

**”Modelo híbrido SOM + regresión lineal para la gestión
del retail: modelos predictivos basados en mapa de
distancia de neuronas”.**

Gregory J. Meléndez A.^a (gjmelendez@uninorte.edu.co), Jesús D. Zamora T.^b
(thowinssonj@uninorte.edu.co)

Corresponding Author:

Gregory J. Meléndez A.

Department of Mathematics and Statistics, Universidad del Norte,
Barranquilla 080001, Colombia

Tel: (+57) 3244737191

Email: gjmelendez@uninorte.edu.co

Corresponding Author:

Jesús D. Zamora Th.

Department of Mathematics and Statistics, Universidad del Norte,
Barranquilla 080001, Colombia

Tel: (+57) 3218298086

Email: thowinssonj@uninorte.edu.co

Modelo híbrido SOM + regresión lineal para la gestión del retail: Modelos predictivos basados en mapa de distancia de neuronas

Gregory J.Melendez A.^a, Jesus D. Zamora Th.^a

^a*Department of Mathematics and Statistics, Universidad del Norte, Barranquilla 080001, Colombia*

Abstract

El análisis de datos en ventas es fundamental para apoyar la toma de decisiones estratégicas en el contexto actual del comercio minorista. Este estudio tiene como objetivo analizar y predecir el comportamiento de las ventas en función de variables clave como el número de factura (**Invoice**), el código del producto (**StockCode**), la descripción del producto (**Description**), la cantidad vendida (**Quantity**), la fecha de transacción (**InvoiceDate**), el precio unitario (**Price**), el identificador del cliente (**Customer ID**) y su país de origen (**Country**). A partir de técnicas avanzadas de aprendizaje automático (**Machine Learning**) y modelos de series temporales, se busca predecir la demanda futura de productos para optimizar la gestión de inventarios. Los resultados indican que el modelo híbrido propuesto, que integra Mapas Autoorganizados (SOM) y regresión lineal, logra reducir sustancialmente los errores de predicción en comparación con modelos individuales, demostrando ser una herramienta eficaz para la gestión estratégica del retail.

Se proponen varios modelos estadísticos y de Machine Learning para abordar este desafío. Estos incluyen **ARIMA**, **Prophet** y **LSTM**, que son cruciales para predecir tendencias y patrones de ventas, evaluar la estacionalidad, detectar anomalías y pronosticar la demanda. La segmentación de clientes mediante

*Corresponding author.

Email addresses: gjmelendez@uninorte.edu.co (Gregory J.Melendez A.),
thowinssonj@uninorte.edu.co (Jesus D. Zamora Th.)

técnicas de clustering como **K-Means** permite un análisis más profundo de los comportamientos de compra, identificando clientes recurrentes y aquellos con alto riesgo de abandono. Además, se evaluará el impacto del precio en la cantidad vendida, optimizando estrategias de precios dinámicos.

El estudio también hace hincapié en la predicción de los productos más demandados en diferentes temporadas, identificando aquellos con baja rotación para generar estrategias promocionales efectivas. En cuanto a la optimización de inventarios, se propondrán métodos para predecir los niveles de stock necesarios, evitando tanto la falta de existencias como el exceso de inventario. Este enfoque integra el análisis de ventas en línea y minoristas, permitiendo una evaluación comparativa de los resultados en diferentes mercados geográficos (países y regiones), con el objetivo de identificar mercados emergentes con potencial de crecimiento.

El uso de algoritmos avanzados de detección de anomalías, como Isolation Forest, proporciona a las empresas una herramienta poderosa para identificar patrones inusuales en sus datos operativos, lo que contribuye a tomar decisiones informadas en áreas clave como la distribución de inventarios, la fijación de precios y el diseño de campañas de marketing personalizadas. A diferencia de modelos predictivos como las redes neuronales recurrentes (LSTM), que se centran en la predicción de tendencias futuras, Isolation Forest se especializa en la detección temprana de comportamientos atípicos que podrían indicar fraudes, errores en la cadena de suministro o fluctuaciones inesperadas en la demanda.

Este estudio adopta un enfoque integral que combina la detección de anomalías con la analítica predictiva, permitiendo a las empresas adaptar sus estrategias en un entorno competitivo y globalizado. Al aprovechar al máximo los datos históricos y la capacidad de identificar desviaciones significativas, las organizaciones pueden optimizar sus procesos, minimizar riesgos y responder de manera proactiva a las dinámicas cambiantes del mercado.

- **Contexto:** La predicción de la demanda de productos en las tiendas minorista Online, como Amazon, Temu o similares, es crucial para mejorar

la gestión de inventarios, la optimización de precios y la planificación de estrategias de marketing. Este estudio se centra en predecir la cantidad de productos que se venderán en el futuro cercano utilizando factores clave como las tendencias de ventas históricas, el comportamiento de los consumidores, las condiciones de mercado y otros datos relevantes. Predecir la demanda con precisión permite a las empresas optimizar su cadena de suministro, reducir costos por sobreinventario y minimizar el desabastecimiento. ¿Cómo pueden las tiendas mayoristas online, como Amazon, Temu o similares, utilizar modelos predictivos de demanda para mejorar la precisión en la gestión de inventarios y la toma de decisiones estratégicas en un mercado competitivo?

- **Métodos:** Para predecir la demanda de productos en el retail online, se utilizan diversos modelos de Machine Learning capaces de captar la complejidad de las relaciones entre los diferentes factores que influyen en las compras de los clientes. Los modelos más adecuados incluyen:
 - **Regresión lineal y no lineal:** Modelos simples que pueden identificar tendencias generales en las ventas a lo largo del tiempo.
 - **Árboles de decisión:** Útiles para clasificar las condiciones de mercado y su impacto en la demanda de productos.
 - **Random Forest:** Un modelo que mejora la precisión mediante la combinación de múltiples árboles de decisión, lo que permite identificar patrones complejos.
 - **Redes neuronales artificiales (ANN):** Modelos avanzados capaces de aprender relaciones no lineales complejas en grandes volúmenes de datos.
 - **XGBoost:** Un algoritmo de boosting eficiente que combina varios modelos débiles para generar predicciones más precisas y robustas.
 - **Máquinas de vectores de soporte (SVM):** Adecuadas para modelar relaciones entre variables con alta dimensionalidad.

El preprocesamiento de los datos, como la limpieza, normalización y selección de características, es fundamental para mejorar la precisión de las predicciones. Los datos de ventas históricos, junto con las variables adicionales como campañas de marketing, temporadas festivas o cambios en los precios, se incorporan para entrenar los modelos y optimizar los resultados.

- **Resultados:** Principales hallazgos del artículo;

1. **Variabilidad y estacionalidad en las ventas**

- El análisis exploratorio de datos (**EDA**) mostró picos y caídas significativas en la cantidad de productos vendidos (**Quantity**) y en los ingresos diarios, lo que sugiere la presencia de estacionalidad y posibles impactos de promociones o eventos específicos (ej., temporada navideña)
- El Reino Unido fue el principal mercado, aunque se identificó un flujo internacional de clientes que puede traducirse en oportunidades de crecimiento global.

2. **Productos más vendidos y estrategias de stock**

- Se detectaron productos con alta rotación y bajo precio, que impulsan el volumen de ventas. El hallazgo sugiere la relevancia de mantener un inventario suficiente y explorar estrategias de precios dinámicos para estos artículos de alta demanda.
- El análisis de artículos con menor rotación permite optimizar promociones dirigidas, reduciendo el riesgo de sobreinventario.

3. **No estacionariedad y presencia de valores extremos**

- Pruebas como KPSS y Kolmogorov-Smirnov evidenciaron que las series de Quantity y Price no eran estacionarias y, además, no seguían

una distribución normal. Esto obliga al uso de métodos más complejos (diferenciación, transformaciones, o modelos robustos) para un modelado adecuado.

- Modelos de aprendizaje profundo (**MLP, RNN, LSTM**) y métodos híbridos mostraron mejor capacidad para adaptarse a las relaciones no lineales y a la complejidad de la serie, aunque requieren un cuidadoso ajuste de hiperparámetros y un preprocesamiento adecuado.

4. Comparación de modelos estadísticos y de machine learning

- Modelos tradicionales de series de tiempo (**como ARIMA o Holt-Winters**) capturan cierta estructura de tendencia y estacionalidad, pero exhiben alta varianza en los residuos y un ajuste insuficiente ante la gran volatilidad de los datos.
- Modelos de aprendizaje profundo (MLP, RNN, LSTM) y métodos híbridos mostraron mejor capacidad para adaptarse a las relaciones no lineales y a la complejidad de la serie, aunque requieren un cuidadoso ajuste de hiperparámetros y un preprocesamiento adecuado.

5. Uso de algoritmos de detección de anomalías

- El aislamiento de datos anómalos mediante Isolation Forest agrega valor en la identificación temprana de comportamientos atípicos (fraudes, errores de facturación, fluctuaciones repentinas de demanda), complementando la perspectiva de predicción.

6. Eficacia del enfoque híbrido (SOM + regresión lineal)

- El experimento con Mapas Autoorganizados (**SOM**) sirvió para agrupar y reducir la dimensionalidad, y posteriormente aplicar regresión lineal.
- Comparado con SOM usado de forma aislada, el modelo híbrido arrojó una reducción sustancial en el error cuadrático medio (**MSE y RMSE**). Sin embargo, el **MAPE** se mantuvo elevado, indicando

sensibilidad a valores extremos de la serie. Aun así, se evidenció una promesa clara para capturar patrones complejos y luego afinar la predicción a través de una regresión supervisada.

- **Conclusión:** El enfoque híbrido propuesto, que integra Mapas Autoorganizados (**SOM**) y regresión lineal, logró una reducción significativa en el error de predicción (**MSE y RMSE**) comparado con el uso aislado del SOM. Aunque el error porcentual medio absoluto (**MAPE**) permaneció relativamente alto, debido a la sensibilidad a valores extremos, el método híbrido demostró una alta capacidad para capturar patrones complejos en los datos de ventas. Este resultado indica que el modelo híbrido puede ser una herramienta efectiva para la gestión estratégica del retail, especialmente en la optimización de inventarios y en la identificación anticipada de patrones en la demanda.

Keywords: predicción de demanda, comercio minorista online, optimización de inventarios, modelos de Machine Learning, análisis de ventas, gestión de inventarios, comportamiento del consumidor, técnicas predictivas, análisis de datos, demanda de productos.

