

# 1er Trabajo De Investigación / Informe Ejecutivo: “Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models”

Víctor Álvarez  
Universidad Técnica  
Federico Santa María  
Vitacura, Chile  
victor.alvarezb@usm.cl

Constantino Mavrakís  
Universidad Técnica  
Federico Santa María  
Vitacura, Chile  
constantino.mavrakís@usm.cl

Matías Otth  
Universidad Técnica  
Federico Santa María  
Valparaíso, Chile  
matias.otth@sansano.usm.cl

Gabriel Vergara  
Universidad Técnica  
Federico Santa María  
San Joaquín, Chile  
gabriel.vergara@usm.cl

## I. INTRODUCCIÓN

El presente informe busca explicar y replicar los resultados obtenidos en el artículo “*Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models*” utilizando nuevos datos. Este artículo publicado el 6 de marzo del 2018 fue aceptado por el Journal “*Expert Systems With Applications*” perteneciente al cuartil Q1.

En el artículo referenciado, los autores se proponen crear distintos modelos híbridos entre redes neuronales recurrentes con varios modelos autoregresivos condicionales generalizados heterocedásticos (GARCH) para pronosticar la volatilidad del índice bursátil de Corea del Sur. Adicionalmente, se sensibiliza la predicción al variar los días de pronóstico y el horizonte de tiempo de la información.

La predicción de la volatilidad en el mercado financiero permite tomar decisiones anticipadas con respecto a compra/venta de activos financieros, para disminuir el riesgo en portafolios u otros. A partir de lo anterior, varias líneas de investigación se han desarrollado y se ha visto una mayor utilización de redes neuronales, dado su potencial en el desarrollo de modelos y pronósticos.

Para esta investigación, se van a utilizar los mismos modelos que en el artículo referenciado, estos son: “Generalized Autoregressive Conditional Heteroskedasticity” (GARCH) (G), “Exponential GARCH” (EGARCH) (E) y “Exponential Weighted Moving Average” (EWMA) (W). Cada uno de estos presentan distintas características que logran capturar. Por ejemplo, GARCH captura la magnitud y persistencia de la volatilidad; EGARCH, adicionalmente, logra medir la dirección de su impacto y capturar el “Leverage Effect” (LE); y EWMA propone ver los cambios a corto plazo.

Además de estos modelos, se va a utilizar redes neuronales. Estas permiten la identificación de patrones no lineales en un conjunto de datos. Para este artículo, “Deep Forward Network” (DFN) y “Long-Short Term Memory” (LSTM) son las redes de interés para obtener la volatilidad.

Por otra parte, se va a realizar una combinación de modelos de redes neuronales con las técnicas financieras. De esta manera, surge: E-DFN, E-LSTM, G-DFN, G-LSTM, W-DFN, G-LSTM, EW-LSTM, GE-LSTM, GW-LSTM y GEW-LSTM. Cada uno de estos modelos, van a ser desarrollados con dos funciones de pérdidas: MSE y MAE para ver su impacto. Posteriormente, se calcularán sus errores utilizando más métricas MSE, MAE, HMAE y HMSE; y los datos serán validados estadísticamente.

Con respecto a los datos que se desea pronosticar, se va a utilizar un nuevo conjunto de datos que corresponde a los precios ‘intra day’ de la crypto moneda Binance (BNB). Este activo es de gran interés ya que, permitirá poner a prueba los modelos descritos en un índice de alta frecuencia y volatilidad. Por lo tanto, se espera que algunos resultados difieran del artículo de referencia.

A partir de todo lo anterior, las contribuciones que se desean realizar son las siguientes:

- Se desea verificar si EGARCH sigue siendo el modelo con mejor capacidad de predicción, entre los tipos Garch, al utilizar un conjunto de datos más volátil. Además, que los modelos híbridos que contengan EGARCH, sobresalgan frente a sus pares.
- Se desea confirmar que LSTM logra un mayor rendimiento que DFN y los modelos Garch. Según la teoría y el artículo de referencia, LSTM debiese ser capaz de capturar patrones temporales no lineales y utilizar/descartar datos pasados según su conveniencia.

