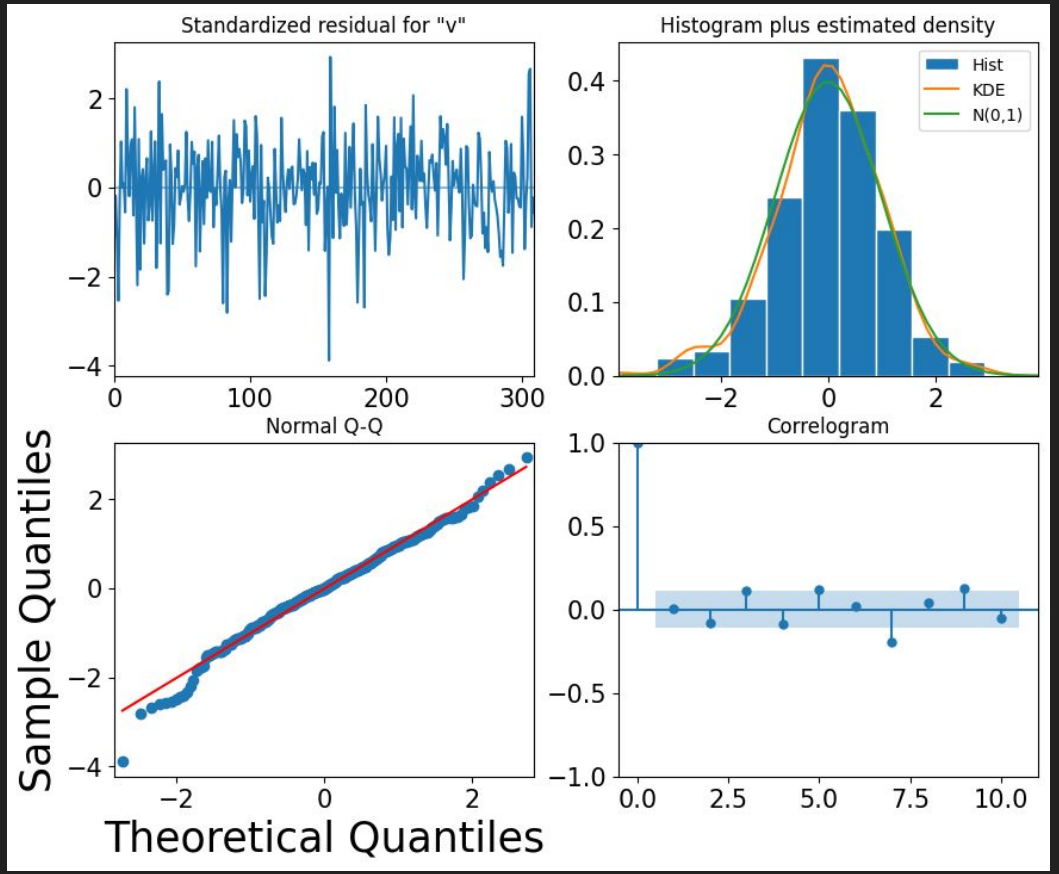
Predicción



Modelo SARIMA

Análisis de residuos

Ljung-Box (L1) (Q)	0.02
Prob(Q)	0.89



Modelo SARIMA

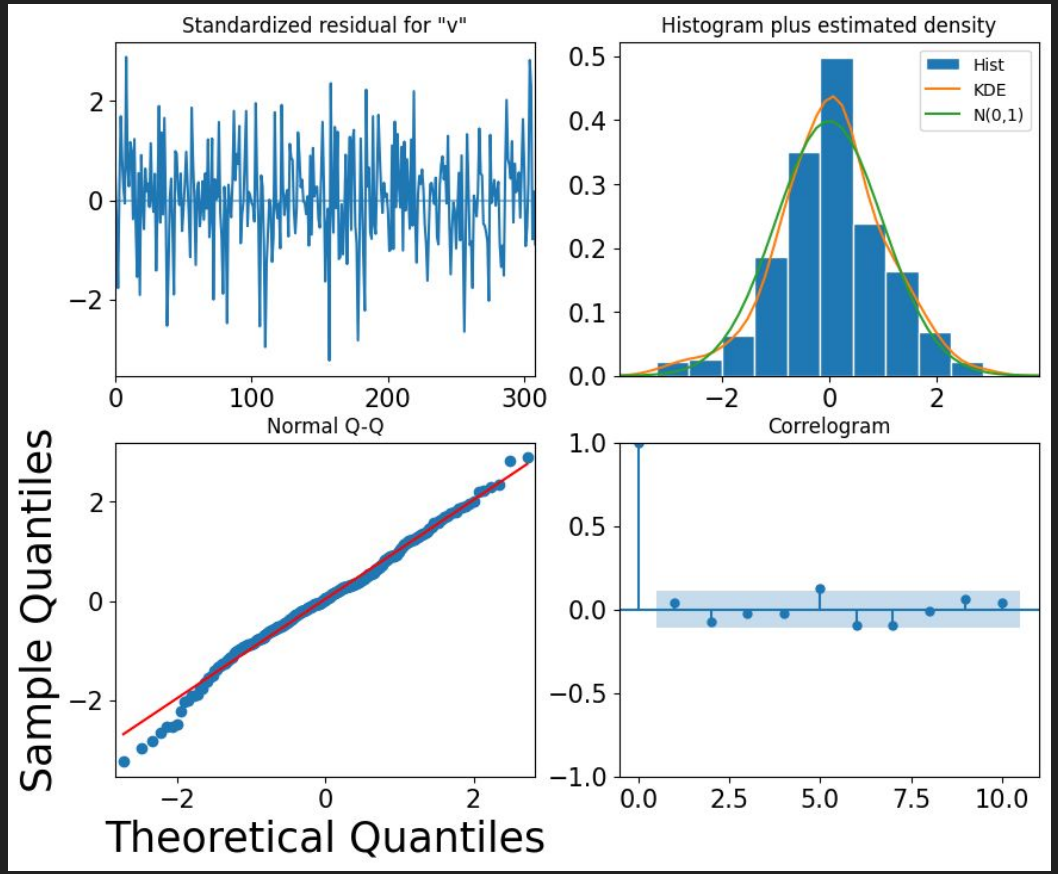
```
# Set parameter range
p = range(1,4)
q = range(1,4)
d = range(0,2)
s = [12]

pdq = list(product(p, d, q))
seasonal_pdq = list(product(p, d, q, s))
# Create SARIMA model for each order and seasonal order
aics = []
for order in pdq:
    for seasonal_order in seasonal_pdq:
        try:
            model = ARIMA(serie, order=order, seasonal_order=seasonal_order)
            results = model.fit()
            if results.llf != 0:
                aics.append((order, seasonal_order, results.aic))
        except:
            print('SARIMA{},{} - Skipped'.format(order, seasonal_order))

# Check for smallest AIC
aics.sort(key=lambda x: x[2])
```

