IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 2 January 2025, accepted 7 March 2025, date of publication 14 March 2025, date of current version 1 April 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3551307

RESEARCH ARTICLE

A Dual Output Temporal Convolutional Network With Attention Architecture for Stock Price Prediction and Risk Assessment

ARINDAM KISHOR BISWAS[1], MD SHARIFUL ALAM BHUIYAN[2], MD NAZMUL HOSSAIN MIR[3], ASHIFUR RAHMAN[4], M. F. MRIDHA[5], (Senior Member, IEEE), MD RASHEDUL ISLAM[6], (Senior Member, IEEE), AND YUTAKA WATANOBE[7], (Member, IEEE)

[1]Master's in Business Analytics, University of the Cumberlands, Williamsburg, KY 40769, USA
[2]Ketner School of Business, Trine University, Angola, MI 48101, USA
[3]Master of Science in Cybersecurity, Washington University of Science and Technology, Alexandria, VA 22314, USA
[4]Department of Computer Science and Engineering, Bangladesh University of Business and Technology (BUBT), Dhaka 1216, Bangladesh
[5]Department of Computer Science and Engineering, American International University–Bangladesh, Dhaka 1229, Bangladesh
[6]Department of Computer Science and Engineering, University of Asia Pacific, Dhaka 1205, Bangladesh
[7]School of Computer Science and Engineering, The University of Aizu, Aizuwakamatsu 965-8580, Japan

Corresponding authors: M. F. Mridha (mridha@aiub.edu) and Md Rashedul Islam (rashed.cse@gmail.com)

ABSTRACT In this paper, we propose a novel deep learning model that integrates a Temporal Convolutional Network (TCN) with an Attention mechanism to predict stock prices and assess risk for MasterCard (MA) and Visa (V). The model is designed with a dual output to forecast future stock prices (Open, Close, High, Low) while simultaneously predicting risk metrics, including volatility and the Sharpe Ratio. By utilizing dilated convolutions, the TCN captures both short-term and long-term dependencies in the stock price data, and the attention layer focuses on critical time steps for enhanced predictive accuracy. On a dataset covering over 15 years (2008-2024), the TCN with attention model achieved a mean absolute error (MAE) of 1.23 for MasterCard and 1.45 for Visa, outperforming LSTM and ARIMA baselines. Additionally, the model demonstrated strong performance in predicting risk metrics, with an MAE of 0.012 for volatility and 0.065 for the Sharpe Ratio. Backtesting on unseen data further validated the model's robustness, with a backtest MAE of 1.25 for MasterCard and 1.50 for Visa. This dual-output architecture provides an accurate and interpretable solution for both stock price forecasting and risk assessment, offering a valuable tool for investors and risk managers.

INDEX TERMS Stock price prediction, temporal convolutional network (TCN), attention mechanism, dual output model, financial time series, risk assessment, deep learning in finances.

I. INTRODUCTION

Stock price prediction and risk assessment are two pivotal tasks in the field of financial markets. Accurate predictions can empower investors to make informed decisions, while effective risk management helps mitigate potential losses [1], [2], [3]. The inherent complexities of financial time series data, including nonlinearity, noise, and long-term dependencies, make these tasks highly challenging [4], [5]. Over the years, various methods have been explored to address these challenges, including statistical models such as AutoRegressive Integrated Moving Average (ARIMA) [6] and machine learning models like Long Short-Term Memory (LSTM) [7], [8]. However, these models often struggle to capture long-term dependencies and non-stationary patterns in stock data, leading to suboptimal performance in forecasting stock prices and evaluating risk metrics [9], [10].

The motivation for choosing Temporal Convolutional Networks (TCN) stems from their ability to address limitations of recurrent architectures like LSTM and GRU,

The associate editor coordinating the review of this manuscript and approving it for publication was Janmenjoy Nayak.

particularly in capturing long-range dependencies without the inefficiencies of recurrent connections. Financial time series data often exhibit complex patterns with dependencies across various temporal scales, which TCNs model effectively through dilated convolutions. Additionally, the efficiency and parallelization capabilities of TCNs are crucial for handling large datasets in financial applications. To further enhance predictive accuracy and interpretability, we integrate an attention mechanism, which selectively focuses on critical time steps in the sequence, addressing the noise and nonlinearity inherent in financial data. Together, TCN and attention provide a robust framework tailored to the challenges of stock price prediction and risk assessment.

TCN have emerged as a powerful alternative to recurrent architectures like LSTM and GRU for sequence modeling tasks [11], [12]. Unlike traditional recurrent networks, TCNs use dilated convolutions to capture long-range dependencies without the need for recurrent connections, enabling faster training and the ability to model longer temporal relationships. Furthermore, the parallelization capabilities of TCNs allow them to be computationally efficient, a critical factor when dealing with large datasets in financial applications [13], [14]. Recent research has shown that including attention mechanisms can enhance the performance of deep learning models by enabling them to focus on the most relevant portions of the input sequence, improving both interpretability and predictive accuracy [15], [16].

In this paper, we propose a novel deep learning architecture that combines the strengths of TCN with an attention mechanism [17] to predict stock prices and evaluate risk for two major financial service companies, MasterCard (MA) and Visa (V). Our model incorporates a dual output structure, where one output forecasts future stock prices, and the other predicts risk metrics, such as volatility and the Sharpe Ratio. By utilizing the dilated convolutions of TCN, our model is capable of capturing both short-term and long-term dependencies in the stock data, while the attention mechanism focuses on the most critical time steps for accurate predictions.

The key contributions of this paper are as follows:

- We propose a state-of-the-art hybrid architecture combining Temporal Convolutional Networks with an attention mechanism for stock price prediction, which improves performance by focusing on critical time steps in the stock data.
- We introduce a dual-output model that simultaneously predicts both stock prices and risk metrics (volatility and the Sharpe Ratio), enabling comprehensive financial forecasting.
- We demonstrate the superiority of our model over traditional models such as LSTM and ARIMA, achieving a mean absolute error (MAE) of 1.23 for MasterCard and 1.45 for Visa in stock price prediction, and a MAE of 0.012 for volatility and 0.065 for the Sharpe Ratio in risk evaluation.
- We validate the robustness of the model by performing backtesting on unseen data, showing its capability to generalize well in real-world market conditions.

To the best of our knowledge, this is the first application of a TCN with attention in a dual-output architecture for simultaneous stock price prediction and risk assessment. We test our model using 15 years of historical stock price data (2008-2024) for MasterCard and Visa. Our results demonstrate that the TCN with attention model significantly outperforms traditional LSTM and ARIMA models. Specifically, the proposed model achieves an MAE of 1.23 for MasterCard and 1.45 for Visa in stock price prediction, while also demonstrating strong performance in risk forecasting, with an MAE of 0.012 for volatility and 0.065 for the Sharpe Ratio. Furthermore, backtesting on unseen data highlights the robustness of our model in real-world scenarios.

The remainder of this paper is organized as follows. Section II reviews the related work in stock price prediction and risk assessment using deep learning techniques. Section III describes data preprocessing and the proposed TCN with attention model and the dual output architecture. Section IV presents the experimental setup and results. Finally, Section V presents discussions and Section VI concludes the paper.

II. RELATED WORKS

The challenge of stock price prediction and risk assessment has long been studied, with early approaches rooted in traditional statistical methods [18], [19]. Over time, as financial markets became increasingly complex and volatile, machine learning (ML) and, more recently, deep learning (DL) techniques have emerged as powerful alternatives [20], [21]. This section reviews the evolution of these methods and highlights their strengths and limitations, providing the foundation for our proposed approach.

Traditional approaches for stock price prediction have primarily relied on statistical techniques such as AutoRegressive Integrated Moving Average (ARIMA), Exponential Smoothing (ETS), Prophet, and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models [22], [23], [24]. In one study, the Facebook Prophet and ARIMA models were applied to forecast stock prices from 2012 to 2020, emphasizing the importance of pre-processing techniques and trend analysis. The ARIMA (2,1,2) model demonstrated high accuracy but was limited by the scope of data and market volatility [25]. Another study compared ARIMA with Artificial Neural Networks (ANN), concluding that ANN generally outperforms ARIMA in forecasting accuracy, though its complexity can lead to overfitting, especially in volatile markets [26]. For the Colombo Stock Exchange, Backpropagation Neural Networks (BPNN) outperformed ARIMA, with an MAE of 4.708423 compared to ARIMA's 29.6975. This study highlighted the "Bank and Finance" sector as a key influence, though it faced limitations due to non-stationarity and focus on specific timeframes (2008-2017) [27].

Hybrid models such as ARIMA-MLP and MLP-ARIMA have also been explored, with the latter showing superior performance in stock forecasting accuracy. These models demonstrated improvements in key metrics like MAE, MSE, and RMSE, but faced challenges with dataset specificity and model complexity [28]. Finally, sector-specific analyses using ARIMA models across Indian stocks revealed consistent accuracy, though generalizability was limited by the reliance on statistical assumptions and dataset specificity [29].

ML models have been extensively used in stock price prediction due to their ability to handle non-linear relationships and complex patterns in financial data [30], [31]. Bansal et al. [30] compared various ML models, including K-Nearest Neighbors (KNN), Linear Regression, Support Vector Regression (SVR), Decision Tree Regression, and LSTM for stock price forecasting. SVR achieved the best performance, with a Symmetric Mean Absolute Percentage Error (SMAPE) of 1.59. However, the study's limitations included a small dataset and a limited timeframe, impacting its broader applicability.

In a separate study, Emioma et al. [32] applied least-squares linear regression to predict stock prices, finding that while the model captured seasonal variations, it struggled with MAPE issues due to zero values in the dataset. Similarly, Ampomah et al. [31] focused on tree-based ensemble methods and found that AdaBoost performed well, achieving high AUC scores, though the study's restriction to tree-based models limits its generalizability. Nassif et al. [33] highlighted KNN as the most accurate model for stock price prediction, with the lowest MAE, outperforming traditional models like Multiple Linear Regression (MLR). Lastly, Elminaam et al. [34] demonstrated that KNN, Random Forests (RF), LR, and Gradient Boosting (GB) effectively predicted closing prices, with low RMSE and MAE scores across different sectors.

In recent years, DL models have become increasingly popular for stock price prediction, as they can automatically learn intricate non-linear patterns from raw data, eliminating the need for extensive feature engineering [35], [36]. Recurrent Neural Networks (RNNs) [37], particularly LSTM networks [38], have shown great promise in capturing temporal dependencies in stock prices. Shaban et al. [39] introduced SMP-DL, a hybrid model combining LSTM and Bidirectional Gated Recurrent Units (BiGRU), achieving an RMSE of 0.2883 and R^2 of 0.9948. However, its complexity may hinder training efficiency, and its reliance on historical data limits predictive accuracy and generalizability across markets. In another study, Sharaf et al. [40] presented StockPred, a framework utilizing various ML and DL models, where the CNN model achieved superior accuracy with an R^2 of 0.95 and RMSE of 1.2. Nonetheless, the model faces challenges with overfitting and potential generalizability issues across different datasets.

Billah et al. [41] compared moving average techniques with LSTM, finding LSTM superior, but faced challenges in prediction reliability for long-term forecasting. The author of the paper [42] proposed a hybrid model using a single-layer RNN, outperforming CNN and LSTM, yet noted inefficiencies in CNN feature extraction. Similarly, Gülmez et al. [43] introduced the LSTM-ARO model, optimized with the Artificial Rabbits Optimization algorithm, achieving low MAE scores. However, it struggles with prediction variability across stocks. Finally, Chowdhury et al. [44] compared various deep learning models with ARIMA, emphasizing deep learning's ability to capture non-linearity but highlighting issues like high computational demands and overfitting risks.

TCNs have emerged as a viable alternative for time series forecasting, particularly in financial applications. Recent studies have demonstrated their effectiveness when combined with attention mechanisms, allowing models to focus on critical time steps for better predictions. Li et al. [45] introduce the CEEMDAN-TCN-GRU-CBAM model, which integrates advanced techniques for improved stock index forecasting. Evaluated on indices from emerging and developed markets, the model outperforms benchmarks, demonstrating enhanced forecasting accuracy, efficiency, and robustness. However, it is sensitive to hyperparameter selection, requiring empirical tuning. Chen et al. [46] proposed a two-stage attention-based hybrid model that combines TCNs with LSTM networks, achieving an average prediction accuracy of 89.35%. While effective for long-term forecasts, this model faces limitations in short-term performance and dataset dependency, which can complicate implementation and increase computational demands. Similarly, Zhang et al. [47] integrated financial news headlines with stock price data using an attention mechanism, resulting in an accuracy of 85%. However, this model's focus on short-term predictions and reliance on high-quality news data restricts its applicability to minor price movements.

