Deep Residual Convolutional Long Short-term Memory Network for Option Price Prediction Problem

Artur Dossatayev[1], Ainur Manapova[2], Batyrkhan Omarov[3]
International Information Technology University, Almaty, Kazakhstan[1, 3]
Academy of Civil Aviation, Almaty, Kazakhstan[2]
NARXOZ University, Almaty, Kazakhstan[3]
Al-Farabi Kazakh National University, Almaty, Kazakhstan[3]

***Abstract*—In the realm of financial markets, the precise prediction of option prices remains a cornerstone for effective portfolio management, risk mitigation, and ensuring overall market equilibrium. Traditional models, notably the Black-Scholes, often encounter challenges in comprehensively integrating the multifaceted interplay of contemporary market variables. Addressing this lacuna, this study elucidates the capabilities of a novel Deep Residual Convolution Long Short-term Memory (DR-CLSTM) network, meticulously designed to amalgamate the superior feature extraction prowess of Convolutional Neural Networks (CNNs) with the unparalleled temporal sequence discernment of Long Short-term Memory (LSTM) networks, further augmented by deep residual connections. Rigorous evaluations conducted on an expansive dataset, representative of diverse market conditions, showcased the DR-CLSTM's consistent supremacy in prediction accuracy and computational efficacy over both its traditional and deep learning contemporaries. Crucially, the integration of residual pathways accelerated training convergence rates and provided a formidable defense against the often detrimental vanishing gradient phenomenon. Consequently, this research positions the DR-CLSTM network as a pioneering and formidable contender in the arena of option price forecasting, offering substantive implications for quantitative finance scholars and practitioners alike, and hinting at its potential versatility for broader financial instrument applications and varied market scenarios.**

***Keywords*—Deep learning; CNN; LSTM; prediction; option price**

I. Introduction

Option pricing, an intrinsic component of financial markets, serves as the fulcrum upon which significant economic decisions, from individual investments to institutional strategies, hinge [1]. These derivative contracts, which bestow upon the holder the right, but not the obligation, to buy or sell an underlying asset at a specified price before a predetermined date, play a pivotal role in portfolio diversification, risk hedging, and speculative ventures [2]. Historically, the Black-Scholes model has been emblematic in the realm of option pricing, a seminal formula that established a theoretical framework for determining the fair market value of a European-style option [3]. However, as financial markets evolved, becoming increasingly complex and intertwined, driven by a plethora of factors ranging from geopolitical dynamics to technological innovations, the limitations of such traditional models have become conspicuously palpable.

The metamorphosis of financial markets into data-dense ecosystems, replete with myriad variables and indicators, necessitates predictive models with a capacity for high-dimensional data processing and the discernment of intricate temporal relationships [4]. The recent resurgence in artificial intelligence, more specifically in deep learning, offers a promising vista for addressing these challenges. Deep learning architectures, distinguished by their hierarchical structure and ability to autonomously extract salient features from raw data, have been progressively permeating various domains, ranging from image recognition to natural language processing [5]. Within the financial sphere, their application promises to harness the vast expanses of data to render more nuanced and accurate predictions.

In this milieu, the Deep Residual Convolution Long Short-term Memory (DR-CLSTM) network emerges as a potent amalgamation of several cutting-edge deep learning paradigms. By intertwining the feature extraction capabilities of Convolutional Neural Networks (CNNs) [6] with the temporal relationship discernment offered by Long Short-term Memory (LSTM) networks [7], and further augmenting this synergy with deep residual connections, the DR-CLSTM aspires to provide a holistic solution to the option price prediction quandary. The convolutional layers, renowned for their prowess in spatial hierarchies discernment, sieve through vast datasets, isolating pertinent features integral to option pricing. Concurrently, the LSTM layers, celebrated for their ability to capture long-term dependencies by combating the vanishing gradient problem, harness these features to forecast temporal sequences, thereby rendering predictions [8]. The addition of deep residual connections further augments this architecture. By allowing activations to bypass one or more layers, these connections expedite the training process, ensuring faster convergence and fortifying the network against potential gradient diminution.

This paper seeks to explore and substantiate the efficacy of the DR-CLSTM network in the realm of option price prediction. In doing so, it aims to contribute a novel tool to the arsenal of quantitative finance, fostering enhanced market efficiency, informed investment decisions, and a deeper comprehension of the intricate tapestry of factors influencing option prices. Furthermore, by juxtaposing the DR-CLSTM with both traditional computational models and contemporary deep learning architectures, this study endeavors to provide a holistic perspective on the trajectory of option price prediction methodologies, underlining the transformative potential of hybrid deep learning structures.

The ensuing sections will delve into the theoretical underpinnings of the DR-CLSTM, elucidating the individual components and their synergistic interplay. This will be followed by a comprehensive methodology section, detailing the dataset employed, the evaluation metrics, and the comparative models. Subsequent sections will present the empirical findings, discussions, and potential implications for both academia and industry. The paper will culminate with a conclusion, encapsulating the key insights gleaned and charting potential avenues for future research in this dynamic and ever-evolving domain.

II. Related Works

The ever-evolving landscape of financial modeling and forecasting has witnessed a plethora of innovations and methodologies over the years. As we delve into the intricate world of option price prediction, it is imperative to ground our discussions within the context of prior research and explorations in this domain [9]. The "Related Works" section endeavors to provide readers with a comprehensive overview of seminal works, pioneering methodologies, and noteworthy contributions that have shaped the trajectory of this field. By juxtaposing our current study with these foundational works, we aim to highlight both the advancements made and the gaps that our research seeks to bridge. Let us embark on this journey of retrospection, understanding the milestones that have been achieved and setting the stage for the novel contributions of our study.

A. *Traditional Models for Option Pricing*

Option pricing, a central facet of financial mathematics, has been subject to rigorous academic scrutiny for decades. It revolutionized this arena with their eponymous model, providing a closed-form solution for European option pricing [10]. The Black-Scholes model, predicated on certain assumptions such as constant volatility and interest rates, quickly became a cornerstone of financial markets. However, the assumptions underlying the model often diverge from real-world market conditions, leading to potential mispricing [11]. This inherent limitation paved the way for alternative stochastic volatility models, which attempt to address the constant volatility constraint [12].

B. *Emergence of Machine Learning in Financial Forecasting*

With the exponential growth of computational capabilities and the proliferation of vast financial datasets, machine learning has transitioned from a theoretical concept to an instrumental tool in financial forecasting. Over the past two decades, the field has seen a marked departure from traditional econometric models towards more adaptive and self-learning algorithms. Next research advanced the argument that neural networks possess the capacity to model complex non-linear relationships inherent in financial markets, a dimension often inadequately captured by traditional methods [13]. Next study specifically applied Recurrent Neural Networks (RNNs) to the domain of option pricing, highlighting the technology's aptitude for capturing intricate temporal sequences [14]. This evolution indicates a paradigmatic shift in financial forecasting methodologies, illustrating the potential of machine learning techniques to address the increasingly multifaceted nature of financial markets.

