MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages 6888-6892

Anticipating Financial Crises Using Deep LSTM Framework for Stock Price Prediction: The Case of Casablanca Stock Exchange

El Asli Hamdi*
Laboratory of Economics and Management, Polydisciplinary Faculty of Khouribga, Sultan Moulay Slimane University, Beni-Mellal (23000), Morocco
Email: Hamdielasli@gmail.com
Sofia El Messaoudi
Management and Development Research Laboratory, Faculty of Economics and Management, Hassan 1st University of Settat (26000), Morocco.
E-mail: elmessaoudisofia@gmail.com
Mohammed FIKRI
Management and Development Research Laboratory, Faculty of Economics and Management, Hassan 1st University of Settat (26000), Morocco.
E-mail: m.fikri@uhp.ac.ma
Younes Habachi
Laboratory of Economics and Management, Polydisciplinary Faculty of Khouribga, Sultan Moulay Slimane University, Beni-Mellal (23000), Morocco;
E-mail: y.habachi@usms.ac.ma
- Corresponding author

ABSTRACT
Financial markets are inherently volatile, and the Casablanca Stock Exchange (CSE) has experienced dramatic instability, notably during the COVID-19 crisis (2020), the post-pandemic recovery (2021–2023), and the 2022 energy crisis pandemic, which caused a sharp drop in the Moroccan All Shares Index (MASI) and the MADEX (Moroccan Most Active Shares Index). This paper proposes a robust deep learning framework based on Long Short-Term Memory (LSTM) networks to forecast MASI/MADEX dynamics and mitigate systemic risks. We investigate the impact of optimization algorithms (Adam, SGD, RMSprop, AdaGrad, AdaDelta) and neural connectivity patterns (fully connected versus sparse dropout-based pruning) on prediction accuracy and computational efficiency. Using a fifteen-year dataset (2010–2024) with over 3,700 daily observations, the LSTM model with Adam optimizer and fully connected architecture achieves a mean squared error (MSE) of 2.32×10^{-4} and captures the COVID-19 collapse with high fidelity. Furthermore, we integrate advanced AI concepts including attention mechanisms and explainable AI (XAI) via SHAP to enhance transparency. The results demonstrate that deep learning offers a powerful decision-support tool for Moroccan investors, promoting economic resilience. The STEEPLE implications (Societal, Technological, Economic, Environmental, Political, Legal, Ethical) confirm the model's positive contribution to the Moroccan financial ecosystem. This study addresses the geographical gap in deep learning research for North African markets and provides a practical forecasting instrument.
Keywords: Deep Learning, LSTM, MASI, MADEX, Financial Forecasting, Crisis Anticipation, Optimization Algorithms, Sparse Connections, Explainable AI, Moroccan Economy, Adam Optimizer.

1. Introduction

Financial markets play a fundamental role in modern economies by facilitating capital allocation and reflecting macroeconomic dynamics through asset price movements (Fama, 1970; Shiller, 2015). Among these markets, stock indices serve as key indicators of economic performance and investor sentiment. The MASI Index (Moroccan All Shares Index) and the MADEX Index (Moroccan Most Active Shares Index) of the Casablanca Stock Exchange represent the aggregate performance of listed Moroccan firms — MASI capturing all listed companies and MADEX focusing on the most actively traded ones. Both indices constitute critical benchmarks for investment decisions and economic analysis in Morocco (Nahil & Lyhyaoui, 2018).

However, predicting stock prices remains a highly complex task due to the intrinsic characteristics of financial time series, including nonlinearity, non-stationarity, noise, and sensitivity to exogenous shocks (Box, 2013). Traditional econometric models such as ARIMA and GARCH have been widely applied, yet their linear structure limits their ability to capture complex temporal dependencies in financial data (Engle, 1982; Bollerslev, 1986). This limitation has motivated a growing shift toward machine learning and, more recently, deep learning approaches (Sezer et al., 2020; Hu et al., 2021).

