axioms

MDPI

Article

Novel Design of Compact Data Learning Frameworks for Time-Series Forecasting

Song-Kyoo Kim

Faculty of Applied Sciences, Macao Polytechnic University, Macao SAR, China; amang@mpu.edu.mo

Abstract

This paper focuses on optimizing machine learning-based time-series forecasting models by constructing compact data. Compact Data Learning for Time Series (CDL-TS) is a novel framework aimed at minimizing forecasting model errors. By utilizing reduced sampling and robust comparison procedures, CDL-TS addresses the challenges of forecasting models on extensive real-time data systems. By strategically minimizing data size while maintaining accuracy, CDL-TS presents an innovative framework that facilitates robust predictions and enhances operational efficiency. The combined bivariate performance measure-based optimization effectively balances sampling frequency with the mean square error (MSE) to improve the trade operation performance in stock markets. Through a series of empirical applications, particularly involving the M/M/1 queuing system and various stock trade optimizations with global high-tech companies, the CDL-TS framework has proven its effectiveness by significantly minimizing both forecasting errors and operational costs. This accomplishment highlights the robust capabilities of CDL-TS in enhancing predictive accuracy while facilitating operation cost savings across different domains, including complex systems like queues and real-time financial markets.

Keywords: compact data learning; data reduction; machine learning; time series; continuous time domain; queuing system

MSC: 62A99; 62H12; 62H20; 68T99; 68Q32

1. Introduction

Pattern recognition tasks, spanning from image processing to natural language processing, are areas where artificial intelligence (AI) and machine learning (ML) show exceptional performance. Modern AI and ML models are essential in a variety of applications [1–3]. Supervised learning, which is one type of ML task, involves creating a function that maps inputs to outputs based on example input–output pairs [4]. This approach generates a function from labeled training data composed of training examples [5]. As the most widely adopted ML training method, supervised learning is utilized in most AI applications. The convolutional neural network (CNN) is a frequently employed ML model in diverse computer vision applications. Few-shot learning (FSL) has been developed to tackle the limitation of small dataset sizes in machine learning training [6]. Acquiring large training datasets is a significant challenge in ML training [7]. To alleviate this, many ML algorithms require that each training dataset aligns with a prototype, but this creates a major practical obstacle in learning thousands of object categories if optimal points are hard to determine [7,8]. Conversely, FSL is capable of efficiently generalizing to new tasks

Check for updates

Academic Editors: Guangying Lv and Jian Wang

Received: 15 April 2026
Revised: 7 May 2026
Accepted: 8 May 2026
Published: 11 May 2026

Copyright: © 2026 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

that contain only a handful of labeled samples, decreasing the necessity for large-scale supervised data gathering by making use of prior knowledge [6].

Compact Data Learning (CDL) in the context of machine learning aims to enhance the efficiency of model training for clarifications [9,10]. Alternatively, the tailor-made compact data should capture the maximum amount of knowledge patterns on a microscopic scale, allowing for effective and customized utilization of large data systems [10]. These CDL approaches aim to optimize runtime speed by reducing sample size and minimizing data features. CDL not only boosts runtime efficiency but also ensures that the original accuracy performance is minimally impacted [9]. Compact Data Learning for Time Series (CDL-TS) is newly designed to reduce errors in forecasting models [11]. This innovative framework offers unique and groundbreaking strategies for creating compact data in a wider time-series context. By employing reduced sampling to decrease data size and implementing a robust comparison and selection procedure to optimize sampling duration, we have successfully addressed the challenges associated with training models on expansive real-time data collection systems. Notably, this groundbreaking methodology enhances runtime efficiency and ensures that the impact of the differences between original and forecasting data is minimal. The results of this research provide insightful knowledge and practical techniques for boosting model training efficiency, particularly in contexts where there is data size reduction in continuous time-series data. As a demonstration of the application of this innovation, this research provides an analytic solution for the M/M/1 queuing system with continuous time parameters under a finite time limit. Generally, the infinite time horizon condition is considered when solving the M/M/1 queuing system with continuous time parameters, but the limited time condition has not been resolved before, even for the basic M/M/1 queuing system. Additionally, the innovative optimization framework designed for CDL-TS has been presented in this paper. The combined bivariate performance measure (CBPM) is a groundbreaking metric developed to evaluate systems marked by conflicting performance metrics [12]. By effectively capturing the intricate interactions between these performance measures, the CBPM enriches our understanding of how they collectively influence overall system performance. This comprehensive approach enables more thoughtful decision-making when optimizing systems that require a delicate balance of conflicting objectives [12]. This research introduces an optimization method through the adaptation of the CBPM. This optimization method formulates the objective function by integrating two conflicting performance metrics while relying on a single variable. In essence, the CBPM plays a pivotal role in optimizing the balance between system efficiency and accuracy, allowing system designers to better manage and enhance complex systems that possess conflicting goals. Such an approach unveils new possibilities for improving system performance across various fields, particularly within data analysis and decision-making processes.

