Çukurova Üniversitesi
Mühendislik Fakültesi Dergisi
Çukurova University
Journal of the Faculty of Engineering
ISSN: 2757-9255 CİLT/VOLUME: 41 SAYI/ISSUE: 2 HAZİRAN/JUNE 2026

Development of an Artificial Intelligence-Based Prediction Engine for the Cryptocurrency Markets

Onur AYGÜN [1,a], **Ceren ULUS** [2,b], **Mehmet Fatih AKAY** [3,c]

[1]*Innovance Information Technologies, Department of Software Development, Istanbul, Türkiye*
[2]*EFA Innovation Consultancy and Technology, Department of R&D, Adana, Türkiye*
[3]*Cukurova University, Department of Computer Engineering, Adana, Türkiye*

[a]***ORCID**: 0009-0008-4534-3783;* [b]***ORCID**: 0000-0003-2086-6381;* [c]***ORCID**: 0000-0003-0780-0679*

Article Info
Received : 16.12.2025
Accepted : 23.01.2026
DOI: 10.21605/cukurovaumfd.1843158
Corresponding Author
Ceren ULUS
f.cerenulus@gmail.com
Keywords
Cryptocurrency markets
Bitcoin closing price prediction
Machine learning
Transformer architecture
Risk management and portfolio optimization
***How to cite:** AYGÜN, O., ULUS, C., AKAY, M.A., (2026). Development of an Artificial Intelligence-Based Prediction Engine for the Cryptocurrency Markets. Cukurova University, Journal of the Faculty of Engineering, 41(2), 285-298.*

ABSTRACT
This study aims to present an innovative solution that will help analyze market trends and price dynamics, identify risk factors, and provide investors with predictions using big data analytics and advanced artificial intelligence algorithms. Bitcoin closing price prediction models for short-term, medium-term, and long-term have been developed using Long-Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and transformer architecture based models and Monte Carlo scenarios have been generated for short-term Bitcoin closing price predictions; and tail-risk metrics such as Value at Risk (VaR) and Conditional Value at Risk (CVaR) have been calculated. A dataset for BTC/USDT symbol has been prepared using Binance API. The performance of developed models was evaluated using Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). The results showed that most successful results have been obtained with prediction models developed for short term. Prediction models developed with GRU achieved superior performance.

Kripto Para Piyasaları için Yapay Zeka Tabanlı Tahmin Motorunun Geliştirilmesi

Makale Bilgileri
Geliş : 16.12.2025
Kabul : 23.01.2026
DOI: 10.21605/cukurovaumfd.1843158
Sorumlu Yazar
Ceren ULUS
f.cerenulus@gmail.com
Anahtar Kelimeler
Kripto para piyasaları
Bitcoin kapanış fiyatı tahmini
Makine öğrenimi
Transformer mimarisi
Risk yönetimi ve portföy optimizasyonu
***Atıf şekli:** AYGÜN, O., ULUS, C., AKAY, M.A., (2026). Kripto Para Piyasaları için Yapay Zeka Tabanlı Tahmin Motorunun Geliştirilmesi. Çukurova Üniversitesi, Mühendislik Fakültesi Dergisi, 41(2), 285-298.*

ÖZ
Bu çalışma, büyük veri analitiği ve gelişmiş yapay zeka algoritmaları kullanarak piyasa trendlerini ve fiyat dinamiklerini analiz etmeye, risk faktörlerini belirlemeye ve yatırımcılara tahminler sunmaya yardımcı olacak yenilikçi bir çözüm sunmayı amaçlamaktadır. Kısa, orta ve uzun vadeli Bitcoin kapanış fiyatı tahmin modelleri, Uzun-Kısa Vadeli Bellek (LSTM), Kapılı Tekrarlayan Birim (GRU) ve transformatör mimarisine dayalı modeller kullanılarak geliştirilmiş ve kısa vadeli Bitcoin kapanış fiyatı tahminleri için Monte Carlo senaryoları oluşturulmuştur; ayrıca Değer Riski (VaR) ve Koşullu Değer Riski (CVaR) gibi kuyruk riski metrikleri hesaplanmıştır. Binance API kullanılarak BTC/USDT sembolü için bir veri seti hazırlanmıştır. Geliştirilen modellerin performansı, Kök Ortalama Kare Hatası (RMSE), Ortalama Mutlak Hata (MAE) ve Ortalama Mutlak Yüzde Hatası (MAPE) kullanılarak değerlendirilmiştir. Sonuçlar, en başarılı sonuçların kısa vadeli tahmin modelleriyle elde edildiğini göstermiştir. GRU ile geliştirilen tahmin modelleri üstün performans göstermiştir.

1. INTRODUCTION

Cryptocurrency markets, a new financial space where investors and institutions transact digital assets, tokens, and smart contracts, are based on blockchain technology like Bitcoin and Ethereum, and represent a significant area for the current economy. The buying and selling of digital assets take place within this ecosystem. These markets are not governed by a single authority; transactions are secured with cryptographic verification and distributed ledger technologies, ensuring high reliability [1].

Despite their high volatility, cryptocurrencies have become increasingly popular due to features such as faster money transfers and lower transaction costs. According to available information, the total market capitalization of the cryptocurrency market exceeded $2.5 trillion USD as of June 2024. Bitcoin, introduced in 2008, is the first open-source digital currency, managed by an open-source software algorithm operating over a global internet network. Currently, over 2,000 different cryptocurrencies are traded on the market. Despite the large number of cryptocurrencies, Bitcoin still accounts for approximately half of the total market capitalization. Bitcoin's price movements can generally be explained by supply and demand dynamics. This, however, is quite significant in determining the overall direction of the cryptocurrency market [2].

While the market is more mature than in previous years due to increased liquidity, cryptocurrencies still exhibit high price volatility. Macroeconomic developments such as interest rates, inflation, and regulation strongly influence prices. However, this high volatility and dynamic nature of cryptocurrency markets make it difficult for investors to make informed decisions. In this context, investors can be vulnerable to instantaneous changes. The dramatic price fluctuations in a short period, driven by market sentiment, social media influences, and macroeconomic factors, make it challenging for investors to take precautions against such fluctuations. The unpredictable nature of the market leads investors to focus more on risk management strategies.

