



Predicción y Simulación del Precio de Bitcoin

Autores: Alberto Perea González, Marcos Dorado Muñoz

Curso: Ciclo de Especialización en Inteligencia Artificial y Big Data (CEIABD)

1. Introducción

Este proyecto representa el trabajo final del curso CEIABD, en el cual se ha desarrollado un flujo completo de análisis y modelado de datos. A lo largo de varios hitos, se ha construido un pipeline de ciencia de datos que abarca desde la generación de datos hasta la creación de un modelo predictivo, siguiendo una metodología clara y estructurada.

2. Justificación

El mercado de las criptomonedas, y en particular Bitcoin, ha ganado una relevancia significativa en los últimos años. La alta volatilidad y el interés creciente en estas

monedas digitales hacen que la predicción de sus precios sea un desafío atractivo y de gran utilidad para inversores y analistas financieros.

3. Objetivos

Desarrollar un pipeline completo de análisis de datos para la predicción del precio de Bitcoin.

Generar datos simulando un mercado ficticio pero realista.

Aplicar técnicas de limpieza y transformación de datos.

Implementar modelos de machine learning para la predicción de series temporales.

Evaluar y simular el comportamiento del modelo en escenarios realistas.

4. Metodología

El proyecto se ha dividido en cuatro hitos principales:

Hito 0: Presentación del problema, flujo, ELT y obtención de datos

- **Descripción:** Se define el problema a abordar y se establece el flujo de trabajo del proyecto. Se utiliza Kafka para producir datos ficticios que simulan el comportamiento de un mercado realista de Bitcoin. Estos datos se almacenan por separado en PostgreSQL y MongoDB, lo que permite un enfoque distribuido en la persistencia de la información.
- **Herramientas utilizadas:** Python, Apache Kafka, PostgreSQL, MongoDB.

Hito 1: Limpieza, preparación y análisis exploratorio de datos

- **Descripción:** Se lleva a cabo la limpieza de los datos generados, eliminando valores nulos y atípicos. Se realiza un análisis exploratorio para entender las características principales de los datos.
- **Técnicas aplicadas:** Estadísticas descriptivas, visualización de datos con matplotlib y seaborn.

Hito 2: Modelado predictivo

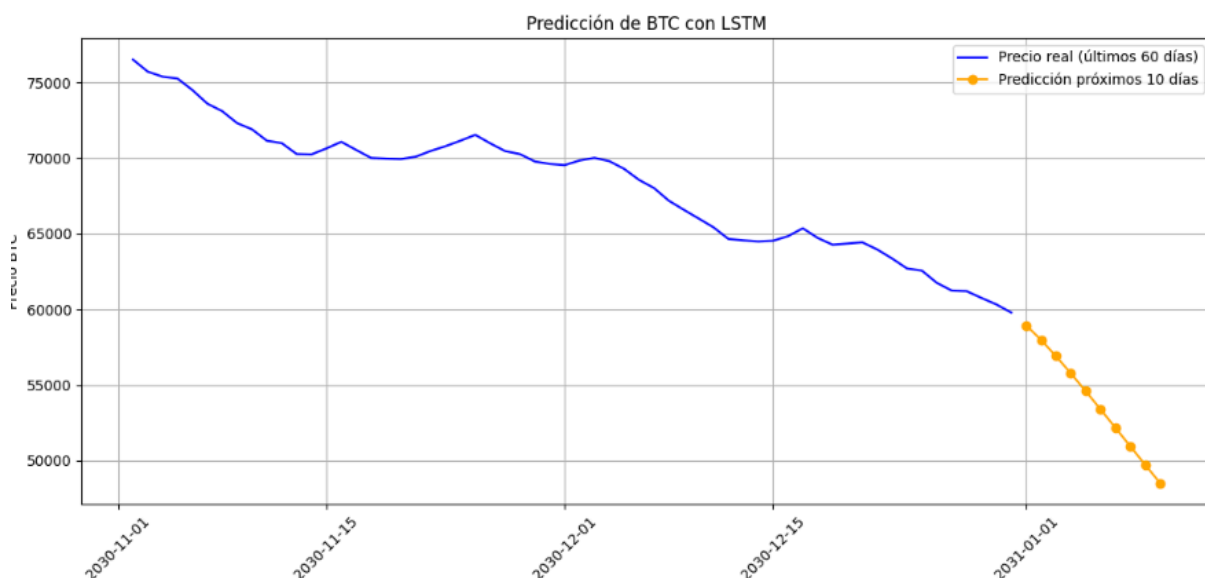
- **Descripción:** Se implementan modelos de machine learning para la predicción del precio de Bitcoin. Se prueban diferentes algoritmos y se selecciona el que ofrece mejores resultados.

- **Modelos utilizados:** Modelo de red neuronal, modelo LSTM, modelo de regresión lineal.

Hito 3: Implementación del modelo y despliegue en Hugging Face

- **Descripción:** En esta fase se implementa una interfaz en Gradio que permite la interacción con el modelo predictivo. El modelo se publica en la plataforma Hugging Face para facilitar su uso online. Se configura el entorno para que el modelo pueda descargarse, cargarse en memoria y realizar predicciones sobre secuencias de 30 días de precios simulados de Bitcoin.
- **Tecnologías utilizadas:** Gradio, Hugging Face.

5. Resultados



El modelo de predicción basado en redes neuronales LSTM predijo una tendencia descendente del precio de Bitcoin durante los próximos 10 días. En el gráfico presentado, se observa cómo los precios predichos continúan la trayectoria bajista mostrada por el precio real de los 60 días anteriores. Esta coherencia visual con la tendencia histórica indica que el modelo ha aprendido correctamente la dinámica temporal del mercado simulado. La predicción es suave y consistente, sin saltos bruscos, lo cual sugiere que el modelo no ha sobreajustado los datos.

6. Anexos

Código fuente de los notebooks:

- [Hito 0](#)
- [Hito 1](#)
- [Hito 2](#)
- [Hito 3](#)

Repositorio del proyecto: [Repositorio GitHub](#)