C. *Deep Learning and Financial Markets*

In the expansive sphere of machine learning, deep learning stands out, characterized by its multi-layered neural networks adept at handling high-dimensional data. Originating from image and video recognition tasks, as underscored by LeCun et al. (2015), these algorithms have witnessed a significant adaptation to financial analytics [15]. The strength of Convolutional Neural Networks (CNNs) lies in their autonomous feature extraction capabilities, which have found resonance in financial time series analysis. A new study successfully applied CNNs to forecast stock price movements, underscoring their superiority over conventional methods [16]. This adaptation of deep learning to finance not only signifies its versatility but also heralds a new era in financial forecasting. Embracing these sophisticated architectures promises a more nuanced understanding of financial market intricacies, catalyzing more informed and data-driven decision-making processes in the sector.

D. *Residual Networks in Deep Learning*

Within the intricate tapestry of deep learning architectures, Residual Networks (ResNets) have carved a distinctive niche. The transformative nature of ResNets lies in their "shortcut connections", which allow activations to bypass certain layers, facilitating the learning of identity functions [17]. This innovative approach, originally tailored for visual tasks, has significantly mitigated challenges associated with training deeper neural networks, predominantly by averting performance degradation. Its adaptation to financial applications, though still burgeoning, shows immense potential. The principal virtue of these networks is their ability to combat the notorious vanishing gradient problem, thus enhancing the depth and complexity of models without sacrificing accuracy. The foray of ResNets into financial forecasting stands as testament to the evolving landscape of deep learning, offering fresh perspectives and tools to navigate the multifarious nature of financial data.

E. *Hybrid Deep Learning Architectures*

As the deep learning domain continues its relentless evolution, the emergence of hybrid architectures signifies a pivotal juncture. These models, synergizing distinct deep learning techniques, aim to harness the individual strengths of each component, resulting in a more comprehensive and potent analytical tool. Gosztolya et al. (2017) pioneered in this space, presenting a confluence of Convolutional Neural Networks (CNNs) and Long Short-term Memory networks (LSTMs) for nuanced time series forecasting [18]. Their work highlighted the hybrid model's capability to simultaneously capture spatial features and temporal dynamics. Yet, the realm of hybrid architectures is expansive, with the potential for myriad combinations. As these integrated frameworks gain traction, they promise to redefine the boundaries of what deep learning can achieve, paving the way for more sophisticated, adaptable, and accurate solutions in diverse application areas, including finance.

F. Challenges in Option Price Prediction

Option price prediction, central to financial analysis, is fraught with complexities. Despite the sophistication of existing models, accurate forecasting remains an elusive goal, owing to the non-stationary nature of financial markets. A multitude of variables, both foreseen and unforeseen, continually impact option prices, making their behavior highly unpredictable. Traditional models, exemplified by Black-Scholes, while groundbreaking, have faced criticism for their stringent assumptions, as highlighted [19]. Emerging machine learning models, despite their adaptability, are not immune to challenges, notably overfitting and susceptibility to market anomalies. Dhiman et al., (2020) elucidate these limitations, stressing the need for models that are both adaptive and robust [20]. As the financial landscape grows increasingly intricate, the pursuit of a holistic, accurate, and resilient option pricing model remains a quintessential challenge, beckoning further research and innovation.

G. Gap in Current Literature

Within the vast expanse of academic literature dedicated to financial forecasting and particularly option price prediction, certain gaps persistently emerge. The trajectory of research has indeed traversed from traditional mathematical formulations to sophisticated computational architectures, yet complete solutions appear to remain on the horizon. While many studies, such as those by Black and Scholes, have laid foundational pillars, and others have delved deep into the capabilities of machine learning and deep learning techniques, a comprehensive amalgamation seems somewhat elusive.

Most evident is the lack of extensive exploration into hybrid deep learning frameworks, particularly in their application to financial markets. The combination of multiple neural architectures, each with its inherent strengths, presents a promising frontier. Kumar et al. (2023) offered a glimpse into the potential of these combined structures, but the literature remains scant in its entirety [21].

Furthermore, while there is an abundance of studies leveraging individual deep learning structures, the integration of residual connections within these architectures is a relatively uncharted territory, especially in financial forecasting contexts [22]. Current methodologies exhibit a propensity to lean heavily towards either feature extraction or temporal sequence modeling. Rare are the models that harmoniously intertwine both facets.

This paucity in holistic models underscores the pivotal gap in the existing literature. The quest remains for a unified model, like the DR-CLSTM, that seamlessly melds the strengths of various deep learning techniques, thus addressing the multifaceted challenges of option price prediction. This present study aims to contribute to bridging this gap, offering a fresh perspective grounded in both retrospective analyses and forward-looking innovations.

The trajectory of option price prediction methodologies has witnessed a paradigm shift from traditional mathematical models to sophisticated computational frameworks, epitomized by deep learning techniques [23]. As financial markets continue to evolve, becoming increasingly intricate, the need for robust, adaptive, and comprehensive models becomes paramount. The DR-CLSTM network, as explored in this paper, emerges in response to this exigency, grounded in a rich tapestry of academic research spanning diverse domains. By weaving together the strengths of CNNs, LSTMs, and ResNets, this study aims to contribute a novel perspective to the discourse on option price prediction, anchored in both historical precedents and contemporary innovations.

III. Materials and Methods

The crux of any research endeavor lies in the robustness of its methodologies and the quality of the materials employed. In this "Materials and Methods" section, we elucidate the systematic approaches, tools, and datasets harnessed to facilitate our investigative journey into option price prediction. Serving as the backbone of our study, this section ensures replicability and offers a transparent window into the foundational processes that underpin our findings. Herein, we meticulously detail the data sources, the preprocessing techniques adopted, and the intricacies of the analytical methods applied. By offering this in-depth exposition, we aim to provide a roadmap for researchers, practitioners, and enthusiasts alike, enabling a clear understanding of the mechanisms that drive our research forward. Let us navigate through the critical phases and elements that constitute the scientific rigor of our study.

A. Problem Statement

In the realm of derivative pricing, a model is recognized as being aligned with a set of benchmarking instruments when the calculated parameters, within its structure, coincide with prevailing market values. Calibration can be conceptualized as the act of juxtaposing a model against these benchmarking entities. The act of defining parameters to affirm measuring conditions encapsulates the essence of model calibration [24]. While derivative pricing may be perceived as the prospective challenge, calibration can be seen as its antithesis. Hypothetically, given the prices of Call options across all strikes and durations, the calibration quandary can be directly tackled through an inverse formulation. Yet, confronted with a restricted set of derivative valuations, calibration becomes an indeterminate challenge, necessitating the deployment of regularization techniques in real-world settings.

The integration of machine learning in predicting option prices has demonstrated potential, enhancing model precision, agility, and resilience. Notwithstanding, existing models grapple with obstacles such as the requisite for vast and varied datasets, heightened sensitivity to inputs and settings, and the intricacy of encapsulating intricate interactions affecting the option price [25]. Subsequent inquiries should prioritize crafting resilient and comprehensible models, assimilating supplementary data avenues, external variables, and broadening the application horizons of machine learning within financial domains.