Li et al. [48] introduced a hybrid model that combines TCNs with a Channel Enhanced Attention mechanism for short-term load forecasting, achieving mean square errors of 0.056 for 24-hour forecasts. Despite this success, the model struggles with capturing local features and generalizing across different datasets. Dai et al. [49] applied TCNs to ultra-high-frequency transaction price data, demonstrating improved prediction accuracy compared to traditional models; however, the findings may lack generalizability and are limited in comparison to other advanced models. Deng et al. [50] developed a Knowledge-Driven TCN (KDTCN) for stock trend prediction, enhancing responsiveness to abrupt market changes but needing broader evaluations. Fan et al. [51] introduced the PSTA-TCN framework, which enhances multivariate time series predictions through parallel spatiotemporal attention, though it faces challenges with long-term dependencies. Lastly, Wan et al. [52] presented the Multivariate Temporal Convolutional Attention Network (MT-CAN), which enhances feature extraction and outperforms models like LSTM, but its complexity and data quality dependence remain limitations.

Liu et al. [53] introduces a channel-temporal dual attention module (CTAM) and a market preference perception module, significantly improving stock prediction performance when combined with TCN. The Information Retrieval Rate (IRR) shows up to 57.8% improvement over TCN alone. However, model limitations and data selection concerns suggest further enhancements are needed. The CL-TCN [54] model combines contrastive learning with temporal convolutional networks, improving implied volatility predictions by enhancing feature differentiation and accuracy over traditional models. Results show it closely tracks volatility trends with linear regression, especially on large datasets. However, the model still faces challenges with high-frequency prediction efficiency and limited impact on prediction time.

Dual output models have become increasingly popular for simultaneously predicting stock prices and assessing risk. These models provide more comprehensive insights, combining price forecasts with risk evaluations, which is essential for making informed decisions in volatile financial markets. Gu et al. [55] introduced the DIA-LSTM model, which employs feature and temporal attention mechanisms with dynamic meteorological data for agricultural price prediction. The model achieved a MAPE reduction of 2.8%-5.5%, with a MAPE of 4.39 for cabbage prices. Despite its success, the model's complexity and reliance on dynamic data limit its broader applicability in stock markets.

Duan et al. [56] proposed a deep neural network (DNN) approach for risk assessment in peer-to-peer (P2P) lending, utilizing a large feature set and one-hot encoding. Their model achieved 93% accuracy on over 277,000 observations. However, the misclassification rate of up to 0.04 and reliance on outdated data present challenges for future applications in dynamic market environments. Protasiewicz et al. [57] developed two neural systems for electricity demand forecasting, incorporating Multilayer Perceptron (MLP) networks and a Self-Organizing Map (SOM)-MLP hybrid. The models achieved a maximum error of 1.75%, but struggled on atypical days, with errors increasing over time. These findings highlight the limitations of using static models in dynamic markets. Wang et al. [58] applied convolutional LSTM models to predict price and risk in the Chinese financial derivatives market, achieving an 11.5% reduction in Mean Squared Error (MSE). However, the model's reliance on data quality and limited indicators may hinder its generalizability across different financial markets.

Table 1 summarizes the key findings of relevant studies on stock price prediction, outlining the models employed, their year of publication, and the main insights derived from each study.

In summary, traditional models such as ARIMA and GARCH have laid the foundation for time series analysis in finance but are limited in their ability to handle complex, non-linear relationships. Machine learning models have improved upon this, yet they rely heavily on feature engineering and lack the ability to capture temporal dependencies. Deep learning models, particularly RNNs like LSTM, have demonstrated their effectiveness in stock price prediction but remain computationally expensive and struggle with long sequences. TCNs and attention mechanisms offer a promising solution by enabling parallel computation and selectively focusing on the most critical time steps. Despite the advances in price prediction models, there is a notable lack of research combining TCN and attention mechanisms in a dual output model for both price prediction and risk assessment. This paper addresses this gap by proposing a novel architecture that predicts both stock prices and risk metrics, providing a comprehensive solution for financial forecasting.

III. METHODOLOGY

In this section, we present the data preprocessing and the proposed architecture combining TCN with an Attention mechanism to predict stock prices and evaluate risk metrics. Our model is designed to output both future stock prices and risk metrics, such as volatility and the Sharpe Ratio, using a dual output framework. We begin by describing the TCN, followed by the attention mechanism, and conclude with the dual-output design, as shown in Figure 1. Finally, we summarize the overall methodology in a step-by-step algorithm.

A. DATA PREPROCESSING

Effective data preprocessing is essential for building reliable and precise predictive models in financial time series analysis. In this section, we outline the preprocessing steps applied to the historical stock price data of MasterCard (MA) and Visa (V) from 2008 to 2024 (Collected from Kaggle [59]). The primary tasks include handling missing values, scaling the data, generating additional features, and preparing the dataset for time series modeling.

1) MISSING VALUE HANDLING

Financial datasets often contain missing values due to holidays, trading suspensions, or incomplete records. To handle missing values in our dataset, we employed forward filling, which replaces a missing value with the last available observation. Mathematically, if x_t is the stock price at time step t and it is missing, the imputed value x'_t is given by:

$$x'_t = x_{t-1}, \quad \text{if } x_t \text{ is missing.} \tag{1}$$

For time periods where no prior data is available (e.g., at the start of the time series), we initialize missing values with the mean of the available values in the dataset. This approach ensures that the time series remains continuous for subsequent analysis.

2) DATA NORMALIZATION

To ensure that all features contribute equally to the learning process and to improve model convergence, we normalize the stock prices and risk metrics using Min-Max scaling. Let x_t

TABLE 1. Some previous work on stock price prediction is critically summarized in this table.

Ref.	Year	Model	Contribution	Limitation
[39]	2024	LSTM and Bi-GRU	This SMP-DL model integrates LSTM and BiGRU for enhanced stock price prediction, achieving an RMSE of 0.2883 and R^2 of 0.9948 across datasets like IBM, Google, and Apple, outperforming traditional methods in accuracy.	This model's complexity may require significant computational resources, and its reliance on historical data limits adaptability to unforeseen market changes, while generalizability across different markets remains unverified.
[41]	2024	LSTM	This paper analyzes SMA, EMA, and LSTM for stock price prediction, demonstrating LSTM's superiority with an RMSE of 12.312 and MAPE of 2.06%. The study provides insights into the effectiveness of these models, achieving an average prediction accuracy of 89.7% in financial market predictions.	This study acknowledges the inherent unpredictability of stock markets, limiting the reliability of any predictive model. LSTM is less effective for long-term price movements compared to SMA and EMA, highlighting a constraint in its application for long-term forecasting scenarios.
[47]	2024	LSTM-BERT	This paper presents a novel model that combines financial news and stock price data through an attention mechanism, achieving an accuracy of 85% and outperforming benchmarks like LSTM-BERT in precision and recall. This integration enhances market movement predictions, providing valuable insights for short-term trading strategies.	This model primarily focuses on short-term predictions, limiting its applicability for long-term strategies. Its performance is heavily reliant on the quality of news data, as inaccurate or outdated headlines can adversely affect predictions. Additionally, it overlooks minor price movements, which may restrict opportunities for certain traders.
[48]	2024	TCN	This study introduces a hybrid model integrating Temporal Convolutional Networks with a Channel Enhanced Attention mechanism, achieving mean square errors of 0.056 for 24-hour predictions and 0.146 for weekly forecasts. This model outperforms baseline models, enhancing prediction accuracy by effectively capturing both short-term and long-term dependencies.	Despite improvements, the model struggles to fully capture local features in predictive curves, indicating a need for further enhancements. Additionally, while effective on the tested dataset, the model's generalizability to other datasets or contexts remains uncertain, and its complex implementation may require extensive tuning for optimal performance.
[45]	2024	CEEMDAN-TCN-GRU-CBAM	This paper introduces the CEEMDAN-TCN-GRU-CBAM model, integrating advanced techniques to improve stock index forecasting accuracy. Evaluated on indices from China, India, America, and Japan, the model outperformed benchmarks, demonstrating enhanced efficiency, robustness, and reliability across various markets.	This model is sensitive to hyperparameter selection, affecting performance. Further improvements could be made by exploring alternative decomposition methods, optimization algorithms, and attention modules to enhance prediction accuracy. Additionally, future applications could extend to other predictive areas like wind speed and traffic flow.
[53]	2023	TCN with attention	This paper introduces a channel-temporal dual attention module (CTAM) and a market preference perception module, enhancing stock prediction performance. By combining these modules with TCN, prediction accuracy improves significantly, with the Information Retrieval Rate (IRR) showing notable improvements of up to 57.8% over TCN alone.	Despite improvements, deep neural networks for stock prediction still face constraints, such as inherent model limitations. Additionally, concerns about data selection and the applicability of hybrid models for multi-source data highlight the need for more effective feature mining, especially beyond stock price data, for better prediction performance.
[43]	2023	LSTM-ARO	This paper introduces the LSTM-ARO model, optimized with the Artificial Rabbits Optimization algorithm, achieving superior performance with the highest R^2 scores and lowest MAE across various stock tickers, demonstrating its effectiveness in stock price predictions compared to traditional models.	This LSTM-ARO model exhibits prediction accuracy variability across different stocks and relies heavily on historical data, which may overlook sudden market shifts. Additionally, the complexity of integrating the ARO algorithm with LSTM could pose challenges for practical implementation in real-world scenarios.
[42]	2023	RNN	This paper presents a hybrid deep learning model utilizing a single-layer RNN to forecast close and high stock prices, outperforming CNN and LSTM models. The RNN achieved a lower MAE of 1.23 and RMSE of 1.54, with an R^2 value of 0.95, indicating high prediction accuracy.	A key limitation of this study is the CNN's inability to effectively extract optimal features from input data, which may hinder performance. Additionally, while the RNN reduces computational complexity, future work must broaden performance evaluation metrics to include accuracy, MSE, F1 score, precision, and recall for a comprehensive assessment.
[30]	2022	SVM	This paper compares K-Nearest Neighbors, Linear Regression, Support Vector Regression, Decision Tree Regression, and LSTM for stock price prediction, highlighting Support Vector Regression's superior performance with a SMAPE of 1.59 and an $R^2$of -0.11, demonstrating the effectiveness of these algorithms in stock market forecasting.	This study is limited by its focus on only five algorithms, potentially excluding other effective methods. Additionally, the dataset comprises only twelve Indian companies, which may not reflect broader market dynamics. The historical data from 2015 to 2021 may also hinder the relevance of predictions in a changing market.
[40]	2021	CNN	This paper introduces StockPred, a framework for stock price prediction using various ML and DL models. The CNN model outperformed others, achieving R^2 values of 0.95, RMSE of 1.2, and MAPE of 5%, demonstrating the effectiveness of deep learning in enhancing prediction accuracy.	This study acknowledges the challenges of overfitting and underfitting due to inappropriate hyperparameters. It also notes issues with excessive feature extraction, leading to increased processing times without significant accuracy improvements. Additionally, the results may not generalize across all stock datasets, as model performance can vary based on data characteristics.
[27]	2020	ARIMA and BPNN	This paper compares ARIMA and BPNN for stock price forecasting, finding BPNN superior with an MAE of 4.71 versus ARIMA's 29.70. It highlights the "Bank and Finance Sector" as key and emphasizes data preprocessing, including PCA and K-means clustering for improved analysis.	This study faces non-stationarity issues in the dataset, requiring differencing for stabilization. It focuses on the 2008-2017 period, potentially missing recent market dynamics. The findings may not generalize to other stock markets due to varying behaviors across different regions and timeframes.

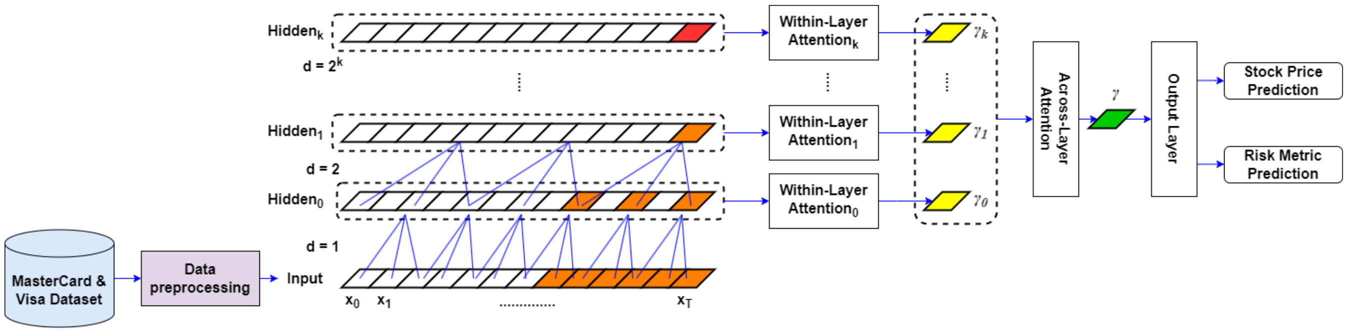


FIGURE 1. Architecture of our proposed stock price prediction and risk assessment system.

represent a stock price at time step t, and $x_{\min}$ and $x_{\max}$ be the minimum and maximum values observed in the dataset. The normalized stock price x'_t is computed as:

$$x'_t = \frac{x_t - x_{\min}}{x_{\max} - x_{\min}}. \tag{2}$$

This transformation ensures that all stock price features are rescaled to the range [0, 1], preventing large values from dominating the learning process.