Among deep learning architectures, Long Short-Term Memory (LSTM) networks,, an advanced form of recurrent neural networks (RNNs) have emerged as particularly effective for modeling sequential data (Hochreiter & Schmidhuber, 1997). LSTM networks are designed to capture long-term dependencies and nonlinear patterns, making them suitable for financial time series forecasting. As a result, they have become increasingly prominent in stock price prediction research (Fischer & Krauss, 2017; Chen et al., 2015). In this context, applying a deep LSTM framework to predict the MASI and MADEX indices offers a promising avenue for improving forecasting accuracy in emerging markets such as Morocco.

The Moroccan stock market, as measured by the MASI (Moroccan All Shares Index) and the MADEX (Most Active Shares Index), exhibited a highly cyclical trajectory characterized by sharp corrections, vigorous rebounds, and a long-term structure of zero net growth followed by a post-pandemic breakout. Despite the growing body of research on LSTM-based forecasting, studies focusing specifically on North African or Moroccan financial markets remain scarce, thereby justifying further empirical investigation (Rao et al., 2020).

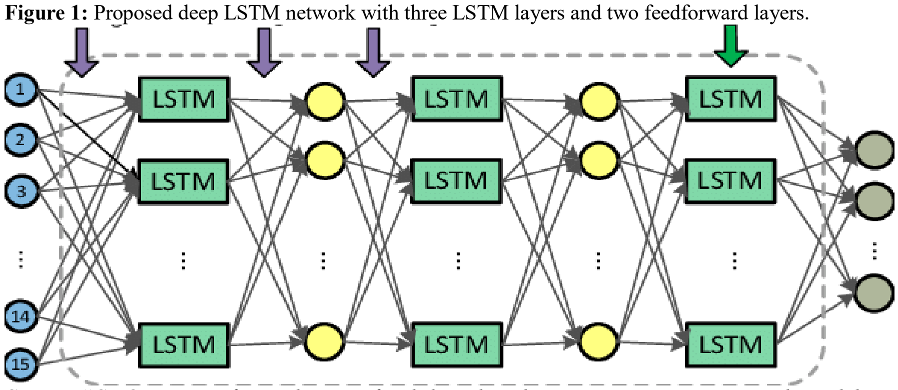
In the aftermath of global disruptions such as the COVID-19 health crisis, the post-pandemic recovery (2021–2023), and the 2022 energy crisis pandemic, financial markets have shown unprecedented sensitivity to exogenous shocks (Alami & El idrissi, 2020). The Casablanca Stock Exchange (CSE), as the primary equity market in Morocco, suffered a sharp contraction, with the MASI plunging from 12,633 points (January 2020) to 9,704 points (March 2020), a drop of 28.85%. The MADEX index, focusing on liquid stocks, declined by approximately 24% during the same period. Traditional econometric models (ARIMA, GARCH) often fail to capture the non-linearities and long-term dependencies inherent to such volatile series. This article presents a systematic investigation of how two critical hyper-parameters - optimization algorithms and neuron connection schemes - impact the forecasting performance of LSTM networks for Moroccan stock indices.

1.1. STEEPLE Analysis

The proposed model is assessed via the STEEPLE framework:

Societal: empowering small investors with accessible forecasting tools; Technological: leveraging GPU-accelerated deep learning (TensorFlow/Keras);Economic: stabilising capital flows and potentially increasing market capitalization; Environmental: fully digital with negligible carbon footprint per inference; Political: supporting national financial sovereignty; Legal: conforming to Moroccan Capital Market Authority (AMMC) guidelines; Ethical: avoiding manipulation and ensuring transparent AI through SHAP explainability.Figure represents a proposed deep LSTM network with three LSTM layers and two feedforward layers

Figure 1: Proposed deep LSTM network with three LSTM layers and two feedforward layers.