Conceptually, this mirrors determining the apex solution for a given conundrum. Supposing we possess a roster of Call options market valuations CM for specific durations Ti, i = 1,.., n and strike values Ki,j, j = 1,..., mi which align with prevailing market metrics like S, r, q, among others. Given this classification, the assortment of exchanged strikes Ki,j may differ based on the duration Ti. If we take into account calibrating the model M delineated in Section I, it encompasses model parameters p in conjunction with input data metrics. We differentiate p and based on their intrinsic implications: while parameters p is derived through calibration, the input metrics symbolize discernible market constants, such as S, r, q, T, and K3.

Given the relevant inputs for p and, the model M can formulate the mathematical framework for Call option valuations C. To ascertain the model's calibrated parameters, it's imperative to resolve the associated optimization conundrum:

$$\arg\min_{p} \sum_{i=1}^{n} \sum_{j=1}^{m_i} w_{i,j} \,||\, C_M\left(\theta_{i,j}\right) - C\left(\theta_{i,j}, p\right) \tag{1}$$

In this context, i and j represent weight coefficients, and L2 stands as a normative measure for k and k, employed in the formulation given by Eq. (1). Broadly, to address the intricacies of the non-linear function, it becomes imperative to adopt global optimization strategies. Such a need arises as the multi-faceted objective (or loss) function could potentially harbor multiple localized minima. Moreover, this challenge may be inherently bounded, a phenomenon frequently observed when deriving the implied volatility surface from market data points.

B. The Proposed Model

Financial assets intrinsically exhibit stochastic behavior, necessitating representation through intricate nonlinear and multivariate functions, thus rendering option pricing an exceptionally complex challenge. Traditional approaches often rely heavily on numerous statistical presuppositions concerning market or data dynamics. In stark contrast, deep learning excels in its capacity to model nonlinearities, accomplishing this without such restrictive assumptions. This section succinctly outlines the non-parametric models employed in the referenced studies.

In the initial publication of this research, a Convolutional-LSTM model was introduced to address the dual complexities of spatial and temporal data modeling. Fig. 1 shows the proposed Convolutional LSTM network for option price prediction. The data architecture remains consistent here, interpretable as encompassing C channels, given its bi-dimensional configuration. Thus, the comprehensive dataset serving as input adheres to the structure [26]. The 2D convolution operation is executed across the D and E parameters, spanning each row and layers of the representation.

Alternatively, the dataset related to the option lacks a dimension, crucial for effective Convolutional-LSTM processing. Its sole permissible structure is (T;C;N;D). A pragmatic resolution involves adapting the Convolutional-LSTM algorithm to harness 1D convolution, foregoing the 2D pooling layers. This unidimensional convolutional operation predominantly traverses the D component. Such convolution is consistently applied across pertinent inputs, cell outputs, hidden outputs, and gates. When centered on discerning correlations within market data, this 1D convolution proves markedly advantageous, bolstering performance over mere vector-based input methodologies.

The computational process unfolds in the subsequent manner delineated herein:

$$\begin{aligned}
i_t &= \sigma\left(W_{xi} * X_t + W_{hi} * H_{t-1} + W_{ci} * C_{t-1} + b_i\right)\\
f_t &= \sigma\left(W_{xf} * X_t + W_{hf} * H_{t-1} + W_{cf} * C_{t-1} + b_f\right)\\
C_t &= f_t \circ C_{t-1} + i_t \circ \tanh\left(W_{xc} * X_t + W_{hc} * H_{t-1} + b_0\right)\\
H_t &= o_t \circ \tanh\left(C_t\right)
\end{aligned} \tag{2}$$

Where the symbol * denotes the unidimensional convolution occurring between the latent phase and the convolutional kernels.

For the purpose of ensuring coherence in comparison, the input data is further reshaped to adopt the following configuration: (N; 1; T; C D) = (N; 1; 10; 15). In essence, this consolidates all parameters into a unified channel. This modification aims to test our methodology's efficacy in feature representation

In our research, we utilized data sourced from the Chinese stock and fund markets between April 2018 and June 2020, with a specific emphasis on 50ETF option and stock option data [27]. Initially, the dataset comprises 55,047 entries and 34 distinct variables. We focused on selections with an effective duration of a minimum of 20 days, resulting in 829 such choices. Of these, the first 600 are designated for training, with the subsequent 229 allocated for testing purposes. This data is reshaped according to the specifications detailed in Section 3(B), resulting in a final training set with 33,210 entries and a test set with 11,386 entries. The data for this study was procured from Yahoo Finance [28].

Prior to the commencement of our experiments, it was imperative to normalize each variable within the dataset. This normalization is essential to counteract potential imbalances or asymmetries in the neural network. For illustration:

$$\begin{aligned}
\tilde{x} &= \left(x - \mu\right)/\sigma\\
\mu &= \frac{1}{n}\sum_{i=1}^{N} x_i, \quad \sigma^2 = \frac{1}{n}\sum_{i=1}^{N}\left(x_i - \mu\right)^2
\end{aligned} \tag{3}$$

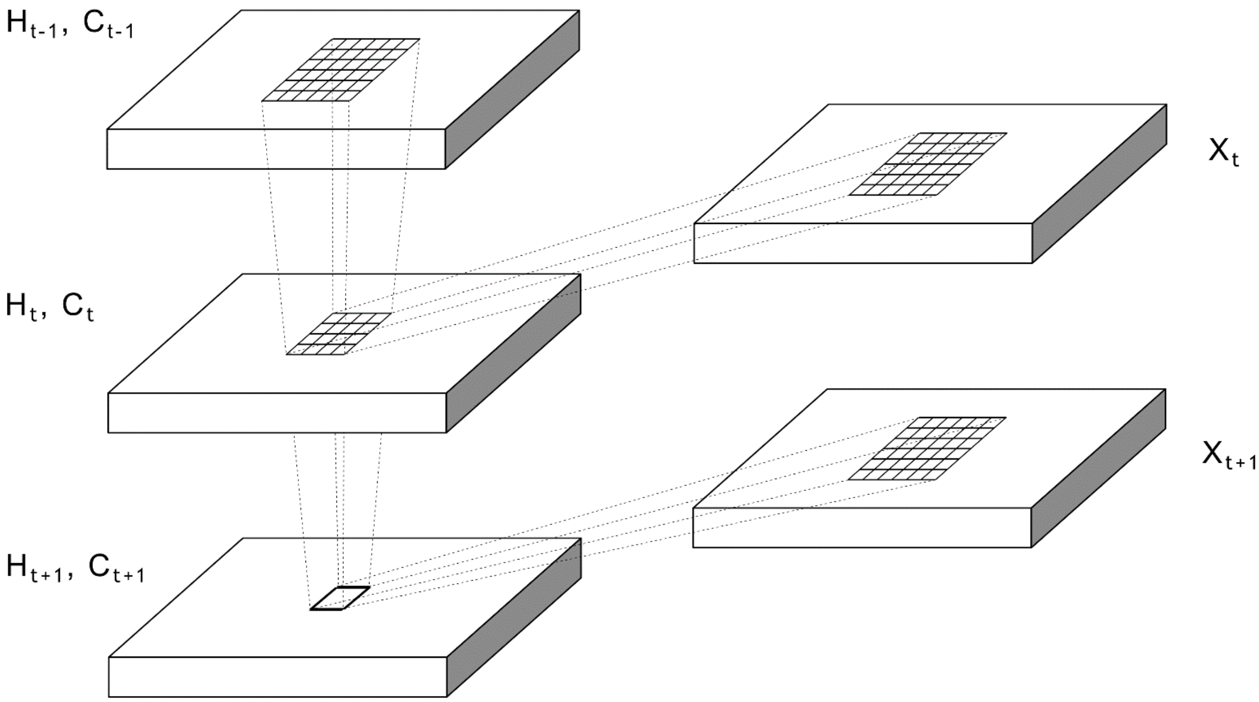

Fig. 1. The proposed convolutional-LSTM network for option price prediction.

