NEVŞEHİR HACI BEKTAŞ VELİ ÜNİVERSİTESİ
Ara, bul.

Kapadokya Akademik Bakış Dergisi
Cilt: 9 Sayı: 1 Periyod
E-ISSN: 2602-3474

MACHINE LEARNING AND QUALITATIVE VARIABLES IN BITCOIN PRICE PREDICTION: AN EMPIRICAL EVALUATION OF MODEL PERFORMANCE

BİTCOİN FİYAT TAHMİNİNDE MAKİNE ÖĞRENİMİ VE NİTEL DEĞİŞKENLER: MODEL PERFORMANSININ UYGULAMALI DEĞERLENDİRMESİ

Dr. Öğr. Üyesi Gökhan SEÇME
Nevşehir Hacı Bektaş Veli Üniversitesi
gsecme@nevsehir.edu.tr
ORCID No: 0000-0002-7098-1583

Geliş Tarihi:
12.06.2025

Kabul Tarihi:
29.06.2025

Yayın Tarihi:
30.06.2025

Anahtar Kelimeler
Tahmin, Bitcoin fiyat tahmini, Makine öğrenmesi, Nitel değişkenler

Keywords
Forecasting, Bitcoin price forecasting, Machine learning, Qualitative variables

ÖZ

Bu çalışmada tahmin yöntemlerinin performans değerlendirmesi için Bitcoin fiyat tahmini problem kullanılmıştır. Bitcoin fiyat tahmini için geleneksel doğrusal regresyon yöntemi ile makine öğrenmesi uygulamalarından neural net fitting ve neural net time series yöntemleri karşılaştırılmıştır. Ayrıca Bitcoin fiyatının volatilitesinin yüksek olması ve sosyal, politik ve davranışsal olayların etkili olması sebebiyle kalitatif faktörlerin de tahmin performansı üzerindeki etkileri incelenmiştir. Bu bağlamda Bitcoin fiyat tahmininde kullanılan kantitatif değişkenlere ilave olarak kalitatif değişkenler olarak korku ve açgözlülük indeksi ve duyarlılık indeks değerleri de kullanılarak oluşturulan modellerin tahmini yapılmıştır. Elde edilen sonuçlara göre, geleneksel çoklu doğrusal regresyon yönteminin tahmin performansının oldukça zayıf olduğu, neural net fitting metodunun ise göreli olarak daha başarılı tahminler yapabildiği görülmüştür. Ayrıca, tek başına kantitatif değişkenlerin kullanıldığı modelin çalışmada kullanılan tüm tahmin yöntemlerinde zayıf performans sergilediği görülmüştür. Kalitatif değişkenlerin kullanıldığı tahmin modelinin ise tüm yöntemlerde en başarılı tahmin sonuçları ürettiği görülmüştür. Dolayısıyla Bitcoin fiyat tahmininde kalitatif değişkenlerin kullanılmasının tahmin performansını arttırdığı, bununla birlikte neural net fitting tahmin yönteminin düşük MSE değerleri ile en başarılı tahmin yönetimi olduğu belirlenmiştir.

ABSTRACT

This study employs the problem of Bitcoin price prediction to evaluate the performance of forecasting methods. Traditional linear regression is compared with machine learning techniques, specifically neural net fitting and neural net time series, to assess their predictive accuracy. Given Bitcoin's high volatility and susceptibility to social, political, and behavioral influences, the study also examines the impact of qualitative factors on prediction performance. In addition to quantitative variables, qualitative variables—such as the Fear and Greed Index and sentiment analysis metrics—are incorporated into the models to enhance forecasting robustness.
The results indicate that traditional multiple linear regression yields relatively weak predictive performance, whereas neural net fitting demonstrates superior accuracy. Furthermore, models relying solely on quantitative variables underperform across all tested methods. In contrast, the inclusion of qualitative variables significantly improves prediction outcomes in all approaches. The study concludes that integrating qualitative variables enhances Bitcoin price forecasting accuracy, with neural net fitting emerging as the most effective method due to its lower mean squared error (MSE) values.

DOI: https://doi.org/ 10.69851/car. 1718027
Atıf/Cite as: Seçme, G. (2025). Machine Learning and Qualitative Variables in Bitcoin Price Prediction: An Empirical Evaluation of Model Performance. *Kapadokya Akademik Bakış Dergisi,* 9(1), 18-30.

1. Introduction

Cryptocurrencies represent a paradigm shift in financial systems, introducing decentralized digital assets that operate independently of traditional banking institutions. The concept of digital currency dates back to cryptographic proposals in the 1980s and 1990s (Chaum, 1983), but it was not until the 2008 financial crisis that Bitcoin (BTC)—the first fully functional cryptocurrency—was introduced by an anonymous entity under the pseudonym Satoshi Nakamoto (Nakamoto, 2008). Bitcoin's underlying technology, blockchain, ensures transparency, immutability, and resistance to censorship through a distributed ledger system. Unlike fiat currencies, Bitcoin has a fixed supply cap of 21 million coins, making it inherently deflationary and attractive as a hedge against inflation and monetary debasement (Böhme et al., 2015).

