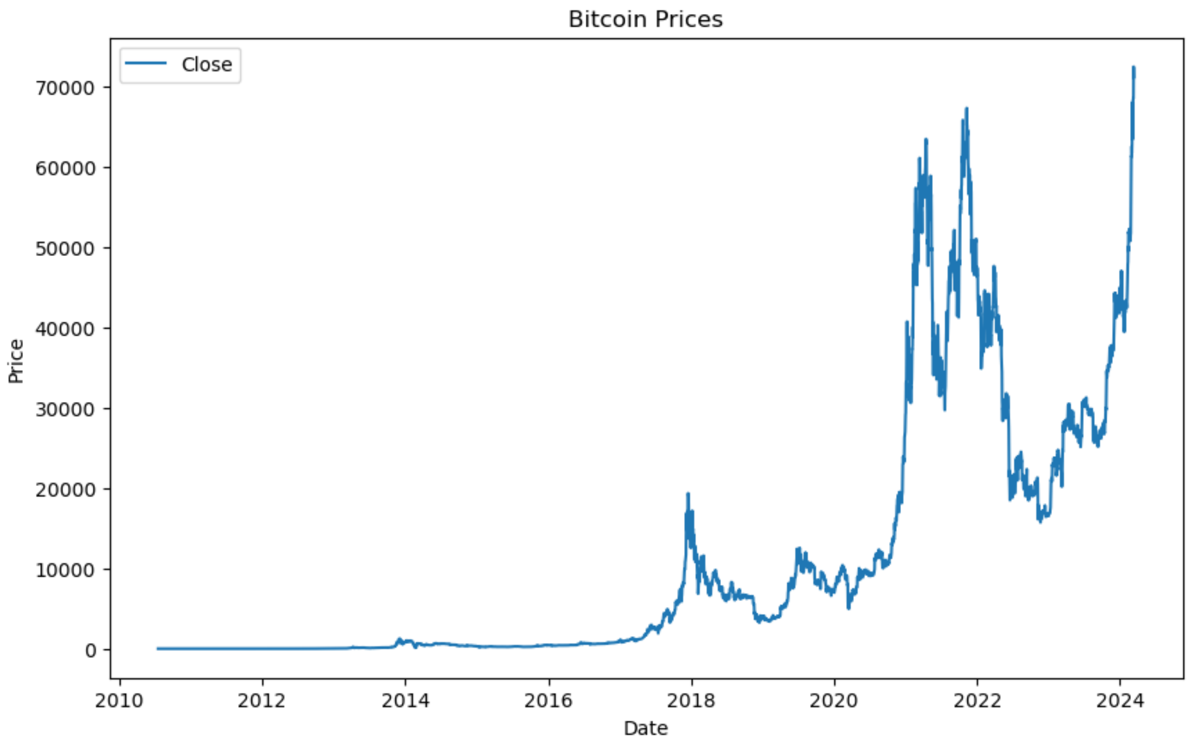
PROJECT 1: PREDICTING BITCOIN (BTC) PRICE

Introduction

The surge in cryptocurrency markets has sparked significant interest among investors, traders, and financial analysts. Among these, Bitcoin remains the most prominent and highly valued cryptocurrency. Predicting Bitcoin prices accurately is invaluable for making informed investment decisions and strategic trading. This project uses historical data to leverage the Long Short-Term Memory (LSTM) network, a robust type of recurrent neural network, to forecast Bitcoin prices.

Data Preparation

The foundation of our analysis is a robust dataset obtained from Kaggle using the Pandas library in Python, comprising daily recorded prices of Bitcoin, specifically the open, high, low, and closing prices (OHLC). For this analysis, we focused on the 'close' prices as they represent the final trading values for each day, which are critical for predicting future price movements.



