

PAPER

Predicción de precios de acciones y criptomonedas.

rafacastillo@ciencias.unam.mx

Resumen

En este proyecto se predijeron los precios de bitcoin y algunas acciones usando el modelo ARIMA (AutoRegresive Integrated Moving Average) y una red neuronal LSTM (Long Short Term Memory). Además se identificó el impacto que tienen celebridades como Elon Musk, Cristiano Ronaldo o comunidades web como los Wall Street Bets en los precios de distintos activos.

Key words: Bitcoin, Elon Musk, Wall Street Bets, ARIMA, LSTM, GME

Introducción

Precio de las acciones

El precio de las acciones de una empresa se ve afectado en general por cuatro factores [1]:

- Oferta y demanda
- Salud financiera de la empresa
- Salud financiera de la industria
- Tendencias económicas

En resumen, cuanto más desee un inversionista poseer una acción, estará dispuesto a pagar un precio más elevado por ella, si muchos inversionistas comparten este deseo la acción subirá de precio. Para crear este deseo en los inversionistas, la empresa debe ser prometedora, es decir, tener un historial saludable que ofrezca grandes ganancias y/o ofrecer un producto nuevo e innovador en una industria que se este expandiendo.

Venta descubierta

Adquirir acciones que van a subir de precio no es la única forma de obtener ganancias con este tipo de activos, existen prácticas entre los inversionistas donde se obtienen ganancias de una acción que está bajando su precio, una de estas prácticas es llamada venta descubierta o posición corta [2].

Para explicar de manera clara la venta descubierta resulta útil poner un ejemplo: Supongamos que un inversionista pide prestadas 10 acciones a un corredor a un precio de 10 dólares cada una. En el momento que el corredor le da las acciones al inversionista, éste las vende en el mercado (a 10 dólares). Luego el inversionista espera un tiempo y las acciones bajan de precio, ahora cuestan 5 dólares cada

una, por lo que el inversionista las vuelve a comprar en el mercado y se las regresa al corredor. En toda esta transacción el inversionista ganó 5 dólares por acción (500 dólares en total). En la figura 1 se muestra el diagrama de esta transacción.



Figura 1. Diagrama de venta descubierta.

En este tipo de prácticas, el riesgo que corre el inversionista es alto, ya que la acción puede subir de precio. En dicho caso el inversionista compraría las acciones más caras en el mercado, dado que el precio de una acción no tiene un límite superior, el inversionista puede terminar pagando cantidades enormes si la acción sube mucho de precio. Es por esto que los inversionistas experimentados realizan estas prácticas únicamente con acciones de empresas que están quebrando.

Criptomonedas

Las criptomonedas (o criptos) como Bitcoin pueden definirse con 3 palabras: moneda digital descentralizada [3]. Las dos principales diferencias que tienen estas criptos con las monedas centralizadas son (entre otras):

- Las monedas centralizadas requieren de un gobierno que genere billetes o monedas certificadas.
- Las monedas centralizadas requieren de uno o varios bancos que lleven las cuentas de los usuarios y completen transacciones entre estos.

Las criptomonedas, por el contrario no tienen estos dos requerimientos y proveen una tecnología que permite llevar las cuentas de los usuarios y completar transacciones sin los costos que conlleva tener un banco o las desventajas de que exista un gobierno con la capacidad de emitir más billetes o efectivo (ya que es digital y descentralizada).

¿Qué le da valor a las criptomonedas?

En resumen las criptomonedas tienen valor gracias a la oferta y la demanda. Criptomonedas como Bitcoin se han vuelto muy populares en los últimos años, por lo que su demanda se ha incrementado de manera radical, sin embargo, existe un número limitado de Bitcoins. Actualmente hay en circulación cerca de 18 millones de Bitcoins de los 21 millones que hay en total [4]. Cada vez se vuelve más complejo que nuevos Bitcoins entren en circulación y no existe un gobierno u organismo que pueda crear más, esto hace que el valor de la criptomoneda suba de precio, cómo se muestra en la figura 2.

