**Abstract**

This project explores the relationship between Bitcoin price movements and key macroeconomic indicators from 2017 to 2021. Macro indicators like inflation rates, GDP growth, interest rates, and stock market indices were examined in five major countries (the United States, China, Germany, the United Kingdom, and India) to establish their predictive capacity for Bitcoin volatility. Feature engineering process generated specialized features such as Bitcoin return rates, volatility indices, and rolling correlations with traditional markets. Three machine learning models - Support Vector Machine (SVM), Long Short-Term Memory (LSTM) networks, and XGBoost were implemented. The Models were evaluated on RMSE, MAE, MAPE, Directional Accuracy and model interpretation metrics i.e. feature importance score and SHAP analysis. With a directional accuracy of 93.5% and MAPE of 0.818%, the XGBoost model outperformed the SVM (directional accuracy: 52.5% and MAPE: 42.09%) and LSTM (directional accuracy: 50.71% and MAPE: 8.23%). SHAP value analysis revealed GDP growth rate as the most influential macroeconomic factor and that high GDP growth rate and low volatility index positively impact Bitcoin prices. With the growing importance of digital currencies in global financial markets these findings provide actionable insights for cryptocurrency investors and regulatory bodies.

# Introduction

Bitcoin, the most significant cryptocurrency, has grown from a specialized digital asset to a major financial instrument on a global level. With more investors looking to invest in crypto, there is growing interest in how macroeconomic indicators affect the price dynamics of Bitcoin. Unlike traditional stock markets, Bitcoin operates in a decentralized environment with minimal intrinsic worth, making its pricing highly sensitive to market sentiment and external economic fluctuations. The project's goal is to investigate how well traditional macroeconomic indicators can explain or forecast changes in Bitcoin prices. As a result, it contributes to the developing field of multidisciplinary research of finance, data science, and economics.

## Research Questions:

1. To what extent do macroeconomic indicators such as inflation, interest rates, and GDP growth impact Bitcoin price volatility?
2. Which of the machine learning models is the most applicable when predicting Bitcoin price trends from macroeconomic indicators?

## Aims and Objectives:

* **Aim:** The aim of this project is to assess the influence of macroeconomic indicators on the Bitcoin price dynamics and evaluate the predictive performance of selected machine learning models
* **Objectives:**
  + Collect and preprocess Bitcoin price data and selected macroeconomic indicators.
  + Conduct exploratory data analysis (EDA) to understand data relationships and detect anomalies.
  + Implement and compare three forecasting models—Support Vector Regression, LSTM and XGBoost.
  + Evaluate and analyse model’s performances to interpret the significance of macroeconomic indicators in forecasting Bitcoin prices.
  + Identify the most suitable machine learning model for predicting Bitcoin price movements.
  + Discuss the impact of the findings on financial predictions, investment strategies, and future study.

# Literature Comparison

The literature comparison is split into 2 sections *Macroeconomic Indicators and Bitcoin Price* and *Machine Learning Models for Bitcoin Forecasting.* This comparison aims to provides a critical overview of existing literature regarding the influence of macroeconomic indicators on Bitcoin price dynamics and evaluates the machine learning models used in that studies.

## Macroeconomic Indicators and Bitcoin Price

Bitcoin is a decentralized asset and multiple studies have explored its sensitivity to traditional macroeconomic variables. Bouri et al. (2017) used VAR and spillover index methods to study Bitcoin’s behaviour under global economic uncertainty, concluding that it exhibits volatility transmission patterns similar to traditional financial assets, particularly during crisis periods. This insight supports the inclusion of macroeconomic indicators like inflation and GDP as potential predictors of Bitcoin price in this project.

Corbet et al. (2018) studied the influence of geopolitical risk and economic policy uncertainty on Bitcoin. They found that Bitcoin is sensitive to macro announcements, with observable lags. This supports time-series approach LSTM used in this project that can capture temporal dependencies.

Ciaian et al. (2016) found that Bitcoin prices are influenced by both market-specific factors and macroeconomic variables such as CPI and exchange rates. This justifies combining both crypto-specific and macroeconomic variables in predictive modelling.

Similarly,Mokhberi et al. (2021) showed using cointegration tests that Bitcoin responds to long-run macroeconomic equilibrium relationships, especially in inflationary periods. This strengthens the hypothesis that inflation influence on Bitcoin price*.*

## Machine Learning Models for Bitcoin Forecasting

With the increasing non-linearity and volatility of Bitcoin prices, recent literature has shown a shift from traditional time-series methods to machine learning and deep learning techniques. McNally et al. (2018) conducted a comparative study of ARIMA and LSTM, showing that LSTM outperformed due to its ability to capture temporal dynamics, justification for its use in this project to model sequential dependencies in macro-financialdata.

Goodfellow et al. (2016) provide a foundational understanding of deep learning architectures like LSTM, highlighting their strength in modelling long-term dependencies and sequential structures. Their discussion reinforces the theoretical justification for using LSTM in this project, especially in the situation of time-series forecasting influenced by lagged macroeconomic effects.

Sebastião and Godinho (2021) used XGBoost and Random Forest to predict cryptocurrency prices based on both market and macroeconomic variables. Their findings revealed XGBoost's superior performance, attributed to its regularization techniques and robustness against overfitting. This aligns with the outcome of this project, where XGBoost provided strong performance and interpretability, making it suitable for use in practical financial applications.