Data

El conjunto de datos **Online Retail II** proporciona una visión integral de las transacciones realizadas por un comercio minorista en línea del Reino Unido entre 2009 y 2011. A partir del análisis exploratorio, se observa una notable variación en las **ventas diarias**, con picos significativos que podrían estar vinculados a eventos estacionales o promociones específicas, como la temporada navideña.

Los **ingresos diarios** reflejan que, en ocasiones, días con menor volumen de ventas generaron mayores ingresos, lo que sugiere la comercialización de productos de mayor valor en esas fechas. En el ámbito geográfico, el **Reino Unido** lidera en ventas, aunque existe una base de clientes internacional que representa oportunidades de crecimiento global. Además, se identifican **clientes recurrentes** con altos volúmenes de compra, lo que podría asociarse a mayoristas o compradores frecuentes, destacando la importancia de implementar estrategias de fidelización.

En conjunto, el análisis revela patrones de comportamiento del consumidor, estacionalidad en las ventas e información valiosa para la toma de decisiones comerciales estratégicas. El conjunto de datos fue tomado de **UC Irvine Machine Learning Repository (Online Retail II Dataset)**.

El **DataFrame** está constituido por **1,067,371** filas y **8** columnas clave, que describen las transacciones realizadas. A continuación se presenta una descripción detallada de cada variable:

Columna	Descripción	Valores Nulos
Invoice	Identificador único de cada transacción.	0
StockCode	Código único asignado a cada producto.	0
Description	Descripción detallada del producto vendido.	4,382
Quantity	Cantidad de unidades vendidas en la transacción.	0
InvoiceDate	Fecha y hora exacta de la transacción.	0
Price	Precio unitario del producto en libras esterlinas (£).	0
Customer ID	Identificador único del cliente que realizó la compra.	243,007
Country	País de origen del cliente.	0

Table 1: Descripción de las variables y valores nulos en el DataFrame Online Retail II

Modelo	MAE	MSE	RMSE	MAPE (%)	Vbds_p_value (BDS)	Valor p (Diebold-Mariano)	Prueba de Ljung-Box (p-value)
ES-WH	7549.27	89593986.58	9465.41	50.48	1.0	0.642889	0.321782
ES-WH(2)	7454.33	87629327.28	9361.05	49.71	0.04025	0.546611	0.0717
ES-HW	11517.40	207611618.40	14408.73	nan	0.552246	0.546611	0.0108
ES-HW ADITIVO	8466.59	115735564.03	10758.05	72.75	0.01855	0.546611	0.0145
ES-HW MULTIPLICATIVO	8466.59	115735564.03	10758.05	nan	0.01855	0.642889	0.0145
ARIMA	9850.47	152041110.28	12330.50	nan	0.06515	0.642889	0.0000
GARCH	9046.29	132709046.46	11519.94	0.43	0.13823	0.642889	0.0000
MLP	12.21	4540.23	67.38	401.95	0.55514	0.000109	0.9994
RNN	1229.30	2374622.21	1540.98	5.39	0.72060	0.000109	0.2946
LSTM	1531.43	3406475.40	1845.66	5.84	0.32906	0.000109	0.2193

Table 2: Resultados de los modelos de predicción con sus métricas y pruebas estadísticas

La tabla presentada resume los resultados de una comparación exhaustiva entre diversos modelos de predicción, evaluados a través de varias métricas estadísticas clave: MAE (Error Absoluto Medio), MSE (Error Cuadrático Medio), RMSE (Raíz del Error Cuadrático Medio), MAPE (Error Porcentual Absoluto Medio), el valor p de la prueba de Diebold-Mariano, la prueba de Ljung-Box, y el valor p de la prueba BDS (Brock, Dechert y Scheinkman). Estas métricas son fundamentales para evaluar la precisión y fiabilidad de los modelos de predicción. Al analizar los resultados, observamos que los modelos más sofisticados, como las redes neuronales MLP (Perceptrón Multicapa) y LSTM (Long Short-Term Memory), muestran un rendimiento sobresaliente, especialmente en cuanto a MAE, MSE y RMSE, lo que indica una alta precisión en la predicción. Sin embargo, es importante notar que aunque el MLP tiene los mejores resultados en términos de MAE (12.21), MSE (4540.23) y RMSE (67.38), su MAPE es significativamente alto (401.95) lo que sugiere que es sensible a los valores atípicos o errores grandes en las predicciones. Por otro lado, el modelo ES-WH, aunque con un rendimiento inferior en MAE y MSE, tiene un MAPE moderado (50.48) y presenta un valor p relativamente alto en la prueba de Ljung-Box (0.321782), indicando que los residuos del modelo se ajustan mejor a la distribución de ruido blanco, lo que sugiere que este modelo es más robusto frente a anomalías en los datos. Es particularmente relevante observar que los modelos clásicos, como ARIMA y GARCH, aunque no presentan los mejores resultados en las métricas de error, muestran valores p muy bajos en la prueba de Ljung-Box, lo que indica que sus residuos no siguen un patrón aleatorio y podrían indicar la

presencia de autocorrelación o heterocedasticidad, características que deben ser consideradas dependiendo del tipo de datos y la aplicabilidad del modelo. En resumen, a pesar de que el MLP destaca en términos de precisión, especialmente en las métricas de error cuadrático y absoluto, su sensibilidad a valores atípicos y su alto MAPE podrían hacer que no sea la mejor opción en contextos donde la variabilidad en los datos sea elevada. Por lo tanto, una evaluación cuidadosa del tipo de datos y de las características específicas del problema de predicción es crucial para seleccionar el modelo más adecuado, y, en este caso, el MLP es el modelo más preciso pero podría ser sensible a anomalías, mientras que modelos como ES-WH podrían ofrecer una opción más robusta dependiendo de los requerimientos específicos del análisis.

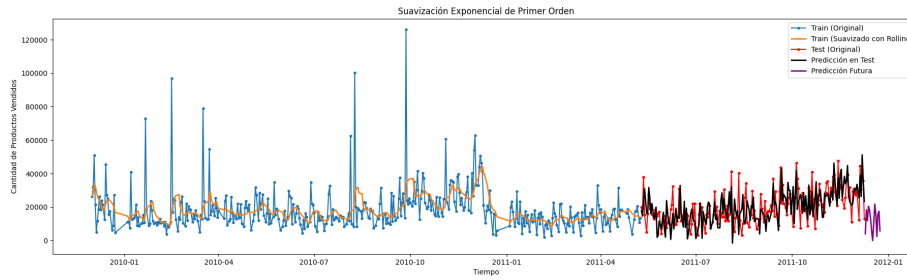


Figure 1: suavización exponencial

la suavización exponencial de primer orden aplicada al conjunto de entrenamiento. En este gráfico, la línea azul representa los valores originales del conjunto de entrenamiento, mientras que la línea naranja muestra los valores suavizados con el método de suavización exponencial, que utiliza un parámetro λ para asignar más peso a las observaciones más recientes. El gráfico también incluye el conjunto de test, mostrado en rojo, que contiene los valores originales para comparar las predicciones del modelo (línea negra) con los valores reales. La línea morada indica las predicciones futuras basadas en los últimos datos suavizados del conjunto de entrenamiento. La suavización exponencial de primer orden ayuda a reducir las fluctuaciones en los datos y permite una mejor captación de las tendencias subyacentes al suavizar los picos y valles

presentes en la serie original. Este enfoque es útil para realizar predicciones más estables y realistas, especialmente cuando se tiene una serie temporal con mucha variabilidad.

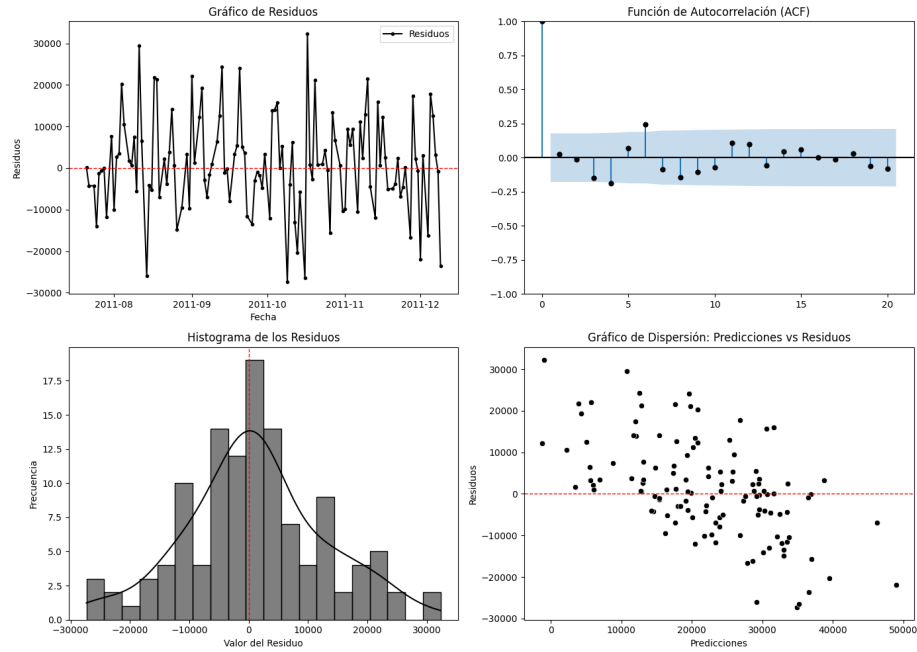


Figure 2: residuos suavización exponencial

El gráfico muestra un análisis detallado de los residuos del modelo, que son las diferencias entre los valores reales y las predicciones realizadas por el modelo. El gráfico superior izquierdo muestra los residuos a lo largo del tiempo. Un patrón aleatorio alrededor de cero sugiere que el modelo está capturando adecuadamente la dinámica de los datos. Si los residuos no tienen una distribución aleatoria, podría indicar que el modelo no ha capturado alguna estructura importante en los datos. El gráfico superior derecho muestra la función de autocorrelación (ACF) de los residuos, que analiza si los residuos están correlacionados en el tiempo. La ausencia de autocorrelación, indicada por valores cercanos a cero, sugiere que el modelo no ha dejado ninguna estructura temporal no explicada. El gráfico inferior izquierdo es un histograma de los residuos, acompañado

de una estimación de la densidad. Este gráfico evalúa la distribución de los residuos; una distribución cercana a una normal es deseable, ya que sugiere que el modelo no está sesgado. Finalmente, el gráfico inferior derecho es un gráfico de dispersión que muestra las predicciones del modelo frente a los residuos. La dispersión aleatoria alrededor de cero en este gráfico también indica que el modelo no presenta sesgos sistemáticos y que los errores son consistentes en función de las predicciones. En resumen, estos gráficos permiten verificar la calidad del modelo y si los residuos cumplen con las suposiciones de independencia, homocedasticidad y normalidad.