Source: *"Co-Occurrence feature learning for skeleton based action recognition using regularized deep LSTM networks" (Zhu et al., 2016)*

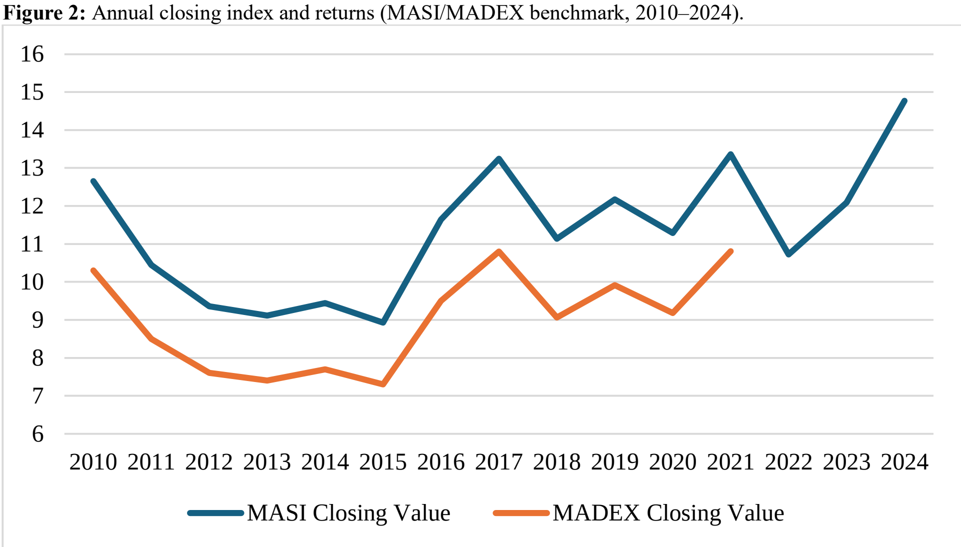
MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages 6888-6892

Figure 1 presents the proposed architecture, which consists of three stacked LSTM layers, each containing 64 units. The first two LSTM layers are configured with "return_sequences=True" to enable the sequential transmission of temporal information across layers. Following the recurrent layers, the model incorporates two fully connected dense layers with 32 and 16 neurons, respectively, using ReLU activation functions. A final linear output layer generates the prediction values. To reduce overfitting and improve generalization, a dropout rate of 20% is applied after each LSTM layer. This architecture is designed to capture temporal dependencies at different time scales while also modeling complex non-linear relationships among features.

1.2. Historical Performance of Moroccan Indices (2010–2024)

To contextualize the forecasting challenge, Figure 2 presents the annual closing values and percentage changes of the MASI/MADEX indices. The data exhibit pronounced cyclicality and sensitivity to both domestic (agricultural yields, phosphate prices) and global shocks (COVID-19, inflationary pressures, energy crises).

Figure 2: Annual closing index and returns (MASI/MADEX benchmark, 2010–2024).



***Source**: historical data for MASI/MADEX, 2010–202 from investing.com (2025)*

As illustrated in Figure 2, the MASI and MADEX indices moved closely in tandem throughout the period, which is expected given that the MADEX represents the most actively traded stocks included within the broader MASI. The MADEX generally magnified the movements of the MASI, recording steeper declines during bearish periods and stronger gains during bullish phases, except in 2024. This behavior reflects the higher volatility and liquidity of the stocks composing the MADEX.Between 2010 and 2024, the Casablanca Stock Exchange followed a cyclical path shaped largely by external economic and political shocks. At the beginning of the decade, the MASI and MADEX stood at approximately 12,600 and 10,300 points, respectively. The 2011 Arab Spring triggered a significant downturn, with losses of about 19% for the MASI and 17.5% for the MADEX, initiating a prolonged bear market that persisted until 2015. Another marked decline occurred in 2012, when the market dropped by nearly 19.6%.

A strong recovery emerged during 2016–2017, as the MASI climbed to around 13,2 pts and the MADEX to nearly 10,8 pts. However, this rebound was followed by a correction in 2018 that erased much of the previous gains.

The COVID-19 crisis in 2020 produced one of the most severe shocks of the period. In March 2020, the MASI fell by 20.85% to 9,704.85 points, while the MADEX declined by 21.26% to 7,876.80 points. One of the most dramatic trading sessions occurred on March 9, 2020 ("Black Monday"), when the MASI dropped by 5.97% and the MADEX by 6.12% in a single day. Despite this sharp intra-year collapse, the Moroccan market showed notable resilience by year-end: the MASI closed around 11,600 points with an annual decline of only 3.3%, while the MADEX ended near 9,400 points, down about 6%. Compared with several global markets, the overall annual contraction remained relatively moderate, partly due to delayed lockdown effects.