- Se desea probar que, combinar múltiples modelos econométricos con redes neuronales, permite la obtención de una mejor la calidad de predicción de la volatilidad. Se espera que mediante un mayor aumento de la cantidad de técnicas, mejor debiese ser el pronostico.
- Se desea identificar posibles diferencias en las magnitudes de los errores, al cambiar la función de perdidas en lo modelos de redes neuronales. Es decir, comparar las funciones MSE y MAE.

Finalmente, este trabajo se desarrolla usando el software R para los modelos econométricos GARCH y; el uso del lenguaje de programación Python, en un entorno Jupyter Notebook, para el desarrollo de los modelos y representación de los datos.

## II. METODOLOGÍA

### II-A. Data

La data utilizada corresponde a precios de cierre con una frecuencia de 15 minutos para el cryptoactivo de BNB en una ventana de 3 meses para un total de 10.000 valores, desde 2022-01-15 hasta 2022-04-30 (**Figura 1**) donde se utilizaron los últimos 1880 valores para la validación.



Figura 1. Precios de cierre a frecuencia de 15 minutos para BNB/USDT, valores en color azul corresponden a datos utilizados para el set de validación de modelos.

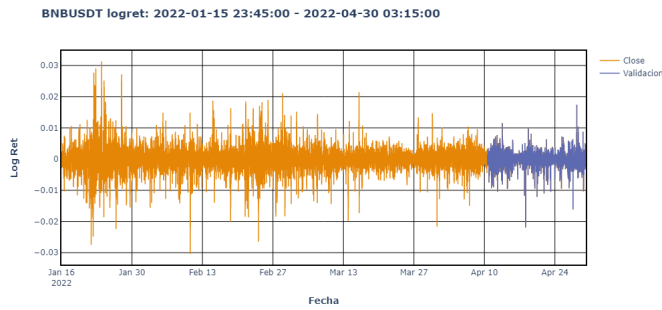


Figura 2. Log retorno de precios de cierre a frecuencia de 15 minutos para BNB/USDT, valores en color azul corresponden a datos utilizados para el set de validación de modelos.

	count	mean	std	min	25%	50%	75%	max	skew	kurt
LogRet	9999.000000	-0.000022	0.003586	-0.030370	-0.001689	0.000000	0.001777	0.031286	-0.089909	7.501829

Figura 3. Resumen de estadísticas para muestra de logretornos de BNBUSDIT utilizados.

luego, la volatilidad realizada se obtuvo realizando la desviación estándar móvil para un intervalo de 10 periodos, esto corresponde a la desviación estándar de las últimas 2.5 horas realizadas.

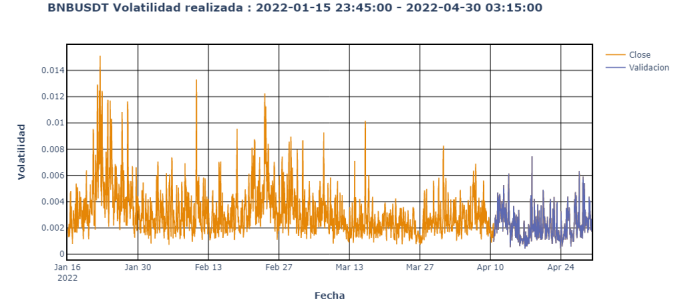


Figura 4. Volatilidad realizada, valores en color azul corresponden a datos utilizados para el set de validación de modelos.

### II-B. Modelos predictivos

#### 1. Modelo GARCH:

En este modelo la varianza condicional se estima en base al error rezagado del estimador inicial y a varianzas estimadas rezagadas (distinguiéndolo con este último punto del modelo ARCH). Donde todos los  $\alpha_i, \beta_j$  y la media  $\alpha_0$  deben ser no negativos.