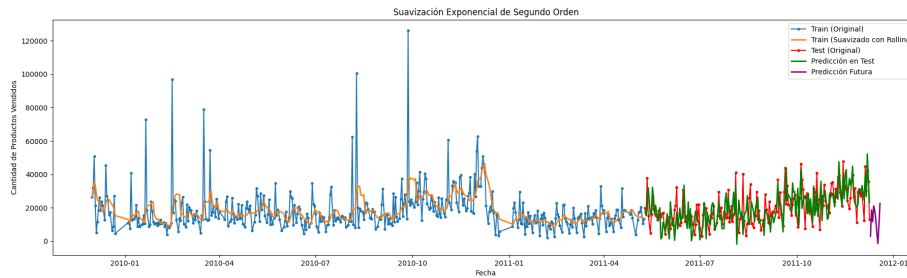


Figure 3: suavización exponencial(2)

En este gráfico se muestra el comportamiento de la suavización exponencial de segundo orden aplicada a los datos de productos vendidos. La línea azul representa los datos originales del conjunto de entrenamiento, mientras que la línea naranja muestra la serie suavizada, la cual captura las tendencias del conjunto de entrenamiento al reducir la variabilidad de los picos y valles. Las predicciones para el conjunto de test, mostradas en verde, se generan utilizando el modelo suavizado y reflejan cómo el modelo ajusta las fluctuaciones de los datos originales del test (línea roja). Además, las predicciones futuras (línea morada) extrapolan este comportamiento hacia el futuro, lo que permite anticipar las tendencias más allá de los datos observados. La suavización exponencial de segundo orden, a diferencia del primer orden, incluye un componente adicional que ajusta las tasas de cambio de las tendencias, permitiendo una mejor captura de dinámicas más complejas en la serie temporal.

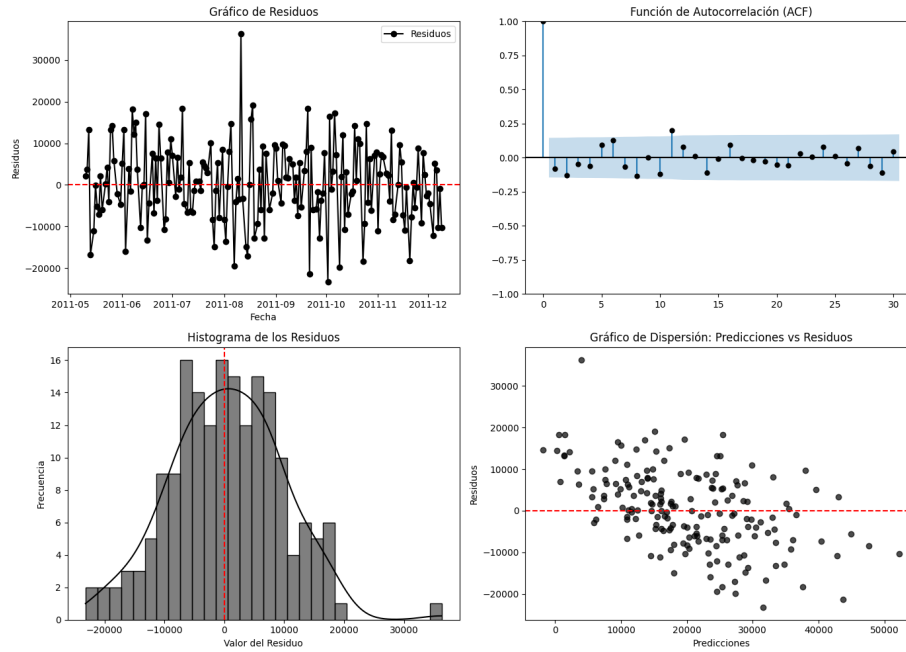


Figure 4: residuos suavización exponencial(2)

El gráfico de residuos es una herramienta esencial para evaluar la calidad del ajuste de un modelo de predicción, ya que muestra las diferencias entre los valores predichos por el modelo y los valores reales observados a lo largo del tiempo. Idealmente, los residuos deberían distribuirse de manera aleatoria alrededor de cero, sin mostrar patrones evidentes ni correlaciones significativas, lo que sugiere que el modelo ha capturado correctamente la dinámica subyacente de la serie temporal. Este comportamiento aleatorio en los residuos implica que el modelo ha sido capaz de extraer toda la información relevante de los datos, y que no quedan patrones sistemáticos que el modelo no haya podido explicar. Desde una perspectiva estadística, si los residuos muestran una distribución normal y no presentan autocorrelación, se puede concluir que el modelo es adecuado, ya que se ajusta a las suposiciones de homocedasticidad (variabilidad constante) e independencia de los errores. Sin embargo, si los residuos muestran patrones sistemáticos, como tendencias, ciclos o correlaciones temporales,

esto sugiere que el modelo no ha logrado capturar alguna estructura importante en los datos, como una posible estacionalidad, heterocedasticidad o dependencia temporal. En este caso, los residuos indican que el modelo debería ser ajustado o mejorado, ya sea incorporando variables adicionales, transformando los datos o utilizando técnicas más complejas que puedan capturar mejor las dependencias temporales.

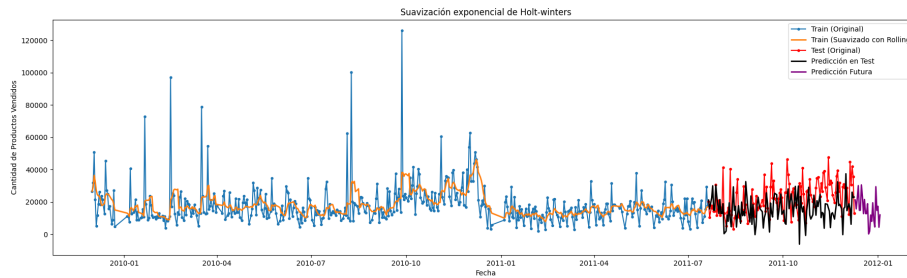


Figure 5: suavización exponencial HW

El gráfico de suavización exponencial de Holt-Winters muestra cómo los datos originales de productos vendidos (línea azul) se ajustan utilizando el modelo de Holt-Winters. La línea naranja representa la serie suavizada que captura tanto la tendencia como la estacionalidad de los datos. Las predicciones para el conjunto de test (línea negra) siguen el comportamiento de la serie temporal original (línea roja), lo que permite una comparación entre las predicciones y los valores reales. Las predicciones futuras (línea morada) muestran la proyección del comportamiento futuro de la serie temporal, basándose en los últimos valores observados. Este enfoque es adecuado cuando se tiene una serie con tendencias y estacionalidades, y el modelo se ajusta a las características de la serie temporal al considerar componentes aditivos o multiplicativos dependiendo de la naturaleza de la estacionalidad.

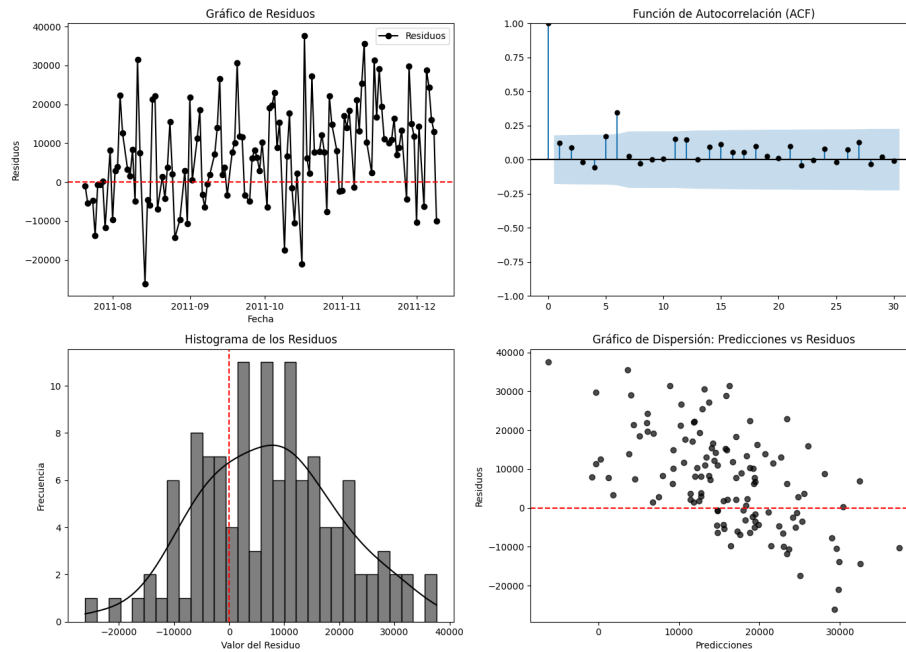


Figure 6: Residuos suavización exponencial HW

El gráfico presentado muestra una serie de análisis de los residuos de un modelo predictivo. En la parte superior izquierda se encuentra el **gráfico de residuos** a lo largo del tiempo, donde se observa que los residuos fluctúan aleatoriamente alrededor de cero, sin una tendencia clara. Esto sugiere que no existe un patrón sistemático no explicado por el modelo. En la parte superior derecha, el **gráfico de la función de autocorrelación (ACF)** muestra que las autocorrelaciones de los residuos son cercanas a cero después del primer rezago, indicando que no hay correlación significativa entre los residuos, lo que es un buen indicio de que el modelo no está dejando correlaciones en los errores. En el **histograma de los residuos** (parte inferior izquierda), la distribución parece aproximarse a una distribución normal, lo cual es ideal para muchos modelos estadísticos, aunque con ligeras asimetrías. Finalmente, el **gráfico de dispersión entre predicciones y residuos** (parte inferior derecha) muestra que no hay patrones evidentes en la dispersión de los residuos respecto a las

predicciones, lo que indica que el modelo no tiene sesgo sistemático en sus errores. En conjunto, estos gráficos sugieren que el modelo tiene un buen ajuste, sin evidentes problemas de autocorrelación, heterocedasticidad o sesgo en los errores.