Bitcoin's emergence marked the beginning of a new asset class, with thousands of alternative cryptocurrencies (altcoins) such as Ethereum, Ripple, and Litecoin following in its footsteps. Despite the proliferation of altcoins, Bitcoin remains the dominant cryptocurrency, accounting for approximately 40-50% of the total crypto market capitalization (CoinMarketCap, 2023). Bitcoin has several advantages. Being a first mover in the market led to the Widest adoption and recognition as "digital gold". Growing interest from hedge funds, corporations (e.g., Tesla and Microstrategy), and ETFs induced institutional investment. Additionally, the largest mining power (hash rate) and liquidity among cryptocurrencies built a network effect that improves the prominence of Bitcoin.

Nowadays, as Bitcoin has become an investment tool, price fluctuations have become very important for decision makers. Bitcoin has experienced periods of extreme volatility, including annual returns exceeding 500% (e.g., 2017) and drawdowns exceeding 80% (e.g., 2018) (CoinMarketCap, 2018; Corbet et al., 2018). This volatility has shown that institutional investors in particular need forecasts that will optimize their hedging strategies and portfolios (Baur & Dimpfl, 2021). Centralized exchanges (e.g., Binance) and lenders (e.g., Celsius) employ price forecasts to dynamically adjust margin requirements and prevent liquidity shortfalls (Baur & Dimpfl, 2021; Cong et al., 2021). "The 2022 Celsius collapse demonstrated the systemic risks of flawed price modeling (In re Celsius, 2022)."

Regulatory bodies (e.g., SEC, FSB) study price dynamics to assess systemic risks in crypto markets (Financial Stability Board, 2022). Bitcoin's growing correlation with traditional assets (e.g., S&P 500) means its price swings can impact broader financial stability (Corbet et al., 2020). Central banks monitor crypto markets to design CBDCs (Central Bank Digital Currencies) and anti-inflationary policies.

Price trends influence miner profitability (via hash rate adjustments) and blockchain security. Startups and developers allocate resources based on market sentiment (e.g., DeFi/NFT booms in bull markets) (Hayes, 2019; Biais et. al. 2023).

All these factors make it critical to accurately forecast Bitcoin's price for multiple stakeholders due to its economic, financial, and technological implications from the perspective of investor decision making, risk management, macroeconomic policy, and technological adoption.

Given Bitcoin's high volatility—driven by speculative trading, regulatory news, macroeconomic trends, and technological developments—accurate price forecasting remains a complex yet valuable endeavor. Traditional financial models, such as ARIMA and GARCH, have been applied to Bitcoin with limited success due to its non-linear and regime-switching behavior (Katsiampa, 2017). Hybrid drivers, such as quantitative factors (e.g., hash rate, liquidity) and qualitative sentiment (e.g., Elon Musk's tweets, FGI extremes) are both respond to Prices More advanced techniques, including machine learning and deep learning, have shown promise in capturing these dynamics (Jang & Lee, 2017).

To forecast Bitcoin price, early studies focused on quantitative factors, including Market capitalization (an indicator of adoption and liquidity), Trading volume (a measure of market activity), Hash rate (proxy for network security and miner confidence), and On-chain transactions (reflecting utility and adoption). However, Bitcoin's price is also heavily influenced by behavioral and sentiment-driven factors. The Crypto Fear and Greed Index (FGI) is a composite metric derived from volatility, social media, surveys, and market momentum (often used to gauge extreme market conditions). Also, the Social sentiment, derived from the Analysis of Twitter, Reddit, and news, has been shown to precede price movements (Shen et al., 2019).

This study contributes to the literature by evaluating the predictive power of both quantitative and qualitative variables in Bitcoin price forecasting. Specifically, we employ Market capitalization, total trading volume, number of transactions, and hash rate as quantitative variables. Qualitative variables are the Crypto Fear and Greed Index (FGI) and social sentiment scores.

In this study, Linear Regression (baseline econometric model), Neural Network Fitting (multilayer perceptron for non-linear patterns), and Neural Network Time Series (LSTM/GRU for sequential dependency) are used as three modeling approaches.

This paper is structured as follows: Section 2 reviews related literature, Section 3 describes the methodology, Section 4 presents empirical results, and Section 5 concludes with implications for investors and future research.

2. Literature Review

Bitcoin (BTC), as the most prominent cryptocurrency, has attracted considerable academic interest due to its high volatility, decentralized nature, and global market influence. Forecasting its price dynamics is critical for investors, policymakers, and researchers. Time series forecasting models offer a structured approach to analyzing the temporal evolution of Bitcoin prices, capturing both short-term volatility and long-term trends. These models range from classical statistical approaches to sophisticated machine learning and deep learning methods. In addition to market fundamentals, the influence of social and behavioral factors has become increasingly significant, prompting researchers to integrate qualitative indicators such as sentiment into forecasting frameworks.

In the study by Songur and Ordu (2023), the relationship between Bitcoin-related news and Bitcoin price and returns was examined using causality analysis. The results indicated a causal relationship between Bitcoin-related news and Bitcoin prices. It was observed that during periods of rising Bitcoin prices, Bitcoin-related news also increased. The study concluded that news has a significant impact on Bitcoin prices.