We implemented several preprocessing steps to prepare this data for the LSTM model. The data was cleaned to fill in missing values and remove possible outliers. We then scaled the 'close' prices using MinMaxScaler to normalize the values between 0 and 1. This normalization is essential for neural network models, as it ensures that the gradient descent algorithm used in training the model operates efficiently and avoids numerical instability.

The LSTM Model

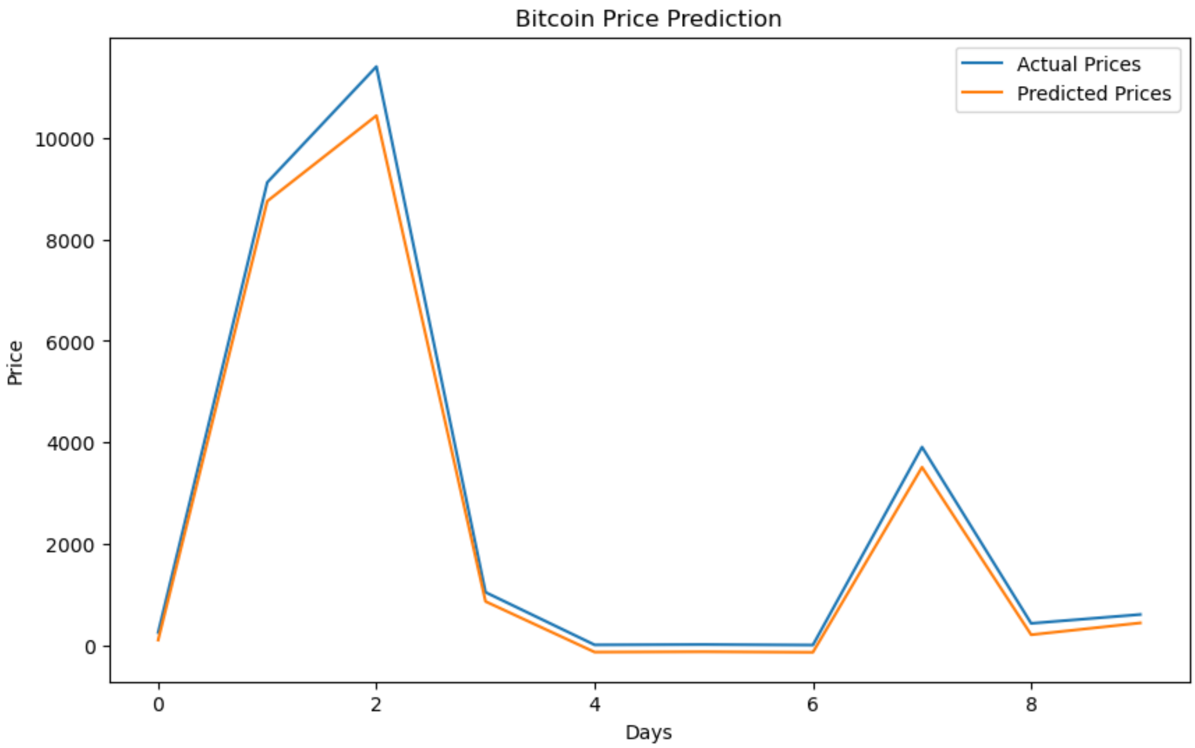
LSTM networks are an advanced type of recurrent neural network designed to handle sequence prediction problems with long input sequences. The critical advantage of LSTM over traditional neural networks is its ability to remember information for an extended period due to its internal mechanisms like gates and cell states. An LSTM unit includes three gates: the input gate, the forget gate, and the output gate, allowing the unit to remember or forget patterns selectively. This capability makes LSTMs particularly suitable for time-series data like stock prices, where past information is crucial for predicting future trends.