Kristjanpoller and Minutolo (2021) explored SVR and neural networks to forecast cryptocurrency volatility. SVR was found effective with smaller datasets and fewer features.

Jang and Lee (2017) developed a Bayesian neural network model including blockchain and economic data to predict Bitcoin prices. They found data preprocessing and feature selection significantly influenced performance. Similar preprocessing logic was used to test ML methods.

The choice of models such as SVR, XGBoost, and LSTM was guided by their suitability for capturing non-linear relationships in financial time series, as discussed in statistical learning literature. *James et al. (2013)* emphasize that flexible non-parametric models can outperform linear models when the true relationship between predictors and response is complex and non-linear, which is characteristic of cryptocurrency markets.

# Dataset

## Dataset Description

## Data Preprocessing

The dataset was formed by merging two sources:

* Bitcoin price data
* Macroeconomic indicators data

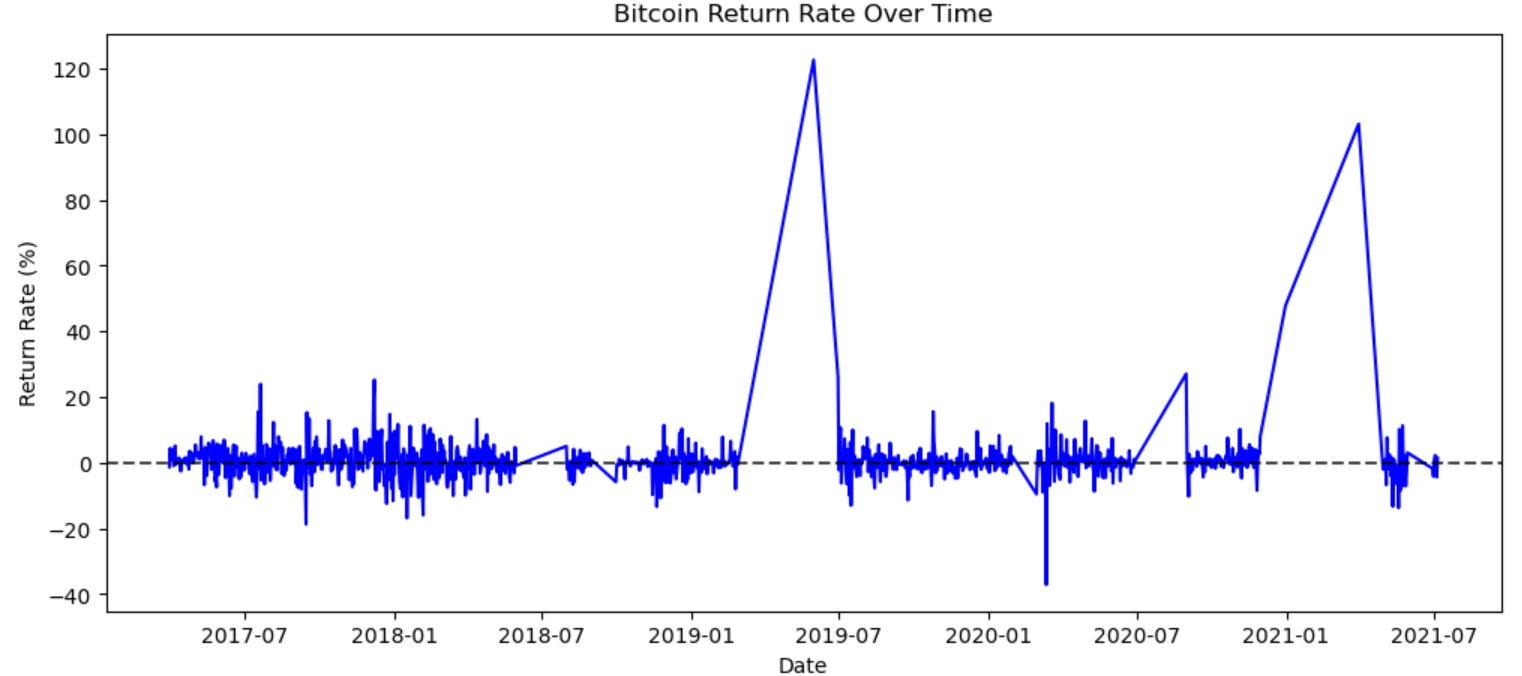
Both datasets were cleaned by:

* Converting the 'Date' column to proper datetime format.
* Filtering for data post-January 1st, 2017.
* Merging datasets on the 'Date' column.
* Selecting key countries (USA, China, Germany, UK, India) for focused analysis.
* Removing irrelevant columns (e.g., unemployment).
* Handling missing values using **forward fill** method to ensure sequential continuity.