This paper is organized as follows: Section 2 provides the mathematical background for constructing the Compact Data Learning for Time Series (CDL-TS) forecasting system. In this section a simple statistical forecasting model is introduced, but it can be replaced with other machine learning-based forecasting models. The M/M/1 queuing system with a continuous time parameter and its analytic solution are introduced in Section 3. This queuing system is used to demonstrate the optimization of sampling duration for minimizing forecasting errors. The mean square error (MSE) performance comparisons with different sampling durations are also explained in this section. Section 4 delves into the groundbreaking optimization method that has been developed for the CDL-TS framework. This section not only outlines the essential elements of the core optimization technique but also describes its practical applications in real-world stock markets, showcasing how these

methodologies can be effectively implemented to enhance trading strategies. Finally, the conclusion is presented in Section 5.

2. Statistical Forecasting Reference Function

The statistical forecasting function within Compact Data Learning (SFF-CDL) provides a straightforward yet highly effective approach, delivering competitive regression performance outcomes across diverse time-series applications [11,13]. Its simplicity facilitates efficient implementation and integration with subsequent augmentation techniques, while maintaining robust predictive capability comparable to more complex models. This balance of computational lightness and accuracy positions SFF-CDL as a reliable baseline for evaluating advanced methods, such as customized convolutional filters and dynamic filter augmentation, in large-scale forecasting tasks [11]. Let $Y(t)$ be an original continuous time-series function, and T_0 is the present moment of selecting samples [11]. $y_k^0, k = 0, 1, \ldots, n$ are the actual sampling values from the original time series $Y(t)$, which can be described as follows:

$$y_k^0(\varsigma) = Y(t_k(\varsigma)), k = 0, 1, \ldots, n(\varsigma), \tag{1}$$

where

$$t_k(\varsigma) = t_0 + \varsigma \cdot k, \tag{2}$$

and ς is the duration between samplings, which is a deterministic value. The discrete function y_k^0 by sampling from the original continuous function $Y(t)$ at the moment of t_k is defined as follows:

$$y_k^0 = \begin{cases} Y(t_0), & k = 0, \\ Y(t_k(\varsigma)), & k = 1, \ldots, n(\varsigma), \end{cases} \tag{3}$$

where

$$n(\varsigma) = \left\lfloor \frac{T_0}{\varsigma} \right\rfloor, t_0 = \mathrm{Max}(0, \delta), \tag{4}$$

$$\delta = T_0 - \varsigma \cdot \left\lfloor \frac{T_0}{\varsigma} \right\rfloor. \tag{5}$$

The basic idea of the statistical forecasting reference model could be described as follows:

$$y_{k+1}^1 = y_k^1 + \Delta_k(\varsigma), y_0^1 = y_0^0, \tag{6}$$

where

$$\Delta_k(\varsigma) = \frac{y_k^0 - y_{k-1}^0}{\varsigma}, \Delta_0(\varsigma) = 0. \tag{7}$$

This basic forecasting reference model could be adjusted based on real present data y_k^0 instead of based on predicted data y_k^1 at the kth epoch. Hence, the reinforcement statistical forecasting model could be constructed from (6) as follows:

$$y_{k+1}^1 = y_k^0 + \Delta_k(\varsigma). \tag{8}$$

From (7)–(8), the reinforcement statistical forecasting model is constructed as follows:

$$y_{k+1}^1(\varsigma) = Y(t_0 + k \cdot \varsigma)\left(1 + \frac{1}{\varsigma}\right) - \frac{Y(t_0 - \varsigma + k \cdot \varsigma)}{\varsigma}, \tag{9}$$

and $y_0^1 = y_0^0$, $k = 1, \ldots, n(\varsigma)$. It indicates the predicted value of the original function based on the kth sampling with the sampling duration ς. The predicted value y_k^1 has the gaps (or errors) with the original value y_k^0, and the mean square error (MSE) could be calculated to measure the real-time-series forecasting system. Let us consider the MSE function $\xi(\varsigma)$, which depends on the sampling duration ς, and this MSE function of the real-time-series forecasting system is constructed as follows:

$$\xi(\varsigma) = \frac{1}{n(\varsigma)} \sum_{k=1}^{n(\varsigma)} \left(y_k^1(\varsigma) - y_k^0(\varsigma)\right)^2, \tag{10}$$

and

$$\varsigma^* = \underset{\varsigma}{argmin}\left\{ {}^{\exists}\varsigma \middle| \xi(\varsigma), \varsigma \in [t_0, T_0] \right\}, \tag{11}$$

which is the optimal sampling duration ς^* for minimizing the MSE of a forecasting system. It is demonstrated that a specific sampling number yields the best performance by minimizing the gaps between predicted and real values.

3. M/M/1 Queuing System with Continues Time Parameters

This section demonstrates Compact Data Learning for a time-series forecasting system [11]. A basic M/M/1 queuing system with a continuous time parameter has been employed for the time-series system. A simple statistical forecasting model has been proposed to illustrate performance differences dependent on the sampling duration, and its optimization has been demonstrated.