In the study by Teker et al. (2020), the impact of news about Bitcoin and cryptocurrencies on Bitcoin's daily closing price, intraday highest price level, and daily trading volumes was investigated. Using data and news from the period between May and December 2018, the study found that positive and negative news about Bitcoin and cryptocurrencies did not lead to any significant differentiation in Bitcoin prices or trading volumes.

Statistical models such as ARIMA (AutoRegressive Integrated Moving Average) and GARCH (Generalized Autoregressive Conditional Heteroskedasticity) have traditionally been used in financial time series forecasting. Chu et al. (2015) applied various GARCH-family models, including EGARCH and GJR-GARCH, to analyze the volatility of major cryptocurrencies, demonstrating their ability to model volatility clustering. Bouoiyour and Selmi (2015) utilized ARIMA models to examine the stochastic behavior of Bitcoin, highlighting the limitations of linear models in capturing nonlinear dynamics inherent in cryptocurrency markets.

Kristjanpoller and Minutolo (2018) extended the analysis by comparing different GARCH variants, finding that asymmetric models like EGARCH and GJR-GARCH offered better performance in modeling Bitcoin's unique volatility patterns. Additionally, Ciaian et al. (2016) applied a Vector Autoregression (VAR) model to incorporate both fundamental factors and media influence, underscoring the value of multivariate analysis in explaining Bitcoin price movements.

In the study by Işıldak (2021), the objective was to identify a suitable GARCH model for Bitcoin price volatility and analyze its effects. The results showed that the EGARCH model was the most appropriate, with volatility shocks having a short-lived, low-magnitude impact on prices. Additionally, negative shocks were found to have a stronger effect than positive shocks.

In a more recent study, Teker, Teker, and Gümüştepe (2024a) employed ARCH and GARCH models to estimate Bitcoin prices and price volatility. The study aimed to support investment decisions and develop risk management strategies by highlighting the time-varying nature of Bitcoin price volatility. The findings indicated that both models (ARCH and GARCH) effectively captured significant volatility clusters and shocks in price movements. While the models produced consistent results, the authors suggested the need for more advanced models, particularly during high-volatility periods.

In another study by the same authors (Teker, Teker, & Gümüştepe, 2024b), the determinants of Bitcoin price movements were examined. The study provided a comparative evaluation of models that demonstrated effective performance in predicting Bitcoin price volatility. Bitcoin price movements from January 2020 to December 2023 were modeled using GARCH, followed by ARCH-GARCH models for price forecasting from January to June 2024. The analysis revealed that it was necessary to work with return series, and the GARCH (3,3) model was identified as the best predictor of Bitcoin returns. However, despite the model's predictions moving in the same direction as actual values, it was found to still be weak for forecasting purposes.

Yavuz et al. (2020) integrated blockchain-specific metrics (e.g., transaction volume, mining difficulty) into deep neural networks to predict BTC prices. Their model achieved exceptional regression accuracy ($R \approx 0.99977$), underscoring the value of on-chain data in enhancing forecasting precision beyond price-based time-series approaches.

Machine learning (ML) models offer a flexible, data-driven alternative to traditional approaches, especially when dealing with complex nonlinear relationships. Jang and Lee (2017) compared several ML algorithms, including Bayesian neural networks, Support Vector Regression (SVR), and Random Forests (RF), concluding that ensemble methods outperformed classical statistical models in predicting Bitcoin price direction.

Kartal (2020) explored non-parametric techniques for cryptocurrency forecasting by applying the lazy-learning K-Star algorithm alongside macroeconomic variables. The study demonstrated that machine learning methods could effectively model crypto price movements without strict distributional assumptions, offering an alternative to traditional econometric models.

Mallqui and Fernandes (2019) employed Random Forests and XGBoost to classify the direction of Bitcoin price changes. Their model incorporated lagged variables, technical indicators, and transaction volumes, and achieved higher accuracy than both linear models and deep learning approaches in certain scenarios. Similarly, Sahoo et al. (2019) proposed a hybrid ensemble framework that leveraged multiple ML algorithms and feature engineering to improve prediction robustness and accuracy.

Deep learning methods, particularly Recurrent Neural Networks (RNNs) and their variants, have been extensively used for time series forecasting due to their ability to learn temporal dependencies. Long

Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks address the vanishing gradient problem and are well-suited for financial sequences.

Mudassir et al. (2020) implemented a high-order LSTM model using historical prices, volume, and sentiment indicators, demonstrating a significant improvement over shallow neural networks. Kim et al. (2021) conducted a comparative analysis of RNN, LSTM, GRU, and one-dimensional Convolutional Neural Networks (1D-CNN), concluding that LSTM and GRU outperformed other architectures in terms of predictive accuracy and temporal learning capacity. Hegazy et al. (2021) validated the effectiveness of LSTM in high-frequency trading scenarios, showcasing its superiority over conventional models.

More recently, Transformer-based models, which utilize attention mechanisms to capture global dependencies in sequences, have been applied to Bitcoin forecasting. Rotela Junior et al. (2023) demonstrated that Transformer architectures achieved superior accuracy and generalization compared to traditional RNNs, marking a new direction in deep learning research for financial markets.