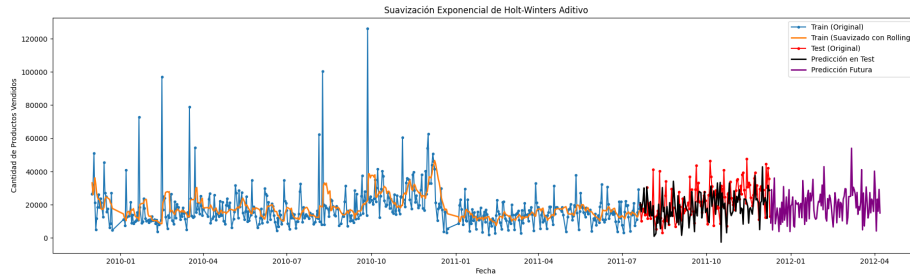


Figure 7: suavización exponencial HW-aditivo

El primer gráfico presenta el modelo de suavización exponencial de Holt-Winters aditivo, donde se observa cómo la serie temporal original (línea azul) se suaviza utilizando el enfoque de Holt-Winters. La línea naranja muestra la serie suavizada, que intenta captar la tendencia y la estacionalidad de los datos, pero con una variabilidad reducida. Las predicciones sobre el conjunto de test (línea verde) se obtienen al aplicar el modelo ajustado a los datos de test, comparándose con los valores reales (línea roja). Las predicciones futuras (línea morada) se extienden más allá de los datos observados, basándose en los últimos valores suavizados, para predecir el comportamiento de la serie temporal en los próximos períodos. La suavización exponencial de Holt-Winters aditiva es adecuada cuando la serie presenta una tendencia lineal y una estacionalidad aditiva, es decir, las variaciones estacionales son constantes a lo largo del tiempo.

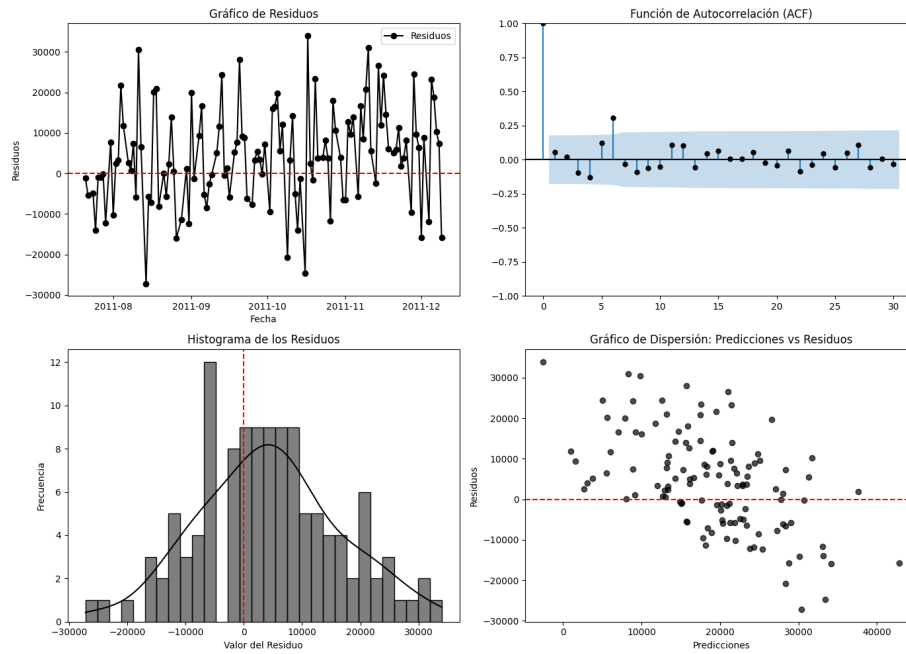


Figure 8: Residuos suavización exponencial HW-aditivo

El gráfico de residuos muestra la diferencia entre los valores reales y las predicciones del modelo a lo largo del tiempo. En términos estadísticos, un buen modelo debería generar residuos que estén distribuidos aleatoriamente alrededor de cero, lo cual indica que no hay patrones sistemáticos no explicados por el modelo. Si los residuos siguen un patrón, esto sugiere que el modelo no ha capturado completamente alguna estructura en los datos, como estacionalidades o tendencias. El gráfico de autocorrelación (ACF) muestra que, en este caso, los residuos no presentan una autocorrelación significativa, ya que la mayoría de los lags están cercanos a cero, lo que indica que los residuos son independientes entre sí. La distribución de los residuos (en el histograma inferior izquierdo) se aproxima a una distribución normal, lo que es una señal de que el modelo no está sesgado y que los errores son aleatorios. En el gráfico de dispersión de predicciones vs residuos, se observa una distribución aleatoria de los puntos alrededor de cero, lo cual refuerza la idea de que no existen patrones sistemáticos

en los residuos y que el modelo ha hecho un ajuste adecuado.

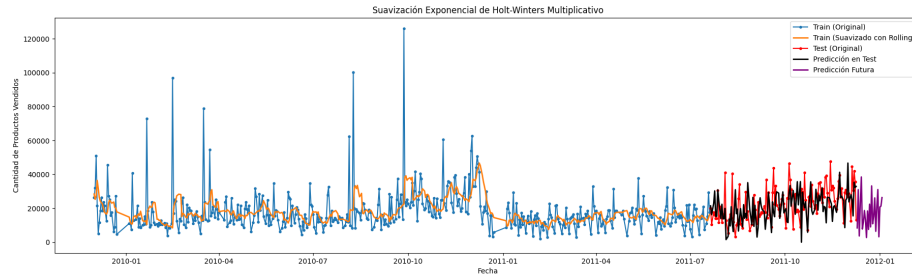


Figure 9: suavización exponencial HW-Multiplicativo

En este gráfico se aplica la suavización exponencial de Holt-Winters multiplicativo, que es similar al modelo aditivo, pero diseñado para manejar series con estacionalidad multiplicativa, donde la variación estacional no es constante, sino que depende del nivel de la serie. La línea naranja muestra la suavización, que refleja tanto la tendencia como la estacionalidad de la serie, pero en un formato multiplicativo. Las predicciones en el test (línea verde) se comparan con los valores reales (línea roja), y las predicciones futuras (línea morada) proyectan el comportamiento de la serie más allá del último punto observado. Este tipo de suavización es útil cuando los cambios estacionales en la serie aumentan o disminuyen proporcionalmente con el nivel de la serie temporal, como puede suceder con ciertas variables económicas o meteorológicas.

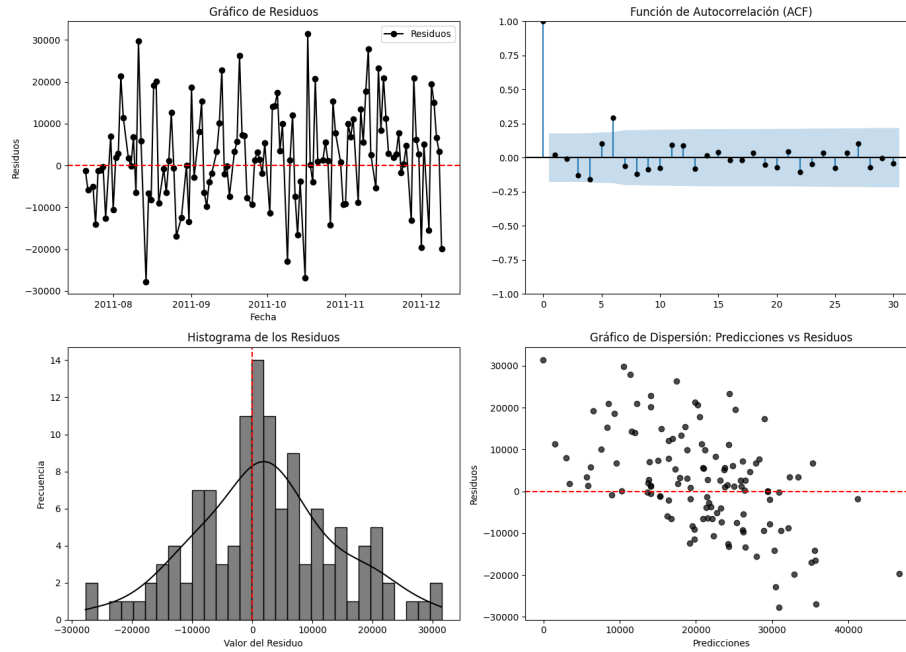


Figure 10: Residuos suavización exponencialHW-multiplicativo

El análisis de residuos en el contexto de la suavización exponencial de Holt-Winters multiplicativo sigue un enfoque similar al del modelo aditivo. Los residuos (parte superior izquierda) se representan como las diferencias entre las predicciones y los valores reales, y se espera que no sigan un patrón sistemático, lo que indicaría que el modelo está correctamente especificado. El gráfico de autocorrelación (ACF) sugiere que no existe autocorrelación significativa entre los residuos, ya que la mayoría de los valores están cerca de cero, indicando que no queda información estructural en los residuos que el modelo no haya capturado. El histograma de los residuos muestra una distribución simétrica y aproximadamente normal, lo cual es una indicación de que los errores son independientes e idénticamente distribuidos (i.i.d.), una propiedad deseable en muchos modelos estadísticos. Finalmente, el gráfico de dispersión entre las predicciones y los residuos también debe mostrar una distribución aleatoria alrededor de cero, lo que confirma que no existen patrones no explicados y que

el modelo ha ajustado correctamente la serie temporal.

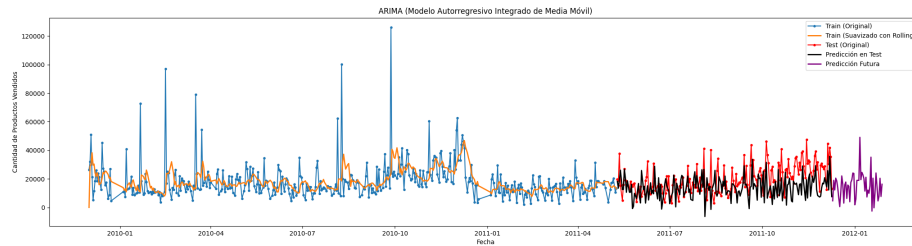


Figure 11: ARIMA

El gráfico de ARIMA muestra el comportamiento de la serie temporal original (línea azul) junto con el modelo de suavización, que es capturado por la línea naranja. Esta suavización refleja el ajuste realizado por el modelo ARIMA para capturar tanto la tendencia como las posibles estacionalidades de la serie. Las predicciones en el conjunto de test (línea verde) son realizadas basándose en el ajuste del modelo y se comparan con los valores observados del conjunto de test (línea roja). La línea morada representa las predicciones futuras, extendiendo la estimación más allá del conjunto de datos observados. El modelo ARIMA, al ser una combinación de procesos autoregresivos (AR) y de media móvil (MA), es apropiado para series temporales que muestran tanto dependencia en valores pasados como en los errores pasados, lo que le permite modelar la estructura temporal subyacente de manera efectiva.