3) FEATURE ENGINEERING

To improve model performance, we generate additional features based on the stock price data. One common approach is to compute the daily returns, which capture the relative change in stock prices between consecutive time steps. The daily return r_c for a given stock at time c is defined as:

$$r_c = \frac{x_c - x_{c-1}}{x_{c-1}} = \frac{\Delta x_c}{x_{c-1}}, \tag{3}$$

where $\Delta x_c = x_c - x_{c-1}$ represents the price change. Daily returns provide important insights into stock volatility and are used as features in our predictive model.

In addition to returns, we compute the moving average over different windows to capture short-term and long-term trends in stock prices. The moving average for stock prices over a window of size w is given by:

$$\mathrm{MA}_t^w = \frac{1}{w}\sum_{i=0}^{w-1} x_{t-i}, \tag{4}$$

where MA_t^w represents the moving average at time step t and w is the window size. In our experiments, we calculate moving averages over 7-day, 30-day, and 90-day windows to capture both short-term and long-term trends.

4) VOLATILITY CALCULATION

Volatility is a key indicator for assessing risk and is typically calculated as the standard deviation of stock returns over a defined time period. Let r_{t-i} represent the return at time step $t-i$, and the volatility σ_t^w over a window size w is calculated as:

$$\sigma_t^w = \sqrt{\frac{1}{w}\sum_{i=0}^{w-1}\left(r_{t-i} - \overline{r_t^w}\right)^2}, \tag{5}$$

where $\overline{r_t^w}$ is the mean of the returns over the window of size w. In our model, we use a 30-day window to compute volatility, which provides a measure of how much the stock prices fluctuate over time.

5) DATA PREPARATION FOR TIME SERIES MODELING

To prepare the data for time series forecasting, we reshape the data into overlapping sequences of fixed length. Given a sequence length L, the input sequence at time t consists of the stock prices and engineered features from $t-L$ to $t-1$, and the target output is the stock price at time t. Formally, the input sequence $\mathbf{X}_t$ and the corresponding target y_t are defined as:

$$\mathbf{X}_t = [x_{t-L}, x_{t-L+1}, \ldots, x_{t-1}], \tag{6}$$
$$y_t = x_t. \tag{7}$$

We use a sliding window approach to generate multiple input-output pairs from the historical stock price data. For example, if the sequence length $L = 30$, then the input consists of the previous 30 days of stock prices, and the output is the stock price on the next day.

6) TRAIN-TEST SPLIT

Finally, we divide the data into training and testing sets to assess the model's performance. A time-based split is used, where the model is trained on data from earlier time periods and tested on subsequent periods, reflecting a real-world stock price forecasting scenario. Let T_{train} and T_{test} represent the training and testing durations, respectively. The model is trained on the data in the interval $[0, T_{\text{train}}]$ and evaluated on the interval $[T_{\text{train}} + 1, T_{\text{test}}]$.

B. TEMPORAL CONVOLUTIONAL NETWORKS (TCN)

TCNs are convolution-based models tailored for sequence modeling, aimed at capturing both short-range and long-range dependencies within time series data. Unlike RNNs such as LSTM and GRU, TCNs use dilated convolutions that allow the model to have a large receptive field without increasing the number of parameters significantly.

1) DILATED CONVOLUTIONS

Given an input sequence $\mathbf{X} = [x_1, x_2, \ldots, x_T]$, where T is the sequence length, the output of a 1D convolution with filter size k, dilation rate d, and weights $\mathbf{W} = [w_1, w_2, \ldots, w_k]$ at time step t is given by:

$$y_t = \sum_{i=1}^{k} w_i \cdot x_{t-d\cdot(i-1)}. \tag{8}$$

Here, the dilation rate d controls the spacing between the elements in the input sequence used by the convolution, allowing the receptive field to grow exponentially as the number of layers increases. This enables the TCN to capture long-term dependencies without increasing the depth of the network.

2) RESIDUAL CONNECTIONS

To avoid vanishing gradient problems and improve the flow of information through deep TCN architectures, we employ residual connections. The residual connection at layer l is defined as:

$$y_t^{(l)} = f(\mathbf{X}_t^{(l)}) + \mathbf{X}_t^{(l-1)}, \tag{9}$$

where $f(\mathbf{X}_t^{(l)})$ is the output of the dilated convolution at layer l, and $\mathbf{X}_t^{(l-1)}$ is the input to the previous layer. The residual connection allows the network to directly pass input information across layers, facilitating the learning of both short-term and long-term dependencies.

C. ATTENTION MECHANISM

The attention mechanism [17] allows the model to concentrate on the most important time steps in the input sequence, improving its capability to generate precise predictions. After processing the input through the TCN layers, we apply attention to the output of the final TCN layer.

Let the output of the TCN at each time step be denoted as $\mathbf{h}_t$. The attention mechanism computes a context vector $\mathbf{c}$, which is a weighted sum of the hidden states $\mathbf{h}_t$ across all time steps. The attention weights α_t are computed as follows:

$$e_t = \mathbf{v}^\top \tanh(\mathbf{W}_h \mathbf{h}_t + \mathbf{b}_h), \tag{10}$$

$$\alpha_t = \frac{\exp(e_t)}{\sum_{i=1}^{T} \exp(e_i)}, \tag{11}$$

$$\mathbf{c} = \sum_{t=1}^{T} \alpha_t \mathbf{h}_t. \tag{12}$$

Here, $\mathbf{v}$, $\mathbf{W}_h$, and $\mathbf{b}_h$ are learnable parameters, e_t is the unnormalized attention score, and α_t is the normalized attention weight for each time step t. The context vector $\mathbf{c}$ captures the most important information from the entire sequence and is used as input for the final prediction layers.

D. DUAL OUTPUT MODEL

Our model is designed with a dual output structure, where one output predicts future stock prices and the other predicts risk metrics such as volatility and the Sharpe Ratio. This allows the model to simultaneously provide both price predictions and risk assessments, making it highly useful for financial forecasting.

1) STOCK PRICE PREDICTION

The first output layer predicts future stock prices, including Open, Close, High, and Low prices. The predicted price $\hat{y}_t$ at time step t is computed as:

$$\hat{y}_t = \mathbf{W}_y^\top \mathbf{c} + b_y, \tag{13}$$

where $\mathbf{W}_y$ and b_y are the learnable parameters for the stock price prediction output layer, and $\mathbf{c}$ is the context vector generated by the attention mechanism.

2) RISK METRIC PREDICTION

The second output layer predicts risk metrics, specifically volatility and the Sharpe Ratio. The predicted volatility $\hat{\sigma}_t$ and Sharpe Ratio $\hat{S}_t$ are computed as:

$$\hat{\sigma}_t = \mathbf{W}_\sigma^\top \mathbf{c} + b_\sigma, \tag{14}$$

$$\hat{S}_t = \mathbf{W}_S^\top \mathbf{c} + b_S, \tag{15}$$

where $\mathbf{W}_\sigma$, $\mathbf{W}_S$, b_σ, and b_S are the learnable parameters for the risk metric prediction layers.

3) LOSS FUNCTION

To train the dual output model, we minimize a weighted sum of the losses for stock price prediction and risk metric prediction. The total loss $\mathcal{L}$ is given by:

$$\mathcal{L} = \alpha \cdot \mathcal{L}_{\text{price}} + \beta \cdot \mathcal{L}_{\text{risk}}, \tag{16}$$

where $\mathcal{L}_{\text{price}}$ is the loss for stock price prediction, $\mathcal{L}_{\text{risk}}$ is the loss for risk metric prediction, and α and β are weights that control the relative importance of each task. We use MSE as the loss function for both tasks:

$$\mathcal{L}_{\text{price}} = \frac{1}{N} \sum_{t=1}^{N} (\hat{y}_t - y_t)^2, \tag{17}$$

$$\mathcal{L}_{\text{risk}} = \frac{1}{N} \sum_{t=1}^{N} (\hat{\sigma}_t - \sigma_t)^2 + \frac{1}{N} \sum_{t=1}^{N} \left(\hat{S}_t - S_t\right)^2. \tag{18}$$

E. ALGORITHM: TCN WITH ATTENTION AND DUAL OUTPUT

The integration of the Temporal Convolutional Network with the attention mechanism is a key component of the proposed model. The TCN is employed as the foundational architecture for temporal feature extraction, leveraging dilated causal convolutions to capture both short-term and long-term dependencies in the stock price series. The output of the TCN is then passed through the attention mechanism, which assigns dynamic weights to the temporal features. Specifically, the attention mechanism computes a context vector by generating alignment scores for each time step, enabling the model to focus on the most relevant temporal patterns. Mathematically, the alignment scores are calculated as Eq. 11 and 12.

This context vector, enriched with attention-modulated temporal information, is then passed to the output layers for stock price and risk metric prediction. This integration ensures that the model dynamically prioritizes influential time steps, significantly improving predictive accuracy. The algorithmic implementation highlights how the TCN's ability to process temporal data is enhanced by the attention mechanism's capacity to focus on critical time dependencies, offering a robust solution for financial forecasting. The complete architecture is summarized in Algorithm 1.

F. TRAINING PROCESS

We train the proposed TCN with an attention model with a dual output architecture using a time-based split of the data. The training set consists of the historical stock prices and risk metrics for the first portion of the dataset, while the remaining data is reserved for testing and backtesting. The model parameters $\mathbf{W}_y$, $\mathbf{W}_\sigma$, $\mathbf{W}_S$, and others are updated using backpropagation with the Adam optimizer [60], a variant of stochastic gradient descent that adapts the learning rate during training.

The training objective is to minimize the total loss $\mathcal{L}$, as defined in Equation (16), by adjusting the learnable parameters. We monitor the model's performance on both the price prediction and risk prediction tasks, ensuring that the dual output architecture optimally balances both objectives.

G. HYPERPARAMETER TUNING

The performance of the TCN with attention model is highly dependent on several hyperparameters. We perform a grid search to optimize key parameters, including:

- **Filter size** k for the dilated convolution layers.
- **Dilation rate** d to control the receptive field.
- **Number of residual layers** L in the TCN.
- **Attention vector size.**
- **Learning rate** η for the Adam optimizer.
- **Batch size** B during training.

Although grid search involves a computational cost, it is essential to ensure the optimal performance of the model. Given the complexity of stock price prediction and risk metric forecasting tasks, fine-tuning hyperparameters through grid search allows us to balance the trade-off between computational expense and model accuracy effectively. Additionally, the grid search process is performed once during model development, and the selected hyperparameters can be reused across similar datasets and applications, reducing the need for repeated tuning.

By evaluating the model's performance on the validation set, we select the hyperparameters that minimize the loss for both stock price prediction and risk metric forecasting. The explored hyperparameter space, along with the best-fitted values, is presented in Table 2.

Algorithm 1 TCN With Attention and Dual Output for Stock Price Prediction and Risk Metrics

1: **Input:** Stock price data $\mathbf{X} = \{x_1, x_2, \ldots, x_T\}$, risk metrics (volatility σ, Sharpe Ratio S)

2: **Output:** Predicted stock prices $\hat{y}_t$, predicted risk metrics $\hat{\sigma}_t$, $\hat{S}_t$

3: Normalize the stock price data $\mathbf{X}$ using Min-Max scaling:

$$\mathbf{X}_{\text{norm}} = \frac{\mathbf{X} - \min(\mathbf{X})}{\max(\mathbf{X}) - \min(\mathbf{X})}$$

4: Generate additional features: daily returns $r_t = \frac{x_t - x_{t-1}}{x_{t-1}}$, moving averages $MA_k = \frac{1}{k}\sum_{i=t-k+1}^{t} x_i$.

5: Initialize TCN with L residual layers, filter size k, and dilation rate d.

6: **for** each time step $t = 1, \ldots, T$ **do**

7: Apply dilated convolution:

$$\mathbf{h}_t = f_{\text{TCN}}(\mathbf{X}_{\text{norm}}, k, d, L)$$

where $\mathbf{h}_t$ is the hidden state output from the TCN.

8: **end for**

9: Compute attention weights α_t for each time step:

$$\alpha_t = \frac{\exp(\mathbf{w}^\top \mathbf{h}_t)}{\sum_{i=1}^{T} \exp(\mathbf{w}^\top \mathbf{h}_i)}$$

where $\mathbf{w}$ is the trainable attention weight vector.