3.1. Preliminaries

The M/M/1 queue describes the queue length in a single-server system, with arrivals governed by a Poisson process and service times following an exponential distribution. As the most basic queuing system, it is an appealing subject of study due to its ability to yield closed-form expressions for many relevant metrics [14]. The particulars of the M/M/1 queuing system for this demonstration are as follows: the number of customers arriving in a given time interval t follows a Poisson process with parameter λ. The time between any two arrivals is exponentially distributed with parameter λ. The time to serve a single customer follows an exponential distribution with parameter μ with one server. Although not specifically stated, both the population and the queue size are assumed to be infinite, and the service is conducted in a FIFO order. Let us consider a classical M/M/1 queuing system [14] with a continuous time parameter, and $N(\theta)$ is the number of customers in the queue. The probability generating function (PGF) of this queue without time parameters $Q(z, t)$ is defined as follows [14]:

$$Q(z, t) = \sum_{k \geq 0} P\{N(t) = k\} z^k, |z| \leq 1. \tag{12}$$

The Laplace transform of (12) is applied as follows:

$$\mathcal{L}(\bullet)(\theta) = \int_{t=0}^{\infty} e^{-\theta \cdot t}(\bullet) dt, \ \mathrm{Re}(\theta) > 0. \tag{13}$$

From (12)–(13), we have

$$Q^*(z, \theta) = \int_0^{\infty} e^{-\theta t} Q(z, t) dt = \mathcal{L}\{Q(z, t)\}(\theta), \tag{14}$$

with the inverse

$$Q(z,t) = \mathcal{L}^{-1}\{Q^*(z,\theta)\}(t),\ \mathrm{Re}(\theta) > 0, \tag{15}$$

which is the PGF (probability generating function) of the M/M/1 queuing system, which could be analyzed by Rouche's theorem, which determines the number of zeros of holomorphic functions inside a closed contour. This principle facilitates efficient root analysis for polynomials and stability assessments in control systems without explicit computation of individual zeros. By Rouche's theorem,

$$Q^*(z,\theta) = \frac{1}{\lambda z_1(1-z_0)\left(1-\frac{z}{z_1}\right)}, \tag{16}$$

where

$$z_0 = \frac{\theta+\lambda+\mu-\sqrt{(\theta+\lambda+\mu)^2-4\lambda\mu}}{2\lambda}, |z_0| \leq 1, \tag{17}$$

and

$$z_1 = \frac{\theta+\lambda+\mu+\sqrt{(\theta+\lambda+\mu)^2-4\lambda\mu}}{2\lambda}, |z_1| \leq 1. \tag{18}$$

The Laplace transform of the average length of the queue $\mathcal{N}$ could be defined as follows:

$$\mathcal{N}(\theta) := \mathcal{L}\{\mathbb{E}[N(t)]\}(\theta) = \left.\frac{\partial Q^*(z,\theta)}{\partial z}\right|_{z=1}. \tag{19}$$

From (16)–(19), we have

$$\mathcal{N}(\theta) = \left(\frac{1}{\lambda}\right)\left(\frac{1}{1-z_0}\right)\left(\frac{1}{(z_1-1)^2}\right), \tag{20}$$

where

$$1-z_0 = \frac{\sqrt{(\theta+\lambda+\mu)^2-4\lambda\mu}-(\theta-(\lambda-\mu))}{2\lambda}, \tag{21}$$

$$(z_1-1)^2 = \left(\frac{\sqrt{(\theta+\lambda+\mu)^2-4\lambda\mu}+(\theta-(\lambda-\mu))}{2\lambda}\right)^2. \tag{22}$$

From (21) and (22), the Laplace transform of the time-dependent average queue length $\mathcal{N}(\theta)$ from (20) could be calculated as follows:

$$\mathcal{N}(\theta) = \left(\tfrac{1}{\lambda}\right)\left(\frac{2\lambda}{\sqrt{(\theta+\lambda+\mu)^2-4\lambda\mu}-(\theta-(\lambda-\mu))}\right)\cdot \left(\frac{4\lambda^2}{\left(\sqrt{(\theta+\lambda+\mu)^2-4\lambda\mu}+(\theta-(\lambda-\mu))\right)^2}\right),$$

$$= \left(\frac{2\lambda}{\theta}\right)\left(\frac{1}{\left(\sqrt{(\theta+\lambda+\mu)^2-4\lambda\mu}+(\theta-(\lambda-\mu))\right)}\right)$$

$$= \left(\frac{1}{2\theta}\right)\left(\frac{\sqrt{(\theta+\lambda+\mu)^2-4\lambda\mu}-(\theta-(\lambda-\mu))}{\theta}\right)$$

$$= \left(\frac{1}{2\theta}\right)\left(\sqrt{1+\left(\frac{2(\lambda+\mu)}{\theta}+\frac{(\lambda-\mu)^2}{\theta^2}\right)}-\frac{\theta-(\lambda-\mu)}{\theta}\right),$$

where (by adapting the Taylor series)

$$\sqrt{1+\left(\frac{2(\lambda+\mu)}{\theta}+\frac{(\lambda-\mu)^2}{\theta^2}\right)} \simeq 1+\frac{1}{2}\left(\frac{2(\lambda+\mu)}{\theta}+\frac{(\lambda-\mu)^2}{\theta^2}\right)-\frac{1}{8}\left(\frac{2(\lambda+\mu)}{\theta}+\frac{(\lambda-\mu)^2}{\theta^2}\right)^2.$$