Application to Bitcoin Price Data

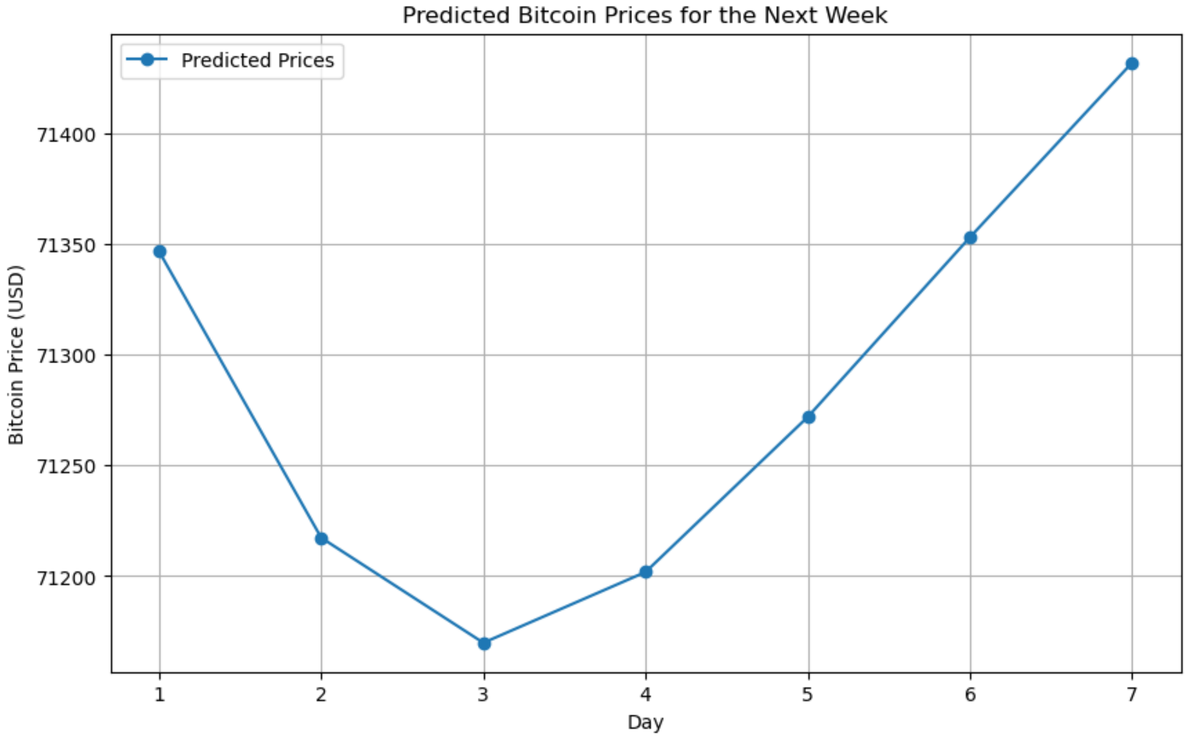
In this project, we constructed a sequential model using Keras and TensorFlow with two LSTM layers to capture the complexities in the historical price data of Bitcoin. Each LSTM layer was followed by a dropout layer, which helps prevent overfitting by randomly dropping units in the neural network during training. The final layer of the model is a dense layer that outputs the predicted price value. The model was trained on a dataset comprising several years of Bitcoin prices, using sequences of 60 consecutive days 'close' prices to predict the following day's price.

Results

After training, the model's performance was evaluated by comparing its predicted prices against the actual prices in the test dataset. The results were visually represented through plots that displayed the predicted prices closely following the actual price trends of Bitcoin.



This close alignment between the predicted and actual prices indicates that the LSTM model has successfully captured the underlying patterns and dynamics in the historical price data, making it a potent tool for forecasting Bitcoin prices.



Conclusion and Future Implications

The success of the LSTM model in predicting Bitcoin prices highlights its potential as a predictive tool in the financial technology sector. This model can be adapted to predict the prices of other cryptocurrencies, offering a broad array of applications for investors and financial analysts looking to gain insights into future market behaviours. As we look towards the next decade, using LSTM models could revolutionize the approach to trading and investment strategies in the cryptocurrency market, providing a competitive edge to those who leverage this technology.