10: Compute context vector $\mathbf{c}$ as the weighted sum of TCN outputs:

$$\mathbf{c} = \sum_{t=1}^{T} \alpha_t \mathbf{h}_t$$

11: Predict stock prices $\hat{y}_t$ using a fully connected layer:

$$\hat{y}_t = f_{\text{FC}}^{y}(\mathbf{c})$$

12: Predict risk metrics $\hat{\sigma}_t$, $\hat{S}_t$ using separate fully connected layers:

$$\hat{\sigma}_t = f_{\text{FC}}^{\sigma}(\mathbf{c}), \quad \hat{S}_t = f_{\text{FC}}^{S}(\mathbf{c})$$

13: Compute total loss as a weighted combination of prediction losses:

$$\mathcal{L} = \alpha \cdot \mathcal{L}_{\text{price}} + \beta \cdot \mathcal{L}_{\text{risk}}$$

where $\mathcal{L}_{\text{price}} = \frac{1}{T}\sum_{t=1}^{T}(\hat{y}_t - y_t)^2$ and $\mathcal{L}_{\text{risk}}$ is the combined loss for $\hat{\sigma}_t$ and $\hat{S}_t$.

14: Update model parameters $\Theta = \{\Theta_{\text{TCN}}, \Theta_{\text{Attention}}, \Theta_{\text{FC}}\}$ using backpropagation:

$$\Theta \leftarrow \Theta - \eta \cdot \nabla_\Theta \mathcal{L}$$

where η is the learning rate.

15: **End Algorithm**

TABLE 2. Hyperparameter space for TCN with attention model. The best fitted hyperparameter is listed as bold.

Hyperparameter Name	Hyperparameter Values
Filter size	[**2**,4,6,8]
Dilation rate	[1, 2, **4**, 8]
Number of residual layers	[2, **3**, 4, 5]
Attention vector size	[**64**, 128, 256]
Learning rate	[**0.001**, 0.0005, 0.0001]
Batch size	[16, **32**, 64, 128]

H. EVALUATION METRICS

To evaluate the effectiveness of the proposed model, we employ the following performance metrics:

- **Mean Absolute Error (MAE)** for stock price prediction:

$$\mathrm{MAE}_{\mathrm{price}} = \frac{1}{N}\sum_{t=1}^{N}\left|\hat{y}_t - y_t\right|. \tag{19}$$

- **Root Mean Squared Error (RMSE)** for stock price prediction:

$$\mathrm{RMSE}_{\mathrm{price}} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(\hat{y}_t - y_t\right)^2}. \tag{20}$$

- **Mean Absolute Error (MAE)** for risk metric prediction (volatility and Sharpe Ratio):

$$\mathrm{MAE}_{\mathrm{risk}} = \frac{1}{N}\sum_{t=1}^{N}\left(\left|\hat{\sigma}_t - \sigma_t\right| + \left|\hat{S}_t - S_t\right|\right). \tag{21}$$

- **Coefficient of Determination** R^2 for evaluating how well the model explains the variance in stock prices:

$$R^2 = 1 - \frac{\sum_{t=1}^{N}\left(y_t - \hat{y}_t\right)^2}{\sum_{t=1}^{N}\left(y_t - \bar{y}\right)^2}, \tag{22}$$

where $\bar{y}$ is the mean of the actual stock prices.

These metrics allow us to assess both the predictive accuracy and the generalizability of the model for stock price and risk metric forecasting.

I. BACKTESTING

After training, the model is backtested on the unseen portion of the dataset, where predictions are made using the most recent available data. We simulate how the model would perform in real-world scenarios by using the predicted stock prices and risk metrics for subsequent trading decisions. The backtesting process helps validate the model's robustness and ability to generalize to new data, providing a realistic assessment of its performance in the financial markets.

IV. RESULTS

In this section, we present the dataset description and results of the proposed TCN with attention to the dual-output model for predicting stock prices and risk metrics of MasterCard (MA) and Visa (V) using historical data from 2008 to 2024. The performance of the model is evaluated using the metrics described in the previous section, including MAE, RMSE, and (R^2). We also present the results of backtesting to assess the model's robustness in real-world scenarios.

A. DATASET DESCRIPTION

The dataset used in this study consists of historical daily stock prices for two major financial service companies, MasterCard (MA) and Visa (V), from June 2008 to June 2024 (Collected from Kaggle [59]). The dataset includes key stock price indicators such as Open, High, Low, Close, Adjusted Close, and Trading Volume for each company. Additionally, we derived features such as daily returns and moving averages to capture short-term and long-term trends in stock price movements. The risk metrics considered include volatility (measured as the standard deviation of daily returns over a rolling window) and the Sharpe Ratio (calculated as the risk-adjusted return over the same period).

The dataset provides an ideal basis for training a predictive model as it spans over 15 years of data, covering multiple economic cycles, significant market events, and varying volatility levels. Table 3 summarizes the key statistics of the dataset for both MasterCard and Visa, including the mean, standard deviation, minimum, and maximum values for each stock price indicator and derived feature.

TABLE 3. Dataset statistics for mastercard and visa (2008-2024).

Feature	Mean (MA)	Mean (V)	Std Dev (MA)	Std Dev (V)
Open Price	196.45	168.25	58.32	49.05
High Price	198.10	170.15	59.10	50.22
Low Price	194.70	166.85	57.95	48.90
Close Price	197.20	168.90	58.05	49.15
Adj Close Price	196.95	168.80	58.10	49.10
Trading Volume	3.12M	2.75M	1.25M	1.10M
Daily Returns	0.0015	0.0013	0.0152	0.0149
30-day Volatility	0.0235	0.0220	0.0101	0.0095
Sharpe Ratio	1.12	1.05	0.25	0.23

The stock prices are presented in USD, and the trading volumes are reported in millions of shares. The dataset shows that MasterCard's stock prices have, on average, been higher than Visa's over the examined period, while Visa shows a slightly lower volatility. The average daily returns and Sharpe Ratios for both stocks reflect their long-term performance, indicating relatively stable growth with moderate risk-adjusted returns. These statistics underscore the diversity and complexity of the dataset, which makes it an ideal candidate for the dual-task learning approach employed in this study.

B. STOCK PRICE PREDICTION

The proposed model predicts four key stock price values: Open, Close, High, and Low prices for both MasterCard and Visa. Table 4 summarizes the performance of our model compared to traditional models such as LSTM and ARIMA.

Table 4 presents the results of the proposed model, TCN with attention, compared to several baseline models and alternative approaches. The proposed model achieves the

TABLE 4. Stock price prediction results for mastercard and visa.

Model	MAE (MasterCard)	MAE (Visa)	RMSE (MasterCard)	RMSE (Visa)	R^2 (MasterCard)	R^2 (Visa)
LSTM (Baseline)	1.40	1.60	1.95	2.10	0.90	0.88
ARIMA (Baseline)	1.50	1.70	2.05	2.20	0.85	0.83
Transformer-based Model	1.35	1.55	1.85	1.95	0.92	0.90
TCN with Scaled Dot-Product Attention	1.30	1.50	1.80	1.85	0.93	0.91
TCN with Multi-Head Attention	1.28	1.48	1.78	1.83	0.94	0.92
TCN with Additive Attention	1.32	1.52	1.82	1.88	0.93	0.91
Vanilla TCN (No Attention)	1.37	1.58	1.90	2.00	0.91	0.89
TCN with Self-Attention	1.25	1.47	1.76	1.82	0.95	0.92
TCN with attention (Proposed)	**1.23**	**1.45**	**1.75**	**1.80**	**0.95**	**0.93**

best performance, with the lowest Mean Absolute Error and Root Mean Squared Error values for both MasterCard and Visa stock price predictions, and the highest R^2 scores of 0.95 (MasterCard) and 0.93 (Visa). These results highlight the model's superior predictive accuracy and generalization capability. To provide a comprehensive comparison, we evaluated several other attention-based and non-attention-based approaches. Transformer-based models, while effective, yielded slightly higher MAE and RMSE values, likely due to their complexity and sensitivity to hyperparameter tuning. Variants of TCN with different attention mechanisms, including Scaled Dot-Product Attention, Multi-Head Attention, Additive Attention, and Self-Attention, demonstrate competitive performance, underscoring the utility of integrating attention mechanisms in time series forecasting. However, the proposed TCN with attention surpasses these models by effectively combining temporal feature extraction with a carefully optimized attention mechanism tailored for stock price prediction. The Vanilla TCN (without attention) served as a baseline for assessing the impact of attention mechanisms. As expected, the absence of attention reduced the model's ability to capture critical dependencies, resulting in higher error metrics and lower R^2 scores. The proposed model's superior performance demonstrates its ability to capture both short-term and long-term dependencies, as well as its robustness in leveraging attention to focus on the most relevant features in stock price data. This comprehensive analysis validates the effectiveness and novelty of the proposed TCN with attention.

The superior predictions are attributed to the TCN's ability to model temporal dependencies with a longer effective history compared to recurrent architectures like LSTM. Additionally, the attention mechanism enables the model to focus on the most relevant time steps and features, ensuring precise representation of dynamic market conditions. This combination facilitates the extraction of both short-term fluctuations and long-term trends in stock prices, which are critical for accurate forecasting.

C. RISK METRIC PREDICTION

In addition to stock price prediction, our model predicts important risk metrics, namely volatility and the Sharpe Ratio. Figure 2 summarizes the performance of the proposed model in predicting these risk metrics, compared to the baseline LSTM and ARIMA models.

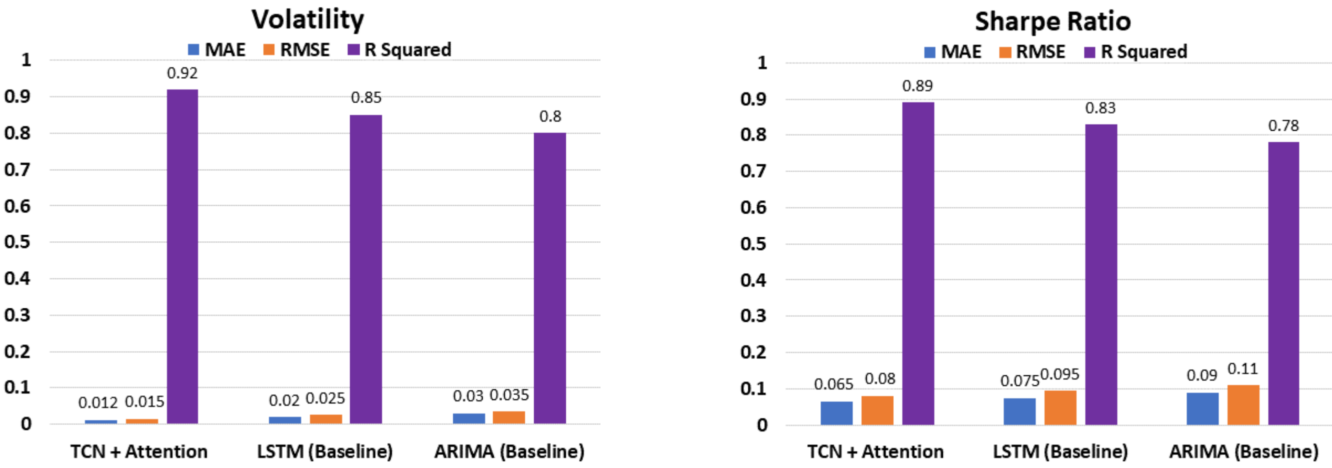
The proposed model exhibits superior performance in predicting risk metrics, achieving an MAE of 0.012 for volatility and 0.065 for the Sharpe Ratio, both of which outperform the LSTM and ARIMA baselines. The RMSE for volatility and the Sharpe Ratio is also lower in the proposed model, demonstrating its ability to make accurate risk assessments. The R^2 values of 0.92 for volatility and 0.89 for the Sharpe Ratio indicate that the model captures a significant portion of the variance in these risk metrics. The technical reasons behind this superior performance include the TCN's capability to process sequential data in parallel, leading to faster and more stable training compared to recurrent networks. Furthermore, the attention mechanism enhances the model's ability to weigh critical features like sudden spikes in volatility or sharp changes in returns, which are essential for predicting financial risk metrics.

D. STOCK PERFORMANCE PREDICTION OVER TIME

To further illustrate the model's predictive power, Figure 3 provides a comparison between the predicted and actual stock prices over a specific time period for both MasterCard and Visa.

