Project Proposal

Bitcoin Price Predictor

Data Mining – AI306

Team Members

Abubakar Waziri - 4220056 (Leader)

Mohammed Sattar - 4310129

Youssef ElNahas - 4311779

Supervisor: Dr. Mohammed Temraz

Problem

Cryptocurrencies, especially Bitcoin, have emerged as a popular form of investment in recent years. However, the extremely volatile nature of Bitcoin's price has made it difficult for both new and experienced investors to make informed decisions. The frequent and sometimes drastic fluctuations can lead to significant financial losses, causing many to lose confidence in the cryptocurrency market altogether.

The core problem we aim to address is the uncertainty and risk associated with Bitcoin price prediction. If we can build a model that reasonably forecasts future Bitcoin prices, it would provide valuable insights for investors, traders, and analysts. Such a tool could potentially reduce the impact of emotional or impulsive investment decisions and increase confidence in the market.

Input and Outcomes

Dataset: Bitcoin Historical Data

Our model will use a historical dataset of Bitcoin prices dating back to 2012, including daily open, close, high, low values, and trading volumes. This data captures key market trends and behaviors over time.

The final output will be the predicted future prices of Bitcoin, accompanied by visualizations that compare the actual and predicted values to better illustrate the model's performance.

Scope and Stakeholders

This project focuses on building and evaluating a short-term Bitcoin price prediction model using historical data. Our approach will involve experimenting with various prediction techniques, including traditional machine learning models like linear regression, as well as deep learning approaches such as Long Short-Term Memory (LSTM) networks and Transformer architectures, which are well-suited for handling sequential data.

Stakeholders:

- Cryptocurrency investors
- Financial analysts,
- Fintech researchers

Success Criteria and Constraints

Success Criteria:

- We will use Mean Squared Error (MSE) and Huber loss as the primary metric to evaluate model performance.
- The model should outperform a naïve baseline (e.g., assuming tomorrow’s price equals today’s).

Constraints:

- Data quality and availability
- Limited computing resources for training more complex models like deep learning.
- Difficulty in capturing external factors (like news events) which can influence price.