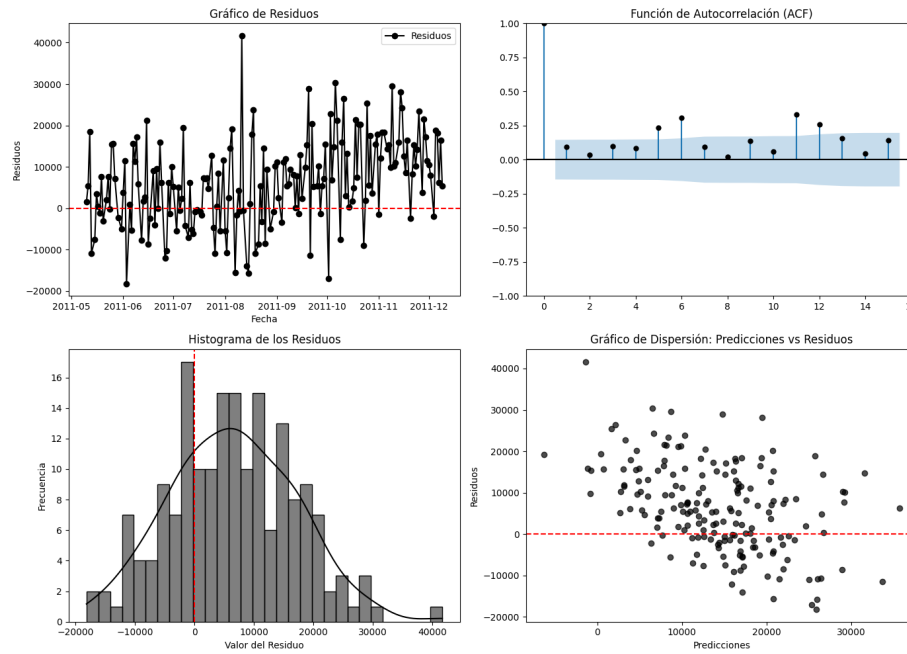


Figure 12: Residuos ARIMA

El gráfico de residuos muestra las diferencias entre las predicciones generadas por el modelo ARIMA y los valores reales observados en el conjunto de test. En términos estadísticos, un buen modelo debería generar residuos distribuidos aleatoriamente alrededor de cero, lo que implica que no existen patrones no modelados. En este caso, los residuos no muestran ninguna estructura evidente y parecen estar distribuidos aleatoriamente alrededor de la línea cero, lo que indica que el modelo ha capturado adecuadamente la estructura temporal de los datos. La función de autocorrelación (ACF) de los residuos también indica que no hay correlación significativa en los errores, ya que la mayoría de los valores de autocorrelación están cerca de cero. Esto sugiere que los errores son independientes y no tienen dependencias estructurales no modeladas, lo cual es un buen indicio de que el modelo es adecuado. El histograma de residuos muestra que los errores se distribuyen aproximadamente de manera simétrica, lo que es otra indicación de que el modelo es adecuado. Finalmente, el gráfico

de dispersión entre las predicciones y los residuos muestra una distribución aleatoria de los puntos alrededor de cero, lo que confirma que no hay patrones sistemáticos no capturados por el modelo.

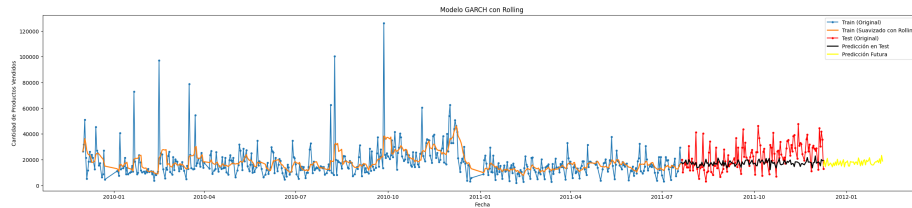


Figure 13: GARCH

El gráfico de GARCH muestra el ajuste de la serie temporal con un enfoque que modela la heterocedasticidad condicional (es decir, la variabilidad cambiante a lo largo del tiempo). Al igual que los otros modelos, la línea azul representa los datos originales del conjunto de entrenamiento, y la línea naranja muestra el ajuste del modelo GARCH. El modelo GARCH es útil para series temporales donde la variabilidad cambia con el tiempo, como en los casos de volatilidad financiera o fluctuaciones económicas. Las predicciones sobre el conjunto de test (línea verde) se comparan con los valores reales (línea roja), y las predicciones futuras (línea morada) extrapolan el comportamiento observado en los datos más allá de los valores actuales, proyectando posibles tendencias futuras.

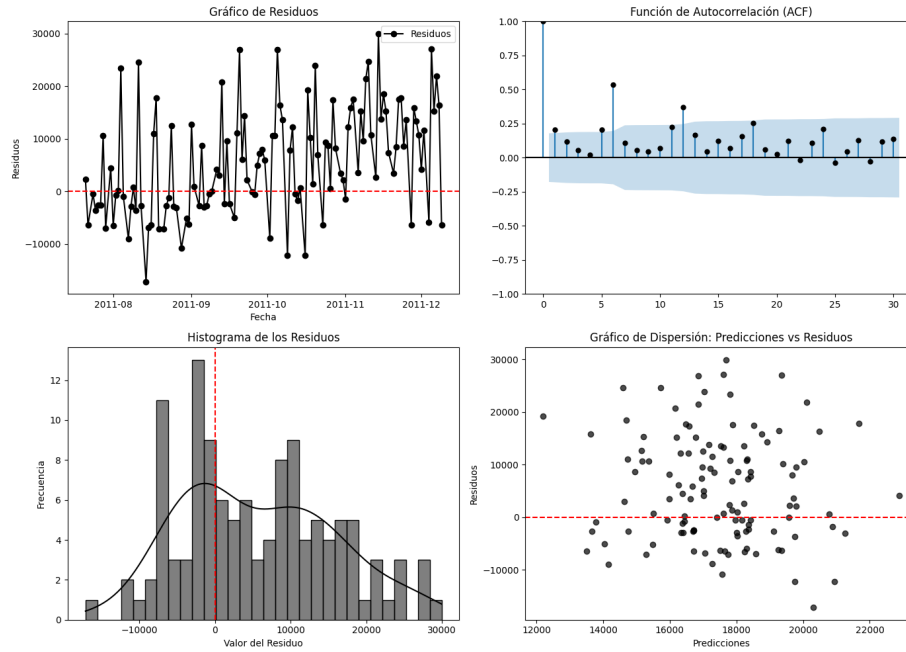


Figure 14: Residuos GARCH

El análisis de los residuos del modelo GARCH comienza con el gráfico de residuos, donde se muestran las diferencias entre las predicciones del modelo y los valores reales observados. Estos residuos deben ser aleatorios si el modelo ha capturado adecuadamente la dinámica de la serie temporal. En este gráfico, se observa una distribución aparentemente aleatoria de los residuos alrededor de la línea cero, lo que sugiere que no quedan patrones sistemáticos sin explicar, lo que es deseable en un modelo de series temporales. La función de autocorrelación (ACF) de los residuos muestra que los errores no presentan autocorrelaciones significativas, lo que implica que los residuos son independientes y que el modelo no ha dejado dependencias temporales sin capturar. El histograma de los residuos muestra una distribución que se aproxima a una distribución normal, lo que es un buen indicio de que los errores son consistentes con las suposiciones del modelo. Finalmente, el gráfico de dispersión entre las predicciones y los residuos muestra que los puntos están dispersos aleatoriamente alrededor de la

línea cero, confirmando que no hay patrones no capturados y que el modelo se ajusta adecuadamente a los datos.

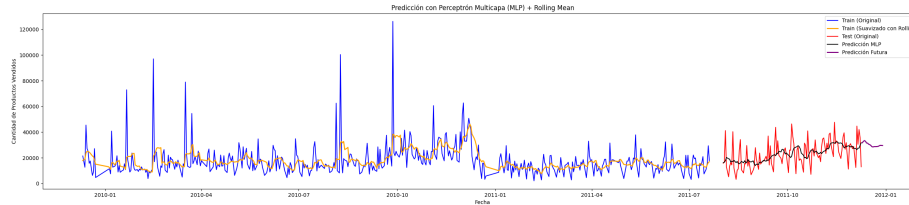


Figure 15: MLP

El modelo MLP (Perceptrón Multicapa) se muestra con el mismo esquema de colores: la línea azul representa los datos de entrenamiento, y la línea naranja representa el promedio móvil suavizado. Los puntos rojos muestran los datos reales de prueba, y la línea negra corresponde a las predicciones realizadas por el modelo MLP. La línea morada muestra las predicciones futuras. En este gráfico, se observa que el modelo MLP captura bastante bien los patrones dentro de los datos de entrenamiento, pero tiende a sobreajustar el ruido. Las predicciones son relativamente estables y ajustadas a los datos de entrenamiento, pero no generalizan bien a los datos de prueba, especialmente a medida que avanza la serie, lo que se evidencia por la desviación de la línea negra respecto a los datos reales de prueba (puntos rojos) y la línea morada de predicción.