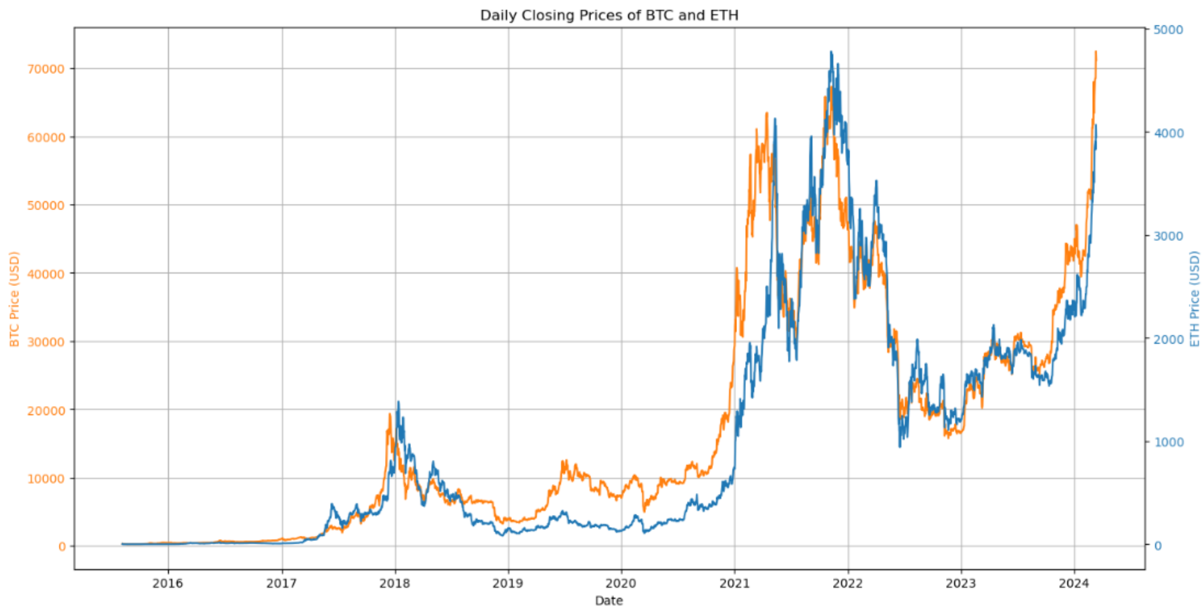
PROJECT 2: ANALYZING CORRELATIONS BETWEEN BITCOIN (BTC) AND ETHEREUM (ETH) PRICES

Introduction

Cryptocurrencies have revolutionized the financial landscape since Bitcoin's inception in 2009. As the first decentralized cryptocurrency, Bitcoin introduced a new era of digital currency away from central governmental control. Ethereum, launched in 2015, extended beyond the digital currency model by incorporating smart contracts, which are self-executing contracts with the terms of the agreement directly written into lines of code.

Bitcoin (BTC): Designed as a digital payment system, Bitcoin remains the most widely recognized and valued cryptocurrency. It operates on a blockchain-based decentralized network where transactions are verified by network nodes through cryptography and recorded in a publicly dispersed ledger called a blockchain.

Ethereum (ETH): While it shares blockchain technology with Bitcoin, Ethereum features an expansive development platform that enables developers to build and deploy decentralized applications (DApps) and smart contracts. Ethereum aims to function both as a digital currency and a platform for blockchain-based projects, making it a pivotal figure in the broader application of blockchain technology.



Methodology

The analysis utilizes daily closing price data for BTC and ETH, from Ethereum's inception in August 2015. This period captures a range of market behaviors, including bull and bear markets, regulatory changes, and technological advancements in the cryptocurrency sector.