Finally, we have:

$$\mathcal{N}(\theta) = \left(\frac{\lambda}{\theta^2}-\frac{\lambda\mu}{\theta^3}-\frac{(\lambda+\mu)(\lambda-\mu)^2}{4\theta^4}-\frac{(\lambda-\mu)^4}{16\theta^5}\right), \tag{23}$$

and the mean of the queue length is as follows:

$$\mathbb{E}[N(t)] = \mathcal{L}^{-1}\{\mathcal{N}(\theta)\}(t). \tag{24}$$

From (23) and (24), the average queue length could be determined as follows:

$$\mathbb{E}[N(t)] = \lambda t - \frac{\lambda\mu t^2}{2} - \frac{(\lambda+\mu)(\lambda-\mu)^2 t^3}{24} - \frac{(\lambda-\mu)^4 t^4}{384}, \tag{25}$$

and the average waiting time could be found from Little's law [11] that the long-run average number of customers present equals the product of the average arrival rate and the average time a customer spends in the system (i.e., $N(t) = \lambda W(t)$). From this law, the average waiting time $W(t)$ is as follows:

$$\mathbb{E}[W(t)] = t - \frac{\mu t^2}{2} - \left(\frac{(\lambda+\mu)(\lambda-\mu)^2}{24\lambda}\right)t^3 - \left(\frac{(\lambda-\mu)^4}{384\lambda}\right)t^4. \tag{26}$$

It is noted that the final result based on an optimized sample duration is based on the waiting time of the queuing system, and the visualization of this case will be shown in the next section.

3.2. Optimizing the Sampling Duration

This section focuses on simulating the M/M/1 queuing system within the context of continuous time-series data. The example presented in this paper aims to optimize the sampling duration, or observation period, under the specific conditions defined for this study. The input rate for this system λ is established at 0.12, while the reciprocal of the service time $(=\frac{1}{\mu})$ is quantified as 7.69. Additionally, the maximum time parameter considered in this simulation is set to 20 (see Table 1).

Table 1. Initial conditions for the simulation [11].

Name	Value	Description
T_0	20	Domain of time
λ	0.12	Input rate
μ	0.13	Service rate
ϵ	0.01	Increasing step
$N(t)$	–	Queue length
$W(t)$	$N(t)/\lambda$	Waiting time
ς	–	Sampling duration

It is noted that the numerical values presented in Table 1 have been arbitrarily selected solely for the purpose of illustrating the concepts and facilitating visual interpretation within the context of the simulation. The sampling duration is systematically increased by an increment of 0.01 (i.e., $\epsilon = 0.01$), leading to variations in error levels when compared to the original continuous time series. A visualization of the error gaps associated with different sampling durations has been provided in Figure 1. In this figure, the bold black line represents the original time-series function that depicts the average waiting time within the M/M/1 queuing system. Each alteration in the sampling duration directly impacts the accuracy of the resultant time-series data, thereby underscoring the importance of selecting an appropriate sampling duration for reliable forecasting.

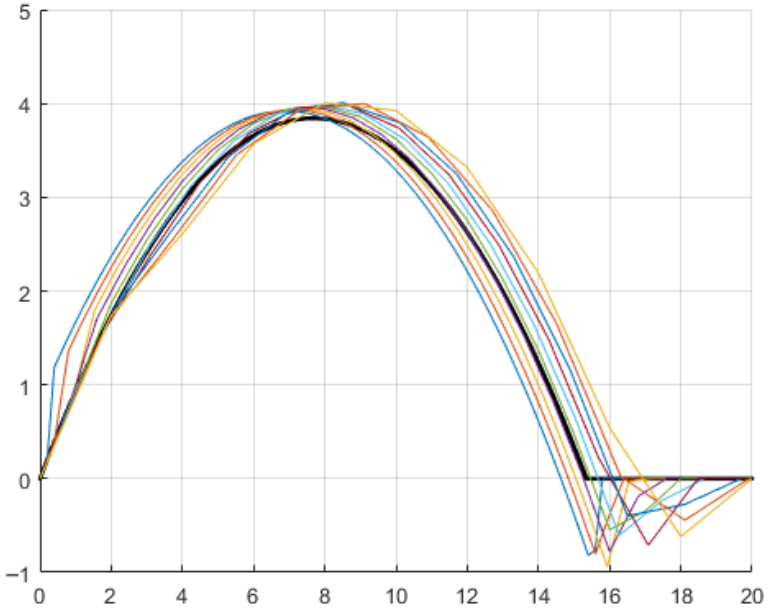


Figure 1. Visualization of the error gap within different sampling durations [11].

The optimal duration for sampling can be determined from (10) by calculating the mean square errors (MSEs) associated with various sampling durations. Figure 2 visually illustrates how the MSE varies as a function of these different sampling durations, specifically for the M/M/1 queuing system. Based on the experimental findings, it has been established that the optimal sampling duration is 0.87, which corresponds to a total of 22 samplings taken during the evaluation period (i.e., $n(\varsigma^*) = 22$), and the initial starting time t_0 is 0.989 (i.e., $t_0 = 0.989$) from (4). This highlights the critical role of selecting an appropriate sampling duration to effectively minimize forecasting errors in the analysis of this queuing system. As shown in the figure, the MSE of the time-series data of the M/M/1 queuing system has been reduced to 4.815×10^{-4}, which is more than eight times better than the worst sampling duration (i.e., 4.8×10^{-4} vs. 3.9×10^{-3} in Figure 2).