Risk management strategies also primarily focus on price analysis. Analyzing data sources such as price movements, trading volume, and market sentiment is a strategic approach that allows for a more accurate, holistic, and predictable assessment of the rapidly changing nature of cryptocurrency markets. However, accurate analysis is crucial. Manual methods cannot fully capture the sensitivity of this analysis. Consequently, these methods do not achieve high success rates. Investors need to be able to accurately analyze market trends and price dynamics. Therefore, providing reliable and real-time predictions is essential for investors to make informed and strategic decisions in the cryptocurrency market. In this context, price prediction for the cryptocurrency market is a crucial strategic action for managing the risks of these highly volatile assets, optimizing investment strategies, and making more informed trading decisions to minimize potential losses and maximize returns. With price prediction in cryptocurrency markets, risks can be managed more effectively [3]. It becomes possible to make buy-sell decisions at the right time, maximize profits, and limit losses. Furthermore, it becomes possible to shape portfolio strategies more rationally and to identify market trends and turning points in advance.

This study aims to present an innovative solution that will help analyze market trends and price dynamics, identify risk factors, and provide investors with predictions using big data analytics and advanced machine learning algorithms. To this end, an AI-based prediction engine has been developed. The goal is to predict future market movements, and Bitcoin closing price prediction models have been developed using prediction algorithms powered by advanced machine learning techniques such as LSTM, GRU, and Transformer-based models.

The study also included the generation of Monte Carlo scenarios for short-term price forecasts and the calculation of tail-risk measures such as VaR and CVaR. Position sizes were dynamically adjusted using threshold and budget rules, and automatic warning mechanisms have been activated against sudden price changes.

The most significant difference between this study and the existing literature is that it simultaneously evaluates short-term (1 min), medium-term (5 min), and long-term (30 min) crypto closing prices. The methodological approach, which includes transformer-based architecture rarely seen in literature, is quite diverse, incorporating time-series-based methods such as LSTM and GRU and the performance of developed models was evaluated using RMSE, MAE, and MAPE. Walk-forward cross-validation has been used in the data preparation and evaluation process. By retraining each fold and maintaining time-series dependency, reproducibility and realistic performance have been ensured. Most studies in academic literature are far from risk optimizations. The literature review does not include any studies that conduct risk analysis. This constitutes an important innovative feature of the study. In this study, Monte Carlo scenarios have been generated; VaR and CVaR tail-risk measures have been calculated. By dynamically adjusting position sizes with threshold and budget rules, automatic warning mechanisms against sudden regime changes have been activated. This has made it possible to make investment decisions more sustainable by basing them not only on expected returns, but also on expected loss and excessive risk measures.

This study is organized as follows: Section 2 includes relevant literature. Dataset generation is presented in Section 3. Methodology is presented in Section 4. Development of models is presented in Section 5. The procedures and analyses related to Risk Management and Portfolio Optimization are presented in Section 6. Results and discussion are given in Section 7. Section 8 concludes the paper.

2. LITERATURE REVIEW

Alnami et al. [4] using datasets of Bitcoin, Ethereum, Binance Coin, and Litecoin cryptocurrencies, they employed machine learning and deep learning to predict closing prices for these cryptocurrencies. A Random Forest (RF) community learning algorithm, a Gradient Boosting model, and a feedforward neural network have been used. A Z-score-based anomaly detection strategy has been adopted to classify closing prices as normal or abnormal and to help identify significant market events. The performance of the models has been analyzed using Mean Squared Error (MSE), RMSE, MAE) and and Determination of Coefficent (R^2) metrics. Experimental results showed that the RF and Gradient Boosting models have been successful.

Asmat and Maiyama [5] emphasized using the Neural Basis Extension Analysis Time Series (N-BEATS) deep learning architecture, which has high performance in modeling complex patterns in time series data to predict Bitcoin prices. 729 days of hourly Bitcoin price data collected from Yahoo Finance have been used. The model has been evaluated by comparing it to Linear Regression and LSTM. In the experimental results, the model has been analyzed using MAE and R^2. N-BEATS provided an R^2 score of 0.00240 and an MAE score of 0.9998.

Kaur et al. [6] Ethereum (ETH), Bitcoin (BTC), and Litecoin (LTC) have been studied for exchange rate prediction. Two categories of Recurrent Neural Networks (RNNs) have been used: LSTM and GRU. The GRU model has been optimized with Adam with a 20% dropout rate to prevent overfitting and implemented as a two-layer deep learning network. Historical price datasets have been obtained from the CryptoDataDownload source and normalized. The dataset has been split into training and testing segments of 80%–20%. Prediction performance has been evaluated using MSE, MAE, MAPE, and RMSE. Experimental results showed that the GRU model performed better than the LSTM. The GRU provided MAPE values of 0.03540, 0.08703, and 0.04415 for BTC, LTC, and ETH, respectively.

Makri et al. [7] aimed to predict Ethereum prices for short-term and limited prediction scenarios, using Large Language Models (LLMs) for this purpose. A new approach has been proposed that adapts pre-trained LLMs to the specifics of the Ethereum price time series problem. This approach has been analyzed for performance across several metrics, including MSE, MAE, and RMSE. The results show that the proposed approach provides the most advanced performance in this area.

Mazinani et al. [8] aimed to predict the prices of Bitcoin (BTC), Ethereum (ETH), and Ripple (XRP) and analyzed the performance of deep learning algorithms in both short-term and long-term prediction. Models have been evaluated using RMSE, MAE, MAPE, and R^2 metrics. Experimental results showed that Transformer has been generally more effective for short-term predictions and also performed well for long-term predictions. Convolutional Neural Network-Iterative Neural Network (CNN-RNN) achieved the lowest complexity in terms of the number of Multiplication and Addition (MAC) operations. The SimpleRNN method has been observed to have the fewest parameters and the smallest FLASH memory requirement. Overall, CNN Gate Iterative Unit (CNN-GRU) provided the best common accuracy-complexity for predicting BTC and ETH prices, while CNN-RNN yielded superior results for XRP price prediction.