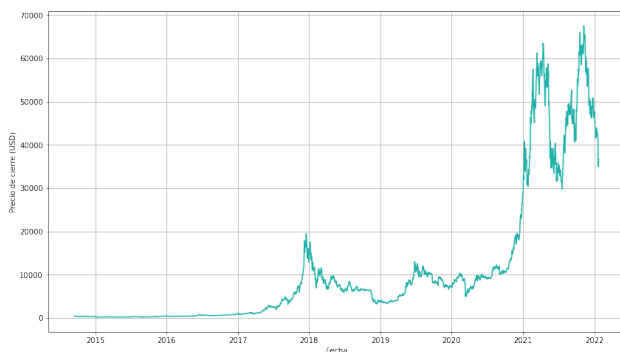


Figura 2. Precio de cierre de Bitcoin en dólares desde 2014.

Es claro que el valor de las criptomonedas, por lo anteriormente expuesto, es volátil y puede tener fluctuaciones altas en su precio en intervalos cortos de tiempo. Por lo que como inversión, presentan un riesgo alto, pero también ganancias considerables.

El problema

A lo largo del 2021 se presentaron diversas situaciones en redes sociales que aparentemente tuvieron un efecto directo en los precios de diversas acciones y criptomonedas. Estos sucesos dieron pie a mitos y creencias sobre lo inseguro que es invertir en criptomonedas u otros activos, con un argumento basado en que los posts en twitter o rechazos públicos de algunas celebridades puede alterar sus valores de manera considerable. A continuación se describirán algunos de esos casos.

Cristiano Ronaldo y Coca Cola

¿Quién es Cristiano Ronaldo?

Cristiano Ronaldo es un futbolista portugués considerado por algunos como el mejor jugador de fútbol soccer del mundo, dada su popularidad, sus comentarios o gestos tienen un impacto en las tendencias de redes sociales y noticieros.

El 15 de junio del 2021, el jugador se encontraba en una rueda de prensa previa a un partido y apartó de su zona dos botellas de Coca Cola (de manera algo despectiva) e inmediatamente tomó una botella de agua, volteó a ver a las cámaras y afirmó con un ademán que había que escoger el agua siempre.



Figura 3. Noticias sobre el rechazo de Cristiano Ronaldo a Coca Cola [5].

Tras el suceso, diversos noticieros anunciaron una caída en el precio de las acciones de Coca Cola, que pasaron de 56.10 a 55.22 dólares y atribuyeron este hecho al rechazo del jugador a los refrescos, afirmando una pérdida para la compañía de 4,000 millones de dólares.

Elon Musk y las criptomonedas

¿Quién es Elon Musk?

Elon Musk ocupa la segunda posición en la lista de personas más ricas del mundo [6] y es el director, subdirector y cofundador de distintas compañías importantes a nivel mundial, como Tesla Motors y PayPal. Al ser un personaje tan importante, sus comentarios tienen un alto impacto en redes sociales y noticieros.

Elon Musk tuiteó en diversas ocasiones acerca de distintas criptomonedas como Bitcoin y Dogecoin, pero dos tweets en particular generaron un desconcierto en noticieros y redes. El 12 de mayo del 2021, Elon Musk informó desde su cuenta de Twitter que ya no iba a ser posible comprar vehículos Tesla con Bitcoin. Durante los días posteriores hubo un decremento en el precio de la criptomoneda de cerca del 40 %, como se muestra en la figura 4.

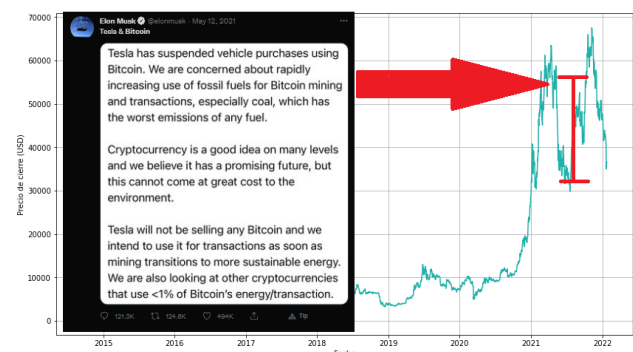


Figura 4. Caída en el precio de Bitcoin tras tweet de Elon Musk.

Diversas páginas de noticias afirmaban que el precio de Bitcoin podía ser controlado por los tweets de Elon [7] e infundieron miedo a los inversionistas pequeños.