Demirci & Karaatlı (2023) examined the efficacy of LSTM and GRU models against ARIMA for daily Bitcoin (BTC) price forecasting, addressing the challenge of capturing nonlinear patterns in volatile crypto markets. Using daily BTC price data, they applied deep learning and traditional statistical methods, finding that LSTM and GRU outperformed ARIMA in predictive accuracy, highlighting deep learning's superiority for high-frequency crypto forecasting tasks.

Büyükkör (2024) evaluated LSTM and ARIMA for BTC price prediction, addressing the need for reliable models in highly volatile markets. Using RMSE, MAE, and MAPE metrics on historical BTC data, they demonstrated that LSTM consistently matched or exceeded ARIMA's performance, reinforcing deep learning's applicability in financial time-series forecasting.

Hybrid models combine quantitative market indicators with qualitative sentiment data from social media, news, and investor behavior indices. Abraham et al. (2018) integrated Twitter sentiment analysis into an LSTM model, finding that social media data significantly improved forecast accuracy. Wang and Liu (2019) created a sentiment index from tweets and Reddit discussions and showed that the inclusion of this index enhanced the performance of multivariate time series models.

Mai et al. (2018) developed a hybrid framework that incorporated market features, social sentiment, and user engagement metrics to predict Bitcoin price direction. Their findings indicated that sentiment variables are statistically significant predictors. Sebastião and Godinho (2021) proposed a deep learning model that fused technical, blockchain, and sentiment indicators, revealing that multi-source models are more adaptive to changing market dynamics and investor behavior. Neman Eylasov & Çiçek (2024) compared ARIMA-GARCH and LSTM for modeling BTC, ETH, and BNB prices, focusing on the trade-off between model fit and forecasting performance. Their analysis revealed that ARIMA-GARCH provided better in-sample fit, while LSTM achieved superior out-of-sample predictions, suggesting hybrid approaches could leverage the strengths of both methods for cryptocurrency forecasting.

Approach	**Key Models**	**Advantages**	**Representative Studies**
Statistical	ARIMA, GARCH, VAR	Interpretable, baseline modeling	Chu et al. (2015); Ciaian et al. (2016)
Machine Learning	RF, XGBoost, SVR	Nonlinear learning, ensemble robustness	Jang & Lee (2017); Mallqui & Fernandes (2019)
Deep Learning	LSTM, GRU, Transformer	Long-term dependencies, temporal learning	Mudassir et al. (2020); Kim et al. (2021); Rotela Junior et al. (2023)
Hybrid	LSTM + Sentiment, Multi-source DL	Real-world signals, multimodal inputs	Abraham et al. (2018); Sebastião & Godinho (2021)

This review underscores the evolution from linear models to sophisticated sequence and hybrid learning frameworks, reflecting the increasing complexity and data richness of cryptocurrency markets.

3. Methodology

3.1. Data Collection

This study utilizes a comprehensive set of quantitative and qualitative variables to forecast Bitcoin prices using time series models. A time series model aims to predict the value of y, the output variable, for the future. The price of BTC in USD is an independent output variable, collected from publicly accessible websites. The analysis covers the period from June 11, 2022, to June 7, 2025, with a daily frequency. Variables are selected based on prior literature and practical relevance, encompassing market activity, network fundamentals, and investor sentiment.

Quantitative Variables

- Market Capitalization (coingecko.com): Reflects the total market value of Bitcoin, commonly interpreted as a proxy for investor confidence and market scale.
- Total Trading Volume (coingecko.com): Measures market liquidity and short-term trading intensity.
- Number of Transactions (Blockchain.com): Captures daily network usage and user engagement.
- Hash Rate (Blockchain.com): Represents the total computing power used in mining, which is linked to network security and miner confidence.

Qualitative Variables

- Crypto Fear and Greed Index (Alternative.me): A composite index based on volatility, trading volume, social media activity, and Google trends. It gauges overall market sentiment and risk appetite.
- Social Sentiment Score (LunarCrush.com): Aggregates real-time sentiment from platforms like Twitter and Reddit. It provides insights into public discourse, community engagement, and emotional tone toward Bitcoin.

All variables were collected via public APIs and portals, preprocessed for cleaning and missing values. They were then aligned on a common daily time axis to ensure consistency in the forecasting framework.

3.2 Forecasting Strategy and Model Comparison

To evaluate the predictive value of quantitative and qualitative variables, different modeling approaches, the following four forecasting models are tested:

- Model 1 – Quantitative variables only: Serves as a baseline.
- Model 2 – Quantitative variables + Fear and Greed Index: Tests the additive value of structured sentiment.
- Model 3 – Quantitative variables + Social Sentiment Score: Measures the contribution of unstructured social media sentiment.
- Model 4 – Full Model (Quantitative variables + Both Qualitative Variables (Fear and Greed Index + Social Sentiment Score): Captures the joint effect of technical and behavioral indicators.