Pandey and Sharma [9] aimed to create a deep learning architecture using RNN, CNN, and LSTM to predict Bitcoin and Ethereum prices. CNNs have been used to extract features from historical price data. A dataset containing major cryptocurrencies has been used as the dataset. Experimental results showed that the proposed model predicted cryptocurrency prices with high accuracy. Consequently, the bidirectional model outperformed traditional time series forecasting methods and one-way models, revealing the potential to improve price prediction in the cryptocurrency market.

Kang et al. [10] aimed to integrate deep learning-based price prediction with technical indicators. Twelve deep learning models have been developed. The success of the developed models and their performance in generating trading signals have been analyzed across various cryptocurrencies and prediction periods. Backtesting has been conducted to determine the most suitable strategy, evaluated using the Sharpe ratio. In the experimental results, it has been observed that SegRNN performed better than other models in price prediction, while a strategy combining TimesNet and Bollinger Bands (BB) provided the highest trading performance in the ETH market with a return of 3.19, a maximum decline of -7.46, and a Sharpe ratio of 3.56. With the integration of models and technical indicators, significant improvements have been shown in the mid-range 4-hour timeframe, while no performance improvement has been achieved in shorter timeframes such as 30 minutes.

Saeed et al. [11] presented a systematic review of existing commodity price forecasting models for housing prices in the real estate market. Different commodity price forecasting models used in academic and commercial settings have been identified and analyzed. The study included forecasting of housing prices, stock prices, and natural gas prices. In each category, ensemble methods, neural networks, support vector machines, time series analysis, and regression analysis have been used. The study also presented an analysis of the strengths and weaknesses of the existing models and the key variables affecting their performance.

Chandra et al. [12] aimed to examine approaches to leverage time series methodology using Prophet for cryptocurrency price prediction and incorporating sentiment analysis from social media, and proposed a CNN-LSTM layer. The approach is trained on real-time financial data streams, and performance is evaluated using accuracy score, MAPE, and R^2. In the experimental results, it has been observed that integrating sentiment analysis with CNN-LSTM provided high performance with an accuracy of 0.94 in the training data and 0.91 in the test data compared to the Prophet algorithm on Bitcoin closing prices.

Golnari et al. [13] focused on BTC and aimed to predict digital currency prices. They used Probabilistic Gated Recurrent Units (P-GRU). For evaluation, a year's worth of BTC price history, sampled at five-minute intervals, has been used. A pre-trained model on BTC data has been used to predict the prices of six major cryptocurrencies: Ethereum, Litecoin, Tron, Polkadot, Cardano, and Stellar. GRU, LSTM, time-distributed, bidirectional, and simple models have been analyzed. A separate model has been created for each cryptocurrency. A recall mechanism has been used for model optimization. The transfer learning paradigm has been adopted in the study.

Kamboj et al. [14] explained the evolution of cryptocurrencies and their global impact. Their aim has been to improve decision-making processes by identifying patterns and indicators in the crypto sector. Using Scikit-Learn and Pandas, they created an adaptive environment.

Kavithra et al. [15] performed cryptocurrency price prediction using deep learning techniques. Focusing on 20 cryptocurrencies, they utilized LSTM, GRU, Bi-Directional Long Short-Term Memory (Bi-LSTM), and Hybrid LSTM-GRU methods, achieving high accuracy in price prediction. Additionally, the study offered strategies to guide users on whether to buy, sell, or hold cryptocurrencies daily. Data preprocessing has been performed using a Min-Max scaler.

Kawli et al. [16] aimed to examine the performance of iterative neural networks, deep learning neural networks, Bayesian regression, k-nearest neighbor, support vector machine, and other algorithms in predicting the prices of cryptocurrencies such as Bitcoin, Ethereum, Dogecoin, and Litecoin. Literature studies on price prediction using machine learning have been utilized, including studies on Non-Fungible Token. (NFT) sales predictability, NFT sales price fluctuation prediction, gold price prediction, and silver price prediction. The study focused on comparing machine learning algorithms with high-dimensional features, time series analysis, and different statistical models. Market liquidity and stock market dynamics have been also analyzed in relation to prediction models.

Rajaei and Mahmoud [17] presented an RF model to predict short-term price jumps in the cryptocurrency market. Binance historical daily data from January 1, 2018, to December 31, 2021, has been used for model training, and the test set date range has been selected as January 2022 to October 2023. An oversampled training data set has been also used. Experimental results showed that using a 1% growth rate on the LSK/USDT test datasets for the periods October 1-31, 2023, July 1, 2023- September 30, 2023, and January 1, 2022-October 31, 2023, profits of 18%, 30%, and 68% have been obtained, respectively.

3. DATASET GENERATION

A dataset for the BTC/USDT symbol has been prepared using the Binance API. BTC is Bitcoin that is traded. USDT is a stable cryptocurrency related to the US dollar. BTC/USDT is a value that shows how much USDT (approximately how many US dollars) 1 Bitcoin is worth. BTC/USDT offers a long and continuous time series, which allows for more stable convergence in prediction models. The dataset contains 45.000 data points. BTCUSDT - 1-minute candlestick data has been retrieved sequentially from the Binance Klines API to reach a total target observation count 45.000. Data has been retrieved with a batch size of 1000 and the endtime has been shifted backward to create causal cumulative data. Derivative price features close_open, high_low, and return_1 have been added. Gaps that might occur due to shift targets and technical indicators to be used in subsequent steps have been cleaned using dropna(). Additionally, technical features have been added.