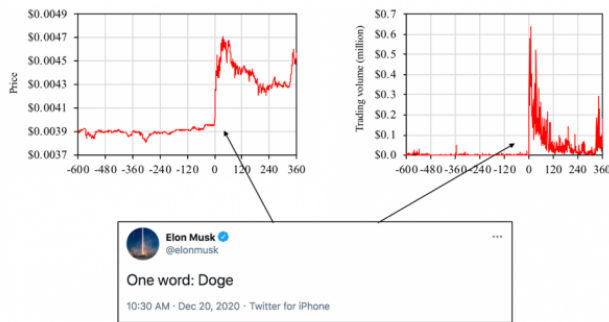


Figura 5. Alza en el precio de Dogecoin tras tweet de Elon Musk.

Anteriormente, Elon Musk había twiteado de manera positiva sobre la moneda Dogecoin, poco después se vió un incremento enorme en su precio, como se muestra en la figura 5, este incremento en el precio también fue atribuido al tweet de Elon.

WallStreetBets

¿Qué es Reddit?

Reddit es un sitio web donde los usuarios pueden añadir textos, imágenes, videos o enlaces. Los usuarios pueden votar a favor o en contra del contenido e incluso comentar acerca de éste, haciendo que aparezcan en las publicaciones destacadas, para funcionar, Reddit ocupa elementos llamados subreddits, que son áreas de interés creadas por el usuario donde se organizan las discusiones.

WallStreetBets es un subreddit sobre inversión en bolsa con casi 7 millones de miembros [8], llegó a convertirse en el subreddit más comentado del mundo.

¿Qué es GameStop?

GameStop es una tienda física estadounidense de videojuegos y consolas y como muchas otras empresas, tuvo una reducción en sus ingresos debido a la pandemia, por lo que el precio de sus acciones cayó. Debido a esto, algunos inversores internacionales pensaron que la empresa quebraría, y decidieron realizar ventas descubiertas con la esperanza de ver un desplome en el precio de las acciones de GameStop. Sin embargo, el panorama de la empresa no era tan desfavorable, ya que en enero del 2021 GameStop tuvo nuevos miembros en el consejo de administración con gran experiencia en el manejo de tiendas en línea [10], por otra parte, en esa temporada salieron a la venta las consolas de PlayStation y de Xbox. Ambos factores marcaban un punto de crecimiento para la empresa: se ofrecían productos nuevos en innovadores con mucha demanda, la industria en la que se basa la empresa estaba en un momento de crecimiento y los nuevos inversionistas aportaban una buena salud financiera.

¿Cómo se relaciona GameStop con WallStreetBets?

Diversas instituciones con ventas descubiertas de la acción no estaban dispuestas a aceptar las pérdidas que conllevaba que la acción de GameStop subiera de precio, por lo que reforzaron su estrategia realizando más ventas en corto sobre dicha acción, usando su capital para hundir los precios, abrieron tantas ventas descubiertas que ya habían pedido prestadas más acciones de las que había disponibles [9], por lo que había más oferta que demanda, lo que forzaba que la acción bajara de precio. Sin embargo, la comunidad de los WallStreetBets comenzó a comprar las acciones de GameStop, compensando la demanda y forzando el precio a la alza, esto dió pie a que las instituciones tuvieran que comprar acciones de GameStop

rápidamente antes de que éstas subieran más de precio y que su pérdida fuera mayor (fenómeno conocido como short squeeze [11]), lo que causó un incremento aún mayor en el precio de la acción.

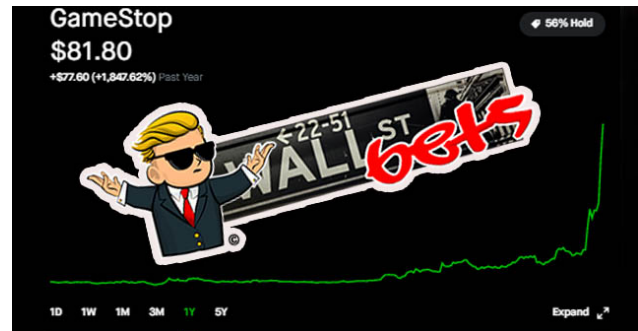


Figura 6. Incremento en el precio de las acciones de GameStop tras múltiples publicaciones en el subreddit de WallStreetBets [12].

Las instituciones se vieron forzadas a cerrar sus ventas descubiertas con pérdidas, ya que si esperaban más, la acción podía subir aún más de precio y su pérdida sería mayor. Las acciones de GameStop subieron de 7 dólares en septiembre del 2020 a 327 dólares en enero del 2021.