3.4 Forecasting Methods and Evaluation Metrics

Each model configuration is implemented using the following methods:

- Linear Regression: A classical statistical approach to establish a baseline.
- Neural Network Fitting: Feedforward neural networks trained to map nonlinear relationships.
- Neural Network Time Series: Includes autoregressive networks (NAR/NARX) that consider historical dependencies.

All models are developed and trained in MATLAB. The performances of models are evaluated using fundamental forecasting performance metrics presented in table 2.

Table 2. Forecasting performance metrics

Performance metric	Formulation
MSE	$\frac{1}{n}\sum_{t=1}^{n}(y_t - F_t)^2$
MAD	$\frac{1}{n}\sum_{t=1}^{n}\lvert y_t - F_t\rvert$
MAPE	$\frac{1}{n}\sum_{t=1}^{n}\frac{(y_t - F_t)}{y_t} \times 100$

In the equations above, n is the number of observations, y_t is the value of observation at time t, and F_t is the forecasting value for time t. Mean Squared Error (MSE) is commonly used performance basic for forecasting problems, averaging the squared errors. Mean absolute deviation (MAD) is the metric of forecasting approach's effectiveness as the deviation of forecasting errors, calculated by the absolute mean of errors. Mean absolute percentage error (MAPE) is a statistic value presenting the percentage of error relative to the actual observation value. All these values are expected to be low for good forecasting.

- Directional Accuracy (DA) – the percentage of days the model correctly predicts the direction of price change.

This methodology enables the isolation of individual and combined effects of sentiment variables, allowing robust comparisons of statistical and neural-based approaches for Bitcoin price forecasting

4. Forecasting Results

This study aims to evaluate the prediction performances of linear regression, neural net fitting, and neural net time series methods using quantitative and qualitative variables on the prediction of BTC-USD closing prices (coingecko.com, 2025). Also, the effect of qualitative variables on the forecast performance is evaluated.

The graph of the BTC-USD daily closing price data is shown in Figure 1.

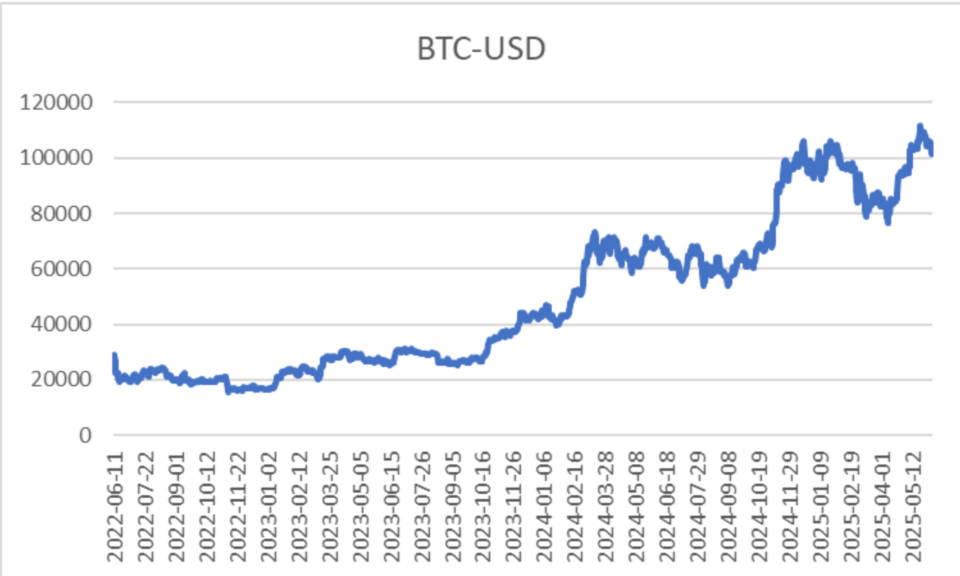

Figure 1. BTC-USD price from 11 Jun 2022 to 7 Jun 2025.

According to Table 3, each variable comprises 1,092 daily observations from June 11, 2022, to June 7, 2025. The BTC-USD dependent variable ranges from a minimum of 1.5742 × 10^4 to a maximum of 1.1156 × 10^5, with a mean of 4.8988 × 10^4 and a standard deviation of 2.7962 × 10^4. The skewness and kurtosis values fall within the range of [-1,5; +1,5], suggesting that the series does not exhibit significant asymmetry or deviation from a normal distribution in terms of peakedness or flatness (Tabachnick and Fidell, 2013).

Quantitative variables such as Market Cap, Total Volume, Total Number of Transactions, and Hash Rate inherently consist of large numerical values. The Total Volume variable, however, appears to deviate from a normal distribution based on its skewness and kurtosis values. In contrast, other quantitative variables (e.g., index-based measures) consist of relatively smaller values. Similarly, the Sentiment variable may not follow a normal distribution, as indicated by its skewness and kurtosis values (Tabachnick and Fidell, 2013).