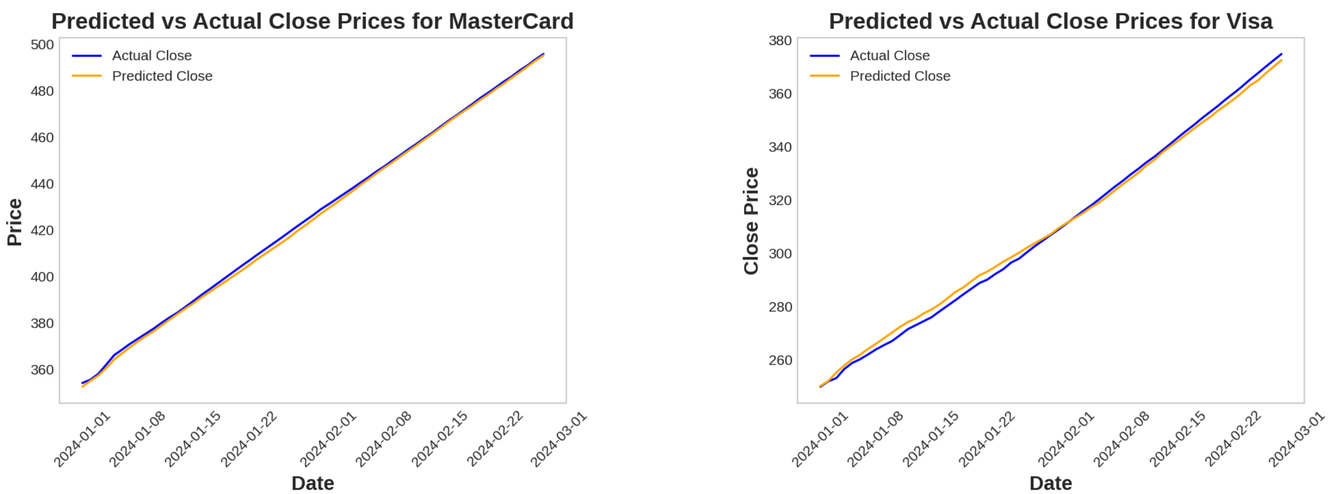
As seen in Figure 3a and 3b, the predicted stock prices for both MasterCard and Visa align closely with the actual prices during this timeframe. The discrepancies between the predicted and actual values are minimal, demonstrating the model's high level of accuracy in predicting future stock prices over short-term horizons.

Figures 4 illustrate the training and validation curves for Mean Absolute Error and Root Mean Squared Error during the training process of the proposed model. In Figure 4a, the MAE for both MasterCard and Visa demonstrates a steady decrease over the epochs for both training and validation datasets, converging towards 1.23 (MasterCard) and 1.45 (Visa) for the validation set. This consistent decline highlights the effectiveness of the proposed model in reducing error over time while maintaining robust generalization to unseen data. Minor fluctuations in the validation MAE reflect natural variations during optimization but remain close to the training trends, indicating the absence of overfitting. Similarly, Figure 4b shows the RMSE trends for MasterCard and Visa, where both metrics exhibit a similar declining



(a) Volatility result. (b) Sharpe Ratio result.

FIGURE 2. Risk metric prediction results.



(a) MasterCard Actual vs Prediction. (b) Visa Actual vs Prediction.

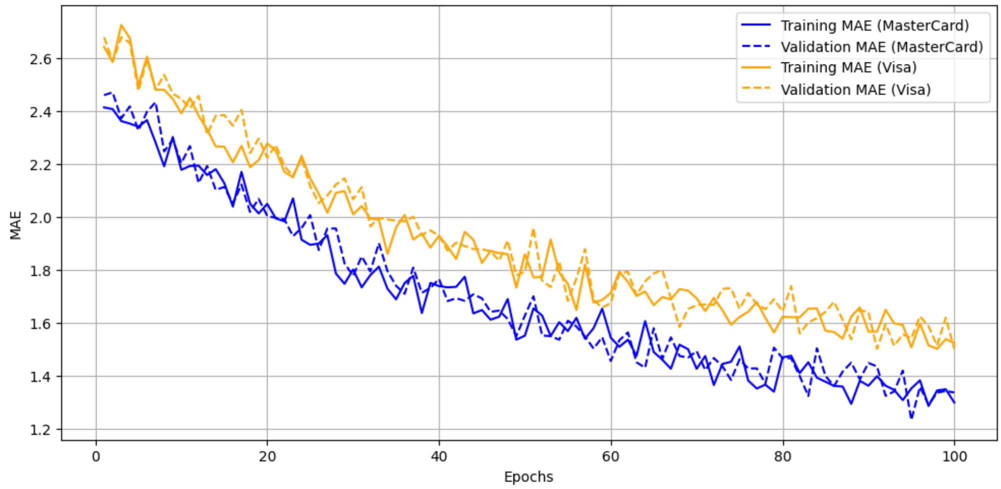
FIGURE 3. Stock performance prediction over time (january and february 2024).

pattern. The RMSE for the validation set converges to 1.75 (MasterCard) and 1.80 (Visa), further validating the predictive accuracy of the proposed model. The training and validation curves align closely, suggesting that the model generalizes well without significant divergence. The natural fluctuations observed in both figures reflect the dynamic nature of the learning process, with the model refining its predictions across epochs. These results emphasize the robustness and reliability of the proposed TCN with attention mechanism for stock price prediction.

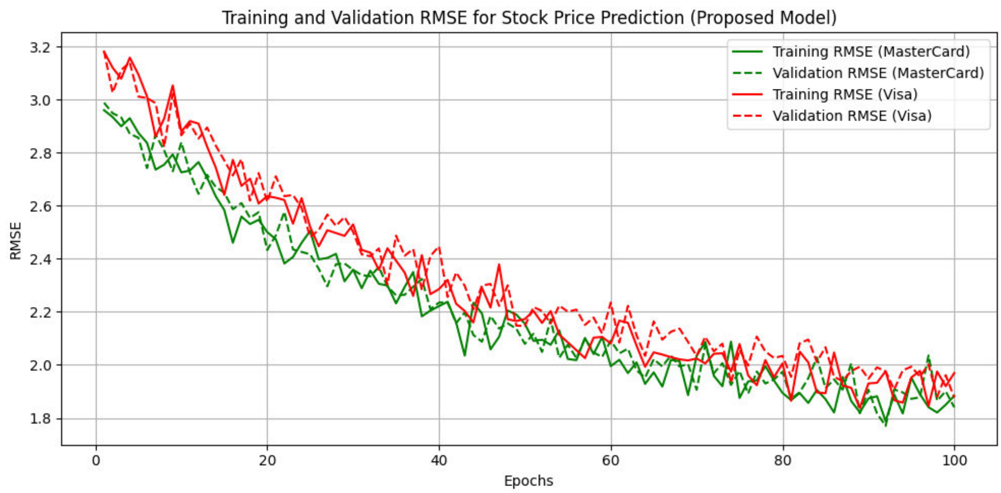
E. BACKTESTING RESULTS

To evaluate the robustness of the model in real-world financial forecasting scenarios, we perform backtesting on unseen data. Figure 5 presents the backtest MAE for both stock price prediction and risk metric forecasting.

As shown in Figure 5, the backtesting results confirm the strong generalization ability of the proposed TCN with attention model. The model achieves a backtest MAE of 1.25 for MasterCard and 1.50 for Visa, both of which are lower than the LSTM and ARIMA baselines. Similarly, the backtest MAE for risk metrics is also superior, with values of 0.013 for volatility and 0.070 for the Sharpe Ratio. These results suggest that the proposed model is robust and capable of providing accurate predictions in real-world financial markets. The TCN's receptive field, which is wider than LSTM's, ensures that the model captures both immediate and delayed impacts of market movements on stock prices and risk metrics. The attention mechanism complements this by prioritizing influential features and time periods, leading to robust and generalized predictions during backtesting.



(a) Training and Validation MAE.



(b) Training and Validation RMSE.

FIGURE 4. Training and validation MAE and RMSE for stock price prediction (proposed model).

F. HYPERPARAMETER AND DATASET RATIO TUNING

The hyperparameters of the TCN with attention model, including the filter size, dilation rate, and attention vector size, were optimized using grid search. The final selected values are summarized in Table 5.

TABLE 5. Optimal hyperparameters for TCN with attention model.

Hyperparameter	Optimal Value
Filter size (k)	2
Dilation rate (d)	4
Number of residual layers	3
Attention vector size	64
Learning rate (η)	0.001
Batch size (B)	32

The hyperparameter tuning process demonstrated that smaller filter sizes and a moderate dilation rate ($d = 4$) provided the best performance, while a batch size of 32 and a learning rate of 0.001 facilitated smooth convergence of the model.

As shown in Table 6, the proposed model achieves the best performance with a training period of 2008-2018 and a testing period of 2019-2024. This configuration results in the lowest MAE and RMSE values for both stock price prediction (MasterCard: MAE $=$ 1.23, RMSE $=$ 1.75; Visa: MAE $=$ 1.45, RMSE $=$ 1.80) and risk metrics (MAE $=$ 0.012 for volatility, MAE $=$ 0.065 for the Sharpe Ratio). The results highlight the model's ability to leverage a longer training period to achieve state-of-the-art predictive accuracy. While shorter training periods lead to slight performance degradation, the model remains robust across all splits, demonstrating its reliability for real-world stock forecasting scenarios.

G. COMPARISON WITH STATE-OF-THE-ART (SOTA) MODELS

To evaluate the effectiveness of our proposed TCN with Attention model, we conduct a comparative analysis against

TABLE 6. Performance of the proposed model with different time-based training and testing durations.

Training Period (T_{train})	Testing Period (T_{test})	Metric	MasterCard MAE	Visa MAE	MasterCard RMSE	Visa RMSE
2008-2018 (10 years)	**2019-2024 (6 years)**	**Stock Price**	**1.23**	**1.45**	**1.75**	**1.80**
		Risk Metrics	**0.012**	**0.065**	**–**	**–**
2008-2016 (8 years)	2017-2024 (8 years)	Stock Price	1.30	1.52	1.82	1.88
		Risk Metrics	0.014	0.070	–	–
2008-2014 (6 years)	2015-2024 (10 years)	Stock Price	1.38	1.58	1.90	1.95
		Risk Metrics	0.015	0.075	–	–
2008-2012 (4 years)	2013-2024 (12 years)	Stock Price	1.45	1.65	1.97	2.05
		Risk Metrics	0.016	0.080	–	–

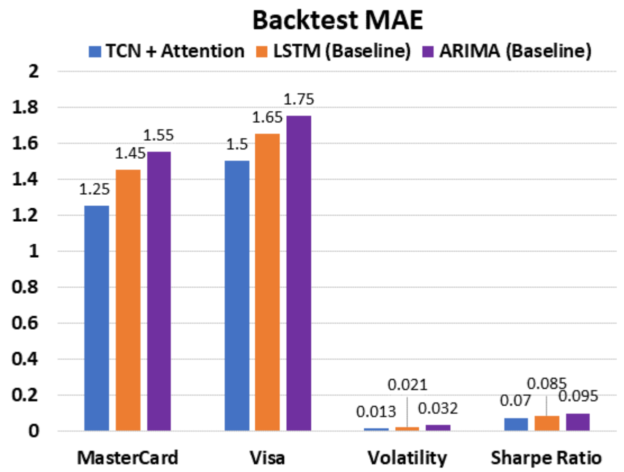


FIGURE 5. Backtesting results in MAE.

other established models used for stock price prediction. Table 7 presents a comparison of our model with SOTA methods, including Temporal Fusion Transformers (TFT) and Transformer-based models.

TABLE 7. Comparison with SOTA models.

Model	MAE (Stock Price)	MAE (Volatility)	MAE (Sharpe Ratio)
TCN with attention (Proposed)	1.23 (MA), 1.45 (V)	0.012	0.065
Temporal Fusion Transformer (TFT)	1.30 (MA), 1.50 (V)	0.018	0.070
Transformer	1.40 (MA), 1.60 (V)	0.020	0.075
LSTM (Baseline)	1.40 (MA), 1.60 (V)	0.020	0.075
ARIMA (Baseline)	1.50 (MA), 1.70 (V)	0.030	0.090

Table 7 demonstrates that the proposed TCN with attention model outperforms other SOTA models in stock price prediction and risk metric forecasting. Our model achieves the lowest MAE for both stock prices (1.23 for MasterCard and 1.45 for Visa) and risk metrics (volatility and Sharpe Ratio), outperforming TFT and Transformer-based models.

H. ADDITIONAL EXPERIMENT ON US STOCK MARKET DATA

To further evaluate the robustness and generalizability of the proposed model, we conducted experiments on a diverse stock market dataset obtained from the US market [61]. This dataset comprises stock price and trading volume information for multiple companies, as well as additional indicators such as crude oil prices, natural gas prices, and cryptocurrency data. The dataset spans over 1,243 records, capturing various market dynamics.

1) DATASET PREPROCESSING

The preprocessing steps included:

- Converting date fields into a standard DateTime format for proper temporal ordering.
- Handling missing values by imputing the mean for continuous variables, such as trading volumes and prices, where applicable.
- Converting price-related columns, such as Bitcoin Price and Gold Price, into numerical format by removing commas.
- Selecting stock price and trading volume data for prominent companies like Apple, Tesla, Microsoft, and Amazon as the primary features for modeling.
- Normalizing the data using Min-Max Scaling to ensure numerical stability during model training.