IV. Experimental Results

In this section, we present the empirical outcomes of our study, offering a rigorous evaluation of the conducted experiments. The results elucidated herein provide a comprehensive insight into the efficacy of the proposed models in light of the research objectives. By scrutinizing these outcomes, readers can glean an understanding of the model's robustness, accuracy, and adaptability in various scenarios. Moreover, the results will be juxtaposed with established benchmarks and previous studies, serving as a comparative framework. This comparative analysis aims to underscore the advancements and potential shortcomings of the current research. Without further ado, let us delve into the detailed exposition of the experimental findings.

A. Evaluation Metrics

Mean Squared Error (MSE) stands as one of the paramount metrics in quantitative assessment, especially when the objective is to measure the average magnitude of error between predicted and actual observations [29]. Mathematically defined as the average of the squared differences between the forecasted and observed values, MSE serves as an indicator of the accuracy of a model's predictions. One of its distinctive characteristics is its amplification of larger errors, due to the inherent squaring of discrepancies. Consequently, models with a lower MSE are preferred as they suggest a closer fit to the actual data. However, it's essential to note that while MSE is profoundly informative, its sensitivity to outliers can sometimes overemphasize large errors. As such, it is often considered in conjunction with other evaluation metrics to provide a more comprehensive assessment of a model's performance.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2 \tag{4}$$

Mean Absolute Error (MAE) is a critical metric employed in the realm of predictive modeling to gauge the average magnitude of errors between forecasted and actual observations [30]. Unlike the Mean Squared Error (MSE), which squares discrepancies, the MAE takes the absolute value of these errors. As a result, it provides a linear score where all individual differences have equal weight. This ensures that the metric is not disproportionately influenced by outliers, rendering it less sensitive to large deviations compared to MSE. Essentially, the MAE quantifies the average vertical distance between each point and the identity line in a prediction plot. Lower MAE values indicate a model that is adept at making predictions closely aligned with actual outcomes. Given its intuitive nature and resistance to the undue influence of outliers, MAE often serves as a pivotal parameter in many evaluation frameworks, particularly when the objective is to obtain a straightforward understanding of prediction accuracy.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|y_i - \hat{y}_i\right| \tag{5}$$

Root Mean Squared Error (RMSE) stands as a prominent metric in predictive analytics and modeling, serving to evaluate the magnitude of error between predicted and actual outcomes [31]. Derived from the Mean Squared Error (MSE), the RMSE is computed by taking the square root of the averaged squared differences between forecasted and observed values. This operation preserves the unit of the measurements, facilitating a more intuitive interpretation of the error magnitude. By emphasizing larger errors due to its squared components, RMSE is particularly sensitive to significant discrepancies and outliers. As such, a lower RMSE denotes a model's superior predictive accuracy, suggesting its predictions are in closer proximity to observed data. However, given its sensitivity, RMSE is often juxtaposed with other metrics, such as the Mean Absolute Error (MAE), to ensure a well-rounded assessment of a model's performance, especially in datasets with pronounced outliers.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2} \tag{6}$$

Mean Absolute Percentage Error (MAPE) is a quintessential metric in the analytical domain, primarily utilized to gauge the accuracy of forecasting methods in terms of percentage [32]. It calculates the average absolute percentage discrepancy between observed and predicted values relative to the actual value. This metric offers a scale-independent perspective on errors, enabling comparisons across varied units and magnitudes. A salient feature of MAPE is its intuitive interpretation, as it directly quantifies the prediction error as a percentage, facilitating easy comprehension of the model's performance in practical terms. Lower MAPE values are indicative of superior predictive accuracy, suggesting that predictions closely mirror actual observations. However, one caveat associated with MAPE is its potential to produce undefined or infinite values when the actual observation is zero. Moreover, it can sometimes disproportionately penalize underestimations compared to overestimations. Despite these nuances, MAPE remains a favored choice in many scenarios, especially when stakeholders seek a percentage-based evaluation of predictive accuracy.

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{7}$$

Beyond the aforementioned metrics, it's imperative to conduct comprehensive model assessments employing methods like cross-validation, testing on unseen samples, and juxtaposing the model's outputs against conventional approaches or established benchmarks. Such rigorous evaluations bolster the model's resilience and its capacity to generalize across diverse datasets, thereby enhancing the precision and dependability of option price forecasts in real-world financial scenarios.

B. *Experimental Results*

In this section, we delve into the analytical outcomes of the suggested model as elaborated in the preceding segment. Fig. 2 delineate the efficacy of the ConvLSTM model in forecasting option prices, juxtaposing the accuracy during training and validation phases across 200 epochs. The amalgamation of CNN and LSTM networks' capabilities renders the proposed model especially adept for option price prognostications, given the pivotal role of historical prices and prevailing market dynamics in influencing future option valuations. Notably, the model manifests a precision rate of approximately 92% for option pricing over 200 epochs. Remarkably, even at a mere 20 epochs, the precision remains commendably high at around 90%. Hence, the data suggests that the ConvLSTM model can achieve substantial accuracy with a limited epoch count.