## Exploratory Data Analysis (EDA)

### Bitcoin Return Rate

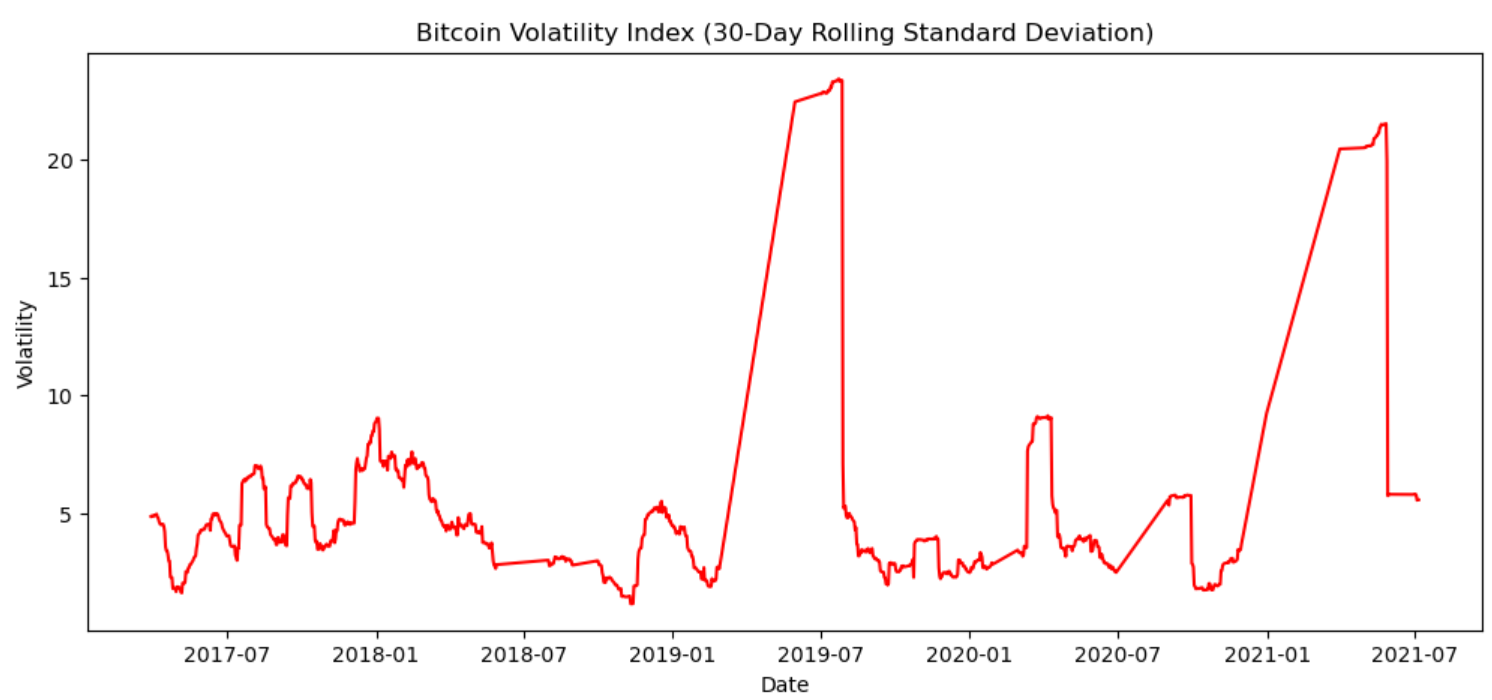
Bitcoin prices alone do not reflect market momentum. The return rate provides a clearer picture of price fluctuations.



* The blue line represents Bitcoin's daily percentage change in price.
* The horizontal dashed line at 0% serves as a reference, showing when Bitcoin’s returns were positive or negative.
* Several spikes indicate times when Bitcoin experienced major price swings.
* The return rate remains mostly stable, but certain extreme points reflect market-moving events.
* The 2019 peak may be linked to Bitcoin halving early DeFi growth, 2020 steep can link to Covid and 2021 peak can be driven by institutional investment like Tesla.
* Pre-2018, Bitcoin was in a more speculative and volatile phase but due to more institutional involvement post-2018, may have led to slightly more stability, though extreme fluctuations still occur.