Sentiment scores have been calculated using feedparser from Really Simple Syndication (RSS) feed sources including Google News (bitcoin/crypto search), CryptoPanic, CoinDesk, Cointelegraph, CryptoSlate, and Finbold. Titles and publication times have been retrieved from each source. Published_at times have been converted to UTC. News articles have been filtered to the time interval intersecting the price data's time window (min_time ≤ published_at ≤ max_time). News publication times have been rounded to the nearest minute using floor('T'). If multiple news articles were available in the same minute, the average compound has been used. A left join has been performed with the price data via open_time. For sentiment analysis, a Valence Aware Dictionary and sEntiment Reasoner (VADER) model has been used to generate a sentiment score for each news headline. VADER, which works at the sentence level using a dictionary-based approach, considers emotionally charged words, negative structures, and intensifiers, especially in short and concise expressions, to produce a polarity_scores(title)['compound'] value. This value provides a single normalized score between -1 and +1. The average of the compound scores of the news headlines has been calculated for each minute, and the sentiment_score has been considered 0 for minutes without any news. To increase the numerical stability of the model, ensure faster convergence, and maintain scale fit between features, all input variables have been scaled to the 0–1 range using MinMaxScaler. Furthermore, since the distributions of the 1, 5, and 30-minute targets differed, each target has been normalized separately using its own min–max values; thus, the model has been able to learn each target in a balanced manner, and the predictions were able to be accurately returned to the actual scale.

The main reason for choosing short-term forecasting horizons of 1, 5, and 30 minutes in this study is that the proposed structure is designed not for long-term price forecasting, but for monitoring high-frequency market behaviors and supporting early warning mechanisms based on model errors.

Cryptocurrency markets, due to their 24/7 availability, high volatility, and frequent exposure to sudden regime changes, cause the loss of significant microstructural information on hourly or daily timescales. Longer time horizons can temporally mask short-term but critical price movements and sudden deteriorations in model performance.

Furthermore, the residual return-based risk and anomaly metrics (percentile-based bounds, VaR, and approximate threshold values) used in this study require a higher observation density to produce statistically significant results. In hourly or daily forecasting horizons, both the decrease in the number of observations and the increase in error accumulation limit the operational effectiveness of these metrics.

In this context, 1, 5, and 30-minute forecasting horizons;

- preserving the effects of the market microstructure,
- enabling local temporal monitoring of model errors, and
- generating early warning and risk signals without delay.

The attributes and descriptions of the dataset are given in Table 1.

Table 1. The attributes and descriptions

Attribute	Description
Open	The opening price of the relevant minute
High	The highest price recorded during that minute
Low	The lowest price recorded during that minute
Close	The closing price of the relevant minute
Volume	The trading volume (amount of BTC) in that minute
Close open	The difference between the opening and closing prices.
High low	The difference between the highest and lowest prices recorded.
SMA-15 (Simple moving average)	The arithmetic means the closing prices of the last 15 minutes.
EMA-15 (Exponential moving average)	The exponential moving average of the last 15 closing prices.
RSI-14 (Relative strength index)	14-period Relative Strength Index (0–100).
Volatility	Standard deviation of the last 15 closing prices.
Return 1	A measure of short-term momentum.
Sentiment score	Average of VADER compound sentiment scores of headings, aligned to open_time (UTC) minutes.
Target close 1	Closing price after 1 minute.
Target close 5	Closing price after 5 minutes.
Target close 30	Closing price after 30 minutes.

4. METHODOLOGY

The literature review revealed that LSTM and GRU methods are frequently used in time series problems for cryptocurrency forecasting [5,6,9,13,15]. Therefore, LSTM and GRU have been chosen as forecasting methods.

4.1. Long Short-Term Memory

LSTM, a type of recurrent deep learning algorithm, is quite significant. Its control flow is similar to that of a Recurrent Network Network (RNN) [18]. It can successfully manage information and data processing processes that propagate forward. An important feature is that, unlike traditional recurrent networks, the feedback loop in an LSTM network makes it possible to process individual data points and entire sequences. In this context, LSTM networks can be used in time series data processing and regression problems [19].

4.2. Gated Recurrent Unit

GRU is an RNN variant that combines entry and forget gates into a single update gate [20]. The update gate is responsible for determining which information should be kept, discarded, or updated with new data. The reset gate controls how much of the past information should be forgotten [21].

4.3. Transformer Architecture

The Transformer model is currently of significant importance as a revolutionary architecture in deep learning. Transformers have achieved remarkable success in various fields and have demonstrated their performance in Natural Language Processing (NLP), including language translation, sentiment analysis, and text summarization. Extending their applicability to image processing, transformers are used in areas such as image classification and object detection and are adapted for visual tasks. They are capable of capturing long-range dependencies and are frequently used in tasks such as predicting stock prices or weather patterns. The model architecture incorporates self-attention mechanisms, parallelism, and multi-heading attention. With self-attention mechanisms, the model weighs the importance of different parts of the input sequence when making predictions. The Transformer architecture is capable of highly parallelized computation, making it more efficient than sequential models. It has faster training time. The model includes an encoder and a decoder. The encoder processes the input sequence by capturing contextual information. The decoder produces an output sequence. The self-attention mechanism is extended with multiple heads, allowing the model to focus on different aspects of the input sequence simultaneously. This enhances the ability to capture complex relationships. Transformers do not naturally understand the order of the input sequence. To address this problem, positional encodings are added to the input embedded vectors, and these encodings provide position information of the tokens in the sequence [22].

5. DEVELOPMENT OF THE MODELS

The last 20% of the data has been allocated as the test set, and the first 80% as the training set. The time order has been preserved. Model input sets have been created using the sliding window approach. The last 60 minutes of price, technical indicator, and sentiment information have been used for each prediction. Sequential sequences have been generated by sliding each window by 1 minute. With this structure, models have been trained to predict closing prices 1, 5, and 30 minutes later, based on a historical 60-minute context. The basic hyperparameter values of the developed prediction models are given in Table 2.