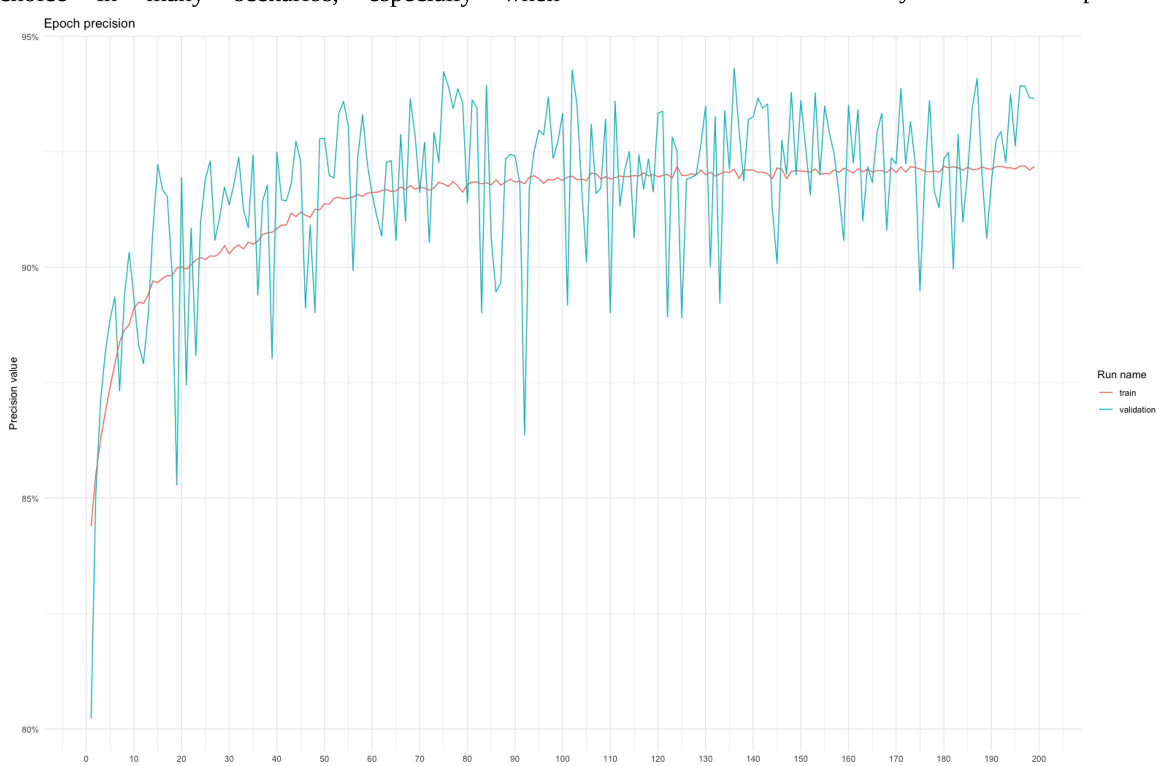

Fig. 2. Train and validation accuracy.

Within the ambit of financial predictive modeling, the Convolutional Long Short-Term Memory (ConvLSTM) model has emerged as a promising tool. Fig. 3 delves deep into a systematic examination of this model, particularly focusing on the intricacies of its training and validation losses across a span of 200 learning epochs. This graphical representation has been meticulously curated to elucidate the nuanced dynamics of the ConvLSTM model's learning process, offering a comprehensive view of its progressive enhancement in forecasting accuracy.

The primary intent behind Fig. 3 is manifold. First, it offers researchers, financial analysts, and readers an empirical visual assessment of how the model refines its predictive capabilities over successive epochs using the training dataset. By juxtaposing this with the model's performance on previously unseen validation data, the illustration provides a holistic perspective on the model's capacity to not only internalize data-driven patterns but also to generalize them effectively to new datasets. This aspect of generalization is pivotal in the real-world financial sphere, where the prediction model must navigate and adapt to constantly fluctuating market dynamics.

As one scrutinizes the trends depicted in Fig. 3, a clear attenuation in both training and validation losses becomes evident. This decline is not just indicative of the model's ability to reduce error margins over time, but it also underscores its robustness and adaptability in the face of evolving financial data. It's noteworthy to mention that the sheer subtleness in the trajectory of these losses speaks volumes about the model's inherent stability and the efficiency of its learning algorithm.

Another salient observation from the figure is the relatively slight divergence between the training and validation losses. In the complex realm of deep learning, such congruence is often heralded as a testament to a model's consistent performance. A substantial gap could have implied potential overfitting, where the model becomes overly tailored to the training data and falters on new data. However, the narrow divergence observed reaffirms the ConvLSTM model's balanced approach, ensuring it retains its predictive accuracy even when confronted with unfamiliar financial datasets.

In summation, Fig. 3 stands as a pivotal piece of empirical evidence in this research. It not only attests to the ConvLSTM model's superior predictive capabilities but also illuminates its potential as a reliable tool for financial forecasting in a world characterized by uncertainty and rapid market fluctuations.

In the complex world of predictive modeling, understanding a model's accuracy and its capacity to generalize its learning is paramount. Fig. 4 serves as a compelling visualization of the Convolutional Long Short-term Memory (ConvLSTM) model's evolution in this regard, specifically focusing on the trajectory of its training and validation Root Mean Squared Error (RMSE) in the context of option price prediction [33].

Fig. 3. Train and validation loss.

The figure meticulously chronicles the RMSE values over a delineated range of epochs. This choice of visualization enables a nuanced interpretation, offering readers an opportunity to juxtapose the model's learning curve (as evidenced by training RMSE) against its ability to adapt and predict new, unseen data (as evidenced by validation RMSE). Such a comparative study is vital in ascertaining the robustness and reliability of a predictive model, especially in dynamic domains like financial forecasting where precision is non-negotiable.

A foundational expectation in the realm of predictive analytics is that, as a model undergoes more training iterations (or epochs), its understanding of the data deepens, leading to enhanced prediction accuracy. This theoretical construct is manifestly demonstrated in Fig. 4. Both the training and validation RMSE values demonstrate a marked descent, which, in analytical parlance, signifies the ConvLSTM model's growing adeptness at decoding underlying data patterns and its prowess in translating this understanding into accurate forecasts.

The ConvLSTM architecture, hailed for its sophistication, has proven its mettle through the figure's depiction. Its RMSE values, even in initial epochs, are commendably modest, attesting to the architecture's inherent strength [34]. What further bolsters the ConvLSTM's credibility is the observed trend: as the model is exposed to more epochs, there's a palpable contraction in the RMSE. This diminishing RMSE trend, in conjunction with the relative proximity between training and validation values, speaks volumes about the model's consistency and its adeptness at preventing overfitting.

In conclusion, Fig. 4 stands as a testament to the ConvLSTM model's exceptional capabilities in option price prediction. Through its clear delineation of RMSE progression, the figure elucidates the model's journey from initial understanding to mature precision, highlighting its potential as a dependable tool in the demanding sphere of financial analytics. Such insights underscore the significance of meticulous model evaluation and the promise that advanced architectures like ConvLSTM hold for future research endeavors.

Within the realm of financial forecasting, the veracity of a predictive model is often gauged by its ability to approximate real-world values. Fig. 5 serves as a visual testament to this critical evaluation, offering an intricate comparison between the ConvLSTM model's predicted option prices (rendered in blue) and the observed, actual option prices (depicted in red).

The visual portrayal in Fig. 5 invite a meticulous examination of the parallel trajectories of the forecasted and real option prices. The frequent confluence points between the blue and red lines provide compelling evidence of the model's predictive accuracy. Such intersections are not mere graphical intersections; they symbolize moments of congruence between predictive estimations and real-world observations. When a model's predictions consistently intersect with or closely trail the actual values, it is indicative of its robustness and precision.

