

Pronóstico de precios de criptomonedas utilizando LSTM, GRU y LSTM bidireccional: un enfoque de Aprendizaje Profundo



UNIVERSIDAD
DE
CÓRDOBA

Artem Mozhegov

Máster Universitario en Inteligencia Computacional e Internet De Las Cosas

Análisis Automático de Datos para las Ciencias Biomédicas, Medioambientales, Agroalimentarias

z22mozma@uco.es

Introducción

El mercado de criptomonedas es un lugar lucrativo para la especulación debido a su alta volatilidad. La aplicación de algoritmos de inteligencia artificial y aprendizaje automático puede ayudar a predecir los precios futuros de las criptomonedas.

El objetivo es mostrar cómo podemos usar los algoritmos de aprendizaje profundo LSTM (Long Short-term Memory), GRU (Gated Recurrent Unit) y Bi-LSTM (Bi-directional Long Short-term Memory) para las predicciones de BTC (Bitcoin), ETH (Ethereum) y LTC (Litecoin). El enfoque está en crear modelos precisos de predicción de precios y abordar preguntas de investigación.

¿Cómo pueden ayudar los algoritmos de Aprendizaje Profundo a pronosticar los precios de las criptomonedas?

¿Cuál es el mejor modelo de red neuronal artificial para pronosticar los precios de las criptomonedas elegidas?

Para evaluar el rendimiento de algoritmos de aprendizaje profundo propuestos, se utilizan el error cuadrático medio (RMSE) y el porcentaje de error absoluto medio (MAPE).

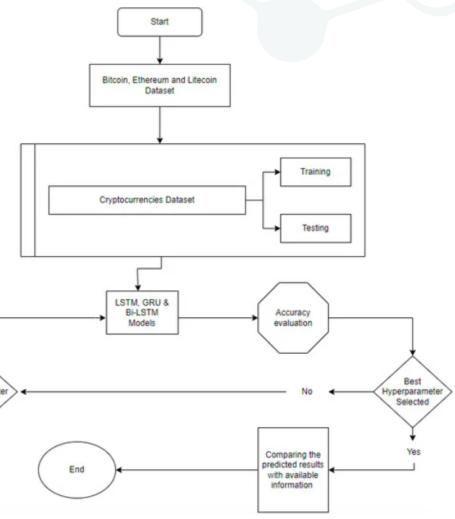
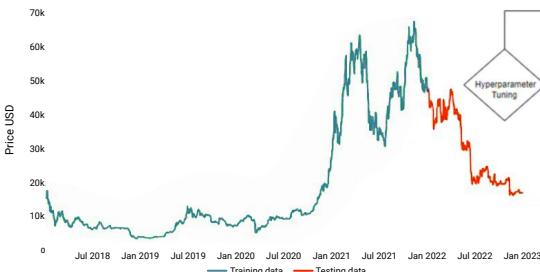
El aprendizaje profundo es una rama de la Inteligencia Artificial que utiliza redes neuronales con múltiples capas para aprender y extraer patrones significativos de los datos. Implica entrenar modelos para aprender y comprender automáticamente representaciones complejas de información, lo que lleva a un desempeño impresionante en tareas como reconocimiento de imágenes y voz, procesamiento de lenguaje natural y más.

Materiales y métodos

El estudio sigue este proceso:

1. Recopilación de datos históricos para BTC, ETH y LTC;
2. Visualización de datos exploratorios;
3. Dividir cada conjunto de datos en conjuntos de datos de entrenamiento y prueba;
4. Entrenamiento de tres tipos de modelos (para BTC, ETH, LTC);
5. Prueba de los modelos;
6. Comparación del rendimiento de cada método de aprendizaje profundo.

Entrenamiento y Prueba para BTC



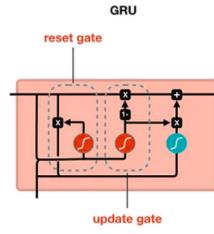
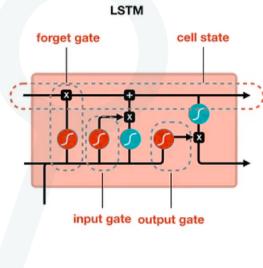
Metodología de procesamiento de datos y selección de modelos

Long Short-Term Memory (LSTM)

Gated Recurrent Unit (GRU)

Bidirectional Long Short-Term Memory (BiLSTM)