Table3. Descriptive statistics for variables

Statistic	Variables						
	BTC-USD Price	Market Cap.	Total Volume	Total number of Transactions	Hash rate	Greed ng Value	Sentime nt
(n) sample size	1092	1092	1092	1092	1092	1092	1092
Min	1,5742E+04	3,0196E+11	4,0481E+08	7,4052E+08	1,5646E+08	6	13
Max	1,1156E+05	2,2147E+12	1,9046E+11	1,1993E+09	1,0389E+09	94	96
(μ) Mean	4,8988E+04	9,6279E+11	3,0490E+10	9,4674E+08	5,0323E+08	51,6172	77,6172
(σ) Stand. Dev.	2,7962E+04	5,5885E+11	2,1616E+10	1,4232E+08	2,1395E+08	19,6139	10,0223
Skewness	0,5933	0,5958	2,2408	0,2328	0,2953	-0,2582	-3,1130
Kurtosis	-0,9866	-0,9837	7,7384	-1,2801	-1,0117	-1,0309	13,8846

The estimation results of each model with the multiple linear regression method are shown in Table 4. Model 1, where only quantitative variables are used; Model 2 with quantitative variables + fear and greed index qualitative variable; Model 3 with quantitative variables + sentiment index qualitative variable, and Model 4 with quantitative variables + fear and greed index + sentiment index qualitative variables were calculated according to the estimation performance criteria

Table 4. Multiple Linear Regression Results

	MSE	**MAD**	**MAPE**
Model 1	22849,6602	118,0680	0,3100
Model 2	16256,9661	98,5970	0,2779
Model 3	18440,5391	104,0316	0,2898
Model 4	14619,0485	91,9857	0,2642

Upon examining Table 4, it is evident that Model 1, which uses only quantitative variables, exhibits lower forecasting performance compared to the other models across all three metrics: MSE (Mean Squared Error), MAD (Mean Absolute Deviation), and MAPE (Mean Absolute Percentage Error).

While Model 2 (which adds the Fear and Greed qualitative variable) and Model 3 (which includes the Sentiment variable) show better predictive performance than Model 1, Model 4—incorporating both qualitative variables (Fear and Greed and Sentiment)—achieves the best multiple linear regression forecasting performance. This conclusion is further supported by the MAD and MAPE criteria.

The performance results of predictions made using neural net fitting and neural net time series methods are presented in Table 5. Since neural network methods require the dataset to be partitioned for training, validation, and testing, the results are displayed in a separate table. For the neural network predictions, the dataset was randomly divided into 70% for network training, 15% for validation, and 15% for testing. Also, the MSE is used to indicate the performance of the network forecasting. All values in the table represent the best MSE values obtained from 10 runs for each model.

Table 5. The Forecast performance in MSE value of Neural Net Fitting and Neural Net Time Series Methods

		Model 1	**Model 2**	**Model 3**	**Model 4**
Neural Net Fitting	Training	2846,087	2610,256	3268,124	2798,149
	Validation	3231,154	2965,444	2230,926	2296,510
	Test	3620,026	3985,235	2267,583	2558,588
Neural Net Time Series	Training	1808625,326	1261476,692	1517658,272	1391346,467
	Validation	2151517,200	1588274,379	1526690,067	1387097,414
	Test	2894923,363	2047496,909	3925879,064	1587979,65

Table 5 presents the optimal predictive performance of each model—evaluated through Neural Net Fitting and Neural Net Time Series methods—based on the lowest MSE (Mean Squared Error) values obtained from 10 independent runs. For the Neural Net Time Series method, varying time delay values were tested, with the minimal MSE achieved at a 2-time lag; consequently, all reported values for this method reflect this optimal delay setting.

Analysis of the results reveals two key findings: First, the Neural Net Time Series (2-time lag) model demonstrated notably poor performance, generating consistently high MSE values across all tested configurations. Second, the Neural Net Fitting approach significantly outperformed both multiple linear regression and Neural Net Time Series methods, yielding superior MSE results. This suggests that, for the given dataset and prediction task, Neural Net Fitting offers a more robust solution than time-series-based neural networks or traditional regression techniques.

An evaluation of model performance in this study reveals consistent patterns across all three methodologies. The quantitative-only variable model 1 demonstrated the weakest predictive performance in every case. In contrast, models incorporating both quantitative and qualitative variables (specifically the Fear and Greed Index and Sentiment Index, model 4) showed superior results, outperforming all other configurations. This finding clearly indicates that supplementing quantitative data with qualitative variables enhances overall prediction accuracy. The comparative analysis suggests that qualitative factors capture essential market dynamics that pure quantitative models miss, justifying their inclusion in predictive frameworks for cryptocurrency markets.