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_q \varepsilon_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_p \sigma_{t-p}^2$$

Finalmente, el modelo utilizado GARCH(1,1) corresponde a:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

#### 2. Modelo EGARCH:

Dado que el modelo GARCH no es útil para interpretar o estimar el efecto de apalancamiento en los activos financiero es que se analiza el EGARCH.

La varianza condicional se determina a través de su valor rezagado y se asume que es positiva independiente de si el valor estimado no lo es. Para determina el apalancamiento se utiliza  $\omega$  y se asume que si se observa que  $\omega + \gamma > 1$  el modelo interpreta la existencia de asimetría en la serie utilizada.

$$\ln(\sigma_t^2) = \alpha'_0 + \beta \ln(\sigma_{t-1}^2) + \omega \left( \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right) + \gamma \left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right|$$

#### 3. Modelo EWMA:

Se plantea como un tipo de suavizamiento exponencial que utiliza la información pasada para ir generando una predicción de la volatilidad condicional en en base al dato pronosticado y al error incurrido.

$$\sigma_t^2 = \rho \sigma_{t-1}^2 + (1 - \rho) \varepsilon_{t-1}^2 \quad (0 < \rho < 1)$$

$\rho$  se define como el grado de decrecimiento del peso relativo que posee la información más reciente en el pronóstico de la volatilidad condicional.

## II-C. Modelos Neuronales

### 1. Deep feed forward:

Corresponde a un modelo de 3 capas, con 10, 5 y 1 neuronas por capa. Las primeras 2 capas son implementadas con funciones sigmoidales mientras que la capa de salida con una función lineal:

$$w_t = \sigma(W_1 x_t + b_1) \quad (1)$$

$$s_t = \sigma(W_2 w_t + b_2) \quad (2)$$

$$y_t = W_3 s_t + b_3 \quad (3)$$

por lo tanto, la predicción de  $RV_t$  se obtiene como:

$$RV_{t+1} = W_3 \sigma(W_2 \sigma(W_1 x_t + b_1) + b_2) + b_3$$

siendo  $x_t$  las variables de input, que corresponden a los últimos 20 valores de retornos, incluyendo los parámetros de los modelos garch en los modelos híbridos.

### 2. Long short-term memory:

$$r_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f) \quad (4)$$

$$i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_f) \quad (5)$$

$$\tilde{C}_t = \tanh(W_c x_t + U_c h_{t-1} + b_f) \quad (6)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (7)$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \quad (8)$$

$$h_t = o_t * \tanh(C_t) \quad (9)$$

finalmente, el output de la celda LSTM se pasa por una capa densa con función lineal:

$$\hat{RV}_{t+1} = W_y C_t + b_y \quad (10)$$

entonces, a esta arquitectura se alimenta una secuencia de 20 valores, es decir, utilizando los 20 retornos pasados (incluyendo los demás parámetros para los modelos híbridos), se alimenta el retorno 1, se obtiene el estado oculto 1, luego viene el retorno 2 junto con el estado oculto 1 y se obtiene el estado oculto 2, hasta ingresar el retorno 20 junto con el estado oculto 19 y se predice la volatilidad realizada en el tiempo 21.

Además, para la estimación de parámetros de la red, se realiza un MinMaxScaling para los valores del conjunto de entrenamiento, es decir, para el set de entrenamiento se obtiene el mínimo  $m$  y máximo  $M$ , por lo tanto para cada valor  $x_t$  del set de entrenamiento se obtiene su valor escalado al intervalo  $[0,1]$  como  $s_t = \frac{x_t - m}{M - m}$  realizando este procedimiento para cada variable de input y para el valor a predecir de volatilidad. Finalmente, para el conjunto de validación se escala utilizando los valores de escalado del set de entrenamiento y posterior a la realización de la predicción se transforman nuevamente a la escala original de la data.