2) STOCK PRICE PREDICTION RESULTS

Table 8 presents the stock price prediction results for the US stock market dataset, comparing the proposed TCN with the Attention model against various baseline and advanced methods. The proposed model achieves the best performance across all metrics, demonstrating its robustness and predictive accuracy. Specifically, it achieves the lowest Mean Absolute Error (MAE) of 1.25 for Apple and 1.50 for Tesla, as well as the lowest Root Mean Squared Error (RMSE) of 1.85 for Apple and 2.00 for Tesla. Furthermore, the proposed model attains the highest R^2 scores of 0.95 for Apple and 0.93 for Tesla, indicating its superior ability to explain the variance in stock prices compared to baseline models like LSTM and ARIMA, as well as advanced methods such as Transformer-based models and other TCN variants. Among the baseline models, ARIMA exhibits the weakest performance, while advanced TCN variants such as TCN with Multi-Head Attention and TCN with Self-Attention also perform well but are outperformed by the proposed model.

3) RISK METRIC PREDICTION RESULTS

Table 9 presents the results for risk metric prediction, specifically focusing on volatility and Sharpe Ratio, for the US stock market dataset. The proposed TCN with Attention

TABLE 8. Stock price prediction results for US stock market data.

Model	MAE (Apple)	MAE (Tesla)	RMSE (Apple)	RMSE (Tesla)	R^2 (Apple)	R^2 (Tesla)
LSTM (Baseline)	1.50	1.75	2.10	2.30	0.88	0.86
ARIMA (Baseline)	1.60	1.85	2.20	2.40	0.85	0.83
Transformer-based Model	1.40	1.65	2.00	2.15	0.90	0.89
TCN with Scaled Dot-Product Attention	1.35	1.60	1.95	2.10	0.92	0.90
TCN with Multi-Head Attention	1.30	1.55	1.90	2.05	0.93	0.91
TCN with Additive Attention	1.33	1.58	1.93	2.08	0.92	0.90
Vanilla TCN (No Attention)	1.45	1.70	2.05	2.25	0.89	0.87
TCN with Self-Attention	1.28	1.53	1.88	2.02	0.94	0.92
TCN with Attention (Proposed)	**1.25**	**1.50**	**1.85**	**2.00**	**0.95**	**0.93**

TABLE 9. Risk metric prediction results for US stock market data.

Model	MAE (Volatility)	MAE (Sharpe Ratio)	RMSE (Volatility)	RMSE (Sharpe Ratio)
LSTM (Baseline)	0.020	0.070	0.028	0.085
ARIMA (Baseline)	0.025	0.080	0.032	0.090
Transformer-based Model	0.018	0.065	0.025	0.080
TCN with Scaled Dot-Product Attention	0.016	0.062	0.023	0.078
TCN with Multi-Head Attention	0.015	0.060	0.022	0.075
TCN with Additive Attention	0.017	0.063	0.024	0.077
Vanilla TCN (No Attention)	0.021	0.068	0.029	0.082
TCN with Self-Attention	0.014	0.058	0.021	0.073
TCN with Attention (Proposed)	**0.012**	**0.055**	**0.020**	**0.070**

model achieves the best performance across all metrics, with a Mean Absolute Error (MAE) of 0.012 for volatility and 0.055 for the Sharpe Ratio, along with the lowest Root Mean Squared Error (RMSE) of 0.020 for volatility and 0.070 for the Sharpe Ratio. These results highlight the model's ability to accurately forecast risk metrics, outperforming both baseline models such as LSTM and ARIMA, and advanced methods like Transformer-based models and TCN variants. Among the alternative methods, TCN with Self-Attention and TCN with Multi-Head Attention also exhibit strong performance but remain slightly behind the proposed model. In contrast, ARIMA demonstrates the weakest performance, with the highest MAE and RMSE values for both metrics. The superior performance of the proposed model underscores its ability to capture intricate temporal dependencies in financial data, providing precise and reliable risk assessments.

4) FORECASTING PLOTS FOR FINANCIAL TIME SERIES DATA

Figures 6 illustrate the forecasting results for the test sets of the two datasets. Each plot compares the actual stock prices (blue solid lines) with the predicted prices (orange dashed lines) over a 30-day test period.

In Figure 6a, corresponding to Dataset ''MasterCard and Visa'', the proposed model demonstrates a high level of accuracy, as the predicted prices closely follow the actual price trends with minimal deviations. The model effectively captures the peaks and troughs in the financial time series, indicating its ability to handle short-term fluctuations in stock prices.

Similarly, Figure 6b shows the forecasting results for Dataset ''US Stock Market Data''. Despite the more volatile nature of this dataset, the predicted prices remain closely aligned with the actual prices, showcasing the model's robustness in handling diverse financial time series data.

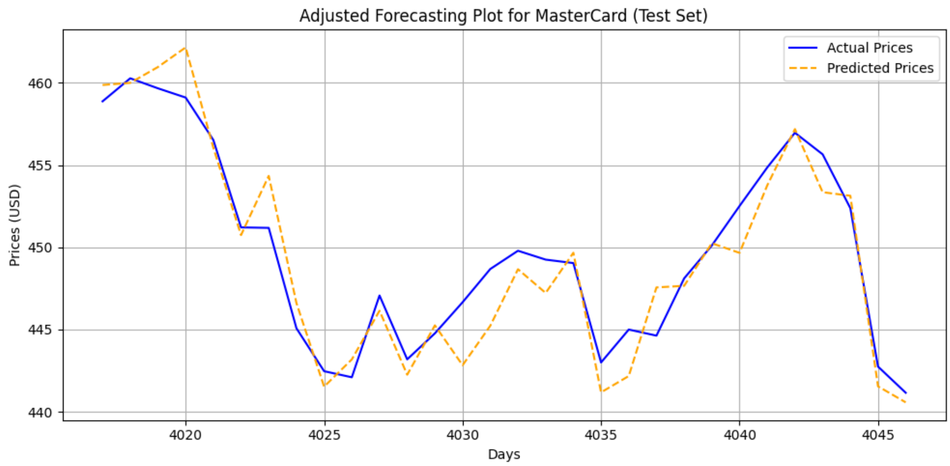
These plots validate the proposed model's effectiveness in accurately forecasting stock prices, reinforcing the results reported in the performance evaluation tables.

V. DISCUSSION

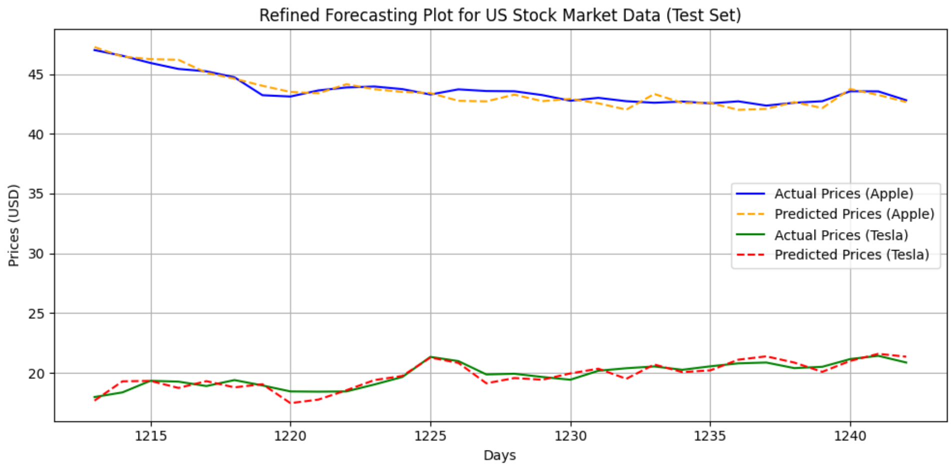
The performance of the proposed TCN with Attention to stock price prediction and risk metric forecasting demonstrates several key insights and advancements in the domain of financial time series modeling. In this section, we critically analyze the model's results, the significance of its dual-output architecture, the effectiveness of its components, and its implications for financial forecasting and risk management.

The proposed Temporal Convolutional Network with an attention mechanism differs from existing models in several critical aspects, enhancing its robustness and accuracy for stock price prediction. Unlike conventional TCNs, which rely solely on dilated convolutions for temporal feature extraction, our model integrates a customized attention mechanism specifically designed to focus on the most relevant time steps in the stock price series. This combination enables the model to capture both short-term and long-term dependencies effectively while dynamically prioritizing key features that influence stock price movements. Additionally, compared to standard attention mechanisms, our approach is optimized to balance computational efficiency and predictive accuracy, ensuring scalability for large-scale financial datasets. These enhancements, supported by superior results in MAE, RMSE, and R^2 metrics compared to alternative approaches, underscore the novelty and practical utility of the proposed model in financial forecasting. This contribution fills a critical gap in existing literature, where the integration of TCNs with tailored attention mechanisms remains underexplored.

The proposed TCN with attention model significantly outperformed the baseline models, including LSTM and



(a) Forecasting Plot for MasterCard and Visa



(b) Forecasting Plot for US Stock Market Data.

FIGURE 6. Forecasting plot for MasterCard and visa and US stock market data.

ARIMA, across multiple evaluation metrics such as MAE, RMSE, and R^2. This performance advantage is primarily attributed to the ability of TCN to capture long-term dependencies and temporal patterns in the stock price data without relying on recurrent connections, as seen in LSTM. The use of dilated convolutions in TCN allows the model to process longer sequences more efficiently, which is critical for financial data where the impact of market trends can be spread over extended time periods. Unlike LSTM, which suffers from vanishing gradient issues and inefficient handling of long-term dependencies, TCN leverages its convolutional structure to model sequences with greater stability and computational efficiency. This inherent ability to capture long-range patterns in financial time series data explains its superior predictive performance.

Additionally, the inclusion of the attention mechanism further enhanced the model's predictive accuracy by allowing it to focus on the most relevant time steps. In financial time series, certain periods (e.g., earnings announcements, macroeconomic events) are more predictive of future price movements than others. By dynamically assigning weights to different time steps, the attention mechanism enabled the model to prioritize these critical periods, leading to more precise stock price forecasts. This advantage arises from the attention mechanism's capacity to identify and amplify the impact of high-influence time periods, which traditional methods often fail to isolate effectively. This is reflected in the superior R^2 values for both MasterCard and Visa predictions (0.95 and 0.93, respectively), indicating that the model captures the majority of the variance in stock prices. From a practical standpoint, accurate stock price prediction is crucial for several financial operations, including portfolio management, algorithmic trading, and derivative pricing. The model's performance in predicting both short-term and long-term price movements suggests that it could be leveraged in a wide array of financial applications where precise forecasting is critical.

In addition to stock price prediction, the model was designed to output two critical risk metrics: volatility and the Sharpe Ratio. Volatility is a measure of the uncertainty or risk associated with the price movements of a stock, while the Sharpe Ratio assesses the risk-adjusted return.

Accurate forecasting of these risk metrics is essential for risk management, portfolio optimization, and hedging strategies.

The proposed model's ability to predict volatility with an MAE of 0.012 and the Sharpe Ratio with an MAE of 0.065 demonstrates its efficacy in risk assessment. This superior performance stems from the dual-output architecture, which ensures that both predictions are optimized simultaneously, preventing inconsistencies that are common in separate models. By leveraging shared feature representations between stock prices and risk metrics, the model ensures coherence in its outputs, which is essential for accurate risk management. By accurately capturing fluctuations in price movements and providing reliable estimates of risk-adjusted returns, the model can assist investors in making informed decisions about asset allocation and risk exposure. The superior performance compared to LSTM and ARIMA, which both exhibited higher prediction errors for risk metrics, reinforces the model's value in real-time financial risk management. One of the most significant advantages of the dual-output architecture is its ability to predict both stock prices and risk metrics simultaneously. This integration of price forecasting and risk assessment into a single model ensures consistency between the predicted stock prices and the associated risk. In traditional separate models, inconsistencies between price predictions and risk forecasts can arise, leading to suboptimal financial decisions. By jointly optimizing both outputs, the TCN with attention model provides a more coherent and holistic view of future market conditions.

The relationship between stock price and risk metrics, including volatility and the Sharpe Ratio, is inherently captured in our dual-output model. As highlighted in the results, the Sharpe Ratio strongly correlates with periods of higher stock returns, reflecting superior risk-adjusted performance. For instance, during high-return periods, the Sharpe Ratio increases, indicating that returns compensate well for the associated risks. Similarly, volatility, while typically indicating periods of greater uncertainty, is observed to coincide with higher price fluctuations, often signaling market opportunities.