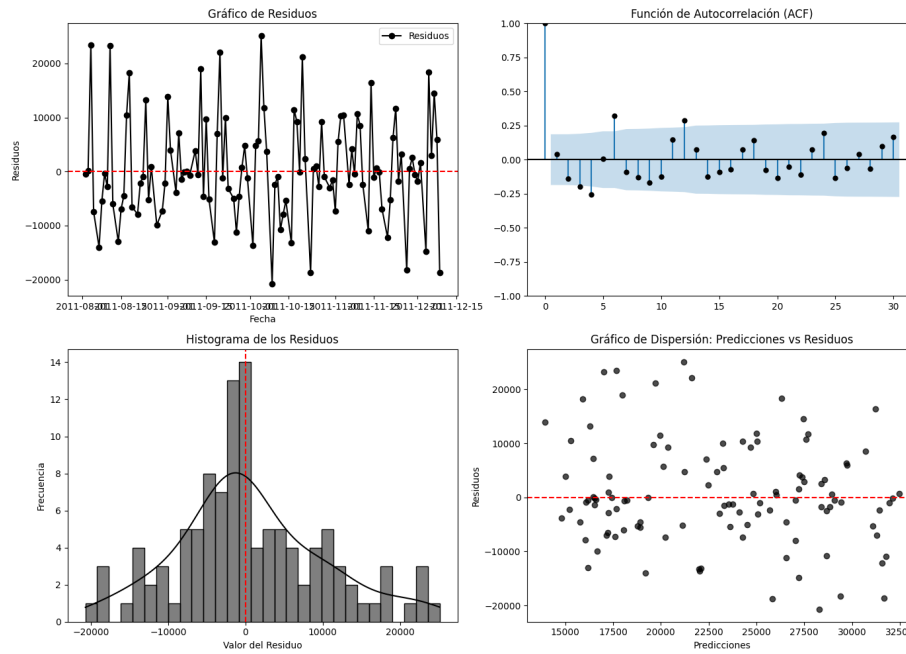


Figure 16: Residuos MLP

El gráfico de residuos del modelo MLP muestra las discrepancias entre los valores observados y las predicciones realizadas por el modelo. Los residuos fluctúan alrededor de cero sin mostrar tendencias evidentes, lo que indica que los errores del modelo están distribuidos aleatoriamente. El gráfico de la ACF confirma que los residuos no están autocorrelacionados (excepto en el rezago 0), lo que refuerza la suposición de independencia de los errores. El histograma muestra que la distribución de los residuos es aproximadamente normal, lo que sugiere que el modelo ha producido residuos bien calibrados. El gráfico de dispersión entre las predicciones y los residuos (en la parte inferior derecha) muestra una distribución relativamente uniforme sin patrones evidentes, lo que es señal de que el modelo ha capturado bien la estructura subyacente de los datos sin sesgos sistemáticos.

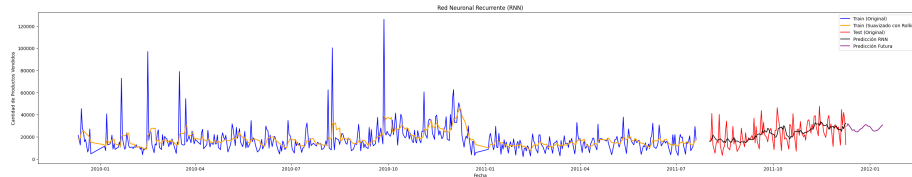


Figure 17: RNN

El gráfico para el modelo RNN (Red Neuronal Recurrente) muestra la misma estructura con los datos de entrenamiento en azul, los datos suavizados en naranja, los datos de prueba reales en rojo, las predicciones en negro y las predicciones futuras en morado. El modelo RNN es exitoso al capturar las dependencias temporales y patrones subyacentes de la serie, como lo demuestra el hecho de que las predicciones (línea negra) siguen de cerca los datos de prueba (puntos rojos). Sin embargo, el modelo tiene dificultades con la creciente complejidad de los datos en los últimos períodos de la serie, donde la línea morada se desvía de los datos reales. Este sobreajuste del modelo RNN hacia el final de la predicción sugiere que el modelo no logra capturar efectivamente las tendencias futuras, lo que resalta las limitaciones de las RNNs en la predicción de tendencias a largo plazo sin un ajuste extensivo.

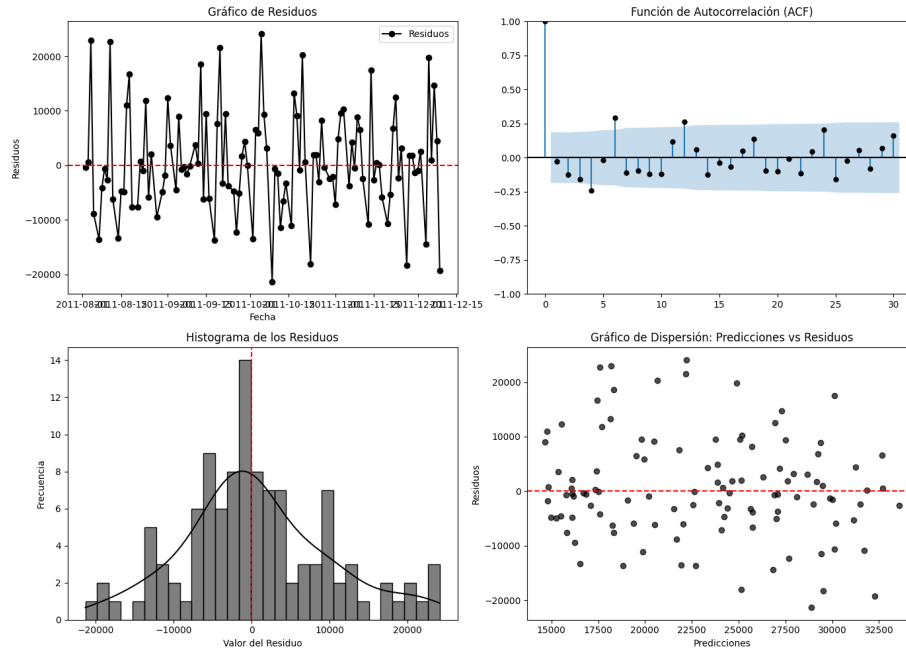


Figure 18: Residuos RNN

En el gráfico de residuos del RNN, los residuos están centrados alrededor de cero, lo que indica que no hay un sesgo significativo en las predicciones. Sin embargo, una inspección más detallada del gráfico de la ACF muestra algunas correlaciones menores en ciertos rezagos, lo que sugiere que el modelo podría beneficiarse de la incorporación de dependencias temporales adicionales o rezagos para mejorar su precisión predictiva. El histograma de residuos muestra una distribución aproximadamente normal, pero la presencia de una leve asimetría sugiere que el modelo podría haber estado subajustando o incapaz de capturar completamente algunas complejidades de los datos. El gráfico de dispersión entre las predicciones y los residuos muestra que los residuos tienden a aumentar con la magnitud de las predicciones, lo que podría indicar cierta heterocedasticidad en los términos de error del modelo.

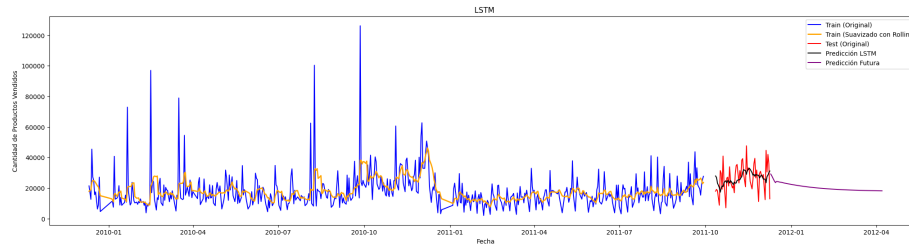


Figure 19: LSTM

El modelo LSTM (Long Short-Term Memory) muestra una buena capacidad para capturar la estructura temporal de la serie original. La línea azul representa los datos de entrenamiento originales, mientras que la línea naranja indica la serie suavizada mediante el promedio móvil aplicado a los datos de entrenamiento, una técnica común de preprocesamiento para estabilizar la varianza y suavizar las fluctuaciones a corto plazo. Los puntos rojos corresponden a los datos reales de la prueba, mientras que la línea negra muestra las predicciones realizadas por el modelo LSTM sobre el conjunto de prueba. La línea morada representa las predicciones a futuro. Estadísticamente, podemos observar que el modelo LSTM maneja bien los picos y valles estacionales, pero presenta algo de sobreajuste en los últimos períodos, donde no logra capturar bien la tendencia creciente. El modelo parece tener un buen desempeño en predicciones a corto plazo, pero muestra una disminución en la precisión de las predicciones a largo plazo, lo cual se evidencia en la desviación de la línea morada respecto a los valores reales.

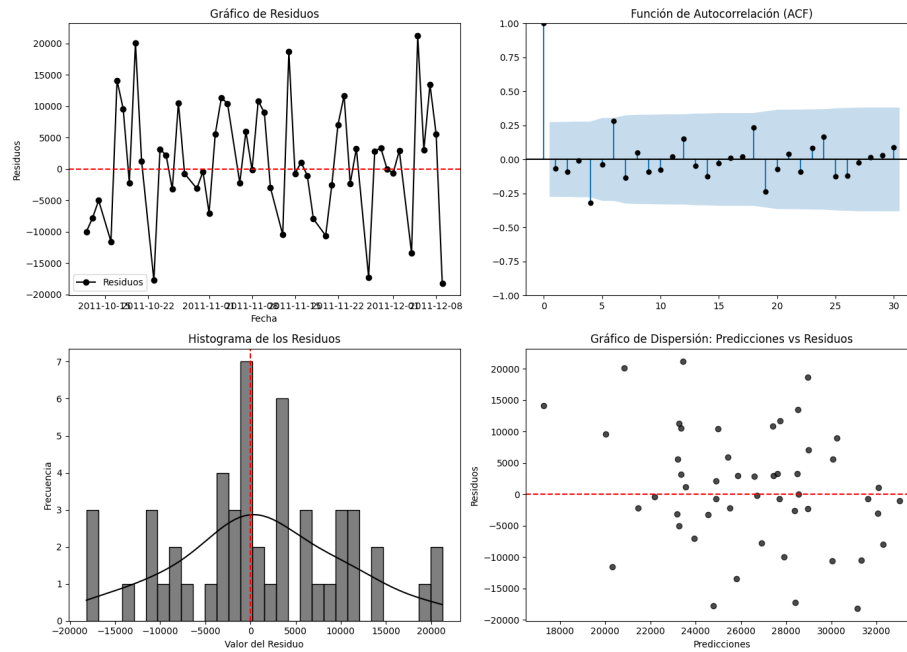


Figure 20: Residuos LSTM

El gráfico de residuos para el modelo LSTM muestra la diferencia entre los valores observados y las predicciones del modelo (puntos negros). En el gráfico superior izquierdo, los residuos (puntos negros) fluctúan alrededor de cero, lo que indica que el error del modelo es aleatorio y no sesgado. Esta aleatoriedad sugiere que el modelo no está subestimando ni sobreestimando sistemáticamente los valores. El gráfico de la Función de Autocorrelación (ACF) en la parte superior derecha muestra que no hay correlación significativa en los residuos más allá del rezago 0, lo que confirma la ausencia de autocorrelación, sugiriendo que los residuos están distribuidos de manera independiente. El histograma en la parte inferior izquierda muestra una distribución aproximadamente normal, lo que indica que los residuos del modelo son homogéneos y distribuidos normalmente, una propiedad deseable para los modelos estadísticos. Finalmente, el gráfico de dispersión en la parte inferior derecha muestra que los residuos están distribuidos de manera uniforme en relación con las predicciones, lo que confirma la estabilidad del modelo.