## II-D. Medidas estadísticas utilizadas

### 1. Volatilidad realizada:

Considerando  $(\hat{v}_t)$  como la volatilidad estimada por la Red Neuronal, se define  $RV_t$  como la volatilidad realizada calculada a través de la siguiente expresión.

$$RV_t = \sqrt{\frac{1}{\rho} \sum_{i=0}^{\rho-1} (r_{t-i} - \bar{r}_t)^2}, \quad \bar{r}_t = \frac{1}{\rho} \sum_{i=0}^{\rho-1} r_{t-i}$$

Con  $\rho$  como el tamaño de la banda, el cual se utiliza con un valor de 20.

### 2. Funciones de pérdida:

$$MAE = \frac{1}{T} \sum |\hat{v}_t - RV_t|$$

$$MSE = \frac{1}{T} \sum (\hat{v}_t - RV_t)^2$$

$$HMAE = \frac{1}{T} \sum |1 - \frac{\hat{v}_t}{RV_t}|$$

$$HMSE = \frac{1}{T} \sum (1 - \frac{\hat{v}_t}{RV_t})^2$$

### 3. Diebold–Mariano (DM)

Este test tiene como objeto comparar de pares los modelos de pronóstico. Se busca demostrar que un mejor desempeño está acompañado de una significancia estadística en la obtención de menor error dado una de las funciones de pérdida utilizadas  $g(\cdot)$ . Demostrando así, que la mejor performance corresponde a la naturaleza de la capacidad predictiva y no a una coincidencia dada la data utilizada.

Diferencial de pérdida  $d_t = g(e_{1,t}) - g(e_{2,t})$ , pérdida media:  $d = \sum d_t / T$

$$DM = \frac{d}{\sqrt{2\pi \hat{f}_d(0)/T}}$$

### 4. Wilcoxon signed rank (WS)

Este test sigue la lógica de las diferencias definidas anteriormente, plantea una hipótesis nula de mediana igual a 0, que equivaldría a que métodos comparados no demuestran diferencias significativas de uno con respecto a otro. En caso de determinarse cierta asimetría de las diferencias esto asumiría la superioridad de la performance de un modelo.

$$WS = \sum_T I_+(d_t) \text{rank}(|d_t|), \quad I_+(d_t) = 1, \text{ si } d_t > 0$$

### III. RESULTADOS

Al ejecutar cada modelo, se obtiene las predicciones de la volatilidad de cada uno. Debido que son hartos, se va a representar gráficamente unos pocos. En la **Figura 5**, se observa el comportamiento de la volatilidad para las técnicas financieras y LSTM. Además, la **Figura 6** muestra el comportamiento de sus errores, la **Figura 7** muestra su distribución y la **Figura 8** estadísticas relevantes de estos.

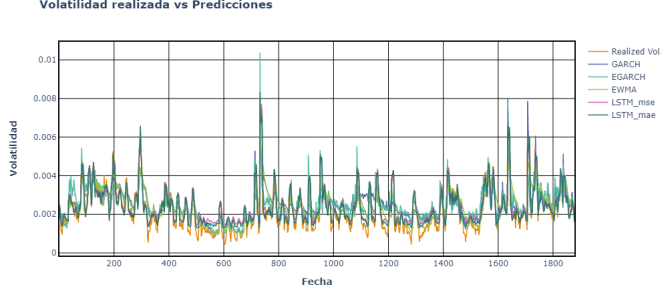


Figura 5. Resultados de predicción de volatilidad realizada para diferentes modelos. Se excluyen los modelos de LSTM utilizando los parámetros de los otros modelos por simplicidad.