Propósito del estudio

En este trabajo se busca construir dos modelos que predigan los precios de acciones y criptomonedas de manera regresiva y verificar si la predicción obtenida debería tomar en consideración las tendencias en redes sociales con respecto al activo en evaluación.

Objetivos

Objetivo general

Predecir el precio de distintas acciones y criptomonedas tomando en consideración las tendencias en redes sociales.

Objetivos particulares

- Construir dos modelos regresivos que predigan el valor de distintos activos a mediano y corto plazo.
- Analizar como afectan el precio de los activos las tendencias en redes sociales.

Hipótesis

El precio de los activos es predecible de manera numérica mediante modelos estadísticos y de machine learning. Sin embargo existen factores no considerados en estos modelos, como la oferta y la demanda, que también influyen en dichos precios. Este trabajo busca demostrar que los comentarios de una figura pública en particular no es parte de estos factores, y no alteran de manera considerable los precios de los activos.

Métodos y modelos

Modelo ARIMA

El modelo ARIMA (Autoregressive Integrated Moving Average) es un modelo estadístico cuyo fin es encontrar patrones para una predicción hacia el futuro, es muy usado en series de tiempo y en econometría. En este modelo las estimaciones futuras vienen explicadas únicamente por datos del pasado [13].

Red neuronal LSTM

Una red neuronal recurrente tiene una función de activación que no actúan en solo una dirección, sino que incluye conexiones que apuntan "hacia atrás", por lo tanto logra realizar una retroalimentación con valores previos, lo cual es sumamente útil cuando se utilizan series de tiempo [14]. LSTM es una extensión de las redes neuronales recurrentes que amplía la memoria para aprender experiencias importantes que han pasado incluso en lapsos grandes de tiempo [15].

Evaluación de modelos

Error porcentual absoluto medio

El error porcentual absoluto medio o MAPE (Mean Absolute Percentage Error) mide el tamaño del error (absoluto) en términos porcentuales. Su expresión matemática es de la forma:

$$\text{MAPE} = \frac{100\%}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right| \quad (1)$$

donde A_t representa el valor real del activo en un instante t y F_t es el valor pronosticado por el modelo [16], el modelo ARIMA fue evaluado mediante este parámetro.

Pérdida y precisión

Para evaluar al modelo de deeplearning se utilizaron las variables pérdida y precisión o loss y accuracy. La variable loss indica que tan bien (o mal) está prediciendo los valores, la función de pérdida utilizada para obtener la variable loss en este trabajo fue *Mean Squared Error*, esta variable detecta que tan grandes son los errores del modelo. La variable accuracy compara las predicciones con los valores reales en términos de porcentajes, así, esta variable es más sensible a detectar la cantidad de errores que comete el modelo. Para sintetizar la interpretación de estas variables, resulta útil observar el cuadro 1:

Cuadro 1. Promedios dependiendo del género

	Valor loss pequeño	Valor loss alto
Valor accuracy pequeño	Muchos errores pequeños	Muchos errores grandes
Valor accuracy alto	Pocos errores pequeños	Pocos errores pequeños

Descripción del proyecto

Para obtener el historial de precios de los activos se utilizó la librería de python *yfinance*, que es una herramienta de código abierto que utiliza las API disponibles públicamente de Yahoo y está destinada a fines educativos y de investigación. Los Tweets de Elon Musk y las publicaciones en el subreddit de WallStreetBets se obtuvieron de las bases de datos de kaggle [17, 18].

Se aislaron los tweets y las publicaciones de reddit dada su fecha de publicación para integrarlos a una gráfica y corroborar de manera visual la correlación de estos con el precio de los activos. Para cada activo en particular se realizó un análisis y se corroboró si la variación en el precio tuvo su origen por las tendencias en redes sociales o si era un comportamiento esperado, es decir, predecible mediante un modelo estadístico como los descritos anteriormente.

Análisis

Cristiano Ronaldo y Coca Cola

Al observar la gráfica del precio de las acciones de Coca Cola en la figura 7, se aprecia, a pesar de las fluctuaciones, una conducta creciente.