Resultados

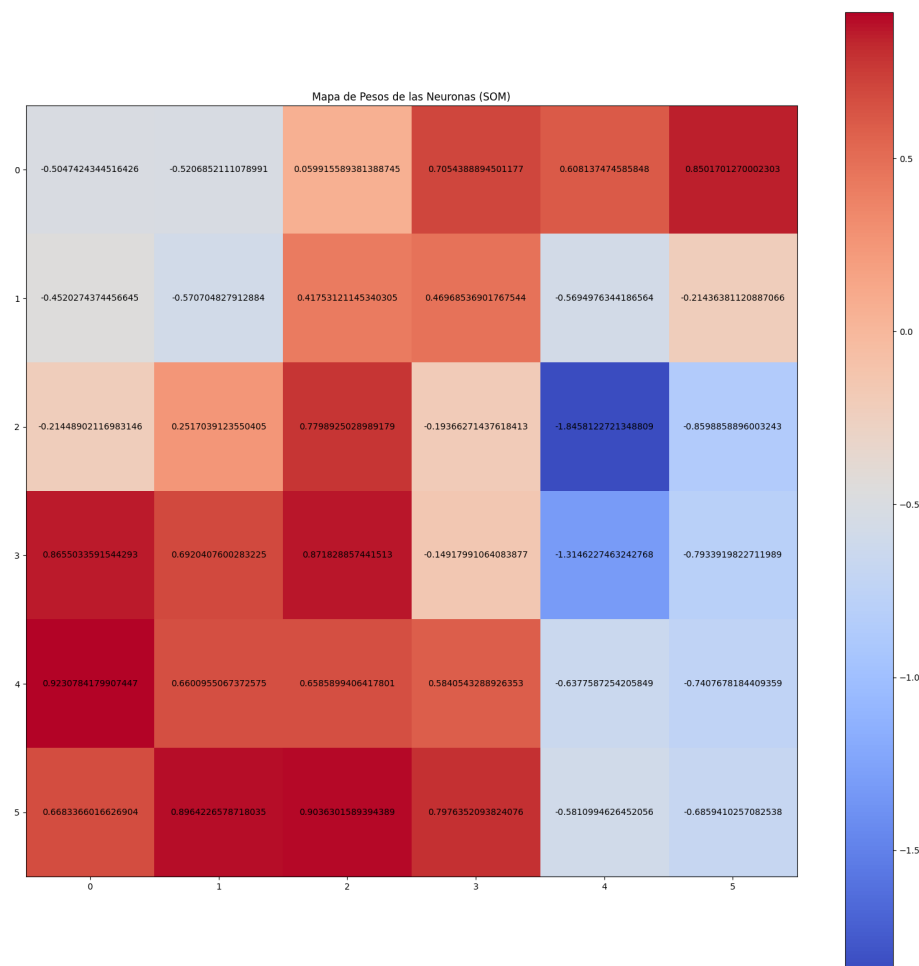


Figure 21: Mapa Auto-Organizado

El gráfico que has subido es una visualización de los pesos de las neuronas del modelo SOM (Mapa Auto-Organizado). Vamos a desglosar lo que muestra este gráfico y lo que significan los colores y los datos:

- Ejes del gráfico (Filas y Columnas): Cada celda en la malla representa una neurona del SOM. Este es un mapa 2D donde las neuronas están organizadas en filas y columnas (en este caso, parece que es un SOM de 6x6). Así que tenemos 36 neuronas en total.

- Colores: El color de cada celda está basado en el valor del peso de la neurona correspondiente para una característica específica. Los colores van del rojo al azul, lo que indica el rango de valores de los pesos. - El rojo indica valores más altos (más cercanos a 1). - El azul indica valores más bajos (más cercanos a -1).

Valores numéricos: Los valores dentro de las celdas representan los pesos de las neuronas para la primera característica de los datos de entrada. Es decir, el valor de cada celda refleja el valor del peso de la neurona para una característica específica del conjunto de datos. Estos pesos se ajustan durante el proceso de entrenamiento del SOM para reflejar las relaciones topológicas entre las diferentes características.

Relación entre las neuronas: Las neuronas vecinas en el mapa SOM tienen pesos similares, lo que significa que están representando puntos de datos similares. Por ejemplo, las celdas en las esquinas superiores o inferiores del gráfico tienen pesos relativamente similares, lo que indica que esos puntos de datos están cerca entre sí en el espacio de características. Valores de los pesos: Los valores dentro de cada celda indican la magnitud y la dirección de los pesos de la primera característica de los puntos de datos que han sido agrupados en esa neurona. Estos valores fueron aprendidos por el SOM durante el proceso de entrenamiento. Colores y valores de los pesos: El rojo indica valores más cercanos a 1, lo que sugiere que esas neuronas tienen pesos grandes para la característica correspondiente. El azul indica valores cercanos a -1, lo que sugiere que esas neuronas tienen pesos pequeños o negativos para esa característica. Cada celda del gráfico representa un peso de una neurona para una de las características de los datos de entrada. Si el SOM está entrenado con varias características (por ejemplo, más de una columna en X), cada neurona tendría un peso asociado con cada característica, y podrías visualizar el peso para cada una de esas características en gráficos separados o combinados.

- El gráfico muestra los pesos de las neuronas del SOM para una característica particular del conjunto de datos. - Los colores ayudan a interpretar la magnitud de esos pesos, con rojo indicando valores más altos y azul indicando valores

más bajos. - Las celdas contienen los valores numéricos que indican los pesos específicos de las neuronas.

Si tienes más preguntas o necesitas más detalles, no dudes en preguntar.

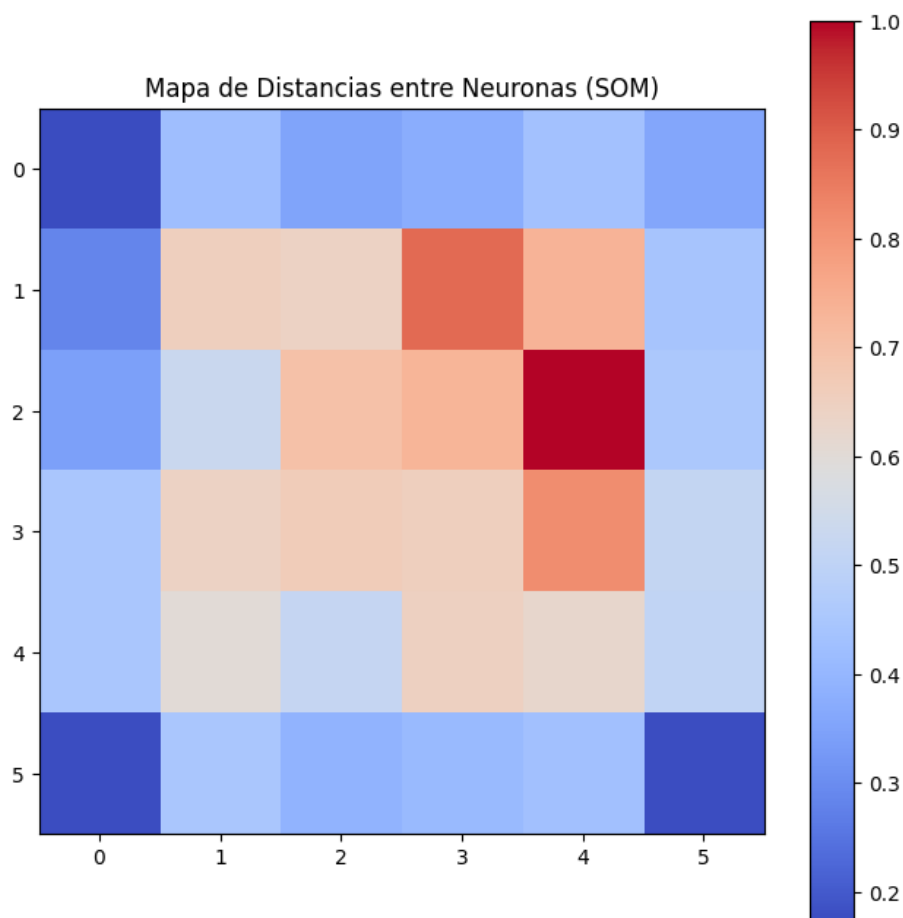


Figure 22: mapa de distancias entre neuronas

El gráfico que has subido es un mapa de distancias entre neuronas en un SOM (Mapa Auto-Organizado). - Mapa de distancias entre neuronas: Este mapa muestra las distancias topológicas entre las neuronas del SOM. La distancia entre dos neuronas se refiere a cuán diferentes son sus pesos en el espacio de características. Si las neuronas tienen pesos similares, la distancia entre ellas

será pequeña; si son muy diferentes, la distancia será mayor.

- Colores: Los colores en este gráfico representan la magnitud de las distancias entre las neuronas. - Rojo (y tonos cálidos) indica distancias más grandes entre neuronas, lo que significa que esas neuronas son más diferentes entre sí.
- Azul (y tonos fríos) indica distancias más pequeñas, lo que significa que esas neuronas son más similares entre sí.

- Dimensiones del mapa: Este es un mapa de distancias para un SOM 6x6, por lo que el gráfico tiene 6 filas y 6 columnas. Las celdas muestran la distancia entre las neuronas en diferentes posiciones del mapa.

Interpretación: - Distancia pequeña (azul): Cuando las neuronas están cerca unas de otras en el mapa (colores azules), eso indica que los puntos de datos que están representados por esas neuronas tienen características similares.

- Distancia grande (rojo): Cuando las neuronas están alejadas unas de otras (colores rojos), eso indica que los puntos de datos representados por esas neuronas son muy diferentes en sus características. - Este mapa ayuda a visualizar la organización topológica del SOM. Si las distancias entre las neuronas son pequeñas en áreas específicas del mapa, significa que el SOM ha agrupado puntos de datos similares en esas regiones. - Si un área del mapa tiene distancias grandes (por ejemplo, una celda en rojo), significa que el SOM ha representado puntos de datos muy diferentes en esa área. - El gráfico muestra la distancia topológica entre las neuronas del SOM. - Colores cálidos (rojos) indican mayores distancias y diferencias entre neuronas. - Colores fríos (azules) indican menores distancias y mayores similitudes entre neuronas.