Table 2. Hyperparameter values

Method	Values
LSTM	"Epoch": [64] "Bach_Size": [64]
GRU	"Earlystopping _Patience": [3] "Reducelronplateau_Factor": [0.5] "Reducelronplateau_ Patience": [2] "Learning_Rate": [0,0000078125 - 0.001]
Transformer Architecture	"T": [60] "F": [13] "D_Model": [128] "Dropout": [0.2] "Layernorm": [128 – 256] "Learning_Rate": [0.01] "Epoch": [20] "Bach_Size": [64]

5.1. LSTM based Bitcoin Closing Price Prediction Models

During the model development phase, separate models have been trained for each prediction (after 1 minute, 5 minutes, and 30 minutes) using the same LSTM architecture. Instead of a single multi-output, three separate models have been used for the 1-minute, 5-minute, and 30-minute targets. This allowed each target to be optimized according to its own dynamics. This architecture has been designed with a 64-cell LSTM with time information transfer (return_sequences = True) in the first layer, a 32-cell LSTM in the second layer, 20% dropout between layers to prevent overfitting, and regression estimation with a single-neuron Dense layer at the output. The models have been trained with the Adam optimization algorithm and the MSE loss function and 10% validation split have been used in the training. In addition, with the Early Stopping mechanism, training has been stopped if the validation loss did not improve for three consecutive epochs, and the best weights have been preserved. The reason we initially used a 64-cell LSTM and then a 32-cell LSTM was to first capture high-level time series patterns and then compress them with smaller cells to learn more abstract representations. Outputs have been normalized using a separate MinMaxScaler for each target, and training has been completed. The predictions generated during the testing phase were subsequently subjected to an inverse transformation to recover their corresponding original price values. This allowed for direct comparison of 1-minute, 5-minute, and 30-minute Bitcoin closing price predictions with actual prices.

5.2. GRU based Bitcoin Closing Price Prediction Models

Without making any changes to the data side, the model has been developed using the same pipeline as in LSTM, but this time with GRU, a lightweight and computationally efficient modeling method.

Since the models developed with GRU provided the most successful prediction performance, various model development attempts have been made based on it. To improve model performance, instead of a fixed 80%–20% train–test split, a progressively expanding Walk-Forward Cross-Validation (WF-CV) has been implemented to maintain temporal dependence. The WF process has been started from the first 60% of the dataset, creating a total of 4 folds. In each fold, approximately 70% of the relevant time interval has been used for training and 30% for validation. For each fold, the model has been retrained only in its own training block and tested on the validation block. If the val_loss value did not improve for two consecutive epochs, the learning rate has been halved.

In model training, the EarlyStopping mechanism has been used, and if the val_loss value did not improve for three consecutive epochs, training has been terminated and the best weights have been restored. The Huber loss function has been preferred to reduce the effect of sudden jumps. It is assumed that news sensitivity is reflected in prices gradually over time, rather than instantly. The sentiment_score variable has been smoothed using an exponentially weighted moving average, and news from the last 10 minutes has been given higher weight with a SENTIMENT_EMA_SPAN of 10. Furthermore, based on the assumption that news is reflected in prices with a delay of a few minutes, the Exponential Moving Average (EMA) output has been shifted by 5 minutes (sentiment_decay_lag), allowing the model to learn about the relationship between news and delayed price effects. The sentiment_decay and sentiment_decay_lag columns have been included in the feature set.

5.3. Transformer based Bitcoin Closing Price Prediction Models

Transformer architecture has been chosen for development because it better models long-range dependencies (self-attention) and cross-feature interactions. However, the additional data-side adjustments adversely affected the error metrics; therefore, these modifications have been reversed, and the original dataset for the GRU/LSTM structures has been retained as the baseline model. Learnable positional embedding has been employed, and 60-length position vectors have been added to the model to integrate time information into the attention mechanism. The Transformer architecture used consists of a doubly repeating encoder block, each containing a LayerNorm, sub-layers, and residual linkage structure. Each encoder block has two sub-layers: Multi-Head Self-Attention (MHA) and position-wise feed-forward network (FFN). Since the number of heads is 4, attention calculations are performed in four parallel heads, and then these outputs are concated and reprojected with W_O (batch, T, 128) to preserve the original dimension. The output of this layer is added to the input via a residual link and passed to the next stage of the encoder. The attention score formula is given in (1).

$$\textit{Attention}\ (Q, K, V) = \textit{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{1}$$

The Position-wise Feed-Forward Network (FFN) consists of a two-layer MLP applied independently at each time step. In the first stage, the input is passed through a Dense layer with 256 neurons and ReLU activation and then transformed into a higher-dimensional representation via a nonlinear transformation. A 20% dropout is then applied to prevent overlearning and increase the model's generalization capacity. In the final stage, a second Dense layer with 128 neurons is used to project the output back to the model dimension of 128. Thus, the FFN, operating independently at each time step, performs a structural transformation that increases the learning capacity of the Transformer encoder block by projecting the input into a 256-dimensional intermediate space and then back to its final 128-dimensional representation, while preserving the output dimension (batch, T, 128). In the encoder block, LayerNorm is applied to both sub-layers, and then the FFN output is summed using residual coupling to ensure stable and rapid convergence of the model along the depth. To transfer the encoder output to the regression head, a batch 128-dimensional representation is obtained by applying GlobalAveragePooling1D on the time domain, and in the final stage, a one-dimensional output predicting the price level is produced with the Dense layer. The Transformer architecture used is built on a backbone structure consisting of two encoder blocks; each block consists of LayerNorm, four-head Multi-Head Self-Attention, residual link, ReLU activated, and 0 position-wise FFN layers respectively. The training configuration includes the MSE loss function, Adam optimization, ReduceLROnPlateau which halves the learning rate when val_loss does not improve for two epochs, and the EarlyStopping mechanism which restores the best weights after three epochs of no improvement.