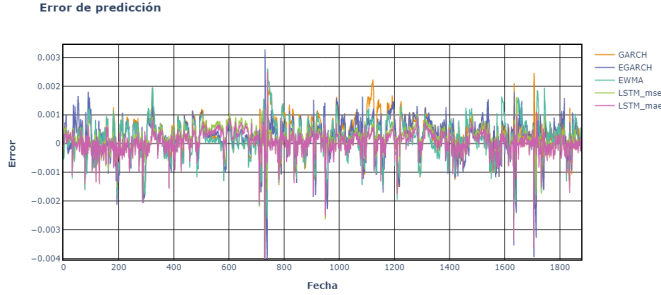


Figura 6. Error de predicción para los modelos GARCH, EGARCH, EWMA, LSTM-MSE y LSTM-MAE respecto a la volatilidad realizada.

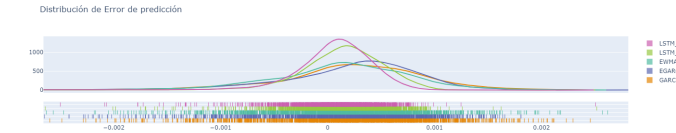


Figura 7. Distribución de resultados en el error de predicción para modelos GARCH, EGARCH, EWMA, LSTM-MSE y LSTM-MAE.

Los resultados de los errores se van a presentar en dos partes las cuales, se diferencian por la función de pérdida utilizada en la estimación de cada modelo. En la primera parte se detallan los resultados al utilizar MSE como función, mientras que en la segunda se utiliza MAE. Para ambas partes, se analizarán los resultados por categoría de modelo (individuales, híbridos simples, dobles y triples).

	GARCH	EGARCH	EWMA	LSTM_mse	LSTM_mae
count	1878.000000	1878.000000	1878.000000	1878.000000	1878.000000
mean	0.000223	0.000192	0.000130	0.000138	0.000065
std	0.000703	0.000704	0.000669	0.000428	0.000395
min	-0.004415	-0.004509	-0.004376	-0.005155	-0.005291
25%	-0.000100	-0.000114	-0.000210	-0.000065	-0.000103
50%	0.000262	0.000317	0.000162	0.000169	0.000100
75%	0.000662	0.000631	0.000547	0.000394	0.000295
max	0.002457	0.003280	0.002619	0.002518	0.002504
skew	-0.788806	-1.518046	-0.580948	-2.287707	-2.860364
kurt	2.636942	5.297601	2.626888	20.025951	27.897484

Figura 8. Estadísticas de resultados en el error de predicción para modelos GARCH, EGARCH, EWMA, LSTM-MSE y LSTM-MAE.

Posteriormente, se van a mostrar los tests de significancia. Sin embargo, a continuación se comentarán los resultados que afectan a ambas partes.

	MSE	MAE	HMAE	HMSE
GARCH	0.000544	0.000569	0.327265	0.240798
EGARCH	0.000533	0.000566	0.311301	0.189155
EWMA	0.000464	0.000517	0.278995	0.170707
LSTM_mse	0.000202	0.000329	0.214805	0.118977
W-LSTM_mse	0.000482	0.000607	0.399975	0.321558
E-LSTM_mse	0.000279	0.000411	0.273796	0.180926
G-LSTM_mse	0.000217	0.000350	0.218994	0.110915
GW-LSTM_mse	0.000208	0.000334	0.213378	0.112149
EW-LSTM_mse	0.000290	0.000417	0.274917	0.174248
GE-LSTM_mse	0.000634	0.000681	0.454264	0.428972
GEW-LSTM_mse	0.000223	0.000334	0.205429	0.105660
LSTM_mae	0.000161	0.000277	0.172599	0.078572
W-LSTM_mae	0.000169	0.000293	0.176553	0.074811
E-LSTM_mae	0.000338	0.000462	0.303788	0.199729
G-LSTM_mae	0.000190	0.000275	0.133232	0.038469
GW-LSTM_mae	0.000289	0.000415	0.275412	0.186244
EW-LSTM_mae	0.000216	0.000324	0.169203	0.061520
GE-LSTM_mae	0.000219	0.000326	0.179966	0.078232
GEW-LSTM_mae	0.000191	0.000298	0.161144	0.059592

Figura 9. Resultado de métricas de rendimiento para diferentes modelos. La métrica de MSE está multiplicada por 1000 por simplicidad de vista.