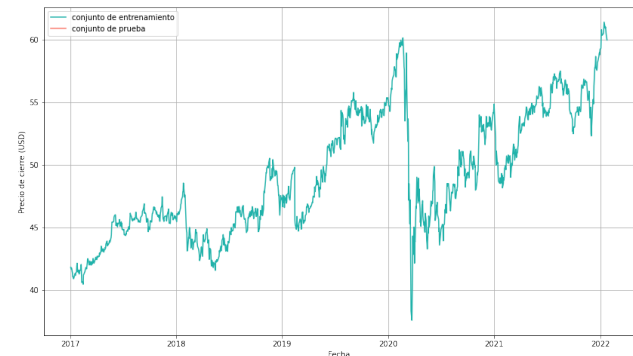


Figura 7. Precio de cierre de Coca Cola en dólares desde 2017.

Al centrar la gráfica cerca del 15 de junio del 2021, se puede apreciar un decremento posterior a la conferencia de prensa donde Cristiano Ronaldo rechazó los refrescos (momento marcado con un punto rojo en la figura 8):

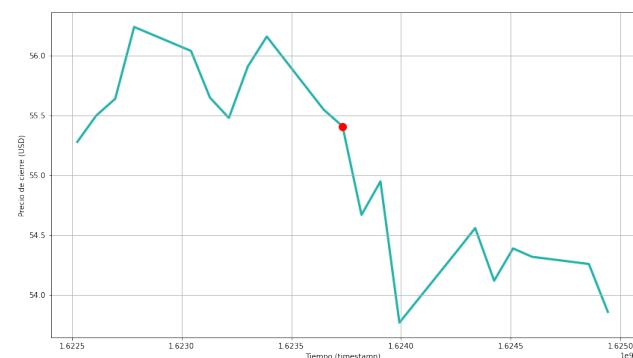


Figura 8. Caída en el precio de cierre de Coca Cola tras el rechazo de CR7.

Al observar en esta escala la gráfica, puede parecer que el suceso tuvo un impacto alto en el precio de las acciones, sin embargo, al observar la gráfica en un intervalo de tiempo mayor (como se muestra en la figura 9) esta idea pierde sustento de manera intuitiva:

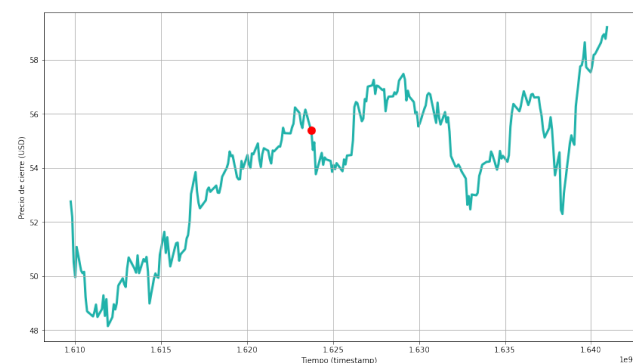


Figura 9. Caída en el precio de cierre de Coca Cola tras el rechazo de CR7, en un mayor intervalo de tiempo.

En la figura 9, se aprecian fluctuaciones aún mayores posteriores al suceso en cuestión. Por otro lado la conducta descendente en el precio de la acción estaba presente desde antes del acontecimiento. Este hecho puede corroborarse al obtener las predicciones de los modelos ARIMA y LSTM.

El modelo LSTM predice un precio de cierre para el 15 de junio de 2021 de 54.56 USD, el valor de cierre real de dicho día fue de 55.41 USD, por lo que el modelo LSTM predecía incluso un declive mayor en el precio de la acción, tomando en cuenta únicamente los precios de cierre anteriores. El modelo ARIMA predice también un precio de cierre más bajo, de 55.06 USD, mostrando que la disminución en el precio de cierre era un suceso que podía predecirse de manera numérica y no necesariamente fue causada por la acción de Cristiano Ronaldo.

Elon Musk y las criptomonedas

Análogamente es posible analizar los tweets de Elon Musk y su correlación con los precios de cierre de la Bitcoin y Dogecoin. En la figura 10 se muestra la gráfica del precio de cierre de Bitcoin, los puntos rojos representan el momento en el que Elon twiteó algo relacionado con la criptomoneda.

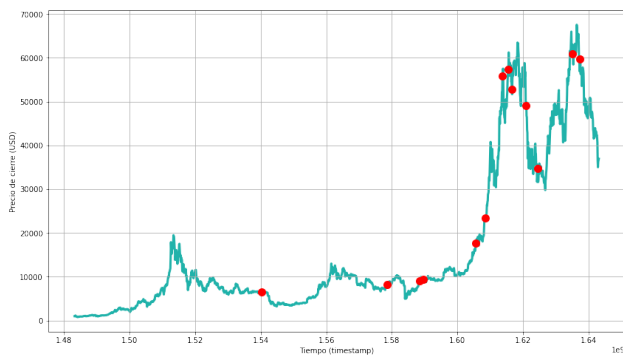


Figura 10. Precio de cierre de Bitcoin y tweets de Elon Musk.