This proposed optimization approach achieves competitive performance relative to alternative methods, as the maximum algorithmic complexity remains bounded by $\mathcal{O}(n^2)$, depending solely on the total sample size n.

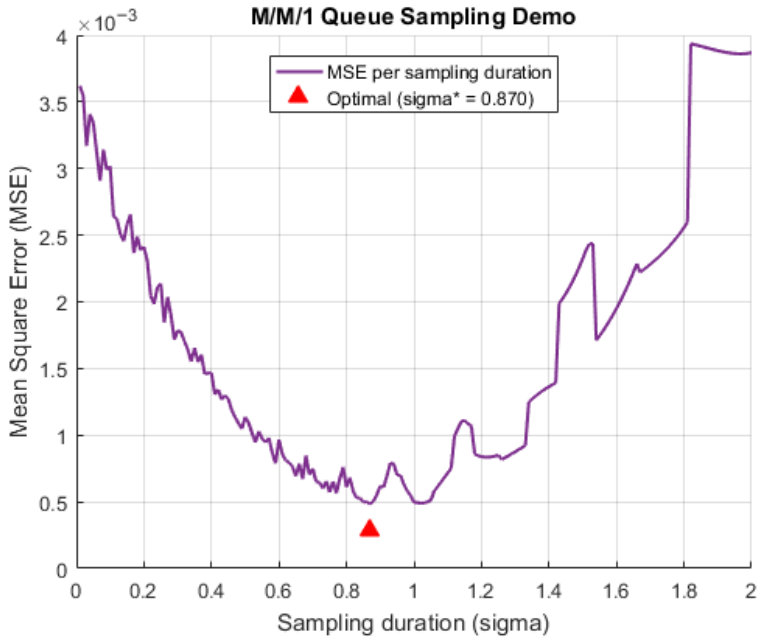


Figure 2. Sampling duration optimization for M/M/1 system [11].

4. Alternative Optimization for Stock Trade Operations

This section has implemented a groundbreaking optimization method through the adaptation of the combined bivariate performance measure (CBPM). The CBPM represents an innovative metric specifically designed to evaluate systems characterized by conflicting performance metrics, such as efficiency and accuracy [12]. The optimization method based on the CBPM formulates the objective function by integrating two conflicting performance metrics, while relying on a singular variable (the sampling duration). The two performance metrics in this objective function are the frequency of samplings and the mean square error (MSE). Both of these performance measures adhere to a smaller-is-better (SIB) trend [12], as a reduced sampling frequency enhances the efficiency of stock trade operations—leading to fewer transactions and lower costs—while a lower MSE indicates improved prediction accuracy.

4.1. Combined Bivariate Performance Measure-Based Optimization

Let us consider two functions $f_0(x)$ and $f_1(x)$, which illustrate the performance of the system across the time-series interval of $x \in [x_0, x_m]$. Given that both performance measures exhibit a SIB (*smaller-is-better*) trend [12], it follows that $f_0(x)$ and $f_1(x)$ must either both be decreasing functions or exhibit an inverse relationship. The objective function for the optimization based on the CBPM is expressed as follows:

$$\Psi(x) = \varphi_0(x) \cdot \varphi_1(x), \tag{27}$$

and

$$\varphi_0(x) = \frac{f_0^1 - f_0(x)}{\Delta f_0}, \varphi_1(x) = \frac{f_1^1 - f_1(x)}{\Delta f_1}, \tag{28}$$

where

$$f_0^0 = min\{f_0(x)\}, f_0^1 = max\{f_0(x)\},$$

and

$$f_1^0 = min\{f_1(x)\}, f_1^1 = max\{f_1(x)\}, x \in [x_0, x_m],$$

$$\Delta f_0 = |f_0^1 - f_0^0|, \Delta f_1 = |f_1^1 - f_1^0|. \tag{29}$$

From (27) and (28), the optimal value x^* of the objective function $\Psi(x)$ could be determined as follows:

$$x^* = max\left\{{}^{\exists}x \middle| \frac{d}{dx}\Psi(x) = 0, x \in [x_0, x_m]\right\}. \tag{30}$$

By maximizing the objective function from (27), the optimal duration for the sampling of the time-series x^* could be accurately determined from (30). Adopting such a holistic optimization strategy is crucial for enhancing the accuracy and efficiency of stock market prediction systems. This improvement not only contributes to more reliable forecasts but also plays a significant role in minimizing potential risks associated with stock trading operations. This innovative optimization method could enhance the efficiency and overall functionality of real-time systems and facilitate the earlier detection of potential defects.