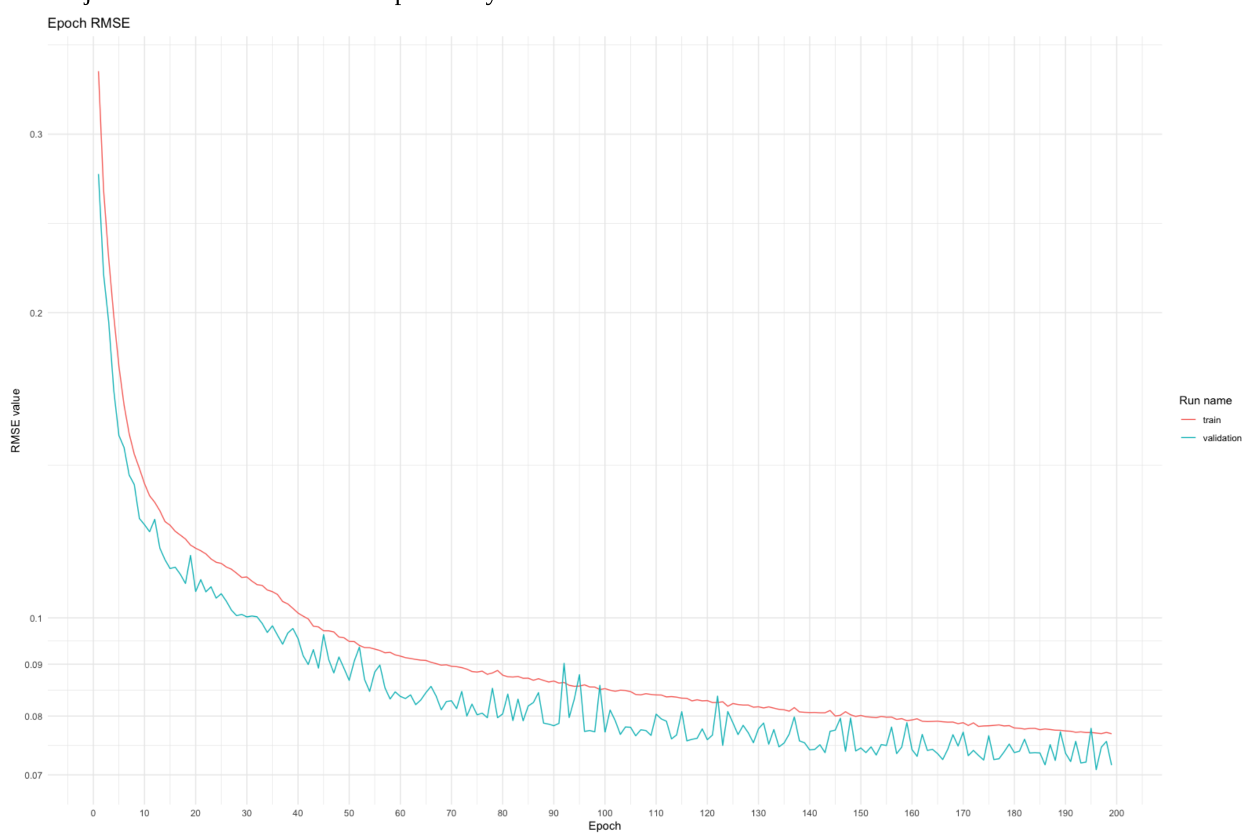

Fig. 4. Root mean squared error changes.

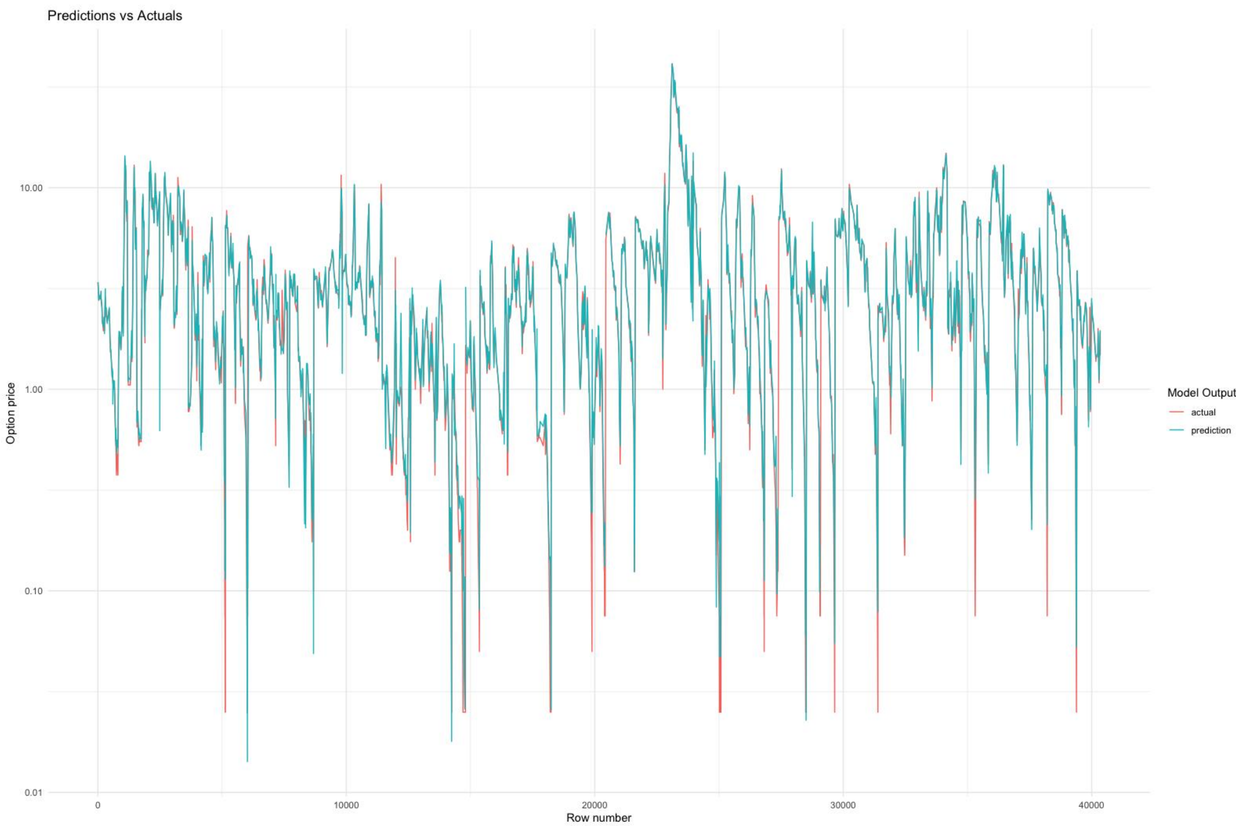

Fig. 5. Predicted and actual results for option pricing.

This observed alignment is not a trivial accomplishment, especially in the volatile domain of option pricing. The inherent unpredictability of financial markets makes the task of forecasting option prices a daunting challenge. Hence, the observed consistency between the model's estimations and the actual values underscores the ConvLSTM's superior modeling capabilities and its adeptness in capturing intricate market dynamics.

Moreover, the frequency of these intersections lends further credence to the ConvLSTM model's reliability. Infrequent matches could be dismissed as fortuitous, but the recurrent overlap observed in Fig. 5, suggests a pattern of sustained accuracy. Such consistency in predictions is emblematic of a model that has successfully internalized complex data patterns and can replicate this understanding in diverse scenarios.

Drawing from the insights garnered from Fig. 5, it becomes palpable that the ConvLSTM architecture's advanced design and computational prowess make it a formidable tool in the sphere of option price prediction. The close alignment between its forecasts and real prices is not merely a favorable outcome; it signifies the model's potential as a reliable instrument for long-term, practical applications in financial forecasting.

In conclusion, Fig. 5 stands as a testament to the ConvLSTM model's empirical efficacy. By providing a visual representation of the model's predictive prowess in the face of real-world data, the figure reinforces the notion that advanced predictive architectures like ConvLSTM are poised to revolutionize the landscape of financial analytics.