### Volatility Index

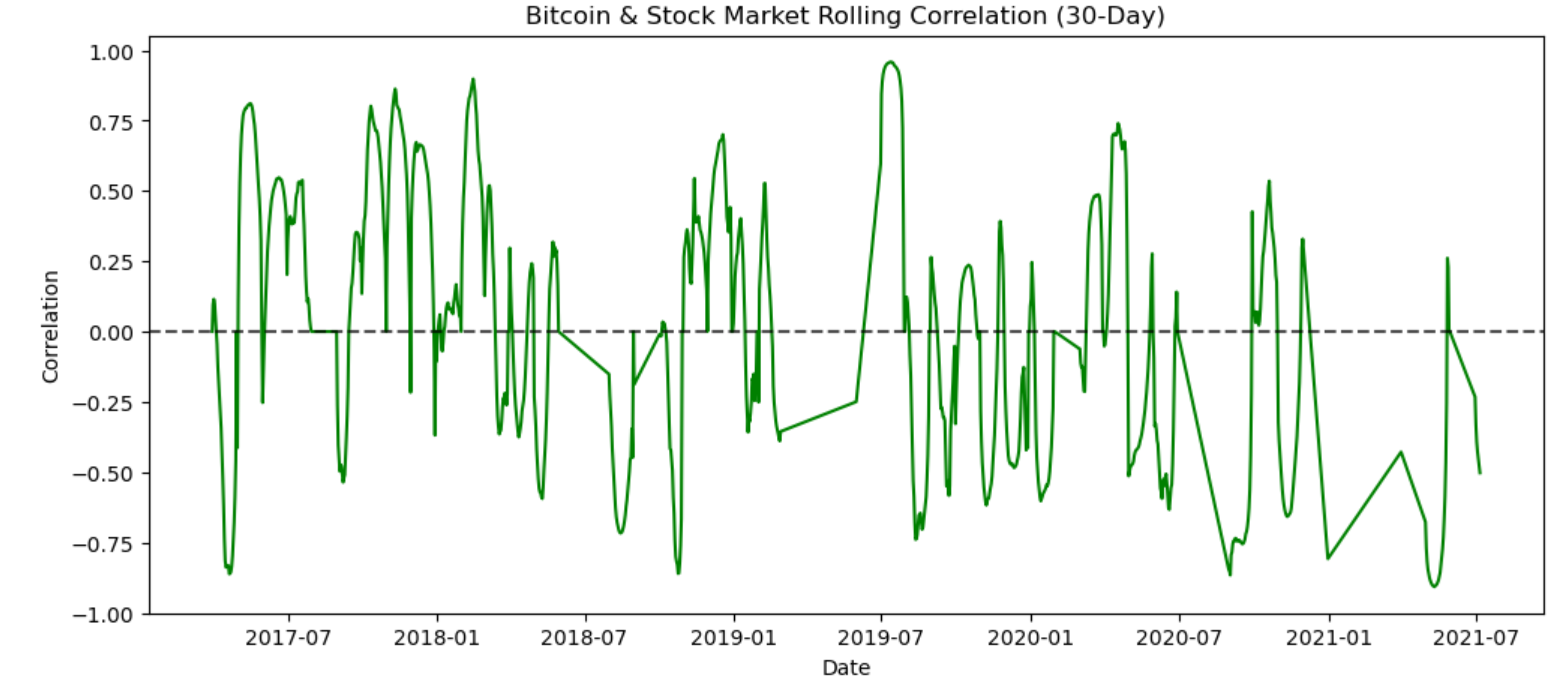
Bitcoin is extremely volatile, so understanding how it interacts with macroeconomic issues is critical.



* This plot represents 30-day rolling standard deviation, a key measure of Bitcoin’s price volatility.
* Volatility spikes indicate periods of uncertainty, often corresponding to major market booms or crashes and sustained low-volatility periods suggest price stabilization
* The two largest spikes (2019 & 2021) show how Bitcoin’s price surged and later corrected.

### Correlation Matrix

Traditional correlation analysis only gives a static relationship, but financial relationships evolve over time.



* This plot measures Bitcoin’s correlation with the stock market over a 30-day rolling window.
* The fluctuating correlation shows that Bitcoin does not consistently move with or against the stock market.
* At times, Bitcoin is highly correlated (~0.75), meaning it moves with stocks, while in other instances, it is negatively correlated (~-0.5), acting as a hedge.
* Bitcoin became highly correlated with equities during the COVID-19 crisis (2020-2021) and subsequent liquidity-driven bull run.
* The wild fluctuations in correlation post 2021 suggest that Bitcoin is still finding its identity as an asset class.

### Summary

* Return Rate Analysis shows Bitcoin’s price swings remain extreme, though potentially stabilizing.
* Volatility Index suggests periods of clustered uncertainty, reinforcing Bitcoin’s speculative nature.
* Stock Market Correlation is unpredictable, making Bitcoin’s role uncertain but evolving.

# Ethical Considerations

# Methodology

This section summaries the technical approach I undertook to investigate the influence of macroeconomic indicators on Bitcoin price dynamics and to compare the predictive performance of three machine learning models: Support Vector Machine (SVM), Long Short-Term Memory (LSTM), and Extreme Gradient Boosting (XGBoost).

## Support Vector Machine (SVM) Methodology

I applied Support Vector Machines (SVM) with the Radial Basis Function (RBF) kernel to predict the Bitcoin closing price. My approach involved detailed data preparation, model optimization via cross-validation, and evaluation through multiple performance metrics.

**Step 1: Data Preparation and Feature Selection:**

I began by loading the processed cryptocurrency and macroeconomic data from a CSV file. These features were used as independent variables (X), and the Bitcoin closing price (Close) was chosen as the target variable (y). The data was then split into training (80%) and testing (20%) sets using train\_test\_split, ensuring no shuffling of the data (shuffle=False) to maintain the sequential structure.

**Step 2: Data Scaling:**

I used StandardScaler from sklearn to standardize the features, applying it to both the training and test sets. This is important for SVM models, as they are sensitive to the scale of the input data.

**Step 3: Model Optimization:**

To optimize the performance of the SVM model, I conducted a hyperparameter search using GridSearchCV. The grid search explored combinations of three key hyperparameters:

* C (Regularization parameter): I tested values: [10, 50, 100, 500, 1000].
* Gamma (Kernel coefficient): I tested values: ['scale', 0.01, 0.1, 1].
* Epsilon (Margin of tolerance): I tested values: [0.01, 0.1, 0.5, 1].

Using 5-fold cross-validation, I trained the SVM model for all combinations of these parameters. After fitting the grid search, I retrieved the best hyperparameters and used them to train the final SVM model.