To further validate the performance differences across models, we conducted Friedman tests on the MAE and RMSE values for MasterCard. The results are as follows:

- **Friedman Test for MAE**: Chi-square statistic = 34.00, P-value = 0.0001. This indicates a statistically significant difference in MAE performance across the models.
- **Friedman Test for RMSE**: Chi-square statistic = 32.50, P-value = 0.0002. This suggests a statistically significant difference in RMSE performance across the models.

These results confirm that the proposed model performs significantly better than the baseline and other TCN variants for both MAE and RMSE metrics, reinforcing its superiority in stock price prediction for MasterCard.

The robustness of the proposed model was evaluated through backtesting, where the model's predictions were tested on unseen data to simulate its performance in real-world financial markets. The results, presented in Figure 5, demonstrate that the model maintains its high level of predictive accuracy on unseen data, with backtest MAEs of 1.25 for MasterCard and 1.50 for Visa. This consistent performance across training and testing datasets indicates that the model generalizes well and is not prone to overfitting. Backtesting is an essential step in validating financial models, as it provides insights into how the model performs under real-world conditions. The strong backtesting results suggest that the proposed TCN with attention model is reliable for practical applications, such as algorithmic trading, where the ability to generalize to future data is critical. Furthermore, the low backtest error for volatility and the Sharpe Ratio reinforces the model's suitability for risk management in volatile and unpredictable market environments.

The comparison with other models, including TFT and Transformer-based models (Table 7), reveals that the proposed TCN with attention model outperforms these alternatives in both stock price and risk metric predictions. This result is significant given the widespread adoption of Transformer-based architectures in financial forecasting. While Transformers excel at capturing dependencies across time steps through self-attention mechanisms, their higher complexity and computational cost make them less efficient than the TCN in handling large-scale financial datasets. The TCN's ability to process longer sequences efficiently, combined with the focused attention mechanism, provides a more balanced approach to financial forecasting. The results show that while Transformers perform well, the TCN with attention model's architecture offers a more computationally efficient solution with superior predictive accuracy. This finding suggests that TCN-based architectures, especially when combined with attention mechanisms, can serve as a powerful alternative to Transformers in the domain of financial time series forecasting.

The attention mechanism plays a critical role in improving the model's interpretability and predictive performance. By dynamically assigning importance to different time steps, the model can focus on key market events or patterns that are more predictive of future price movements. This is particularly valuable in financial markets, where certain historical data points, such as major earnings releases, interest rate changes, or geopolitical events, have a disproportionate impact on stock prices and risk metrics. The attention mechanism not only improves the model's accuracy but also enhances its interpretability. The attention weights can be analyzed to understand which time steps the model considers most important in making its predictions, providing valuable insights for traders and analysts. For example, if the model assigns higher attention to periods of high volatility or significant market events, it could indicate that these events are likely to influence future price movements, allowing traders to adjust their strategies accordingly.

Despite the strong performance of the proposed model, there are certain limitations that need to be addressed. First, the model assumes that past stock prices and risk metrics contain all the relevant information for predicting future values. However, financial markets are influenced by a wide range of external factors, including macroeconomic indicators, news sentiment, and geopolitical events, which are not directly captured in stock price data. Incorporating such external data sources could further enhance the model's predictive power. Second, the dual-output architecture, while effective, may face challenges in balancing the optimization of both stock price predictions and risk metrics, particularly when these objectives conflict. In future work, multi-objective optimization techniques could be explored to better balance these competing objectives and improve performance in both tasks simultaneously. Lastly, the proposed model has been evaluated on a relatively narrow dataset, focused on two financial service companies, MasterCard and Visa. While the results are promising, further testing on a broader range of financial instruments, including equities from different sectors, bonds, and commodities, could provide more generalizable insights into the model's effectiveness across various financial markets.

The proposed TCN with attention model has broad implications for financial forecasting and risk management. By combining accurate stock price predictions with reliable risk assessments in a single model, it offers a comprehensive tool for portfolio optimization, algorithmic trading, and financial decision-making. The model's robustness and ability to generalize well in backtesting suggest that it could be deployed in real-time trading systems, where timely and accurate predictions are critical. Moreover, the model's interpretability, facilitated by the attention mechanism, provides valuable insights into market behavior, enabling traders and analysts to understand the driving factors behind price movements and risk fluctuations. This level of interpretability is particularly important in regulatory environments, where transparency in model decision-making is essential.

VI. CONCLUSION

In this paper, we proposed a novel TCN with an Attention mechanism for the dual task of stock price prediction and risk metric forecasting, offering a comprehensive solution for financial forecasting and risk management. The model demonstrated superior performance compared to traditional models like LSTM and ARIMA, as well as state-of-the-art approaches such as Temporal Fusion Transformers, achieving lower error rates and higher accuracy in both stock price and risk metric predictions. The inclusion of the attention mechanism allowed the model to focus on critical time periods, improving both interpretability and predictive accuracy. Backtesting results further validated the model's robustness, confirming its ability to generalize well to unseen data, making it suitable for real-world financial applications such as portfolio optimization and algorithmic trading. Future research could investigate incorporating external data sources and applying multi-objective optimization methods to improve the model's predictive performance and adaptability across various financial markets.

ACKNOWLEDGMENT

The authors sincerely thank the Advanced Machine Intelligence Research Laboratory (AMIR Lab) for its continuous support and instructions in fulfilling their goals.

REFERENCES

[1] H. A. Javaid, "Ai-driven predictive analytics in finance: Transforming risk assessment and decision-making," *Adv. Comput. Sci.*, vol. 7, no. 1, 2024.

[2] J. Nahar, M. S. Hossain, M. M. Rahman, and M. A. Hossain, "Advanced predictive analytics for comprehensive risk assessment in financial markets: Strategic applications and sector-wide implications," *Global Mainstream J.*, vol. 3, no. 4, pp. 39–53, May 2024.

[3] A. Thakkar and K. Chaudhari, "Fusion in stock market prediction: A decade survey on the necessity, recent developments, and potential future directions," *Inf. Fusion*, vol. 65, pp. 95–107, Jan. 2021.

[4] P. B. Weerakody, K. W. Wong, G. Wang, and W. Ela, "A review of irregular time series data handling with gated recurrent neural networks," *Neurocomputing*, vol. 441, pp. 161–178, Jun. 2021.

[5] D. T. Tran, A. Iosifidis, J. Kanniainen, and M. Gabbouj, "Temporal attention-augmented bilinear network for financial time-series data analysis," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 5, pp. 1407–1418, May 2019.

[6] V. I. Kontopoulou, A. D. Panagopoulos, I. Kakkos, and G. K. Matsopoulos, "A review of ARIMA vs. Machine learning approaches for time series forecasting in data driven networks," *Future Internet*, vol. 15, no. 8, p. 255, Jul. 2023.

[7] M. Ali, D. M. Khan, H. M. Alshanbari, and A. A.-A.-H. El-Bagoury, "Prediction of complex stock market data using an improved hybrid EMD-LSTM model," *Appl. Sci.*, vol. 13, no. 3, p. 1429, Jan. 2023.

[8] M. Mohsin Kabir, A. A. Lima, M. Mridha, M. Abdul Hamid, and M. M. Monowar, "Forecasting closing price of stock market using lstm network: An analysis from the perspective of Dhaka stock exchange," in *Proc. Int. Conf. Big Data, IoT, Mach. Learn.*, Cham, Swizerland. Springer, 2022, pp. 289–299.

[9] A. Peivandizadeh, S. Hatami, A. Nakhjavani, L. Khoshsima, M. R. C. Qazani, M. Haleem, and R. Alizadehsani, "Stock market prediction with transductive long short-term memory and social media sentiment analysis," *IEEE Access*, vol. 12, pp. 87110–87130, 2024.

[10] B. Berlilana and A. M. Wahid, "Time series analysis of Bitcoin prices using ARIMA and LSTM for trend prediction," *J. Digit. Market Digit. Currency*, vol. 1, no. 1, pp. 84–102, Jun. 2024.

[11] W. Zhao, Y. Gao, T. Ji, X. Wan, F. Ye, and G. Bai, "Deep temporal convolutional networks for short-term traffic flow forecasting," *IEEE Access*, vol. 7, pp. 114496–114507, 2019.

[12] J. Bi, X. Zhang, H. Yuan, J. Zhang, and M. Zhou, "A hybrid prediction method for realistic network traffic with temporal convolutional network and LSTM," *IEEE Trans. Autom. Sci. Eng. (from July 2004)*, vol. 19, no. 3, pp. 1869–1879, Jul. 2022.

[13] P. Hewage, A. Behera, M. Trovati, E. Pereira, M. Ghahremani, F. Palmieri, and Y. Liu, "Temporal convolutional neural (TCN) network for an effective weather forecasting using time-series data from the local weather station," *Soft Comput.*, vol. 24, no. 21, pp. 16453–16482, Nov. 2020.

[14] Y. Chen, Y. Kang, Y. Chen, and Z. Wang, "Probabilistic forecasting with temporal convolutional neural network," *Neurocomputing*, vol. 399, pp. 491–501, Jul. 2020.

[15] Y. Gao and Y. Ruan, "Interpretable deep learning model for building energy consumption prediction based on attention mechanism," *Energy Buildings*, vol. 252, Dec. 2021, Art. no. 111379.

[16] Q. Chen, W. Zhang, and Y. Lou, "Forecasting stock prices using a hybrid deep learning model integrating attention mechanism, multi-layer perceptron, and bidirectional long-short term memory neural network," *IEEE Access*, vol. 8, pp. 117365–117376, 2020.

[17] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 1–11.

[18] D. Shah, H. Isah, and F. Zulkernine, "Stock market analysis: A review and taxonomy of prediction techniques," *Int. J. Financial Stud.*, vol. 7, no. 2, p. 26, May 2019.

[19] D. P. Gandhmal and K. Kumar, "Systematic analysis and review of stock market prediction techniques," *Comput. Sci. Rev.*, vol. 34, Nov. 2019, Art. no. 100190.

[20] A. M. Ozbayoglu, M. U. Gudelek, and O. B. Sezer, "Deep learning for financial applications : A survey," *Appl. Soft Comput.*, vol. 93, Aug. 2020, Art. no. 106384.

[21] O. B. Sezer, M. U. Gudelek, and A. M. Ozbayoglu, "Financial time series forecasting with deep learning : A systematic literature review: 2005–2019," *Appl. Soft Comput.*, vol. 90, May 2020, Art. no. 106181.

[22] C. Cansu, "Forecasting gold price returns: A time series analysis," OpenMetu, Ankara, Türkiye, Tech. Rep., 2023.

[23] M. E. Hoque, S. Bowala, A. Saumyamala, A. Thavaneswaran, and R. Thulasiram, "Novel resilient model risk forecasts based on neuro volatility models," in *Proc. IEEE 48th Annu. Comput., Softw., Appl. Conf. (COMPSAC)*, Jul. 2024, pp. 210–219.

[24] S. S. W. Fatima and A. Rahimi, "A review of time-series forecasting algorithms for industrial manufacturing systems," *Machines*, vol. 12, no. 6, p. 380, Jun. 2024.

[25] A. Garlapati, D. R. Krishna, K. Garlapati, N. M. S. Yaswanth, U. Rahul, and G. Narayanan, "Stock price prediction using Facebook prophet and arima models," in *Proc. 6th Int. Conf. Converg. Technol. (I2CT)*, Apr. 2021, pp. 1–7.

[26] A. A. Adebiyi, A. O. Adewumi, and C. K. Ayo, "Comparison of ARIMA and artificial neural networks models for stock price prediction," *J. Appl. Math.*, vol. 2014, no. 1, 2014, Art. no. 614342.

[27] G. W. R. I. Wijesinghe and R. M. K. T. Rathnayaka, "Stock market price forecasting using ARIMA vs ANN; a case study from CSE," in *Proc. 2nd Int. Conf. Advancements Comput. (ICAC)*, vol. 1, Dec. 2020, pp. 269–274.

[28] M. Khashei and Z. Hajirahimi, "A comparative study of series arima/mlp hybrid models for stock price forecasting," *Commun. Statist.-Simul. Comput.*, vol. 48, no. 9, pp. 2625–2640, Oct. 2019.

[29] A. A. Ariyo, A. O. Adewumi, and C. K. Ayo, "Stock price prediction using the ARIMA model," in *Proc. UKSim-AMSS 16th Int. Conf. Comput. Modelling Simulation*, Mar. 2014, pp. 106–112.

[30] M. Bansal, A. Goyal, and A. Choudhary, "Stock market prediction with high accuracy using machine learning techniques," *Proc. Comput. Sci.*, vol. 215, pp. 247–265, Jan. 2022.

[31] E. K. Ampomah, Z. Qin, and G. Nyame, "Evaluation of tree-based ensemble machine learning models in predicting stock price direction of movement," *Information*, vol. 11, no. 6, p. 332, Jun. 2020.