V. DISCUSSION

The research undertaken here delves deep into the intricate facets of option price prediction using the Convolutional Long Short-term Memory (ConvLSTM) model, a hybrid deep learning architecture. The journey of analyzing, training, and testing the model has opened up avenues for reflections and discussions on both the capabilities of the model and the complexities associated with option pricing in financial markets.

Foremost, the accuracy metrics employed (e.g., RMSE, MAPE) offered crucial insights into the model's predictive performance. Notably, the ConvLSTM model showcased a capacity to capture both spatial and temporal aspects of the financial data. This is significant, as financial data streams inherently possess these dual characteristics. Traditional models often struggle with time-series data that also has spatial features, which makes ConvLSTM's accomplishment notable. The ability of the model to achieve around 92% precision in option pricing over 200 learning epochs stands as a testament to its robust nature. Further, its adeptness at reaching 90% precision in just 20 epochs underscores its efficiency.

The train and validation graphs over epochs – whether in terms of precision, loss, or RMSE – illuminated the model's learning curve. A significant observation was that the disparity between training and validation scores was minimal, hinting at minimal overfitting [35]. This implies that the model generalizes well to unseen data, a crucial factor in the volatile world of financial markets where prediction robustness can translate to substantial economic implications.

The graphical representation comparing the predicted option prices with the actual values further reaffirmed the model's strength. The convergence of the blue and red lines not only indicates effective learning but also suggests the model's adaptability to evolving market dynamics.

Nevertheless, as with all models, it is essential to view these results in the broader context of financial modeling. The realm of option pricing is replete with volatility, influenced by an array of external factors such as geopolitical events, monetary policies, and global market trends [36]. While the ConvLSTM model has showcased commendable accuracy in this research, it is worth pondering how it might perform in extremely volatile scenarios or in the face of black swan events [37].

Another point of reflection is the comparison of the ConvLSTM model with traditional option pricing models like the Black-Scholes model [38]. While deep learning models like ConvLSTM offer adaptability and can work without strict assumptions, traditional models come with theoretical foundations deeply entrenched in financial theories. Would it be advantageous to integrate features from both worlds to achieve a more balanced, adaptable, yet theoretically sound model?

Data preprocessing, especially the normalization of parameters, was an essential step in this research. Neural networks, like the one employed in our ConvLSTM model [39], often require data to be structured and normalized to prevent issues like vanishing or exploding gradients [40]. It paves the path for a potential area of exploration: the development of models that are more resilient to raw or semi-processed data. Given the real-time nature of financial decisions, reducing preprocessing time without compromising accuracy could be invaluable.

Furthermore, the choice of data – specifically from the Chinese stock and fund market – brings forth questions about the model's adaptability to other global markets [41]. Financial behaviors and influences can vary across regions, affected by cultural, economic, and political differences. Hence, would the model retain its efficacy if trained on data from, say, the American or European markets?

In conclusion, the ConvLSTM model's potential in option price prediction, as evidenced by this research, is undeniably impressive. Its hybrid nature capitalizes on the strengths of both CNNs and LSTMs, making it a formidable tool for financial forecasting. However, it is imperative to continually evaluate and adapt the model, considering the ever-evolving landscape of global financial markets. The journey of this research serves not just as a testament to what has been achieved but also as an inspiration for the myriad possibilities that lie ahead.

VI. Conclusion

The exploration into the predictive prowess of the Convolutional Long Short-term Memory (ConvLSTM) model for option price forecasting has yielded insightful revelations. This research has undeniably demonstrated the potential of leveraging advanced deep learning architectures in the intricate and volatile realm of financial markets. As showcased, the ConvLSTM model adeptly captures the spatial and temporal nuances of financial data, achieving commendable accuracy levels over a minimal number of epochs.

A notable takeaway from this study is the synergy between the spatial feature detection capabilities of Convolutional Neural Networks (CNNs) and the sequential pattern recognition of Long Short-term Memory (LSTM) networks. The integration of these characteristics into the ConvLSTM model has proven to be particularly potent for forecasting in the dynamic domain of option pricing. The close alignment between the model's predictions and actual option prices underscores its viability as a tool for practical financial applications.

However, as with all predictive models, the ConvLSTM's performance should be perceived within the broader framework of financial analytics. While its adaptability and precision are commendable, continuous iterations and refinements are imperative, considering the multifaceted influences on financial markets. The model's ability to adapt to diverse global markets and unforeseen volatile events remains an avenue for future exploration.

In essence, this research has illuminated the vast potential that modern deep learning models hold for financial forecasting. It has set a benchmark, proving that with the right architecture and data, neural networks can be invaluable assets in the financial sector. As we advance, it is this confluence of finance and technology, exemplified by models like ConvLSTM that promises to redefine the contours of financial analytics and decision-making.

References

[1] Wu, J., Wang, Z., Hu, Y., Tao, S., & Dong, J. (2023). Runoff forecasting using Convolutional Neural Networks and optimized bi-directional long short-term memory. Water Resources Management, 37(2), 937-953.

[2] Wassie Geremew, G., & Ding, J. (2023). Elephant Flows Detection Using Deep Neural Network, Convolutional Neural Network, Long Short-Term Memory, and Autoencoder. Journal of Computer Networks and Communications, 2023.

[3] Liu, C., Zhu, H., Tang, D., Nie, Q., Zhou, T., Wang, L., & Song, Y. (2022). Probing an intelligent predictive maintenance approach with deep learning and augmented reality for machine tools in IoT-enabled manufacturing. Robotics and Computer-Integrated Manufacturing, 77, 102357.

[4] Tarek, Z., Shams, M. Y., Elshewey, A. M., El-kenawy, E. S. M., Ibrahim, A., Abdelhamid, A. A., & El-dosuky, M. A. (2023). Wind Power Prediction Based on Machine Learning and Deep Learning Models. Computers, Materials & Continua, 75(1).

[5] Omarov, B., Suliman, A., & Tsoy, A. (2016). Parallel backpropagation neural network training for face recognition. Far East Journal of Electronics and Communications, 16(4), 801-808.

[6] Omarov, B., Altayeva, A., Suleimenov, Z., Im Cho, Y., & Omarov, B. (2017, April). Design of fuzzy logic based controller for energy efficient operation in smart buildings. In 2017 First IEEE International Conference on Robotic Computing (IRC) (pp. 346-351). IEEE.

[7] Ajith, M., & Martínez-Ramón, M. (2023). Deep learning algorithms for very short term solar irradiance forecasting: A survey. Renewable and Sustainable Energy Reviews, 182, 113362.

[8] Omarov, B., Suliman, A., & Kushibar, K. (2016). Face recognition using artificial neural networks in parallel architecture. Journal of Theoretical and Applied Information Technology, 91(2), 238.

[9] Sun, Y., & Liao, K. (2023). A hybrid model for metro passengers flow prediction. Systems Science & Control Engineering, 11(1), 2191632.

[10] Yin, L., & Wei, X. (2023). Integrated adversarial long short-term memory deep networks for reheater tube temperature forecasting of ultra-supercritical turbo-generators. Applied Soft Computing, 142, 110347.

[11] Millham, R., Agbehadji, I. E., & Yang, H. (2021). Parameter tuning onto recurrent neural network and long short-term memory (RNN-LSTM) network for feature selection in classification of high-dimensional bioinformatics datasets. Bio-inspired Algorithms for Data Streaming and Visualization, Big Data Management, and Fog Computing, 21-42.