**Step 4: Model Training**

I used the Radial Basis Function (RBF) kernel, which is effective for non-linear relationships. After training the model on the scaled training data, I proceeded to make predictions on the test data.

**Step 5: Model Evaluation**

I evaluated the performance of the trained SVM model using several metrics:

* Root Mean Squared Error (RMSE): Measures the average magnitude of error, with greater penalty for larger errors.
* Mean Absolute Error (MAE): Provides the average absolute difference between predicted and actual values.
* Mean Absolute Percentage Error (MAPE): Gives a percentage-based metric that helps assess prediction accuracy in terms of relative error.
* Additionally, I computed the Directional Accuracy which indicates how often the model correctly predicted whether the price increased or decreased. This was done by comparing the sign of the actual and predicted changes in the Bitcoin closing price.

## Long Short-Term Memory (LSTM)

**Step 1: Data Preparation and Feature Selection:**

**Step 2: Data Scaling:**

I normalized both the independent variables (features) and the target variable using the MinMaxScaler to scale the data between 0 and 1.

**Step 3: Time-Series Data Transformation:**

I transformed the dataset into a time-series format suitable for LSTM by creating sequences of 30 previous time steps (i.e., using data from the past 30 days to predict the next day) to capture the temporal dependencies in the data.

The resulting data was split into training (80%) and testing (20%) sets

**Step 4: LSTM Model Architecture:**

I used the Keras Sequential API to define a deep learning architecture. The model architecture consists of:

1. LSTM layers: These layers are the core of the model, designed to capture temporal patterns in the data. The first LSTM layer outputs sequences to feed into the next LSTM layer.
2. Dropout layers: These layers help prevent overfitting by randomly setting a fraction of input units to 0 during training.
3. Dense layer: The final dense layer outputs a single value, which is the predicted Bitcoin closing price.

**Step 5: Model Optimization**

To optimize the model, I used Keras Tuner with the Bayesian Optimization strategy, which is an efficient method for hyperparameter tuning. The hyperparameters I tuned included:

* LSTM units: The number of units in the LSTM layers (ranging from 32 to 128).
* Dropout rate: The dropout rate in each dropout layer (ranging from 0.2 to 0.5).
* Learning rate: The learning rate for the Adam optimizer (choosing from 0.0001, 0.001, and 0.01).
* Batch size: I tested batch sizes of 16, 32, and 64.

**Step 6: Model Training**

The LSTM model was trained using the best hyperparameters obtained from the tuning process. The training included 50 epochs with a batch size of 32, using the training and testing data split earlier. The loss function was Mean Squared Error (MSE), and the optimizer was Adam.

**Step 7: Model Evaluation**

Extreme Gradient Boosting (XGBoost)