4.2. Optimizing the Stock Trade Operations

In the pursuit of enhancing stock trade operations, this section discusses the implementation of the CBPM, which optimally balances conflicting metrics: the sampling frequency and MSE. By minimizing the sampling frequency and MSE, trade operations become more cost-effective and responsive, ultimately improving market predictions. This section also presents real-world examples from the stock market, specifically focusing on major stocks, such as Apple, AMD, and Google. These cases illustrate how the CBPM effectively reduces the MSE while ensuring a lower sampling frequency. This advanced technique not only enhances the efficiency of real-time data systems but also facilitates early detection of defects, significantly improving the overall quality and reliability of stock trading operations. Such innovations provide a pathway to mitigate risks and optimize trading efficiency, thereby cultivating a more resilient market environment. The optimization case for Apple stock has been illustrated in Figure 3. The objective function $\Psi(x)$ for the stock trade operations from (27) could be optimized when the sample duration of the stock is 7.3 h (i.e, $x^* = 7.3$ [hr]). It indicates that both efficiency and prediction accuracy balance best, before again declining as parameters continue to increase. The intersection points and values suggest a diminishing return, emphasizing the trade-off inherent in the sampling frequency versus the error rate. In the figures in this section, the green line represents the frequency performance (SIB), the orange line illustrates the MSE performance (SIB), and the purple line denotes the CBPM objective, which integrates both measures.

An additional example of optimization for AMD stock is presented in Figure 4. This figure depicts the correlation between performance measures in predicting AMD stock. The performance parameter corresponds to the sampling hours. Notably, the bold purple line starts to decline sharply after approximately 23.5 h, which is the optimal performance point where the trade-off between the sampling frequency and prediction accuracy is maximized (i.e., $x^* = 23.5$ [hr]).

As the performance parameter increases beyond this point, the ratio decreases, highlighting diminishing returns and the importance of finding an optimal sampling strategy to enhance prediction reliability while controlling costs. This optimization is essential for effective decision-making in stock trading. The last optimization visualization for Google stock has been illustrated in Figure 5. As the sampling duration x increases from 0 to 10 h, the frequency performance gradually declines, while the MSE performance initially rises

and then falls. The objective function for Google (i.e., purple bold line) oscillates, highlighting the interplay between these two metrics. The optimal performance is marked at 5 h, where the CBPM optimization achieves its maximum value, indicating the most favorable conditions for both the frequency and MSE (i.e., $x^* = 5$ [hr]). This visual representation underscores the importance of balancing these performance measures to improve overall system efficiency.

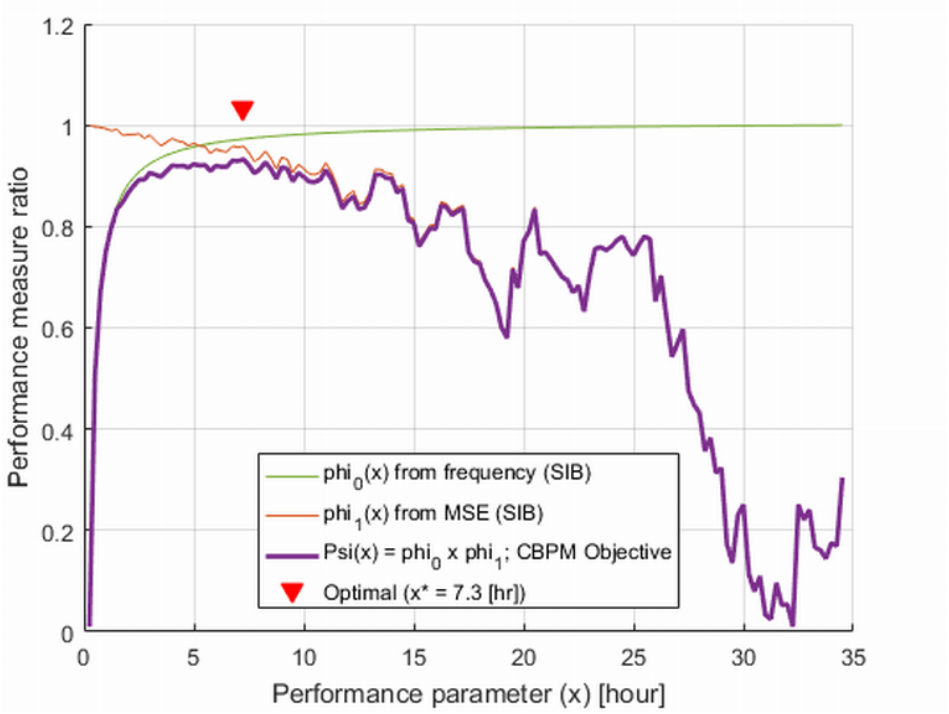


Figure 3. CBPM optimization visualization for Apple stock.