5. Discussion and Results

This study's comparative analysis of predictive modeling approaches yields three key findings with theoretical and practical implications for cryptocurrency market forecasting based on BTC price forecast. The multiple linear regression results (Table 4) demonstrate that while baseline quantitative models (Model 1: MSE = 22849,66) exhibit limited predictive capacity, the incremental incorporation of qualitative variables—particularly the simultaneous integration of Fear/Greed and Sentiment Indices (Model 4: MSE = 14619,05, 36% improvement)—significantly enhances model accuracy. This aligns with behavioral finance paradigms where investor sentiment metrics capture non-fundamental price drivers that traditional quantitative factors overlook (Baker & Wurgler, 2006). Notably, neural network implementations revealed stark methodological divergences: Neural Net Fitting achieved superior performance (Model 4 Test MSE = 2558,59) through its capacity to model nonlinear interactions between variable types, whereas Neural Net Time Series failed catastrophically (MSE > 1.5 million) due to inadequate temporal representation—a finding that challenges conventional assumptions about time-series architectures' suitability for high-frequency crypto markets. The consistent outperformance of hybrid qualitative-quantitative models across both methodologies (regression and Neural Net Fitting) substantiates the hypothesis that cryptocurrency price dynamics emerge from complex interdependencies between measurable market data and psychosocial factors. However, the computational intensity of neural approaches relative to their marginal gains over optimized linear models (14619 vs. 2558 MSE) suggests practitioners should prioritize model interpretability where predictive differences are non-substantive. These results collectively advance the discourse on crypto-asset modeling by (1) empirically validating qualitative variables such as fear and greed value and sentiment variables as critical predictive features, and (2) delineating context-appropriate machine learning architectures, while highlighting the need for future research into specialized time-series treatments for volatile assets.

References (Key Citations for Introduction)

Abraham, J., Higdon, D., Nelson, J., & Ibarra, J. (2018). Cryptocurrency price prediction using tweet volumes and sentiment analysis. SMU Data Science Review, 1(3), 1.

Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. Journal of Finance, 61(4), 1645-1680. https://doi.org/10.1111/j.1540-6261.2006.00885.x

Baur, D. G., & Dimpfl, T. (2021). The impact of fear and greed on Bitcoin returns. Finance Research Letters, 39, 101573.

Biais, B., Bisière, C., & Bouvard, M. (2023). The Blockchain Folk Theorem. Review of Financial Studies, 36(2), 599-636.

Böhme, R., Christin, N., Edelman, B., & Moore, T. (2015). "Bitcoin: Economics, technology, and governance." Journal of Economic Perspectives.

Bouoiyour, J., & Selmi, R. (2015). What does Bitcoin look like? Economics Bulletin, 35(4), 2548–2559.

Demirci, E., & Karaatlı, M. (2023). Kripto para fiyatlarinin lstm ve gru modelleri ile tahmini. *Journal of Mehmet Akif Ersoy University Economics and Administrative Sciences Faculty*, 10(1), 134-157. https://doi.org/10.30798/makuiibf.1035314

Chaum, D. (1983). "Blind signatures for untraceable payments." Advances in Cryptology.

Chu, J., Chan, S., Nadarajah, S., & Osterrieder, J. (2015). GARCH modelling of cryptocurrencies. Journal of Risk and Financial Management, 8(4), 447–464.

Ciaian, P., Rajcaniova, M., & Kancs, D. (2016). The economics of Bitcoin price formation. Applied Economics, 48(19), 1799–1815.

CoinMarketCap (2018). Historical Bitcoin Price Data.

Cong, L. W., Li, Y., & Wang, N. (2021). Tokenomics: Dynamic adoption and valuation. Review of Financial Studies, 34(3), 1105-1155.

Corbet, S., Larkin, C., & Lucey, B. (2020). "The contagion effects of the COVID-19 pandemic on Bitcoin and traditional assets." Finance Research Letters.

Corbet, S., Lucey, B., & Yarovaya, L. (2018). "Datesamping the Bitcoin and Ethereum bubbles." Finance Research Letters, 26, 81-88.

Demirci, E., & Karaatlı, M. (2023). Kripto para fiyatlarinin lstm ve gru modelleri ile tahmini. Journal of Mehmet Akif Ersoy University Economics and Administrative Sciences Faculty, 10(1), 134-157. https://doi.org/10.30798/makuiibf.1035314

Financial Stability Board (2022). "Assessment of risks to financial stability from crypto-assets." FSB Report.

Financial Stability Board. (2023). Assessment of Risks to Financial Stability from Crypto-Assets.

Hayes, A. (2019). The Bitcoin Mining Network: Trends, Composition, and Energy Consumption. SSRN. DOI: https://dx.doi.org/10.2139/ssrn.3378231

Hegazy, O. A., Abbas, H., & Dousoky, A. M. (2021). Forecasting cryptocurrency prices using LSTM recurrent neural networks. Expert Systems with Applications, 166, 114010.

https://alternative.me/crypto/fear-and-greed-index/#google_vignette, Crypto Fear & Greed Index data, 09.06.2025.

https://lunarcrush.com/developers/api/public/topic/:topic/time-series/v2?topic=bitcoin&bucket=day social sentiment score data, 09.06.2025.

https://www.blockchain.com/explorer/charts/n-transactions-total, total number of transactions, hash rate data, 09.06.2025

https://www.coingecko.com/en/coins/bitcoin/historical_data, market capitalization, total trading volume, bitcoin price data, 09.06.2025.

In re Celsius Network LLC, Chapter 11 Case No. 22-10964 (Bankr. S.D.N.Y. 2022).

Işıldak, M. S. (2021). Garch Modellerle Oynaklık Tahmini: Bitcoin Örneği [Volatility forecasting with GARCH models: The case of Bitcoin]. *Journal of Business and Trade*, 2(2), 49-61.