**Step 1: Data Preparation and Feature Selection:**

**Step 2: Data Scaling:**

**Step 4: Model Optimization**

**Step 5: Model Training**

**Step 6: Model Evaluation**

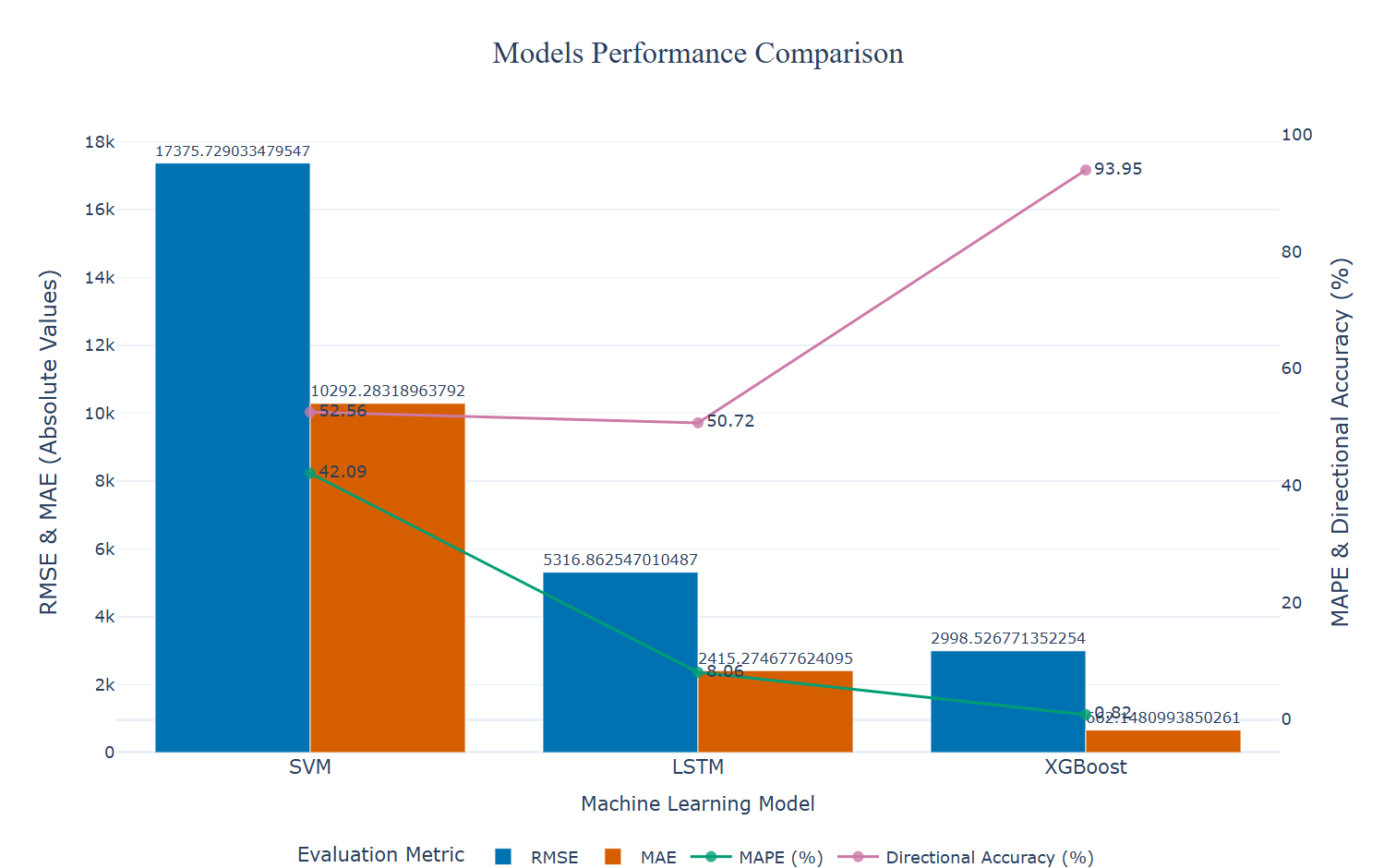
# Results

## Metrics Used:

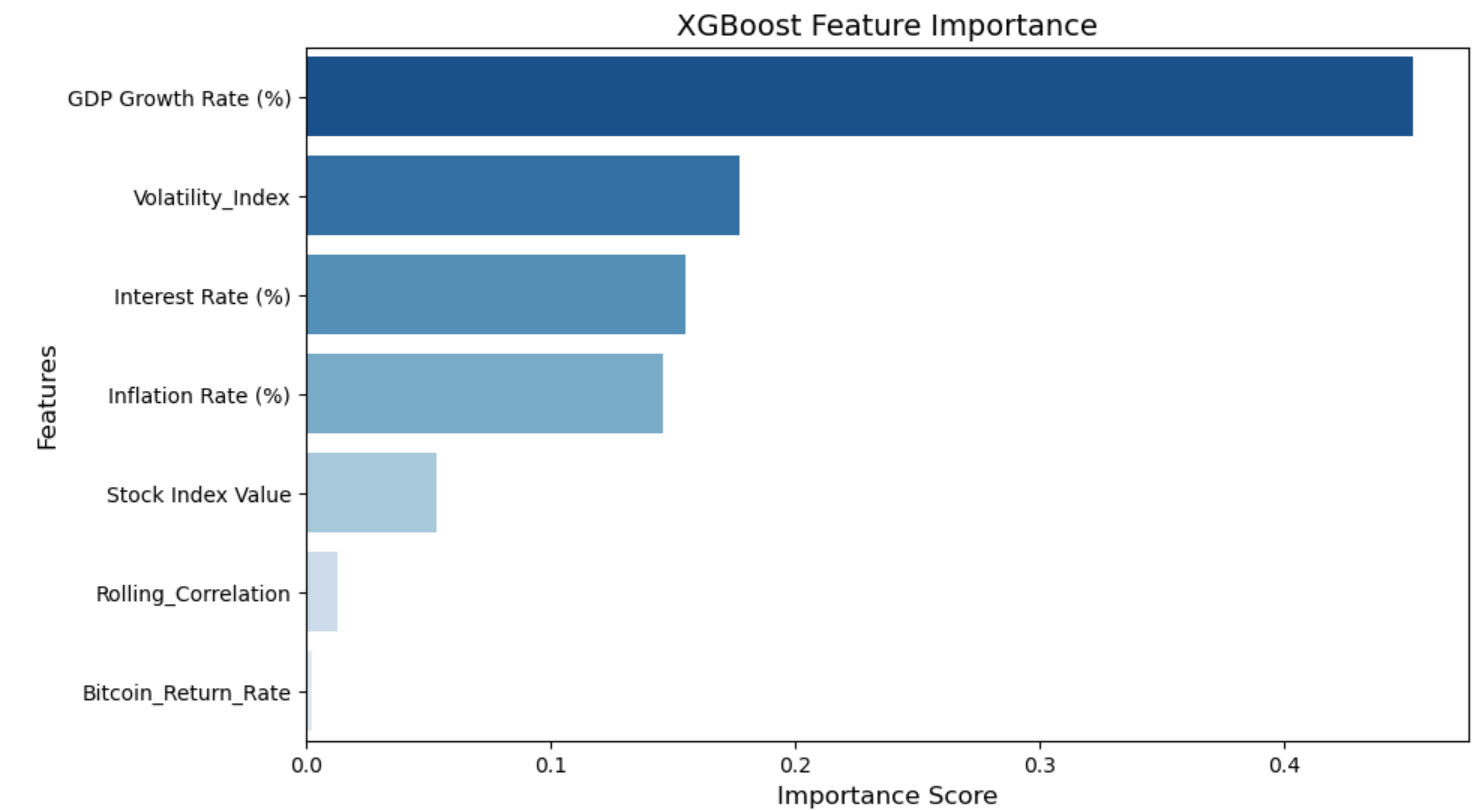
* Error-Based Metrics
  + RMSE
  + MAE
  + MAPE
* Model Interpretation Metrics:
  + Feature Importance Score
  + SHAP Analysis
* Time-Specific Metrics:
  + Auto Correlation Function
  + Partial Autocorrelation Function
  + Directional Accuracy