En la figura 11 se muestran los tweets de Elon Musk:

1. pay by bitcoin capability available outside us later this year
2. tesla is using only internal amp open source software amp operates bitcoin nodes directly bitcoin paid to tesla will be retained as bitcoin not converted to fiat currency
3. you can now buy a tesla with bitcoin
4. btc bitcoin is an anagram of the boring company what a coincidence
5. to be clear i am not an investor i am an engineer i don't even own any publicly traded stock besides tesla however when fiat currency has negative real interest only a fool wouldn't look elsewhere bitcoin is almost as good as fiat money the key word is "almost"
6. tesla's action is not directly reflective of my opinion having some bitcoin which is simply a less dumb form of liquidity than cash is adventurous enough for an smp500 company
7. bitcoin is almost as good as fiat money
8. bitcoin is my safe word
9. 🍷 toss a bitcoin to ur witcher 🍷
10. i still only own 0.25 bitcoins btw
11. pretty much although massive currency issuance by govt central banks is making bitcoin internet 🍷 money look solid by comparison
12. how much for some anime bitcoin
13. bitcoin is not my safe word
14. wanna buy some bitcoin 🍷
15. tesla has suspended vehicle purchases using bitcoin
16. BTC (Bitcoin) is an anagram of TBC(The Boring Company) What a coincidence
17. How many Bitcoin maxis does it take to screw in a lightbulb?
18. Bitcoin cures cancer
19. Out of curiosity, I acquired some ascii hash strings called "Bitcoin, Ethereum & Doge". That's it. As I've said before, do n't bet the farm on crypto! True value is building products & providing services to your fellow human beings, not money in any form.

Figura 11. Tweets de Elon Musk.

En este trabajo se hará énfasis en el tweet número 15, que fue publicado el 12 de mayo del 2021, diversos medios de comunicación afirmaban que la caída de en el precio descrita anteriormente fue debida a dicha publicación, el modelo LSTM predice un valor de cierre para el 12 de mayo de 50,189.55 USD, mientras que el precio real fue de 49,150.53 USD, cabe destacar que el cierre del 11 de mayo (un día antes) fue de 56,704.57 USD, por lo que el deceso en el precio también estaba predicho en este modelo, el modelo ARIMA predice un precio de cierre de 49,587.62 USD. Sustentando de nuevo la hipótesis. Sin embargo en esta ocasión el precio predicho fue más alto que el real, por lo que la publicación pudo inducir un descenso

aún mayor en la caída del precio, de cualquier manera la caída en el precio era predecible de manera estadística.

Con respecto a la criptomoneda Dogecoin (véase figura 5) es posible ver que tras el tweet de Elon Musk la criptomoneda comenzó un periodo de crecimiento, el modelo ARIMA puede usarse para hacer predicciones a largo plazo y predecir una tendencia. En la figura 12 observable que, la tendencia de Dogecoin predicha por el modelo ARIMA estaba muy por debajo de los precios reales de cierre en el tiempo posterior.

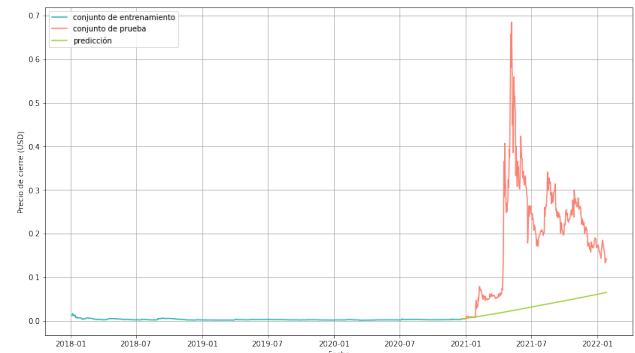


Figura 12. Precio de cierre de Dogecoin y predicción ARIMA.

Esta falla en la predicción del modelo podría atribuirse a que efectivamente el tweet de Elon Musk alteró las tendencias de en los precios de cierre de la criptomoneda si no se analizan otros factores. En la siguiente sección se analizarán otras circunstancias que pudieron ocasionar estas fluctuaciones tan fuertes en el precio de la criptomoneda.