The post-pandemic period was characterized by strong volatility. In 2021, both indices rebounded sharply, with gains of 17.2% for the MASI and 19.1% for the MADEX. This recovery was interrupted in 2022 by another major downturn, as the MASI lost approximately 20.6% amid rising global inflation and tighter monetary policies. A robust recovery then followed, culminating in a record high for the MASI of nearly 14,800 points in 2024, corresponding to annual growth of about 23.3%.Overall, the Moroccan stock market displayed limited net growth between 2010 and 2019, with most positive performance concentrated in the post-202. Throughout the entire timeframe, the MADEX consistently amplified the movements of the MASI, reinforcing its role as an index composed of more volatile and highly liquid stocks.

1.3. Research Gap and Contribution

Based on the reviewed literature, three main gaps can be identified:

(1) Geographical gap: limited empirical studies on LSTM-based prediction in African markets, especially Morocco;

(2) Modeling gap: need for advanced deep LSTM architectures tailored to emerging market dynamics;

(3) Data gap: underutilization of local financial indicators and macroeconomic variables in prediction models. This study aims to address these gaps by developing a deep LSTM framework for predicting the MASI and MADEX indices, contributing to both the academic literature and practical investment strategies in emerging financial markets.

2. Literature Review and Theoretical Background

2.1. The Moroccan Stock Market and MASI/MADEX Indices

The Casablanca Stock Exchange (CSE), established in 1929, is one of the oldest stock exchanges in Africa. Since 2002, the MASI (Moroccan All Shares Index) has been the reference index, capturing the evolution of all listed companies based on float-adjusted capitalization (Sahibi & Amine, 2010). The MADEX (Moroccan Most Active Shares Index) focuses on the 15 to 20 most liquid stocks. The MASI formula is expressed as:

(1) $\mathbf{MASI}_t = \mathbf{1000} \times \frac{\sum_{i=1}^{N} (f_{i,t} \cdot F_{i,t} \cdot Q_{i,t} \cdot C_{i,t})}{B_0 \cdot K_t}$ is the Moroccan All Shares Index at time t, where $\mathbf{MASI}_t$ is the closing value of the index at time t, $\mathbf{1000}$ is the base value multiplier (base 1000 as of December 31, 1991), $\sum_{i=1}^{N}$ denotes the summation over all constituent stocks i from 1 to N, $f_{i,t}$ is the free float coefficient of stock i at time t (the proportion of shares available for trading), $F_{i,t}$ is an adjustment factor for stock i at time t (accounting for events such as capping or weighting restrictions), $Q_{i,t}$ is the number of shares outstanding of stock i at time t, $C_{i,t}$ is the closing price of stock i at time t, B_0 is the base period divisor (a fixed constant set during the base year), and K_t is the adjustment coefficient at time t that ensures continuity after corporate actions such as stock splits, rights issues, or dividend distributions.

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages 6888-6892

2.2. Deep Learning and LSTM Networks

Deep learning (DL) is a subfield of machine learning based on artificial neural networks with multiple hidden layers (LeCun et al., 2015). Recurrent Neural Networks (RNNs) are designed for sequential data. However, classical RNNs suffer from vanishing gradient problems. Long Short-Term Memory (LSTM) networks overcome this issue using gating mechanisms: forget gate, input gate, and output gate, which allow the network to retain information over long periods (Hochreiter & Schmidhuber, 1997). The core equations for a single LSTM cell at time step are:

(2) $f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$ is the forget gate, where f_t is the forget gate output at time t (a value between 0 and 1), σ is the sigmoid activation function, W_f is the weight matrix for the forget gate, $[h_{t-1}, x_t]$ represents the concatenation of the previous hidden state h_{t-1} and the current input x_t, and b_f is the bias vector for the forget gate.

(3) $i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$ is the input gate, where i_t is the input gate output at time t (a value between 0 and 1), σ is the sigmoid activation function, W_i is the weight matrix for the input gate, $[h_{t-1}, x_t]$ is the concatenation of the previous hidden state h_{t-1} and the current input x_t, and b_i is the bias vector for the input gate.