## Models Comparison

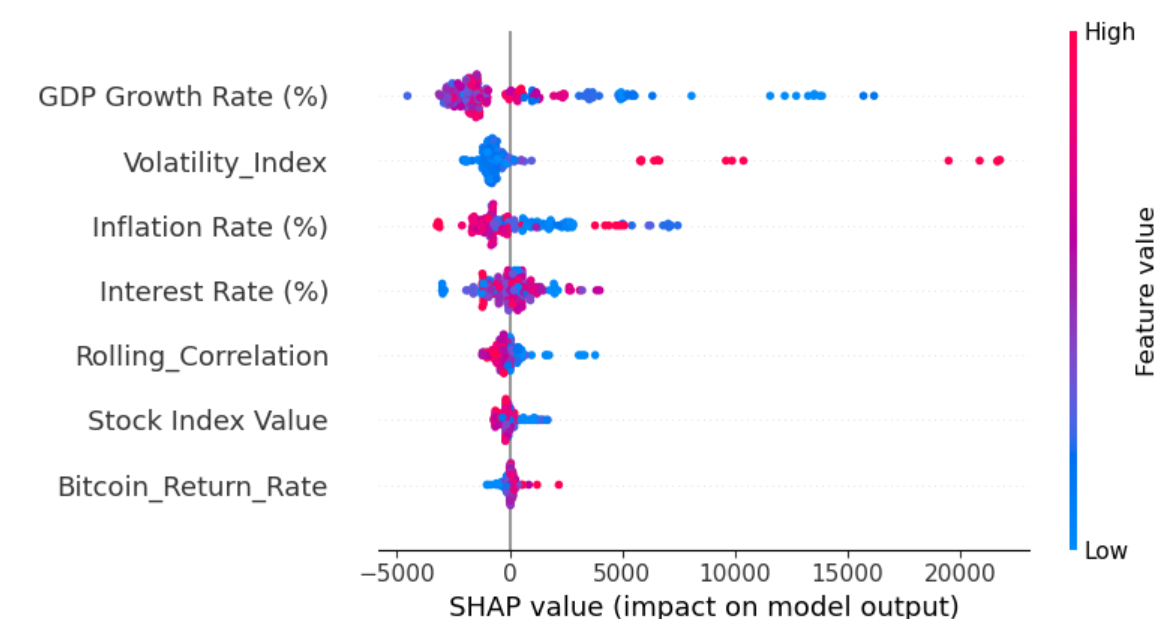
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | RMSE | MAE | MAPE | Directional Accuracy |
| SVM | 17,375.73 | 10,292.28 | 42.09% | 52.56% |
| LSTM | 5,316.86 | 2,415.27 | 8.06% | 50.72% |
| XGBoost | **2,998.53** | **662.15** | **0.82%** | **93.95%** |



Model Interpretation Metrics

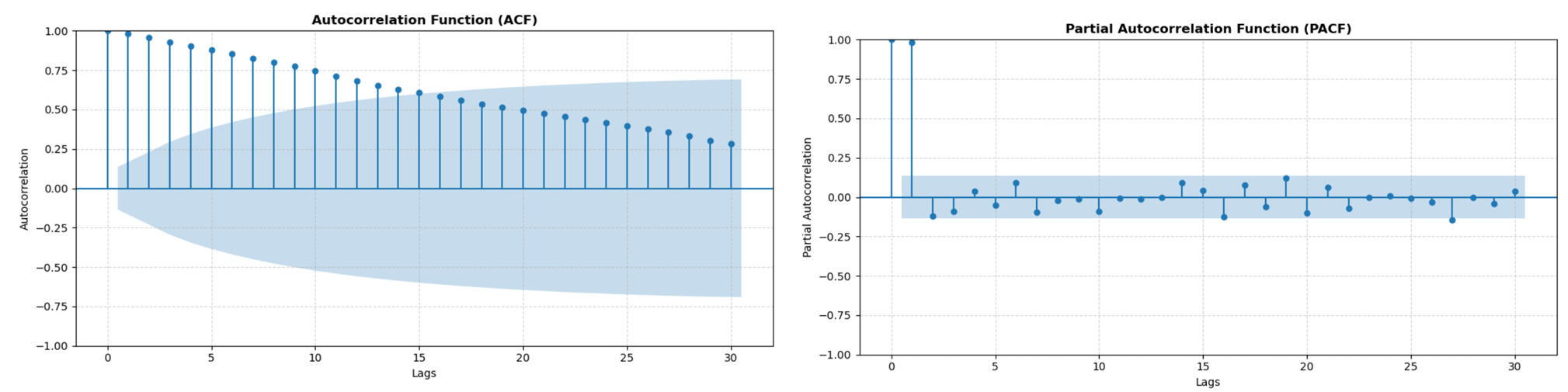


* GDP Growth Rate (%) is the most important feature, indicating that Bitcoin’s price is heavily influenced by macro-economic growth.
* Volatility Index plays a crucial role, showing that market uncertainty directly impacts Bitcoin’s movements.
* Interest Rate (%) and Inflation Rate (%) are significant factors, reinforcing Bitcoin’s role as a hedge against inflation.
* Stock Index Value has moderate importance, suggesting that Bitcoin is partially correlated with equity markets.
* Rolling\_Correlation and Bitcoin\_Return\_Rate are minor contributors, meaning Bitcoin’s past performance is less relevant than macroeconomic trends.