En la **Figura 9** se muestran los errores según la función de pérdida. Es notorio que los modelos DFN no se encuentran presente. Esto se debe a que, ninguno de los modelos fue capaz de entregar una predicción de la volatilidad coherente. Esto puede ser un indicativo de la existencia de ineficiencias en la estructura de estas redes. En consecuencia, no se va a poder comparar el rendimiento de estas con LSTM y los otros modelos. Por partes de los autores, se sabe que existen dificultades al momento de modelar una red y que lamentablemente, no se consigue para los datos que se deseaban estudiar.

Al utilizar la función de pérdida MSE, se observa que para los modelos sin redes neuronales, EWMA presenta un rendimiento superior sobre GARCH (13 % MSE; 8 % MAE;

10 % HMAE; 10 % HMSE) y EGARCH (15 % MSE; 9 % MAE; 15 % HMAE; 29 % HMSE). Este resultado difieren con los que se obtienen en el artículo de referencia ya que, se planteaba que EGARCH es la técnica que sobresale de los modelos GARCH. Sin embargo, ellos utilizan un indicador de mercado el cual, su naturaleza no es tan volátil como la de las crypto monedas. Por lo tanto, para este trabajo, tiene sentido de que EGARCH no sea un buen predictor ya que, exagera la predicción de la volatilidad y de esta manera sufre un mayor castigo que EWMA. Lo anterior se puede observar en el **Figura 10**.

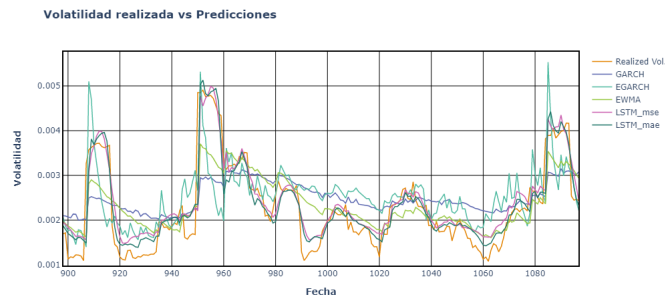


Figura 10. Zoom de resultados de predicción de volatilidad realizada para diferentes modelos. Se excluyen los modelos de LSTM para la simplicidad de la visualización.

Cuando se agregan la red neuronal, se observa que LSTM presenta una mejora significativa sobre EWMA (56 % MSE; 36 % MAE; 23 % HMAE; 30 % HMSE). Esto quiere decir que, LSTM es capaz de capturar eficazmente la volatilidad de BNB el cual, supera a las técnicas financieras tradicionales. Además, este resultado se asemeja al que obtienen los autores del trabajo anterior. Sin embargo, a partir de aquí, se dejan de ver similitudes con su trabajo de investigación

Al comparar los modelos híbrido simples, se visualiza de que el modelo G-LSTM presenta una mejora sobre E-LSTM (22 % MSE; 15 % MAE; 20 % HMAE; 39 % HMSE) y W-LSTM (55 % MSE; 42 % MAE; 45 % HMAE; 65 % HMSE). Sin embargo, al compararlo con LSTM, se obtiene una leve disminución en su rendimiento en la mayoría de sus métricas (−7 % MSE; −6 % MAE; −2 % HMAE; 7 % HMSE). Se esperaba que E-LSTM fuese el mejor modelo híbrido y superior a LSTM. Sin embargo, al visualizar su comportamiento en la **Figura 11**, se aprecia zonas donde E-LSTM sub estima la volatilidad mientras que, G-LSTM se asemeja más en su comportamiento. Por otro lado, se puede apreciar que para los datos de esta investigación, los modelos híbridos no necesariamente son sinónimos de una mejor predicción.