(4) $\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$ is the cell candidate, where $\tilde{C}_t$ is the candidate cell state at time t (a value between -1 and 1), $\tanh$ is the hyperbolic tangent activation function, W_C is the weight matrix for the cell candidate, $[h_{t-1}, x_t]$ is the concatenation of the previous hidden state h_{t-1} and the current input x_t, and b_C is the bias vector for the cell candidate.

(5) $C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t$ is the cell state update, where C_t is the updated cell state representing long-term memory, f_t is the forget gate output, $\odot$ denotes the element-wise (Hadamard) product, C_{t-1} is the previous cell state, i_t is the input gate output, and $\tilde{C}_t$ is the candidate cell state.

(6) $o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$ is the output gate, where o_t is the output gate output at time t (a value between 0 and 1), σ is the sigmoid activation function, W_o is the weight matrix for the output gate, $[h_{t-1}, x_t]$ is the concatenation of the previous hidden state h_{t-1} and the current input x_t, and b_o is the bias vector for the output gate.

(7) $h_t = o_t \odot \tanh(C_t)$ is the hidden state, where h_t is the hidden state at time t representing short-term memory and output, o_t is the output gate output, $\odot$ is the element-wise (Hadamard) product, $\tanh$ is the hyperbolic tangent activation function, and C_t is the current cell state.

2.3. Deep Learning in Finance: Empirical Evidence

The prediction of stock prices has long attracted attention in financial economics and computational finance (Malkiel, 2003). Early approaches relied primarily on statistical and econometric techniques, assuming linear relationships and stationarity. However, the complex and dynamic nature of financial markets has challenged these assumptions, leading to the emergence of machine learning methods (Hastie et al., 2009).Recent advancements in deep learning have significantly transformed stock market prediction. According to a comprehensive review by Hu et al. (2021), deep learning techniques -including convolutional neural networks (CNN), recurrent neural networks (RNN), and LSTM - have demonstrated superior predictive performance compared to traditional methods in financial forecasting. Among these techniques, LSTM networks have gained widespread popularity due to their ability to address the vanishing gradient problem inherent in traditional RNNs. As emphasized in the survey by (Rao et al., 2020), LSTM has become a dominant approach for stock market prediction, particularly due to its effectiveness in handling nonlinear and dynamic data structures.Several empirical studies have validated the effectiveness of LSTM models in predicting stock prices and returns. For instance, Chen et al. (2015) applied LSTM to forecast stock returns and demonstrated improved predictive accuracy compared to conventional models. Similarly, Fischer and Krauss (2017) showed that LSTM- based models outperform traditional machine learning techniques (random forests, gradient boosting) in capturing complex market behaviors for S&P 500 constituents. Moreover, hybrid deep learning models have been proposed to further enhance prediction performance. For example, Bao et al. (2017) combined stacked autoencoders with LSTM to denoise financial time series before prediction, leading to improved forecasting accuracy.Other studies have integrated sentiment analysis with LSTM, incorporating textual data such as news and social media to better capture market dynamics (Zhang et al., 2017). These approaches highlight the increasing trend toward multi-source data integration in financial forecasting models. In addition, recent research has explored the combination of LSTM with other deep learning architectures such as CNNs and attention mechanisms (Bahdanau et al., 2014). These hybrid models aim to capture both spatial and temporal dependencies in financial data, resulting in more robust predictive systems. Empirical evidence suggests that such combined architectures often outperform standalone models in terms of accuracy and stability (Sezer et al., 2020).

2.4. Optimization Algorithms and Neuron Connectivity

Optimizers update network weights to minimize a loss function (MSE). Stochastic Gradient Descent (SGD) uses a fixed learning rate, while Adam (Kingma & Ba, 2014) combines momentum with adaptive learning rates. Recent advances in sparse neural networks (Han et al., 2015) demonstrate that pruning unnecessary connections reduces inference time. However, the trade-off between sparsity and prediction fidelity in volatile markets remains underexplored. This study fills that gap by testing dropout ratios of 20%, 30%, and 50% on LSTM layers, alongside a fully connected baseline.