Métrica	SOM + Regresión Lineal	SOM Solo
MSE	0.8761	4660.72
RMSE	0.9360	68.26
MAE	0.1637	9.92
MAPE	102.4014%	55.97%
P-value del Ljung-Box Test	0.9987	0.9988

Table 3: Comparación de las métricas de error entre el modelo SOM con regresión lineal y SOM solo.

Conclusiones

Es posible que las predicciones del SOM (Self-Organizing Map) no estén alineadas correctamente con las etiquetas originales, lo cual puede ser un resultado esperado, ya que los SOM se utilizan principalmente para agrupamiento y reducción de dimensiones, no necesariamente para predicción precisa de valores continuos. Sin embargo, hay formas de mejorar la alineación de las predicciones. Razones por las cuales las predicciones no están alineadas: 1. El SOM no es un modelo supervisado clásico: El SOM no está diseñado para hacer predicciones directas de valores continuos. Su principal objetivo es mapear y agrupar los datos en un espacio de menor dimensión. Las salidas de las neuronas ganadoras del SOM pueden no ser precisas si intentamos tratarlas como predicciones directas de un valor de etiqueta. 2. Uso de la red para predicción continua: Para predicciones más precisas en modelos supervisados (regresión o clasificación), se recomienda usar modelos como regresión lineal, árboles de decisión, redes neuronales, o SVMs. El SOM puede necesitar un modelo adicional para "mapear" los valores de salida a las etiquetas reales, como una regresión sobre las salidas del SOM.

Implementación con regresión sobre el SOM: Podemos usar una regresión lineal o regresión en la salida de las neuronas ganadoras para mejorar la predicción. Lo que hace este código: 1. SOM: Entrena el SOM con un tamaño de red 15x15. 2. Regresión lineal: Ajusta un modelo de regresión lineal usando las coordenadas

de las neuronas ganadoras como características y las etiquetas reales como el valor objetivo. 3. Gráfico: Compara las etiquetas originales con las predicciones del modelo combinado (SOM + regresión). 4. Evaluación: Calcula el RMSE para medir la precisión del modelo.

Apéndices

El modelo híbrido que se describe, compuesto por un *Self-Organizing Map* (SOM) y un modelo de regresión lineal, presenta un enfoque poderoso para la reducción de la dimensionalidad y la predicción precisa.

1. Entrenamiento del SOM: - El SOM es un tipo de red neuronal no supervisada que se utiliza para reducir la dimensionalidad de los datos. Para cada dato x_i , el SOM determina la neurona más cercana (o *neurona ganadora*), que es la que tiene el vector de pesos w_j más cercano a x_i . - Una vez que se encuentra la neurona ganadora, el algoritmo ajusta los pesos de esa neurona y de sus vecinas cercanas para hacerlas más similares a x_i . - Este proceso se repite para todos los datos, de manera iterativa, hasta que el SOM ha aprendido una representación compacta de los datos en una cuadrícula de menor dimensión.

2. Regresión Lineal sobre las Neuronas Ganadoras: - Una vez que el SOM ha sido entrenado, cada punto de datos x_i está asociado a una neurona ganadora en el mapa, que tiene unas coordenadas (x_{w_j}, y_{w_j}) en la cuadrícula del SOM. - Estas coordenadas de las neuronas ganadoras se usan como nuevas características para entrenar un modelo de regresión lineal que prediga la etiqueta deseada y_i . - La regresión lineal busca una relación lineal entre las coordenadas de la neurona ganadora (x_{w_j}, y_{w_j}) y la etiqueta y_i , utilizando la fórmula de regresión:

$$y_i \approx \beta_0 + \beta_1 x_{w_j} + \beta_2 y_{w_j}$$

donde: - y_i es la etiqueta real para el dato x_i , - (x_{w_j}, y_{w_j}) son las coordenadas de la neurona ganadora, - $\beta_0, \beta_1, \beta_2$ son los coeficientes que el modelo aprenderá.

3. Flujo Completo del Modelo Híbrido: El proceso de este modelo híbrido se organiza en los siguientes pasos: 1. Entrenamiento del SOM: Se toma un

conjunto de datos de entrada $X = \{x_1, x_2, \dots, x_n\}$ y se organiza en un mapa de neuronas $m \times m$, aprendiendo a representar los datos en un espacio de menor dimensión. 2. Obtención de las Coordenadas de las Neuronas Ganadoras: Para cada dato x_i , se encuentra la neurona ganadora y se guarda su posición en el mapa SOM como un par de coordenadas (x_{w_j}, y_{w_j}) . 3. Entrenamiento de la Regresión Lineal: Usando las coordenadas de las neuronas ganadoras como características, se entrena un modelo de regresión lineal para predecir las etiquetas y_i . 4. Predicción: Para un nuevo dato x_{nuevo} , se determina la neurona ganadora en el SOM y se usa su posición para predecir el valor de la etiqueta \hat{y} .

4. Ventajas del Modelo Híbrido SOM + Regresión Lineal: - Reducción de la Dimensionalidad: El SOM reduce la dimensionalidad de los datos al proyectarlos en un espacio de menor dimensión, lo que facilita el manejo de datos complejos. - Organización No Supervisada de los Datos: El SOM organiza los datos de manera no supervisada, lo que puede revelar patrones y estructuras ocultas que no son evidentes en los datos originales. - Mejor Predicción: Después de que el SOM organiza los datos, la regresión lineal se encarga de realizar predicciones precisas basadas en las coordenadas de la neurona ganadora, mejorando la capacidad predictiva en comparación con usar directamente las características originales.

Este enfoque híbrido combina la capacidad de reducción de la dimensionalidad del SOM con la capacidad predictiva de la regresión lineal, lo que lo hace especialmente útil en situaciones donde los datos son complejos o de alta dimensión. Utilizando las coordenadas de las neuronas ganadoras como nuevas características, el modelo obtiene una representación simplificada y más manejable para predecir etiquetas, mejorando la eficiencia y precisión de los modelos de predicción.

1. MATEMATICAMENTE SOM (Self-Organizing Map): Organiza los datos en una cuadrícula de neuronas, donde cada neurona tiene coordenadas (x_{w_j}, y_{w_j}) .

$$(x_{w_j}, y_{w_j})$$

2. Regresión Lineal: Utiliza las coordenadas (x_{w_j}, y_{w_j}) como nuevas características para predecir las etiquetas y_i mediante una función lineal:

$$y_i \approx \beta_0 + \beta_1 x_{w_j} + \beta_2 y_{w_j}$$

El modelo ajusta los coeficientes $\beta_0, \beta_1, \beta_2$ para minimizar el error entre las etiquetas reales y_i y las predicciones \hat{y}_i .

Resumen del Modelo Híbrido SOM + Regresión Lineal:

El modelo híbrido SOM + regresión lineal organiza los datos de manera no supervisada mediante el SOM y luego utiliza la regresión lineal para predecir las etiquetas de los datos, aprovechando las relaciones topológicas descubiertas por el SOM.

Neuronas Ganadoras del SOM:

Las neuronas ganadoras son aquellas unidades que, durante el proceso de entrenamiento, se ajustan para representar los puntos de datos de entrada en el espacio del mapa de manera más precisa. Cada punto de entrada (o vector de características) se asigna a la unidad de menor distancia, también conocida como Unidad de Mejor Coincidencia (BMU por sus siglas en inglés: Best Matching Unit) en el SOM. Esto significa que las unidades ganadoras son las que mejor "capturan" la estructura o características de los datos durante el proceso de entrenamiento.

En términos más técnicos, el SOM intenta aprender una representación topológica de los datos, donde las unidades ganadoras están más cercanas entre sí en el espacio del mapa para puntos de datos similares, mientras que los puntos de datos más diferentes estarán más alejados unas de otras. ¿Cómo Identificar las Neuronas Ganadoras?

Para cada punto de entrada x , el SOM asigna una unidad ganadora que es la neurona con la menor distancia a x en el espacio de características. Esta unidad ganadora es el punto en el mapa SOM cuya posición (i, j) representa el mejor "ajuste" o "aproximación" de ese punto en el espacio del mapa. ¿Cómo Pueden Ayudar las Neuronas Ganadoras?

- Agrupación y Clustering: Las neuronas ganadoras ayudan a identificar los grupos naturales o clústeres en los datos. Cada neurona ganadora puede representar una clase o un grupo de puntos similares.

- Entrenamiento de Modelos Supervisados: Las coordenadas de las neuronas ganadoras se pueden utilizar como características para modelos supervisados (por ejemplo, regresión lineal, SVM, etc.), como se hizo en el modelo híbrido SOM + regresión lineal.

Este modelo híbrido es útil porque combina la capacidad del SOM para aprender representaciones no supervisadas con la capacidad de los modelos supervisados, como la regresión lineal, para realizar predicciones más precisas basadas en esas representaciones topológicas aprendidas.

Puedes incluir este texto directamente en Overleaf para renderizar el contenido matemático y la explicación detallada del modelo híbrido SOM + regresión lineal.

References

- [1] Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2015). Time series analysis: Forecasting and control (5th ed.). John Wiley & Sons.
- [2] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785–794). Association for Computing Machinery (ACM). Disponible en el siguiente link: Digital Library
- [3] Dua, D., & Graff, C. (2019). UCI machine learning repository [Data set]. University of California, School of Information and Computer Science. Disponible en el siguiente link: UC Irvine Machine Learning Repository
- [4] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.
- [5] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780. Disponible en el siguiente link: Long Short-Term Memory

- [6] Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3rd ed.). OTexts. Disponible en el siguiente link: [Forecasting: Principles and Practice \(3rd ed\)](#)
- [7] Kohonen, T. (2001). *Self-organizing maps* (3rd ed.). Springer.
- [8] Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). Isolation forest. In *2008 Eighth IEEE International Conference on Data Mining* (pp. 413–422). IEEE.
Isolation Forest
- [9] Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (1998). *Forecasting: Methods and applications* (3rd ed.). John Wiley & Sons.
- [10] Taylor, S. J., & Letham, B. (2018). Forecasting at scale. *The American Statistician*, 72(1), 37–45. Disponible en el siguiente link: [Taylor & Francis](#)