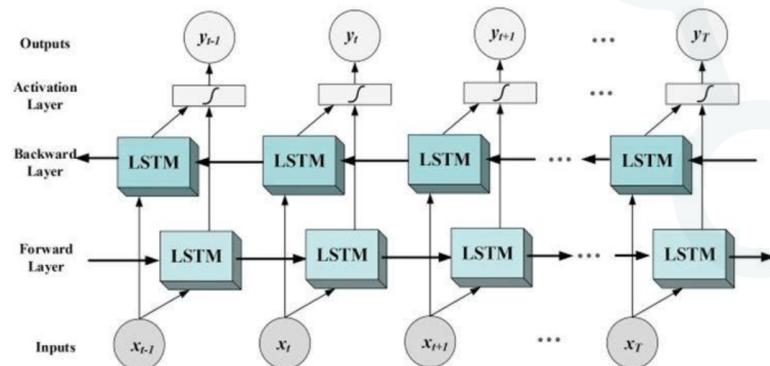
- Capaz de aprender y recordar información sobre secuencias extensas de datos;
- Contiene una celda de memoria, que actúa como una unidad de almacenamiento para mantener la información;
- 3 puertas principales: la puerta de entrada, la puerta de olvido y la puerta de salida:
 - La puerta de entrada (input gate) controla qué datos deben almacenarse en la celda de memoria;
 - La puerta de olvido (forget gate) descarta datos irrelevantes;
 - La puerta de salida (output gate) controla el flujo de información desde la celda de memoria al siguiente estado oculto y salida.
- Es eficaz para capturar y modelar dependencias a largo plazo en datos secuenciales.
- Más complejo computacionalmente en comparación con GRU debido a la presencia de tres puertas y la celda de memoria.



sigmoid tanh pointwise multiplication pointwise addition vector concatenation

- Arquitectura simplificada en comparación con LSTM, que combina las puertas de entrada y de olvido en una única puerta de actualización;
- Sin celda de memoria separada;
- 2 puertas principales: la puerta de actualización (update gate) y la puerta de reinicio (reset gate):
 - La puerta de actualización equilibra la información pasada y presente en el estado oculto;
 - La puerta de reinicio determina qué información pasada olvidar.
- Más eficiente computacionalmente que LSTM.

- Captura información de contextos pasados y futuros al procesar la secuencia de entrada en ambas direcciones hacia adelante y hacia atrás simultáneamente;
- Proporciona una mejor comprensión de la secuencia de entrada al considerar tanto la información anterior como la siguiente, capturando las dependencias de largo alcance de manera efectiva;
- Mejora la precisión de la predicción, particularmente en tareas donde la información futura es crucial;
- Al incorporar información de ambas direcciones, las redes BiLSTM tienen la capacidad de capturar un contexto más amplio y pueden ser más efectivas.



Resultados

Los resultados sugieren que el modelo Bi-LSTM es más eficaz para predecir las tendencias de BTC, ETH y LTC que LSTM y GRU.

Currency	Model	RMSE	MAPE
BTC	LSTM	1031.3401	0.0394
	Bi-LSTM	1029.3617	0.0356
	GRU	1274.1706	0.0572
ETH	LSTM	148.5215	0.2971
	Bi-LSTM	83.9531	0.1243
	GRU	98.3136	0.1479
LTC	LSTM	9.6680	0.0636
	Bi-LSTM	8.0249	0.0411
	GRU	8.1224	0.0458

Métricas de rendimiento

Para evaluar el rendimiento de los esquemas de algoritmos de aprendizaje profundo propuestos, utilizamos el error cuadrático medio (RMSE) y el porcentaje de error absoluto medio (MAPE).

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (A_t - P_t)^2}{n}}$$

$$MAPE = \frac{100}{n} \times \sum_{t=1}^n \frac{|A_t - P_t|}{A_t}$$

Pt - valor previsto

At - valor real

n - el número de pasos de tiempo

Conclusiones

- LSTM, GRU y Bi-LSTM se utilizaron para predecir los precios de Bitcoin, Ethereum y Litecoin;
- El desempeño de los modelos se evaluó utilizando dos métricas: RMSE y MAPE;
- El modelo Bi-LSTM proporcionó las predicciones más precisas para las tres monedas, seguido del modelo GRU;

- La combinación de flujos hacia adelante y hacia atrás en modelos bidireccionales mejora el rendimiento;
- Los algoritmos de aprendizaje profundo son efectivos para predecir los precios de las criptomonedas.