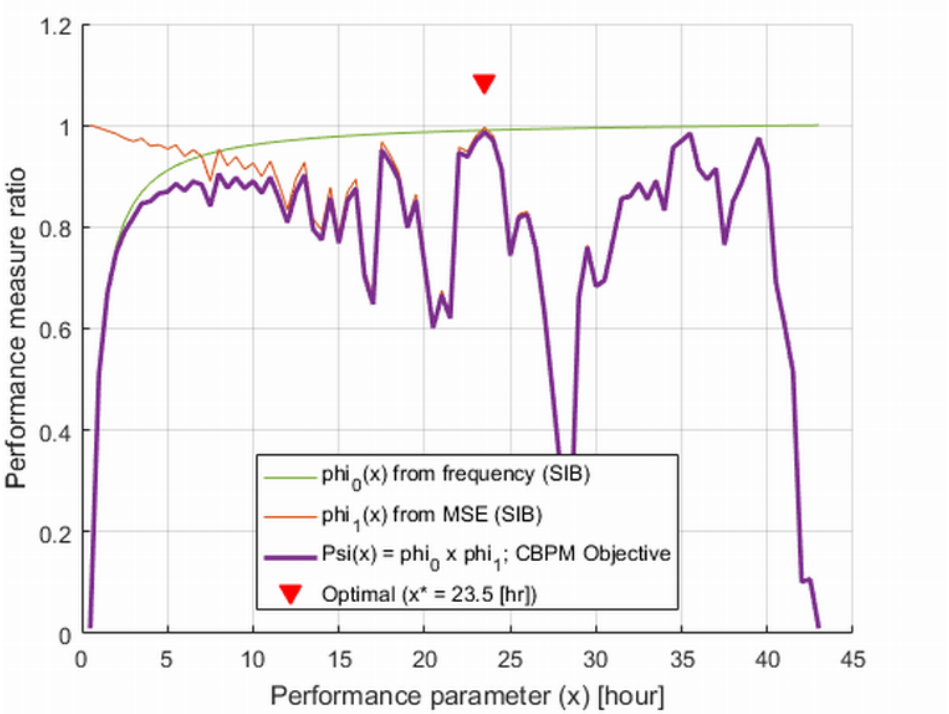


Figure 4. CBPM optimization visualization for AMD stock.

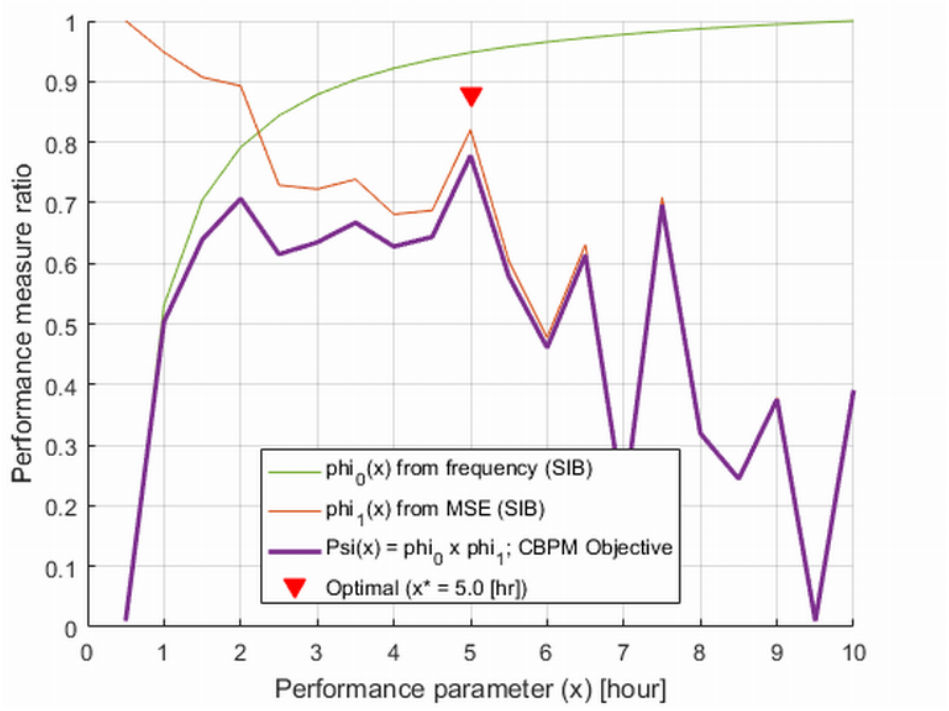


Figure 5. CBPM optimization visualization for Google stock.

Numerous techniques for time-series data reduction and forecasting optimization have been extensively investigated [15–18]. Recent studies report root mean square error (RMSE) ranges for real-world time-series datasets when employing various forecasting models, including SARIMAX (Seasonal AutoRegressive Integrated Moving Average with eXogenous regressors), SARIMA (Seasonal Autoregressive Integrated Moving Average), ARIMA (Autoregressive Integrated Moving Average) and the LSTM (Long Short-Term Memory) machine learning algorithm, spanning from 6.83 (SARIMAX) to 41.68 (LSTM) [18]. Alternatively, RMSE values for stock price time-series datasets obtained through the CDL-TS framework range from 0.869 (Apple) to 1.323 (Google), demonstrating superior forecasting performance compared to existing models despite variations across different time-series datasets.

5. Conclusions

A pioneering approach to improving machine learning training for time-series systems has been introduced, featuring a framework known as Compact Data Learning for Time Series (CDL-TS). This novel framework aims to accelerate the training phase without increasing forecasting error and provides unique strategies for creating compact data in a broader time-series context. Additionally, the integration of the combined bivariate performance measure (CBPM) optimization into the CDL-TS framework represents a substantial advancement in machine learning methodologies. This innovative approach effectively balances the often conflicting objectives of the sampling frequency and mean square error, significantly enhancing predictive accuracy while minimizing operation costs. This newly proposed framework might reveal limitations in high-dimensional, dynamic, non-stationary, or irregular series, multivariate extensions, and noisy conditions with long-range dependencies, alongside dependence on base models and fixed sampling constraints. The generalization capability of the CDL-TS framework to non-stationary, multivariate, or highly irregular time series requires further investigation. Future research should also focus on refining the CBPM methodology, exploring the adaptability of CDL-TS in other sectors, such as health informatics or climate forecasting, which could benefit from optimized data analysis techniques. Moreover, assessing the potential of integrating real-time data streams

with this framework could further enhance its relevance in dynamic applications. Overall, the significant achievements made through the CBPM optimization within CDL-TS lay the groundwork for transformative developments in machine learning, providing an exciting pathway for enhanced efficiency and accuracy in time-series forecasting.