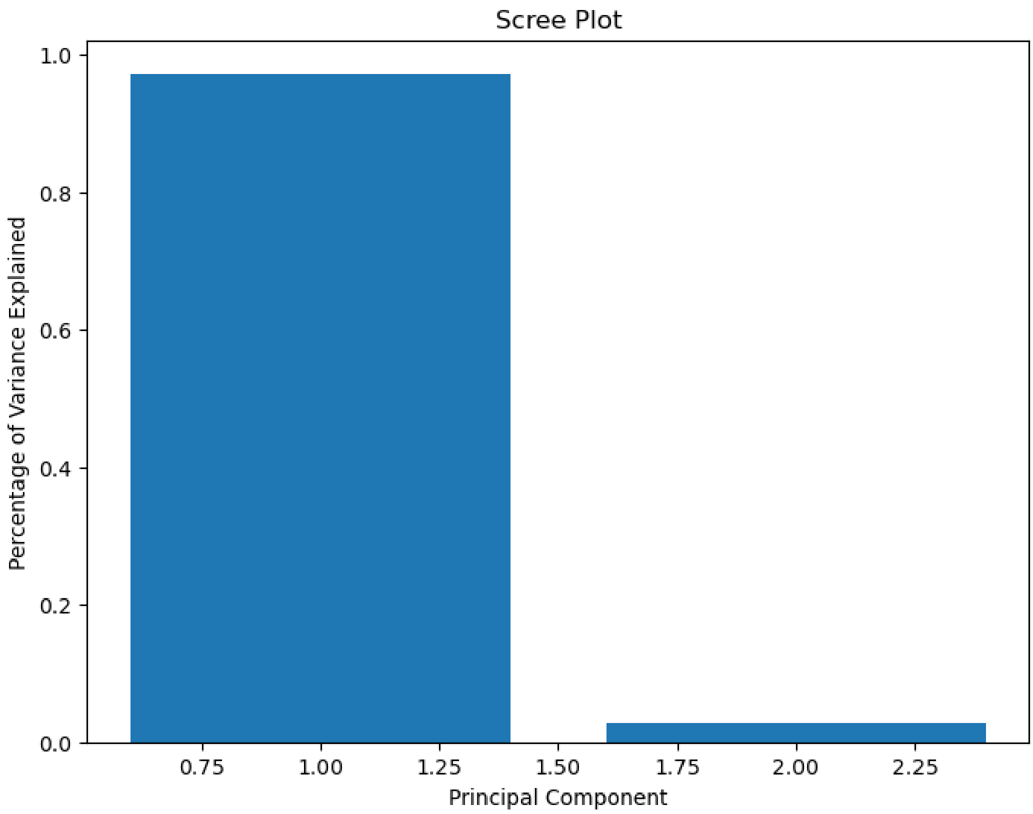
Principal Component Analysis (PCA) is a statistical technique that emphasizes variation and captures strong patterns in a dataset. It transforms the original data into linearly uncorrelated variables known as principal components (PCs). This analysis helps reduce the dimensionality of the data while retaining those characteristics of the dataset that contribute most to its variance, thereby simplifying the complexity of data but keeping the most valuable parts intact.