Modelo SARIMA

p = 1, d = 0, q = 2

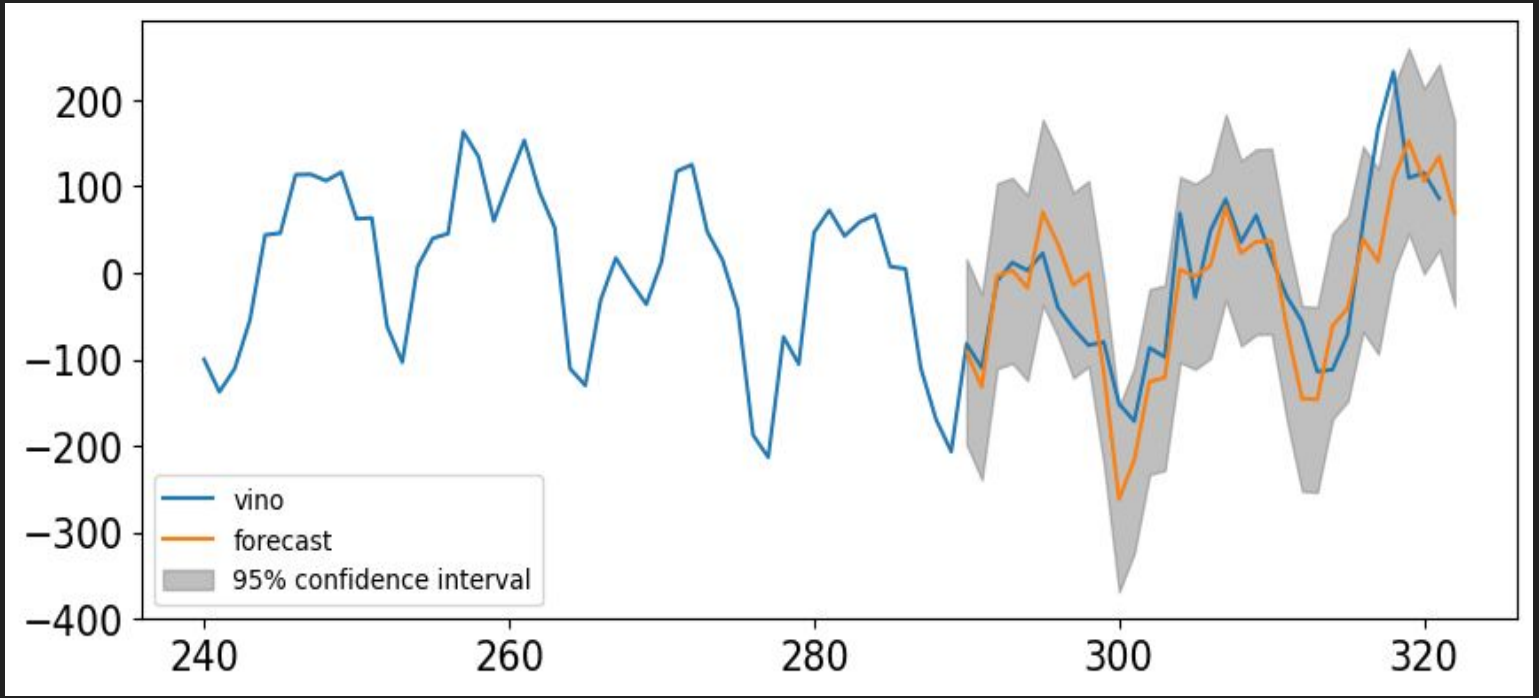
P = 1, D = 1, Q = 2, S = 12

Log Likelihood	-1684.040
AIC	3386.080
BIC	3419.680
Ljung-Box (L1) (Q)	0.59
Prob(Q)	0.44



