

**CENTRO DE INVESTIGACIÓN MATEMÁTICA, UNIDAD
MONTERREY**



PROTOCOLO DE INVESTIGACIÓN:

**Análisis de sentimiento en redes sociales para el
pronóstico en criptomonedas.**

VERANO DE INVESTIGACIÓN CIMAT 2021

PRESENTA: Christopher Román Jaimes

ASESOR: Andrés García Medina

22 de junio del 2021

Introducción	3
Objetivos	4
Metodología	5
Análisis de sentimiento	5
Modelización de la volatilidad	5
Despliegue en la nube	5
Plan de trabajo	6
Visualización de datos(7 días)	6
Preprocesamiento de datos(7 días)	6
Análisis de sentimiento (7 días)	6
Modelización de la volatilidad y precio(18 días)	6
Despliegue en la nube (5 días)	7
Escribir reporte final (5 días)	7
Cronograma	8
Referencias	9

Introducción

En los últimos meses la inversión en criptomonedas (principalmente en Bitcoin) ha aumentado considerablemente a nivel mundial, esto debido principalmente a la actual crisis sanitaria causada por la pandemia de Covid-19 que trajo consigo una crisis financiera mundial que motivó a los inversores la necesidad de diversificar el riesgo en sus carteras por medio de las criptomonedas, esto generó una gran revalorización del Bitcoin y trajo consigo el crecimiento de otras criptomonedas.

Dado lo anterior las redes sociales se plagaron de noticias relacionadas a estos activos lo que aumentó la cantidad de inversores, esto nos da la pauta a pensar que el comportamiento del mercado de criptomonedas es afectado por las tendencias en redes sociales y en efecto las decisiones de los inversionistas se ven influenciados por la información de los medios de comunicación y redes sociales (García, 2017). Por otro lado, estos activos presentan una importante peculiaridad y es que su valor es muy volátil a comparación de otros activos como las acciones, esto trae consigo nuevos retos en la modelización de este factor de riesgo lo cual lo hace un fenómeno muy interesante.

Estos tres rubros: la actual importancia de las criptomonedas para los inversores, la dependencia de las redes sociales en la valorización de los activos y la gran volatilidad presente en el mundo cripto, hacen de gran interés el estudio en este tema es por ello que en este proyecto se realizará análisis de sentimiento en redes sociales con el objetivo de mejorar las predicciones en criptomonedas, para tomar una mejor decisión de inversión.

Objetivos

El presente proyecto está enfocado en el estudio del comportamiento de las criptomonedas.

El objetivo es hacer análisis de sentimiento usando información textual de twitter, desde la extracción hasta el análisis y así poder hacer una predicción con base en la serie de la polaridad. Los resultados serán desplegados en un dashboard en la nube pensado para que los inversionistas puedan utilizarlo para mejorar su toma de decisiones.

Metodología

Para poder realizar nuestra predicción nos enfocaremos en dos cuestiones importantes: el análisis de sentimiento en redes sociales (Twitter) y la modelización de la volatilidad. Una vez hecho esto se procederá a desplegar los resultados en la nube para su uso por cualquier persona interesada.

Análisis de sentimiento

Como se concluye en el artículo de García (2017) el análisis de sentimiento de datos textuales de Twitter aportan información relevante para el análisis de los índices financieros globales más importantes, lo que nos da la pauta para incorporar esta nueva variable en nuestro modelo.

Modelización de la volatilidad

En el mundo cripto sabemos que la volatilidad que las caracteriza ha retrasado su aceptación entre los inversionistas (Altamira, Hernández, López y Santamaría, 2020) por lo cual es de gran importancia modelar este factor de riesgo para llegar a un buen pronóstico. Se propone utilizar los modelos GARCH, LSTM y GARCH-LSTM para lograr este fin y entonces hacer un pronóstico.

Despliegue en la nube

Los resultados serán presentados en un dashboard interactivo por medio de Dash y Plotly en Python, tales resultados serán desplegados en la nube por medio de Azure(o algún otro servicio) para su funcionamiento independiente.

Plan de trabajo

Algunas de las siguientes tareas se harán a la par, en el cronograma de la siguiente sección se muestra cuales.

Visualización de datos(7 días)

- Instalar todo el software necesario (MySQL, Spyder, etc.).
- Leer y practicar con la documentación de Dash y Plotly para hacer dashboards.
- Cargar la base de datos en MySQL con los precios de las criptomonedas y acceder a ella desde Python.
- Visualizar las series de tiempo de las bases de datos en Plotly.

Preprocesamiento de datos(7 días)

- Se limpiará la base de datos si es necesario realizando eliminación de datos faltantes o de criptomonedas que no sean de interés.
- Se extraerá y transformará información textual de Twitter para posteriormente hacer el análisis de sentimiento.
- Se hará selección de variables para el pronóstico.

Análisis de sentimiento (7 días)

- Se realizará el análisis de sentimiento por medio de polaridad con los datos obtenidos de la tarea anterior.

Modelización de la volatilidad y precio(18 días)

- Utilizar algún modelo sencillo como el logístico para poder predecir la dirección del precio
- Buscar los parámetros adecuados para los modelos y usar el mejor para pronosticar tomando en cuenta el análisis de sentimiento ya hecho.
- Implementar la modelización de la volatilidad anterior para pronosticar el precio de los crypto activos por medio de modelos como ARIMA.

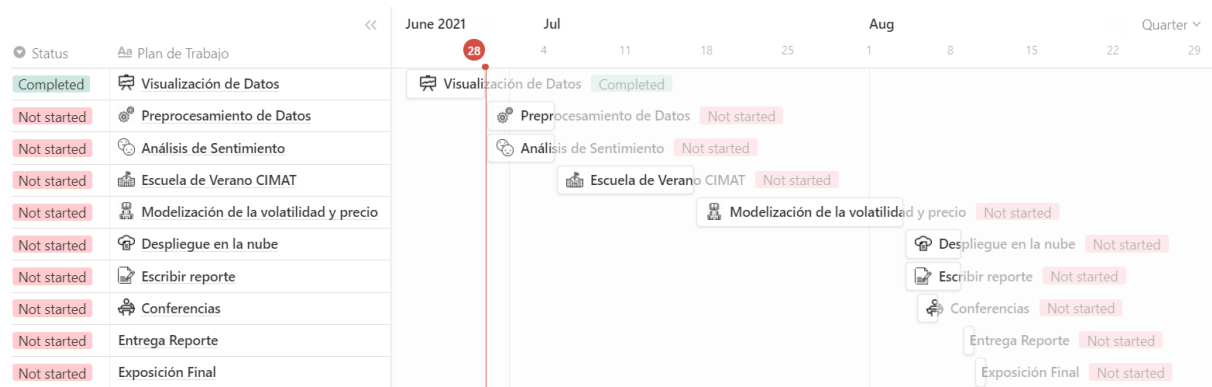
Despliegue en la nube (5 días)

- Desplegar los resultados finales en el servicio de nube Azure para su consulta sin necesidad de un servidor privado.

Escribir reporte final (5 días)

Cronograma

El siguiente cronograma muestra esquemáticamente el plan de trabajo mencionado anteriormente.



Referencias

García Medina, Andrés. (2017). El uso de Twitter en el análisis financiero: aproximación desde la econofísica. *INTERdisciplina*. 5. 23. 10.22201/ceiich.24485705e.2017.12.61469.

Vásquez, Margarita & Castro, Elda & Vicarte, Gloria & Trejo, Lorena. (2020). El impacto de la volatilidad en la funcionalidad de las criptomonedas. *Interconectando Saberes*. 5. 10.25009/is.v0i9.2659.