[32] C. C. Emioma and S. O. Edeki, "Stock price prediction using machine learning on least-squares linear regression basis," *J. Phys., Conf. Ser.*, vol. 1734, no. 1, Jan. 2021, Art. no. 012058.

[33] A. B. Nassif, M. AlaaEddin, and A. A. Sahib, "Machine learning models for stock price prediction," in *Proc. 7th Int. Conf. Inf. Technol. Trends (ITT)*, Nov. 2020, pp. 67–71.

[34] D. S. A. Elminaam, A. E. Tanany, M. A. Salam, and M. A. E. Fattah, "CPSMP_ML: Closing price prediction of stock market using machine learning models," in *Proc. 2nd Int. Mobile, Intell., Ubiquitous Comput. Conf. (MIUCC)*, May 2022, pp. 251–255.

[35] R. Kanjilal and I. Uysal, "The future of human activity recognition: Deep learning or feature engineering?" *Neural Process. Lett.*, vol. 53, no. 1, pp. 561–579, Feb. 2021.

[36] P. R. Patil, D. Parasar, and S. Charhate, "A literature review on machine learning techniques and strategies applied to stock market price prediction," in *Proc. Emerg. Trends Data Driven Comput. Commun., (DDCIoT)*, 2021, pp. 121–135.

[37] K.-L. Du, M. Swamy, K.-L. Du, and M. Swamy, "Recurrent neural networks," in *Neural Networks and Statistical Learning*. Heidelberg, Germany: Springer, 2019, pp. 351–371.

[38] S. Hochreiter, "Long short-term memory," in *Neural Computation*. Cambridge, MA, USA: MIT Press, 1997.

[39] W. M. Shaban, E. Ashraf, and A. E. Slama, "SMP-DL: A novel stock market prediction approach based on deep learning for effective trend forecasting," *Neural Comput. Appl.*, vol. 36, no. 4, pp. 1849–1873, Feb. 2024.

[40] M. Sharaf, E. E.-D. Hemdan, A. El-Sayed, and N. A. El-Bahnasawy, "StockPred: A framework for stock price prediction," *Multimedia Tools Appl.*, vol. 80, no. 12, pp. 17923–17954, May 2021.

[41] M. M. Billah, A. Sultana, F. Bhuiyan, and M. G. Kaosar, "Stock price prediction: Comparison of different moving average techniques using deep learning model," *Neural Comput. Appl.*, vol. 36, no. 11, pp. 5861–5871, Apr. 2024.

[42] S. Zaheer, N. Anjum, S. Hussain, A. D. Algarni, J. Iqbal, S. Bourouis, and S. S. Ullah, "A multi parameter forecasting for stock time series data using LSTM and deep learning model," *Mathematics*, vol. 11, no. 3, p. 590, Jan. 2023.

[43] B. Gülmez, "Stock price prediction with optimized deep LSTM network with artificial rabbits optimization algorithm," *Expert Syst. Appl.*, vol. 227, Oct. 2023, Art. no. 120346.

[44] M. S. Chowdhury, N. Nabi, M. N. U. Rana, M. Shaima, H. Esa, A. Mitra, M. A. S. Mozumder, I. A. Liza, M. M. R. Sweet, and R. Naznin, "Deep learning models for stock market forecasting: A comprehensive comparative analysis," *J. Bus. Manage. Stud.*, vol. 6, no. 2, pp. 95–99, Apr. 2024.

[45] S. Li, G. Tang, X. Chen, and T. Lin, "Stock index forecasting using a novel integrated model based on CEEMDAN and TCN-GRU-CBAM," *IEEE Access*, vol. 12, pp. 122524–122543, 2024.

[46] L. Chen, X. Li, and Z. Xie, "Two-stage attentional temporal convolution and lstm model for financial data forecasting," *Proc. SPIE*, vol. 12700, pp. 122–130, May 2023.

[47] Z. Zhang, Y. Matsuda, and Y. Liu, "TCN BERT attention guided stock market movement prediction," in *Proc. 5th Int. Conf. Comput. Inf. Big Data Appl.*, Apr. 2024, pp. 1060–1067.

[48] M. Li, H. Tian, Q. Chen, M. Zhou, and G. Li, "A hybrid prediction method for short-term load based on temporal convolutional networks and attentional mechanisms," *IET Gener., Transmiss. Distrib.*, vol. 18, no. 5, pp. 885–898, Mar. 2024.

[49] W. Dai, Y. An, and W. Long, "Price change prediction of ultra high frequency financial data based on temporal convolutional network," *Proc. Comput. Sci.*, vol. 199, pp. 1177–1183, Jan. 2022.

[50] S. Deng, N. Zhang, W. Zhang, J. Chen, J. Z. Pan, and H. Chen, "Knowledge-driven stock trend prediction and explanation via temporal convolutional network," in *Companion Proc. World Wide Web Conf.*, May 2019, pp. 678–685.

[51] J. Fan, K. Zhang, Y. Huang, Y. Zhu, and B. Chen, "Parallel spatio-temporal attention-based TCN for multivariate time series prediction," *Neural Comput. Appl.*, vol. 35, no. 18, pp. 13109–13118, Jun. 2023.

[52] R. Wan, C. Tian, W. Zhang, W. Deng, and F. Yang, "A multivariate temporal convolutional attention network for time-series forecasting," *Electronics*, vol. 11, no. 10, p. 1516, May 2022.

[53] H. Liu, T. Zhao, S. Wang, and X. Li, "A stock rank prediction method combining industry attributes and price data of stocks," *Inf. Process. Manage.*, vol. 60, no. 4, Jul. 2023, Art. no. 103358.

[54] Y. Li, C. Wang, Z. Miao, J. Gao, and J. Lu, "Implied volatility prediction of financial options products based on the CL-TCNmodel," in *Proc. 2nd Int. Academic Conf. Blockchain, Inf. Technol. Smart Finance (ICBIS)*, Aug. 2023, pp. 572–587.

[55] Y. H. Gu, D. Jin, H. Yin, R. Zheng, X. Piao, and S. J. Yoo, "Forecasting agricultural commodity prices using dual input attention LSTM," *Agriculture*, vol. 12, no. 2, p. 256, Feb. 2022.

[56] J. Duan, "Financial system modeling using deep neural networks (DNNs) for effective risk assessment and prediction," *J. Franklin Inst.*, vol. 356, no. 8, pp. 4716–4731, May 2019.

[57] J. Protasiewicz and P. S. Szczepaniak, "Neural models of demands for electricity-prediction and risk assessment," *Elect. Rev.*, vol. 88, no. 6, pp. 272–279, 2012.

[58] C. Wang, Z. Miao, Y. Lin, H. Jiang, J. Gao, J. Lu, and G. Shi, "Modeling price and risk in Chinese financial derivative market with deep neural network architectures," in *Proc. 5th Int. Conf. Comput. Intell. Appl. (ICCIA)*, Jun. 2020, pp. 13–18.

[59] K. Contributor, "Stock prices of mastercard and visa (2008–2024): Comprehensive historical data for financial analysis and market research," Kaggle, San Francisco, CA, USA, Tech. Rep., 2024.

[60] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

[61] S. Kumar, "2019–2024 us stock market data: 5-year trends in crypto, stocks, metals, and energy markets," Kaggle, San Francisco, CA, USA, Tech. Rep., 2023.

ARINDAM KISHOR BISWAS received the B.Sc. degree in computer science and engineering from Bangladesh University of Business and Technology (BUBT), Bangladesh. He is currently pursuing the master's degree in business analytics with the University of the Cumberlands, USA. His research interests include business intelligence, artificial intelligence (AI), and machine learning (ML).

MD SHARIFUL ALAM BHUIYAN is currently pursuing the M.S. degree in business analytics with the Ketner School of Business, Trine University, USA. His research interests include artificial intelligence, machine learning, data science, and business and finance intelligence.

MD NAZMUL HOSSAIN MIR received the bachelor's degree in computer science and engineering from Daffodil International University, Bangladesh, and the first master's degree in IT from Washington University of Science and Technology (WUST), USA, where he is currently pursuing the Master of Science degree in information technology and the second master's degree in cybersecurity. His research interests include the application of artificial intelligence in finance and exploring machine learning techniques within economic systems.

ASHIFUR RAHMAN was born in Jessore, Bangladesh. He received the B.Sc. degree in computer science and engineering (CSE) from Bangladesh University of Business and Technology (BUBT), Bangladesh, in 2022. He is always willing to learn new things with full enthusiasm and passion. He is particularly interested in machine learning, deep learning, pattern recognition, biomedical imaging, and computer vision. He has hands-on experience in Python, Keras, TensorFlow, Pandas, Matplotlib, Seaborn, Sklearn, and SciPy. He is currently researching histopathology image-based automated disease detection systems. His enthusiasm for his field has driven him to seek further opportunities in academia. He is currently looking for Ph.D. or M.Sc. positions in the exciting fields of machine learning, deep learning, and biomedical imaging.

M. F. MRIDHA (Senior Member, IEEE) received the Ph.D. degree in AI/ML from Jahangirnagar University, in 2017. He is currently an Associate Professor with the Department of Computer Science, American International University–Bangladesh (AIUB). Before that, he was an Associate Professor and the Chairperson with the Department of CSE, Bangladesh University of Business and Technology. He was a Faculty Member and the Graduate Head with the CSE Department, University of Asia Pacific, from 2012 to 2019. His research experience, within both academia and industry, results in over 120 journal and conference publications. His research work contributed to the reputed journal of *Scientific Reports* (Nature), *Knowledge-Based Systems*, *Artificial Intelligence Review*, IEEE Access, *Sensors*, *Cancers*, and *Applied Sciences*. For more than ten years, he has been with the master's and undergraduate students as a Supervisor of their thesis work. His research interests include artificial intelligence (AI), machine learning, deep learning, natural language processing (NLP), and big data analysis. He has served as a program committee member in several international conferences/workshops. He has served as a Reviewer for reputed journals and IEEE conferences, such as HONET, ICIEV, ICCIT, IJCCI, ICAEE, ICCAIE, ICSIPA, SCORED, ISIEA, APACE, ICOS, ISCAIE, BEIAC, ISWTA, IC3e, ISWTA, CoAST, icIVPR, ICSCT, 3ICT, and DATA21. He served as an Associate Editor for several journals, including *PLOS One*.

MD RASHEDUL ISLAM (Senior Member, IEEE) received the B.Sc. degree in computer science and engineering from the University of Rajshahi, Rajshahi, Bangladesh, in 2006, the M.Sc. degree in informatics from the Högskolan i Borås (University of Boras), Boras, Sweden, in 2011, and the Ph.D. degree in electrical, electronic, and computer engineering from the University of Ulsan, Ulsan, South Korea, in 2016. He is currently a Chief Researcher of computer vision and AI with Chowagiken Corporation, Japan, and also an Associate Professor (on leave) with the Department of Computer Science and Engineering, University of Asia Pacific (UAP), Dhaka, Bangladesh. Previously, he was a Senior Architect with the Research and Development Department, Exvision Corporation, Tokyo, Japan; a Visiting Researcher (a Postdoctoral Researcher) with the School of Computer Science and Engineering, The University of Aizu, Japan; a Graduate Research Assistant with the Embedded System Laboratory, University of Ulsan; an Assistant Professor with the Department of Computer Science and Engineering, UAP; and a Lecturer with the Department of Computer Science and Engineering, Leading University, Sylhet, Bangladesh. Also, he has a good experience in professional IT system analysis and development. His research interests include machine learning, signal and image processing, HCI, health informatics, and bearing fault diagnosis. He is a PC member of several international conferences. He is a member of the IEEE Computer Society and the IEEE Computational Intelligence Society. He has served as the Secretary for the Organizing Committee of the 19th International Conference on Computer and Information Technology 2017 (ICCIT2017); the Organizing Chair for the Organizing Committee of the ACM-ICPC Dhaka Regional Site 2017; the Head for the Self-Assessment Committee (SAC) of the Department of CSE under IQAC, UAP; a Coordinator for the MCSE Program, Department of CSE, UAP; a Convener for the Software and Hardware Club, Department of CSE, UAP; a Coordinator for the Admission Committee, Department of CSE, UAP; and a Treasurer for Bangladesh Advanced Computing Society. He is a Reviewer of several journals, such as IEEE Transactions on Industrial Electronics, IEEE Access, *Applied Science*, *Multimedia Tools and Applications*, *Cluster Computing*, *Shock and Vibration*, *Journal of Information Processing Systems*.

YUTAKA WATANOBE (Member, IEEE) received the master's and Ph.D. degrees from The University of Aizu, Japan, in 2004 and 2007, respectively. He was a Research Fellow with Japan Society for the Promotion of Science (JSPS) and The University of Aizu, in 2007. He is currently a Senior Associate Professor with The University of Aizu. His research interests include visual programming, smart learning, data mining, and robotics.