6. RISK MANAGEMENT

Monte Carlo scenarios have been generated for short-term price predictions; tail-risk metrics such as VaR and CVaR have been calculated. Position sizes have been dynamically adjusted using threshold and budget rules, and automatic warning mechanisms have been activated against sudden regime changes. Thus, investment decisions were made more sustainable by basing them not only on expected return but also on expected loss and tail-risk metrics. The Monte Carlo simulation has been run in the return space; prediction errors have been converted from price to residual return, and residual return values during the training period have been bootstrapped from two separate pools by regime separation (high volatility / normal volatility); in the testing phase, the pool corresponding to the volatility regime of the relevant minute has been used. Edge gating has been applied to increase cost awareness in trading decisions; positions have been not opened, or existing positions have been not maintained in cases where the round-trip transaction cost and additional safety margin have not been met. Position sizing has been done based on CVaR, and a risk budget and maximum position limit have been also defined. The integrity of the risk management process has been strengthened by generating log-based alerts for situations such as high volatility, VaR violations, and approaching the maximum position limit. With the implementation of the risk layer, the gross number of trades has been reduced, and expected loss per trade has been controlled. Edge gating eliminates unnecessary micro-trades, and tail risk is kept under control through CVaR-based sizing.

7. RESULTS AND DISCUSSION

The performance of the developed prediction models has been evaluated using RMSE, MAE and MAPE metrics. Since the problem is a regression problem, MAE, MAPE, and RMSE error metrics have been chosen as evaluation metrics. The formulas for the MAE (2) and MAPE (3) RMSE (4) metrics are given below.

$$MAE=\frac{1}{n}\sum_{i=1}^{n}|y_i - y\hat{}_i| \quad (2)$$

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i - y\hat{}_i}{y_i}\right| \quad (3)$$

$$RMSE= \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - y\hat{}_i)^2} \quad (4)$$

7.1. Results of Models Developed Using the LSTM

The most successful results are presented in Table 3. The graph of the models is given in Figure 1.

Table 3. RMSE, MAE and MAPE values of the developed models with LSTM

Prediction horizon (min)	RMSE	MAE	MAPE (%)
1	156.23	99.46	0.03
5	198.76	130.03	0.11
30	356.09	223.5	0.2

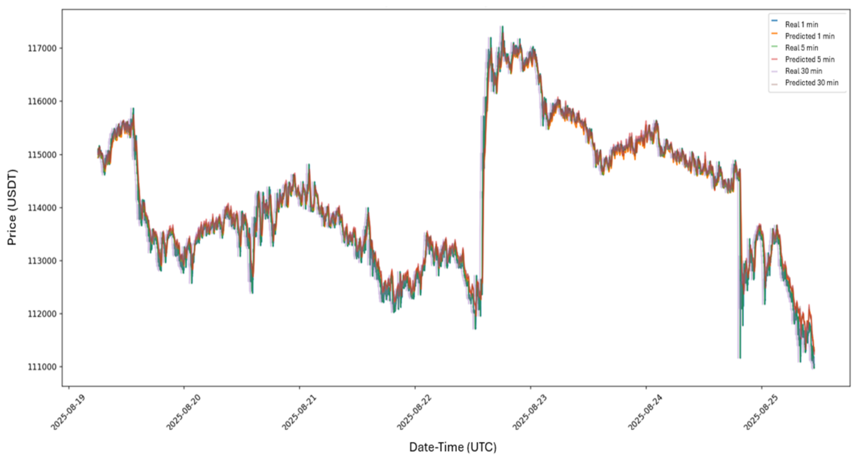

Figure 1. The real and forecast values of BTC/USDT price prediction model developed using LSTM

- It is clearly seen that the error rates decrease as the prediction horizon shortens.
- In 1-minute predictions, a high prediction performance has been achieved with RMSE (156.23), MAE (99.46), and MAPE (0.09%).
- Although the error rates increased slightly in 5-minute predictions, it can be said that the model exhibited strong performance in capturing the trend.
- In 30-minute predictions, the error values increased due to the natural volatility of the market but still remained at reasonable levels.

7.2. Results of Models Developed Using the GRU

Since the models developed with GRU provided the most successful prediction performance, various model development experiments have been carried out based on it. It has been observed that these experiments did not improve performance; on the contrary, they worsened it.

At a 1-minute prediction horizon, it has been observed that the RMSE metric produced approximately 20.4% higher error metric values, the MAE metric 33.0% higher error metric values, and the MAPE metric 28.6% higher error metric values. At a 5-minute prediction horizon, the differences decreased, with RMSE metric showing a 4.6% decrease and MAE metric a 6.0% decrease, while no change has been observed in the MAPE value. At a 30-minute prediction horizon, the differences in prediction performance between the experiments increased again. Results have been obtained that were approximately 9.1% higher in RMSE, 12.6% higher in MAE, and 15.8% higher in MAPE. In light of all this, the experiments conducted performed weaker than the basic model in terms of performance. The most successful results are presented in Table 4. The graph of the models is given in Figure 2.

Table 4. RMSE, MAE and MAPE values of the developed models with GRU

Prediction horizon (min)	RMSE	MAE	MAPE (%)
1	125.15	76.68	0.07
5	178.22	110.13	0.1
30	348.71	219.2	0.19

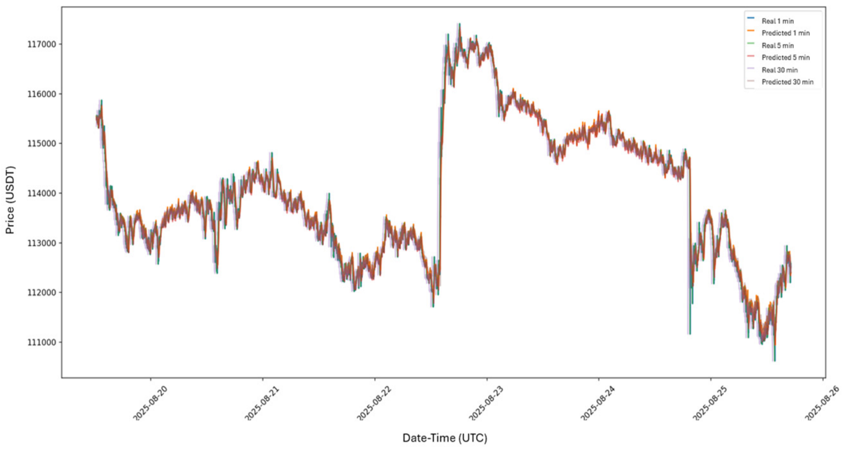

Figure 2. The real and forecast values of BTC/USDT price prediction model developed using GRU – Leakage Fixed

- The superior performance of the developed models has been achieved in the 1-minute predictions.
- The MAPE values of the models were lower than those of LSTM, indicating that model performance improved in GRU.
- These results show that GRU architecture provides significant gains in short and medium-term predictions, but only marginal improvement in the long term.
- It has been observed that these experiments did not improve performance; on the contrary, they worsened it. This increase in error metrics can be explained by several factors: the weakening of the signal's time-matching and amplitude due to the 5-minute delay of the sentiment score using EMA; and the decrease in the signal-to-noise ratio as a result of the model's capacity/regularization remaining the same despite the addition of two new features.