WallStreetBets

Es posible analizar la gráfica de los precios de cierre de Dogecoin como en la figura 10, pero observando los posts en reddit donde se menciona a la criptomoneda, como se muestra en la figura 13.

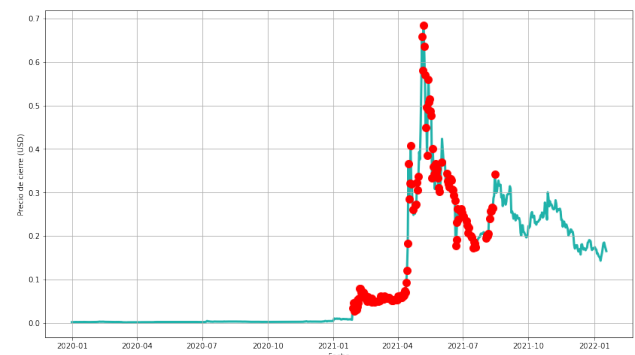


Figura 13. Precio de cierre de Dogecoin y posts en WallStreetBets.

Pueden analizarse las palabras que más se repiten en dichos posts mediante un wordcloud, como se muestra en la figura 14. De esta forma es notorio que, la creciente demanda de la criptomoneda pudo tener su origen en el subreddit y no en la publicación de Elon Musk, este hecho puede analizarse si se repite el análisis con el comportamiento del precio de otro activo, como las acciones de AMC Entertainment Holdings Inc. Dada la constante caída en el precio de sus acciones, los grandes inversionistas tendían a hacer ventas descubiertas con este activo, igual que con GameStop, los usuarios de WallStreetBets realizaron una estrategia similar con las

- bes. Recuperado 26 de enero de 2022, de <https://www.forbes.com/sites/davidewalt/2016/12/14/the-worlds-most-powerful-people-2016/?sh=522952c41b4c>
7. Molla, R. (2021, 14 junio). When Elon Musk tweets, crypto prices move. Recode. Recuperado 26 de enero de 2022, de <https://www.vox.com/recode/2021/5/18/22441831/elon-musk-bitcoin-dogecoin-crypto-prices-tesla>
8. Wallstreetbets. Reddit. Recuperado 26 de enero de 2022, de <https://www.reddit.com/r/wallstreetbets/>
9. GameStop: Hedge Funds vs Amateur Traders? (2021, 24 febrero). HEC PARIS. Recuperado 27 de enero de 2022, de <https://www.hec.edu/en/knowledge/instant/gamestop-hedge-funds-vs-amateur-traders>
10. Ryan Cohen multiplica 25 veces las ganancias de su inversión en GameStop por rally de «meme stocks». (2021, 25 agosto). EL CEO. Recuperado 27 de enero de 2022, de <https://elceo.com/mercados/ryan-cohen-multiplica-25-veces-las-ganancias-de-su-inversion-en-gamestop-por-rally-de-meme-stocks/>
11. Mitchell, C. (2022, 2 enero). Short Squeeze. Investopedia. Recuperado 27 de enero de 2022, de <https://www.investopedia.com/terms/s/shortsqueeze>
12. Reddit Is Seriously F***ing With Wall Street. (2021, 26 enero). eBaum's World. Recuperado 27 de enero de 2022, de <https://gaming.ebaumsworld.com/articles/reddit-is-seriously-fing-with-wall-street/86564611/>
13. ARIMA model. statsmodels. Recuperado 27 de enero de 2022, de <https://www.statsmodels.org/dev/generated/statsmodels.tsa.arima.model.ARIMA.html>
14. Gulli, A. (2018). Deep Learning with Keras. Packt Publishing.
15. LSMT layer. Keras. Recuperado 27 de enero de 2022, de https://keras.io/api/layers/recurrent_layers/lstm/
16. Mean absolute percentage error. Wikipedia. Recuperado 27 de enero de 2022, de https://en.wikipedia.org/wiki/Mean_absolute_percentage_error
17. Posts de WallStreetBets, conjunto de datos de Kaggle, <https://www.kaggle.com/gpreda/reddit-wallstreetsbets-posts>
18. Tweets de Elon Musk, conjunto de datos de Kaggle, <https://www.kaggle.com/andradaolteanu/all-elon-musks-tweets>