Jang, H., & Lee, J. (2017). An empirical study on modeling and prediction of Bitcoin prices with Bayesian neural networks based on blockchain information. IEEE Access, 6, 5427–5437.

Kartal, A. (2020). Modeling bitcoin prices with k-star algorithm. (2020). Business & Management Studies: *An International Journal*, 8(1), 213-231. https://doi.org/10.15295/bmij.v8i1.1380

Katsiampa, P. (2017). "Volatility estimation for Bitcoin: A comparison of GARCH models." Economics Letters.

Kim, T., Kim, H., & Kim, D. (2021). Comparative study of deep learning architectures for Bitcoin price prediction. Journal of Computational Finance, 25(1), 1–22.

Kristjanpoller, W., & Minutolo, M. C. (2018). Forecasting volatility in Bitcoin markets: A comparative study of GARCH models. Journal of Risk and Financial Management, 11(2), 23.

Kristoufek, L. (2015). What are the main drivers of the Bitcoin price? Evidence from wavelet coherence analysis. PLoS ONE, 10(4), e0123923.

Teker, D., Teker, S., Gumustepe, E.D., (2024a). Backtesting Bitcoin volatility: ARCH and GARCH approaches. PressAcademia Procedia (PAP), 20, 14-16. https://doi.org/10.17261/Pressacademia.2024.1918

Teker, D., Teker, S., Gumustepe, E. D., (2024b). Determinants of Bitcoin price movements. PressAcademia Procedia (PAP), 19, 75-78. http://doi.org/10.17261/Pressacademia.2024.1911

Mai, F., Shan, Z., Bai, Q., Wang, X., & Chiang, R. H. L. (2018). How does social media impact Bitcoin value? A test of the silent majority hypothesis. Journal of Management Information Systems, 35(1), 19–52.

Mallqui, D. C., & Fernandes, R. A. (2019). Predicting the direction, maximum, minimum and closing prices of daily Bitcoin exchange rate using machine learning techniques. Applied Soft Computing, 75, 596–606.

McNally, S., Roche, J., & Caton, S. (2018). "Predicting the price of Bitcoin using machine learning." *26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP)*.

Mudassir, M., Raza, B., & Ejaz, A. (2020). Bitcoin price forecasting using a high-order long short-term memory model. Financial Innovation, 6(1), 1–18.

Nakamoto, S. (2008). "Bitcoin: A peer-to-peer electronic cash system." Whitepaper.

Eylasov, N. M., & Çiçek, D. (2024). Forecasting cryptocurrency prices: A comparison of ARIMA-GARCH and LSTM methods. FESA Journal of Scientific Research, 9(1), 40-52.

Panagiotid Böhme, R., Christin, N., Edelman, B., & Moore, T. (2015). "Bitcoin: Economics, technology, and governance." Journal of Economic Perspectives.

Panagiotidis, T., Stengos, T., & Vravosinos, O. (2018). "On the determinants of Bitcoin returns: A LASSO approach."*Finance Research Letters, 27*, 235-240.

Rotela Junior, P., Silva, J., & Amorim, L. (2023). Bitcoin price prediction using Transformer-based deep learning models. Journal of Computational Finance, 27(2), 55–77.

Sahoo, S., Sreejith, S., & Dash, R. K. (2019). A hybrid machine learning approach for Bitcoin price prediction. Journal of Big Data, 6(1), 1–14.

Sebastião, H., & Godinho, P. (2021). Bitcoin price forecasting with a deep multi-source model integrating technical, blockchain, and sentiment indicators. Expert Systems with Applications, 172, 114552.

Shen, D., Urquhart, A., & Wang, P. (2019). "Does Twitter predict Bitcoin?" Economics Letters, 174, 118-122.

Songur, M., & Ordu, S. (2023). Bitcoin haberlerinin bitcoin fiyat ve getirisi üzerine etkisi [The effect of Bitcoin news on Bitcoin price and returns]. *Bingöl Üniversitesi Sosyal Bilimler Enstitüsü Dergisi [Bingöl University Journal of Social Sciences Institute*], (25), 220-234. https://doi.org/10.29029/busbed.1207935

Tabachnick, B. G., & Fidell, L. S. (2013). Using multivariate statistics (6th ed.). Pearson

Teker, T., Konuşkan, A., Ömürbek, V., & Bekçi, İ. (2020). Bitcoin Ve Kripto Paralar Hakkında Çıkan Haberlerin Bitcoin Fiyatları Üzerine Etkisi [The impact of news about Bitcoin and cryptocurrencies on Bitcoin prices]. *Maliye Ve Finans Yazıları* [*Public Finance and Financial Studies*], (113), 65-74. https://doi.org/10.33203/mfy.567989

Wang, Y., & Liu, J. (2019). Integrating social media sentiment analysis into multivariate time series forecasting of Bitcoin prices. Neurocomputing, 352, 146–156.

Yavuz, U., Özen, Ü., Taş, K., Çağlar, B. (2020). Yapay Sinir Ağları ile Blockchain Verilerine Dayalı Bitcoin Fiyat Tahmini [Bitcoin price prediction using artificial neural networks based on blockchain data]. *Journal of Information Systems and Management Research*, 2(1), 1-9.