Modelo ARIMA

Predicción



Red LSTM

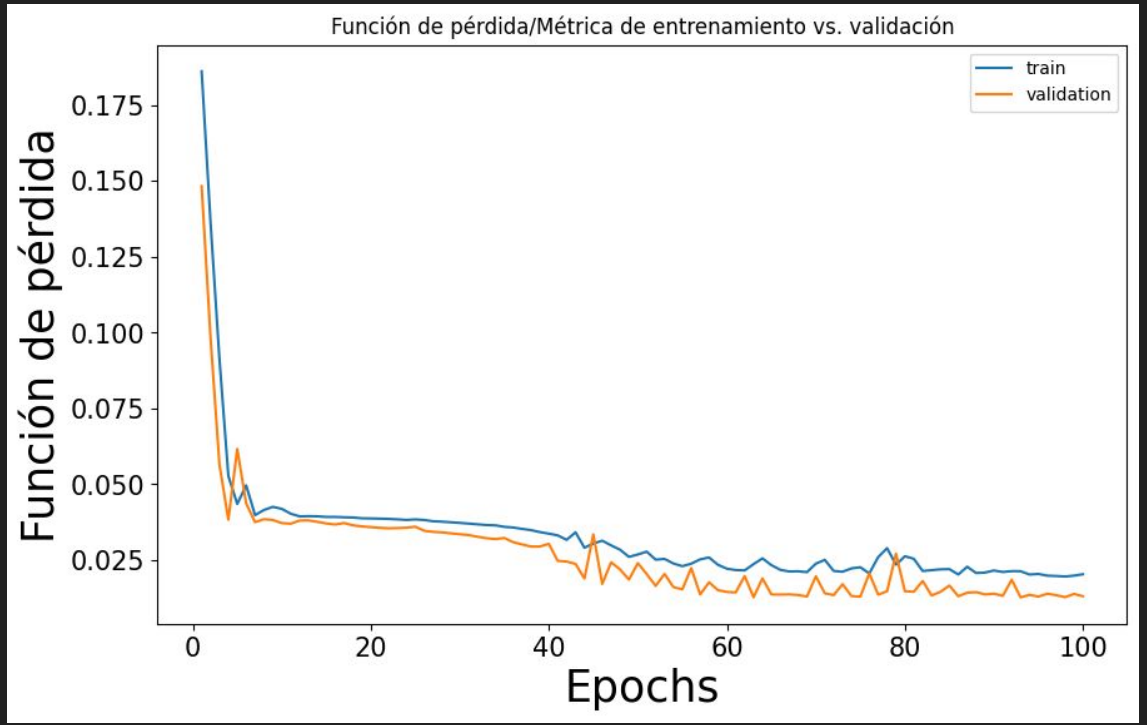
Entrada de 2 ciclos

(entrada de 24 muestras)

50 bloques de

neuronas LSTM

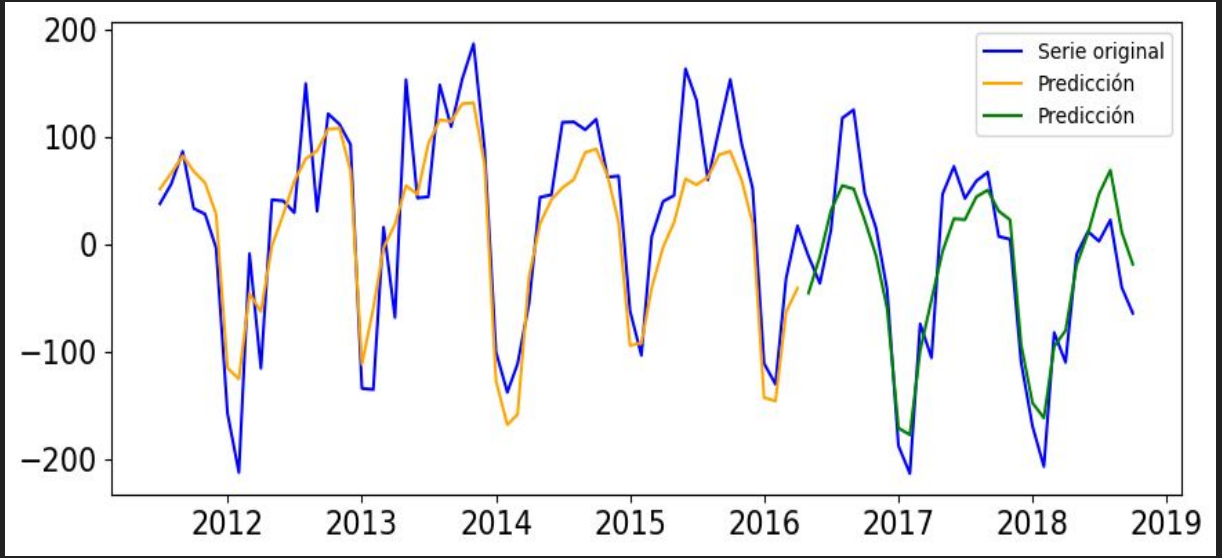
Función de activación Relu



Red LSTM

Train Score: 66.75 RMSE

Test Score: 69.33 RMSE



Conclusiones

	AIC	BIC	RMSE
ARMA(1,2)	3811.276	3830.148	88.60
SARIMA(1,0,2)(3,1,3,12)	3421.922	3459.288	61.09
SARIMA(1,0,2)(1,1,2,12)	3386.080	3419.680	62.10
Red LSTM (50)			66.75