PCA Application

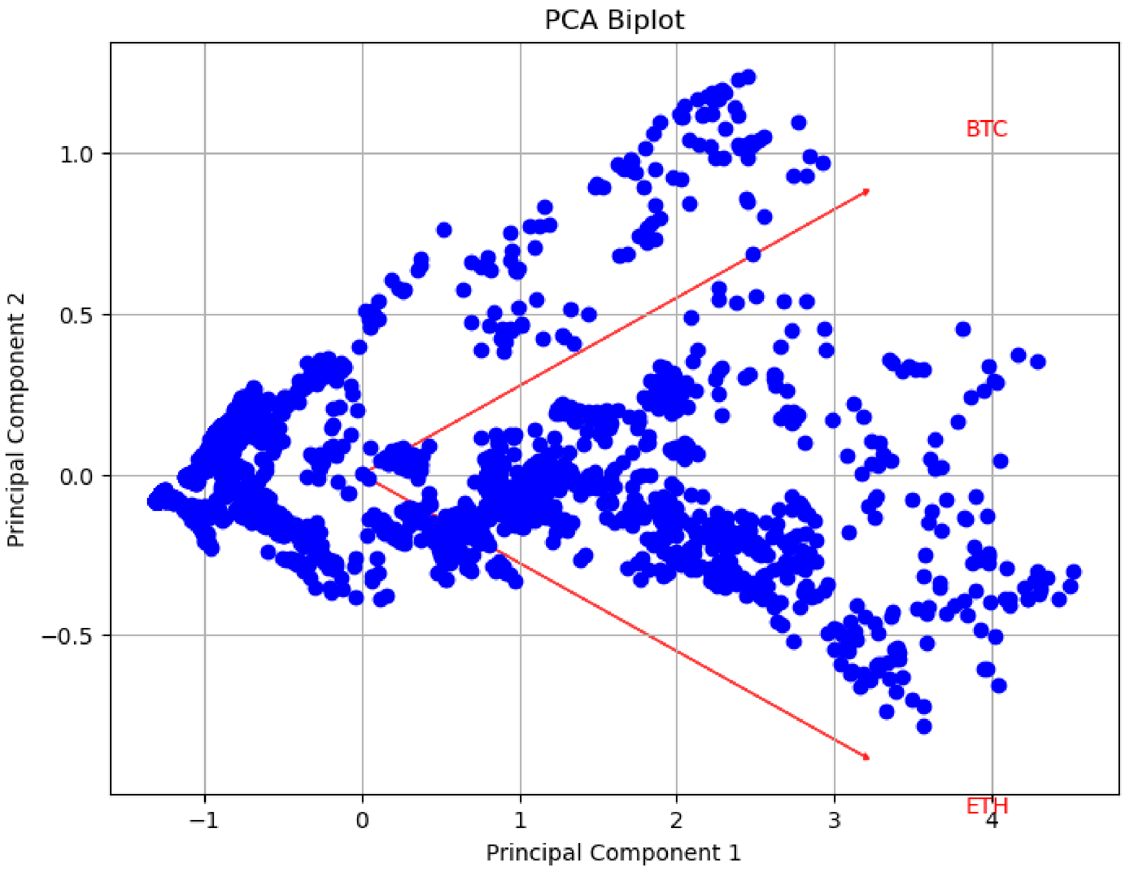
The data was first standardized using a Standard Scaler which normalizes the data to have a mean of zero and a standard deviation of one. PCA was applied to the standardized closing prices of BTC and ETH. The number of components was set to capture the maximum variance with the fewest components.

Results

The scree plot, a visual aid in determining the number of principal components to keep, indicated that the first component alone accounts for approximately 95% of the variance. This suggests a strong shared trend between BTC and ETH.



The biplot for BTC and ETH shows their daily price movements projected onto the first two principal components. The BTC and ETH vectors are separated by about 45 degrees, suggesting a moderate positive correlation. This indicates that while BTC and ETH prices often move together, they do retain distinct individual movements influenced by separate factors.



Conclusions

The PCA analysis confirms a significant but not perfect correlation between Bitcoin and Ethereum, which can be attributed to their shared market sentiments and investor behaviors but also highlights their unique influences. For investors, the implications are critical for diversification strategies; although BTC and ETH move together to some extent, their differences are substantial enough to warrant consideration for risk management.

References

Data source: https://www.kaggle.com/datasets/svaningelgem/crypto-currencies-daily-prices

Sample Project: https://towardsdatascience.com/cryptocurrency-price-prediction-using-lstms-tensorflow-for-hackers-part-iii-264fcdbccd3f

LSTM Wiki: https://en.wikipedia.org/wiki/Long_short-term_memory

BTC vs. ETH discussion: https://www.investopedia.com/articles/investing/031416/bitcoin-vs-ethereum-driven-different-purposes.asp