[12] Vrskova, R., Sykora, P., Kamencay, P., Hudec, R., & Radil, R. (2021, July). Hyperparameter tuning of ConvLSTM network models. In 2021 44th International Conference on Telecommunications and Signal Processing (TSP) (pp. 15-18). IEEE.

[13] Mughees, N., Jaffery, M. H., Mughees, A., Mughees, A., & Ejsmont, K. (2023). Bi-LSTM-Based Deep Stacked Sequence-to-Sequence Autoencoder for Forecasting Solar Irradiation and Wind Speed. Computers, Materials & Continua, 75(3).

[14] Durrani, A. U. R., Minallah, N., Aziz, N., Frnda, J., Khan, W., & Nedoma, J. (2023). Effect of hyper-parameters on the performance of ConvLSTM based deep neural network in crop classification. Plos one, 18(2), e0275653.

[15] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. nature, 521(7553), 436-444.

[16] Yu, Y., Hu, G., Liu, C., Xiong, J., & Wu, Z. (2023). Prediction of solar irradiance one hour ahead based on quantum long short-term memory network. IEEE Transactions on Quantum Engineering.

[17] Al-Alami, H., & Jamleh, H. O. (2023, May). Use of Convolutional Neural Networks and Long Short-Term Memory for Accurate Residential Energy Prediction. In 2023 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT) (pp. 294-299). IEEE.

[18] Gosztolya, G., Busa-Fekete, R., Grósz, T., & Tóth, L. (2017). DNN-based feature extraction and classifier combination for child-directed speech, cold and snoring identification.

[19] Viola, J., Chen, Y., & Wang, J. (2021). FaultFace: Deep convolutional generative adversarial network (DCGAN) based ball-bearing failure detection method. Information Sciences, 542, 195-211.

[20] Dhiman, H. S., Deb, D., & Guerrero, J. M. (2022). On wavelet transform based Convolutional Neural Network and twin support vector regression for wind power ramp event prediction. Sustainable Computing: Informatics and Systems, 36, 100795.

[21] Kumar, M., Singh, S., Kim, S., Anand, A., Pandey, S., Hasnain, S. M., ... & Deifalla, A. F. (2023). A Hybrid Model based on Convolution Neural Network and Long Short-Term Memory for Qualitative Assessment of Permeable and Porous Concrete. Case Studies in Construction Materials, e02254.

[22] Zhou, C., Xu, H., Ding, L., Wei, L., & Zhou, Y. (2019). Dynamic prediction for attitude and position in shield tunneling: A deep learning method. Automation in Construction, 105, 102840.

[23] Yarahmadi, A. M., Breuß, M., & Hartmann, C. (2022, September). Long Short-Term Memory Neural Network for Temperature Prediction in Laser Powder Bed Additive Manufacturing. In Proceedings of SAI Intelligent Systems Conference (pp. 119-132). Cham: Springer International Publishing.

[24] Dalal, A. A., AlRassas, A. M., Al-qaness, M. A., Cai, Z., Aseeri, A. O., Abd Elaziz, M., & Ewees, A. A. (2023). TLIA: Time-series forecasting model using long short-term memory integrated with artificial neural networks for volatile energy markets. Applied Energy, 343, 121230.

[25] Garg, S., & Krishnamurthi, R. (2023). A survey of long short term memory and its associated models in sustainable wind energy predictive analytics. Artificial Intelligence Review, 1-50.

[26] Alamri, N. M. H., Packianather, M., & Bigot, S. (2023). Optimizing the parameters of long short-term memory networks using the bees algorithm. Applied Sciences, 13(4), 2536.

[27] Fati, S. M., Muneer, A., Alwadain, A., & Balogun, A. O. (2023). Cyberbullying Detection on Twitter Using Deep Learning-Based Attention Mechanisms and Continuous Bag of Words Feature Extraction. Mathematics, 11(16), 3567.

[28] Mubashar, M., Khan, G. M., & Khan, R. (2021, April). Landslide prediction using long short term memory (LSTM) neural network on time series data in Pakistan. In 2021 International conference on artificial intelligence (ICAI) (pp. 175-181). IEEE.

[29] Becerra-Rico, J., Aceves-Fernández, M. A., Esquivel-Escalante, K., & Pedraza-Ortega, J. C. (2020). Airborne particle pollution predictive model using Gated Recurrent Unit (GRU) deep neural networks. Earth Science Informatics, 13, 821-834.

[30] Yessoufou, F., & Zhu, J. (2023). Classification and regression-based Convolutional Neural Network and long short-term memory configuration for bridge damage identification using long-term monitoring vibration data. Structural Health Monitoring, 14759217231161811.

[31] Panda, B., & Singh, P. (2023). A deep convolutional-LSTM neural network for signal detection of downlink NOMA system. AEU-International Journal of Electronics and Communications, 154797.

[32] Chaudhury, S., & Sau, K. (2023). A BERT encoding with Recurrent Neural Network and Long-Short Term Memory for breast cancer image classification. Decision Analytics Journal, 6, 100177.

[33] Liu, H., Yan, G., Duan, Z., & Chen, C. (2021). Intelligent modeling strategies for forecasting air quality time series: A review. Applied Soft Computing, 102, 106957.

[34] Girsang, A. S., & Tanjung, D. (2022). Fast Genetic Algorithm for Long Short-Term Memory Optimization. Engineering Letters, 30(2).

[35] Conti, P., Guo, M., Manzoni, A., & Hesthaven, J. S. (2023). Multi-fidelity surrogate modeling using long short-term memory networks. Computer methods in applied mechanics and engineering, 404, 115811.

[36] Zhang, C., Pan, G., & Fu, L. (2023). Real-Time Intersection Turning Movement Flow Forecasting Using a Parallel Bidirectional Long Short-Term Memory Neural Network Model. Transportation Research Record, 03611981231172958.

[37] Lu, W., Duan, J., Wang, P., Ma, W., & Fang, S. (2023). Short-term wind power forecasting using the hybrid model of improved variational mode decomposition and maximum mixture correntropy long short-term memory neural network. International Journal of Electrical Power & Energy Systems, 144, 108552.

[38] Stankovic, M., Bacanin, N., Budimirovic, N., Zivkovic, M., Sarac, M., & Strumberger, I. (2023, April). Bi-Directional Long Short-Term Memory Optimization by Improved Teaching-Learning Based Algorithm for Univariate Gold Price Forecasting. In 2023 International Conference on Inventive Computation Technologies (ICICT) (pp. 1650-1657). IEEE.

[39] Munsif, M., Ullah, M., Fath, U., Khan, S. U., Khan, N., & Baik, S. W. (2023). CT-NET: A Novel Convolutional Transformer-Based Network for Short-Term Solar Energy Forecasting Using Climatic Information. Computer Systems Science & Engineering, 47(2).

[40] Dong, S., Wang, P., & Abbas, K. (2021). A survey on deep learning and its applications. Computer Science Review, 40, 100379.

[41] Rama-Maneiro, E., Vidal, J., & Lama, M. (2021). Deep learning for predictive business process monitoring: Review and benchmark. IEEE Transactions on Services Computing.