- The learning rate with ReduceLROnPlateau has been halved from the 3rd epoch onwards; in some periods, it dropped to very low levels with successive halvings.
- This situation led to insufficient learning, especially in the 5–30-minute horizons.

7.3. Results of Models Developed Using Transformer Architecture

The results of the developed models are presented in Table 5. The graph of the models is presented in Figure 3.

Table 5. RMSE, MAE and MAPE values of the developed models with transformer architecture

Prediction horizon (min)	RMSE	MAE	MAPE (%)
1	195.63	112.39	0.1
5	336.69	226.45	0.2
30	591.96	409.83	0.37

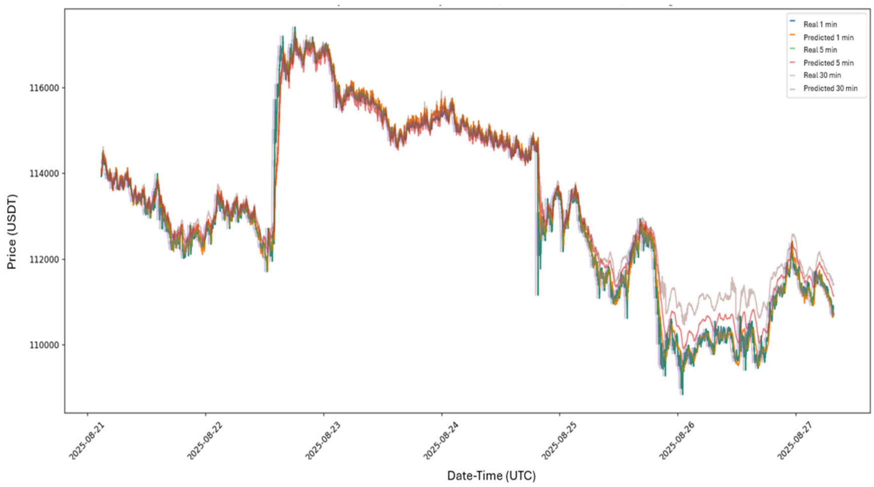

Figure 3. The real and forecast values of BTC/USDT price prediction model developed using Transformer Architecture – Leakage Fixed

- Prediction models developed using the Transformer architecture yielded poor prediction performance compared to GRU and LSTM models.
- The reasons for the Transformer model's failure are explained by readout dilution (GlobalAveragePooling1D), the absence of causal constraints (no look-ahead mask), and the lack of an optimization regimen.

7.4. Risk Management

A sample risk test has been conducted using GRU-based models, and the results obtained, along with performance and risk indicators values, are presented in Table 6. The values for Risk and Distribution Descriptors, used to evaluate risk signals, are presented in Table 7.

Table 6. Performance and risk indicators values

Performance and risk indicators	Description	Values
Cumulative PnL	Total profit or loss obtained	−0.77
Sharpe (annual)	A measure of an investment's risk-reward balance by dividing a portfolio's return by its risk	−76.09
Max Drawdown	The percentage of the maximum loss an investment experiences from a peak value before reaching a new high.	−0.77
Win-Rate	The ratio of profitable trades to total trades	%0.4
Median (edge)	The median value of returns per trade	0.052 bp
Number of trades (position change)	Number of position changes	1

Table 7. The values for risk and distribution descriptors

Risk and distribution descriptors	Description	Values
Residual return pool (n)	A fixed-length observation set consisting of the prediction errors generated by the model and covering the last n time steps.	35,903
Mean	Arithmetic means of values	0.000219 (≈ 2.19 bp)
p1 / p99 (queues)	These represent the 1% and 99% percentiles, respectively, of the residual return distribution.	−0.2037% / +0.2256%
VaR95 range (sample alerts)	A risk measure calculated from the residual return distribution represents the maximum expected margin of error at a 95% confidence level.	71.66 – 287.20
RiskBudget (minutes)	Approximate minimum return level required for the strategy to be considered to have a significant advantage (edge)	50
Edge threshold (approx.)	An approximate threshold value derived from the residual return distribution and used to identify statistically significant deviations.	≈ 30 bp (RT ~10bp + marj ~20bp)

In this round, since the dominant conditions were the edge remaining below the transaction cost and VaR95 exceeding the risk budget, the system chose to limit tail-risk by avoiding opening trades; thus, losses have been kept to a minimum while risk discipline has been maintained at a high level.

Overall, the predicted value results are in the range of [111,000–117,000], and since these are large six-digit numbers, MAPE decreases and MAE increases as the number gets larger. The most successful results have been obtained with short-term forecasting models. GRU achieved the best performance in RMSE and MAE values. LSTM, on the other hand, showed a high success rate with a MAPE value of 0.03. Transformer Architecture had a weaker forecasting performance compared to LSTM and GRU.

With the developed system,

- Investors are enabled to analyze market trends more accurately, protect themselves against sudden market changes, and increase their financial gains by making more informed investment decisions.
- A user-friendly interface has been provided with an easily accessible platform for individual and institutional investors.
- A model has been presented that can analyze large, multi-dimensional datasets such as historical price data, trading volume, market sentiment, and on-chain activity in cryptocurrency markets, and predict market movements and trends with high accuracy.