Funding: This work was supported in part by the Macao Polytechnic University (MPU), under Grant RP/FCA-05/2024.

Data Availability Statement: The datasets and Matlab source codes that are related to this paper are available from the GitHub repository (https://github.com/amangkim/cdlts, accessed on 30 March 2026).

Acknowledgments: This paper was revised using AI/ML-assisted tools (Grok 3). Special thanks to the reviewers who provided valuable advice for improving this paper. This paper is an extended research contribution from the previous research [11].

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Barreno, M.; Nelson, B.; Sears, R.; Joseph, A.D.; Tygar, J.D. Can machine learning be secure? In *ASIACCS '06*; Association for Computing Machinery: New York, NY, USA, 2006; pp. 16–25.
2. Xu, Z.; Saleh, J.H. Machine learning for reliability engineering and safety applications: Review of current status and future opportunities. *Reliab. Eng. Syst. Saf.* **2021**, *211*, 107530. [CrossRef]
3. Olufowobi, H.; Engel, R.; Baracaldo, N.; Bathen, L.A.D.; Tata, S.; Ludwig, H. Data Provenance Model for Internet of Things (IoT) Systems. In *Proceedings of the Service-Oriented Computing—ICSOC 2016 Workshops*; Springer: Cham, Swizerland, 2017; pp. 85–91.
4. Russell, S.; Norvig, P. *Artificial Intelligence: A Modern Approach*, 3rd ed.; Pearson: London, UK, 2009
5. Mohri, M.; Afshin Rostamizadeh, A.T. *Foundations of Machine Learning (Adaptive Computation and Machine Learning)*; MIT Press: Cambridge, MA, USA, 2012.
6. Wang, Y.; Yao, Q.; Kwok, J.T.; Ni, L.M. Generalizing from a Few Examples: A Survey on Few-shot Learning. *ACM Comput. Surv.* **2020**, *53*, 63. [CrossRef] [PubMed]
7. Fei-Fei, L.; Fergus, R.; Perona, P. One-shot learning of object categories. *IEEE Trans. Pattern Anal. Mach. Intell.* **2006**, *28*, 594–611. [CrossRef] [PubMed]
8. Fink, M. Object classification from a single example utilizing class relevance metrics. In Proceedings of the Advances in Neural Information Processing Systems, NIPS'04, Vancouver, BC, Canada, 13–18 December 2004; pp. 449–456.
9. Kim, S.K. Compact Data Learning For ML Classifications. *Axioms* **2024**, *13*, 137. [CrossRef]
10. Kim, S.K. Toward Compact Data from Big Data. In *Proceedings of the 2020 15th International Conference for Internet Technology and Secured Transactions (ICITST), Virtual, 9–11 December 2020*; IEEE: Piscataway, NJ, USA, 2020; pp. 1–5.
11. Kim, S.K. Compact Data Learning For Time-Series Forecasting Systems. In Proceedings of the 2024 6th International Conference on Electrical, Control and Instrumentation Engineering, Pattaya, Thailand, 23–24 November 2024; pp. 1–5.
12. Kim, S.K. Combined Bivariate Performance Measure. *IEEE Trans. Instrum. Meas.* **2024**, *73*, 1009404. [CrossRef]
13. Kim, S.K. Customized Dynamic Filter Augmentation. *IEEE Signal Process. Lett.* **2026**, *33*, 639–642. [CrossRef]
14. Kleinrock, L. *Queueing Systems. Volume 1: Theory*, 1st ed.; Wiley-Interscience: Hoboken, NJ, USA, 1975.
15. Ashraf, M.; Anowar, F.; Setu, J.H.; Chowdhury, A.I.; Ahmed, E.; Islam, A.; Al-Mamun, A. A Survey on Dimensionality Reduction Techniques for Time-Series Data. *IEEE Access* **2023**, *11*, 42909–42923. [CrossRef]
16. Pealat, C.; Bouleux, G.; Cheutet, V. Improved Time-Series Clustering with UMAP dimension reduction method. In Proceedings of the 2020 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 10–15 January 2021; pp. 5658–5665.
17. Debnath, M.; Alamgeer, S.; Ngu, A.H.H. Improving Time-Series Forecasting with Statistical and AI-Driven Feature Optimization. In Proceedings of the 2025 IEEE 49th Annual Computers, Software, and Applications Conference (COMPSAC), Toronto, ON, Canada, 8–11 July 2025; pp. 901–906.
18. Raikar, G.; D, H.; Guruprasad, N. Advanced Forecasting and Analysis Techniques for Optimizing Website Performance: A Comparative Study of Time Series Models. In Proceedings of the 2024 International Conference on IoT, Communication and Automation Technology (ICICAT), Gorakhpur, India, 23–24 November 2024; pp. 414–419.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.