La intuición descrita toma mayor peso al analizar los resultados de los modelos híbridos dobles y triples. Para el caso de los dobles, solo GW-LSTM logra una leve mejora

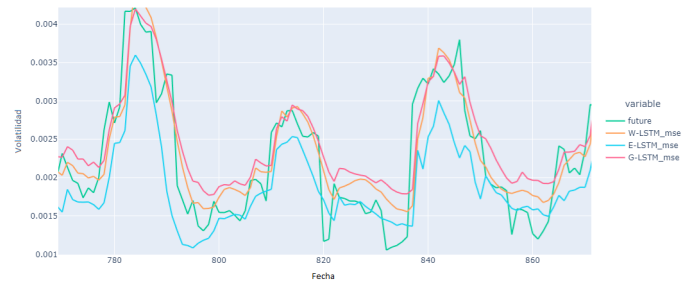


Figura 11. Zoom de resultados de predicción de volatilidad realizada para los modelos GARCH, EWMA y EGARCH.

contra G-LSTM (4 % MSE; 5 % MAE; 3 % HMAE; −1 % HMSE). Sin embargo, al compararlo con LSTM, solo se detectan leves mejoras en las métricas ajustadas a la heterocedasticidad (−3 % MSE; −2 % MAE; 0 % HMAE; 5 % HMSE).

Por parte de los modelos híbridos triples, GEW-LSTM logra mejorar el rendimiento de LSTM solo en las métricas HMAE y HMSE (−11 % MSE; −1 % MAE; 4 % HMAE; 11 % HMSE). Por lo tanto, GEW-LSTM puede ser catalogado como mejor modelo para las métricas ajustadas.

Como se puede apreciar, existe una notable variación de los resultados entre los modelos. Si se ve la investigación desde una perspectiva general, GEW-LSTM logra obtener el menor error en las métricas heterocedasticas. Sin embargo, fue necesario agregar tres técnicas financieras para lograrlo. Bajo una perspectiva particular, combinar una o dos modelos con redes neuronales, no entrega consistentemente la capacidad de mejorar el poder predictivo. Es decir, para ciertos casos si se va lograr un mayor rendimiento y para otros no. Para complementar los resultados, es necesario ver lo que sucede cuando se utiliza otra función de perdida para la predicción.

En la segunda parte, se utilizan los errores obtenidos al utilizar la función de perdida MAE en la **Figura 9**. Es decir, los modelos GARCH, EGARCH y EWMA no son afectados al poseer los mismos resultados expuestos al inicio de la sección. Cuando se agrega el modelo LSTM a la comparación, se obtiene nuevamente una mejora significativa en el rendimiento sobre EWMA (65 % MSE; 47 % MAE; 38 % HMAE; 54 % HMSE). Por lo tanto, nuevamente se confirma que los modelos LSTM son una buena forma de predecir la volatilidad en la crypto moneda BNB.

Con respecto a los modelos híbridos simples, se obtienen resultados que difieren levemente de la primera sección. En el caso de la métrica MSE, se obtiene que W-LSTM es el mejor modelo pero, posee un rendimiento inferior que LSTM (−5 % MSE; −6 % MAE; −2 % HMAE; −5 % HMSE). Para las métricas restantes, G-LSTM logra ser un modelo

Con el fin de verificar si los resultados son estadísticamente significativos, se realizan los test DM y WS los cuales, se presentan en la **Figura 12**. Los valores de las diagonales superiores representan los resultados por DM y la diagonal inferior, representa a WS. Se logra significancia en 320 de 342 casos. Esto quiere decir que el 93 % de los casos son validos para el estudio y respalda los resultados previamente mencionados.

Finalmente, a partir de los resultados, el modelo híbrido G-LSTM con la función de pérdida MAE, es el que obtiene el menor error entre las métricas. Por lo tanto, se puede catalogar como el mejor modelo híbrido en esta investigación.