* The SHAP plot shows how each feature impacts the model’s output.
* High GDP Growth Rate and low Volatility Index positively impact Bitcoin prices.
* High inflation and high interest rates also push Bitcoin prices higher (as expected for an inflation hedge).
* Higher stock index values and lower rolling correlations suggest less reliance on past trends.

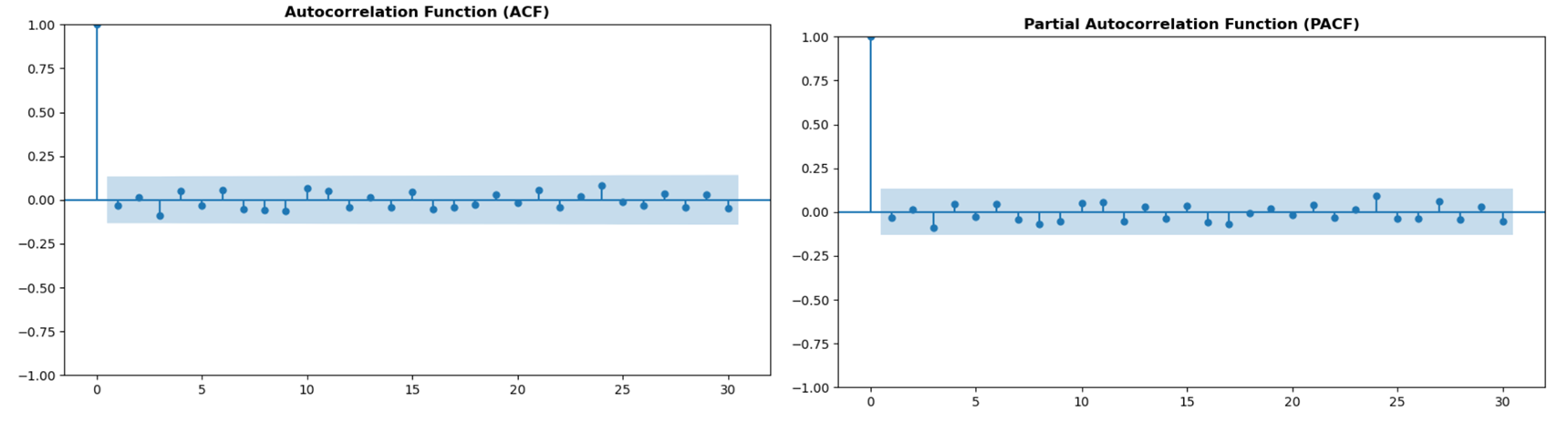
## Time-Specific Metrics

ACF Plot:

Shows that Bitcoin prices have a gradual autocorrelation decay, indicating that past values influence future prices over time. The slow decline suggests Bitcoin prices are not purely random and do exhibit time-dependent relationships.

PACF Plot:

The first few lags have significant values, meaning short-term price dependencies exist. However, after the first 2-3 lags, values diminish, indicating that long-term dependencies are weaker



ACF Plot:

The ACF plot shows no strong autocorrelation over time. Bitcoin prices do not exhibit strong long-term dependencies. Unlike traditional financial assets, Bitcoin movements are highly volatile and less dependent on past values.

PACF Plot:

The PACF plot indicates that only the first lag (short-term price movements) has a significant correlation. This suggests that recent price action has a stronger effect on the next movement, but deeper historical data is not as influential.

# Analysis and Discussion

1. Conclusion

This project investigated the impact of macroeconomic indicators—such as inflation, interest rates, and GDP growth—on Bitcoin price trends, utilizing various machine learning models including SVM, LSTM, and XGBoost. The results indicate that GDP Growth Rate (%) is the most important feature, indicating that Bitcoin’s price is heavily influenced by macro-economic growth. Volatility Index plays a crucial role, showing that market uncertainty directly impacts Bitcoin’s movements. Stock Index Value has moderate importance, suggesting that Bitcoin is partially correlated with equity markets.

Among the models tested, XGBoost outperformed the others in terms of accuracy, robustness, and practical applicability. Its ability to model non-linear relationships and handle multicollinearity proved particularly valuable for financial data. LSTM demonstrated strong potential in capturing temporal dependencies, but it came with higher computational costs and lower interpretability. SVM, although straightforward, lagged in performance due to its limitations in dealing with the complexity of the data.

The key real-world application of this research lies in its potential to help investors and financial analysts in making data-driven decisions, especially in volatile or uncertain economic environments. It also has utility for regulators and economists aiming to understand how digital assets respond to shifts in the global macroeconomy.

Future work could expand in several directions. Incorporating geopolitical risk indicators, social media sentiment, and regulatory events could improve the model’s predictive power. Finally, testing the model across different cryptocurrencies or applying transfer learning could generalize the findings to the broader crypto market.