8. CONCLUSIONS

In today's economic world, cryptocurrency markets, which have emerged alongside developing technology, are experiencing very high trading volumes. However, these markets are dynamic because they depend on many different factors. This dynamism necessitates that investors take action by anticipating certain strategies. In this context, this study presents an innovative solution that will help analyze market trends and price dynamics, identify risk factors, and provide investors with predictions using big data analytics and advanced artificial intelligence algorithms. Short-term, medium-term, and long-term Bitcoin closing price prediction models have been developed using LSTM, GRU, and transformer architecture. Furthermore, Monte Carlo scenarios have been generated for short-term price forecasts; tail-risk measures such as VaR and CVaR have been calculated. Overall, the most successful results have been obtained with short-term forecasting models. GRU achieved the best performance in RMSE and MAE values. Future studies could expand this work in various ways to improve prediction performance. Firstly, different machine learning and statistics-based methods could maximize the accuracy of the study. Similarly, the use of different datasets is another important factor that could affect the applicability of the study. Additionally, exploring ensemble modeling approaches is a significant action that could be implemented in the future. Finally, risk analysis studies could be expanded to include more variables. All of these are important elements that could be covered in future studies.

9. REFERENCES

1. Almeida, J. & Gonçalves, T.C. (2023). A systematic literature review of investor behavior in the cryptocurrency markets. *Journal of Behavioral and Experimental Finance, 37*, 100785.
2. Atree, M.K. & Tripathy, N. (2025). Cryptocurrency research: Bibliometric review and content analysis. *International Review of Economics & Finance, 98*, 103940.
3. John, D.L., Binnewies, S. & Stantic, B. (2024). Cryptocurrency price prediction algorithms: A survey and future directions. *Forecasting, 6*(3), 637-671.
4. Alnami, H., Mohzary, M., Assiri, B. & Zangoti, H. (2025). An integrated framework for cryptocurrency price forecasting and anomaly detection using machine learning. *Applied Sciences*, *15*(4), 1864.
5. Asmat, G. & Maiyama, K.M. (2025). Bitcoin price prediction using N-BEATs ML technique. *EAI Endorsed Transactions on Scalable Information Systems*, *12*(2), 1-8.
6. Kaur, R., Uppal, M., Gupta, D., Juneja, S., Arafat, S.Y., Rashid, J. & Alroobaea, R. (2025). Development of a cryptocurrency price prediction model: leveraging GRU and LSTM for Bitcoin, Litecoin and Ethereum. *PeerJ Computer Science*, *11*, e2675.
7. Makri, E., Palaiokrassas, G., Bouraga, S., Polychroniadou, A. & Tassiulas, L. (2025). Ethereum price prediction employing large language models for short-term and few-shot forecasting. *arXiv preprint arXiv:2503.23190*.
8. Mazinani, A., Davoli, L. & Ferrari, G. (2025). Deep learning algorithms for cryptocurrency price prediction: A comparative analysis. *Distributed Ledger Technologies: Research and Practice*, *4*(1), 1-38.
9. Pandey, P. & Sharma, G. (2025). Effective price prediction of cryptocurrencies using CNN-based dual directional model. *Science & Technology Asia*, 201-219.
10. Kang, M., Hong, J. & Kim, S. (2025). Harnessing technical indicators with deep learning based price forecasting for cryptocurrency trading. *Physica A: Statistical Mechanics and its Applications*, *660*, 130359.
11. Saeed, N., Shafi, I., Pervez, S., Thompson, E.B., Castilla, A.K., Samad, M.A. & Ashraf, I. (2025). Intelligent decision making for commodities price prediction: Opportunities, challenges and future avenues. *Computational Economics*, 1-59.
12. Chandra, D., Tyagi, P., Gupta, R.S., Saxena, A.M. & Kharaliya, S. (2024). Cryptocurrency price prediction using machine learning. In *2024 2nd International Conference on Advancement in Computation & Computer Technologies (InCACCT), IEEE*, 258-264.
13. Golnari, A., Komeili, M.H. & Azizi, Z. (2024). Probabilistic deep learning and transfer learning for robust cryptocurrency price prediction. *Expert Systems with Applications*, *255*, 124404.
14. Kamboj, R., Jindal, A., Dubey, K. & Luthra, M. (2024). Cryptocurrency price prediction using machine learning. In *2024 4th Asian Conference on Innovation in Technology (ASIANCON), IEEE,* 1-5.
15. Kavithra, S., Girishankar, V.P., Iniyan, G. & Logesh, B. (2024). Cryptocurrency price prediction through integrated forecasting techniques. *In 2024 3rd International Conference on Artificial Intelligence For Internet of Things (AIIoT), IEEE*, 1-6.
16. Kawli, D.P., Chaudhari, A.S., Ingale, P.D., Telange, G.A. & Banik, A. (2024). Cryptocurrency price prediction using machine learning. *Asian Journal For Convergence in Technology (AJCT)*, *10*(1), 19-23.
17. Rajaei, M. & Mahmoud, Q.H. (2024). A prediction model for short price jump in cryptocurrency market. In *2024 IEEE 14th Annual Computing and Communication Workshop and Conference (CCWC), IEEE*, 0144-0150.
18. Özbek, A. (2022). Daily sea water temperature forecasting using machine learning approaches. *Cukurova University, Journal of the Faculty of Engineering*, *37*(2), 307-318.
19. Chen, X. & Long, Z. (2023). E-commerce enterprises financial risk prediction based on FA-PSO-LSTM neural network deep learning model. *Sustainability*, *15*(7), 5882.
20. Uluocak, İ. (2025). Comparative study of emission prediction using deep learning models. *Cukurova University, Journal of the Faculty of Engineering*, *40*(2), 337-346.

21. Basavarajaiah, S. & Garudachar, R. (2024). Prediction of global ionospheric TEC using attention based bidirectional long short-term memory and gated recurrent unit. *Bulletin of Electrical Engineering and Informatics*, *13*(4), 2797-2808.
22. Mozaffari, L. & Zhang, J. (2024). Predictive modeling of stock prices using transformer model. In *Proceedings of the 2024 9th International Conference on Machine Learning Technologies*, 41-48.