Durante todo este trabajo, estaba claro que utilizar un índice tan volátil pudiese generar dificultades en los modelos predictivos. A partir de los resultados, se confirma las sospechas ya que, se encuentran algunos comportamientos diferentes a los esperados.

Con respecto a los modelos que se aplicaron correctamente en la investigación, se obtiene que EWMA es el mejor modelo Garch para el pronóstico de la volatilidad de BNB. Este resultado va es una diferencia con los del artículo de

Figura 12. Tabla resumen p-valores para los test de Diebold Mariano y Wilcoxon de signo. Diagonal superior corresponden a valores para test de diebold mariano mientras que diagonal inferior para test de Wilcoxon.



referencia ya que, en este se define que EGARCH es mejor modelo. A partir de esto, se analiza ciertos factores que pueda influir a lo observado y se concluye que: El motivo de que EGARCH tenga un peor desempeño, se debe a sus excesiva sobre-estimación y sub-estimación de la volatilidad lo cual, es castigado por la función de los errores. En consecuencia, se corrobora estadísticamente que EWMA es el mejor predictor que EGARCH y GARCH.

Además, se demuestra que las redes neuronales recurrentes pueden generar mejoras significativas frente a las técnicas tradicionales. En específico. Se concluye que LSTM es un buen pronosticador para gran parte de los casos, se logra aplicarla satisfactoriamente y se obtienen resultados similares a los del artículo de referencia.

Se descubre que distintos factores pueden cambiar los resultados obtenidos. En particular, se observa que cambiar la función de pérdida, desde MSE hacia MAE, logra una disminución promedio del 40 % en los errores.

Por otra parte y a diferencia de los autores, no se observa una mejora sistemática al combinar los modelos Garch con redes neuronales. Si bien, existen ciertos casos particulares donde se observa mejoras considerables, no es la regla. Lo anterior se debe a que LSTM presenta resultados inmejorables en una cantidad de casos no menor. Además, la combinación de modelos, puede sesgar la captura de características y no ajustarse bien a información muy volátil.

Lo anterior no quiere negar la utilidad de generar modelos híbridos sino, que hay que tener cuidado en seleccionarlos. A partir de los resultados, se puede corroborar que, efectivamente, los modelos híbridos pueden logran mejores resultados. En específico, se concluye que G-LSTM es el mejor modelo de todos y se respalda por las pruebas estadísticas. En consecuencia, se comprueba la hipótesis de que, los modelos híbridos pueden generar una mejor captura de información.

## V. TRABAJOS FUTUROS

A partir de los resultados obtenidos, es claro que no se logran resultados comparables con el artículo de referencia. Si bien, el análisis de error entrega ciertas direcciones sobre la precedencia de este, al momento de redactar este trabajo, no queda claro si es uno o múltiples factores.

Por lo anterior, es predominante averiguar si la utilización de los hiperparámetros y otros factores son los adecuados para esta investigación. Por lo tanto, un próximo trabajo se debe centrar en el análisis y sensibilización de los parámetros utilizados, para comprender los comportamientos observados para la métrica MAPE y los modelos que contienen DFN.

Existe también la posibilidad de que, factores externos sean necesarios para un mejor pronóstico de la volatilidad de las crypto monedas. Para el caso de este artículo, solo se

utilizan como entradas el precio de BNB y los parámetros calculados de los modelos financieros. Por esta razón, una posible investigación se debiese centrar en modelar el comportamiento del BNB al considerar datos de entradas exógenos. Por ejemplo, los valores de crypto monedas rivales, tasa de inflación en las economías centrales y la influencia de las redes sociales.

Finalmente, debido a los resultados, existe la posibilidad de que los modelos GARCH no son los óptimos para modelar la volatilidad en este indicador. En consecuencia, nace la inquietud de que la combinación de distinta técnicas de predicción puedan generar resultados más relevantes. Es más, la combinación entre redes neuronales no supervisadas, probabilísticas y supervisadas puede ser de gran interés.