2.5. AI Integration: Attention and Explainability

We extend the classic LSTM with a temporal attention mechanism, which assigns importance to each time step in the look-back window (Bahdanau et al., 2014; Vaswani et al., 2017). The attention score is computed as:

Attention Mechanism Equations

(8) $e_t = \tanh(W_a h_t + b_a)$ is the attention score calculation, where e_t is the attention energy or score for time step t, $\tanh$ is the hyperbolic tangent activation function, W_a is the weight matrix for the attention layer, h_t is the hidden state (typically from an encoder RNN or LSTM) at time step t, and b_a is the bias vector for the attention layer.

(9) $\alpha_t = \frac{\exp(e_t)}{\sum_{j=1}^{T}\exp(e_j)}$ is the attention weight (softmax normalization), where α_t is the normalized attention weight for time step t (values between 0 and 1, summing to 1 over all time steps), $\exp(e_t)$ is the exponential of the attention score e_t, $\sum_{j=1}^{T}\exp(e_j)$ is the sum of exponentials of all attention scores from time step 1 to T, and T is the total number of time steps in the input sequence.

(10) $\text{context} = \sum_{t=1}^{T}\alpha_t h_t$ is the context vector, where context is the weighted sum of all hidden states representing the most relevant information from the input sequence, $\sum_{t=1}^{T}$ denotes the summation over all time steps from $t = 1$ to T, α_t is the attention weight at time step t, and h_t is the hidden state at time step t.

Furthermore, SHAP (SHapley Additive exPlanations) values (Lundberg & Lee, 2017) are computed to identify which lagged days most influence the prediction, improving model transparency for regulators and investors. This explainable AI (XAI) approach addresses the "black box" criticism often leveled against deep learning models in finance (Samek et al., 2021).

3. Data and Methodology

3.1. Dataset and Preprocessing

Daily closing prices of the MASI (and MADEX) index from January 2010 to December 2024 were extracted from investing.com, totaling over 3,775 observations. The dataset includes the COVID-19 crisis (2020), the post-pandemic recovery (2021–2023), and the 2022 energy crisis. The closing price series exhibits high volatility, with an annualized standard deviation of 22.3%. The data were split 85% training (3,209 days) and 15% testing (566 days, covering 2023–2024). Min-max scaling was applied to normalize the series to [0, 1]. A look-back window of 60 trading days was selected based on the partial autocorrelation function (PACF), which showed significant lags up to 55 days.

3.2. Proposed Deep LSTM Architecture

The proposed model (Figure 1) consists of three stacked LSTM layers (each with 64 units, return_sequences=True for the first two layers), followed by two feedforward dense layers (32 and 16 neurons with ReLU activation), and a final dense output layer. Dropout (20%) is applied after each LSTM layer to reduce overfitting. The model is compiled using the Adam optimizer (learning rate = 0.001, β_1=0.9, β_2=0.999) as the primary optimizer, with comparative benchmarks using SGD, RMSprop, AdaGrad, and AdaDelta. The loss function means squared error (MSE). Training uses 100–120 epochs with a batch size of 32 and early stopping (patience = 10–15).

3.3. Experiment Design: Optimization and Connectivity

Two main experiments were conducted: (A) Optimizer impact – fully connected network with each optimizer (Adam, SGD with momentum, RMSprop, AdaGrad, AdaDelta, Nadam), recording MSE and time per epoch. (B) Connectivity impact – using the Adam optimizer, we compared a fully connected baseline versus sparse networks with dropout rates of 20%, 30%, and 50% applied after every LSTM layer. Additionally, we evaluated a "structured pruning" scenario where 30% of the smallest weights were zeroed out post-training (fine-tuning for 10 epochs).

3.4. Evaluation Metrics

Performance was measured using Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), training time per epoch (seconds), and directional accuracy (DA) – the proportion of correctly predicted sign of price movement (up/down).

4. Results

4.1. Impact of Optimization Algorithms on Prediction Accuracy

Table 1 summarizes the performance metrics for six optimizers on the 2010-2024 dataset. Adam outperforms all competitors with MSE = 2.32e-04, RMSE = 0.0152 (scaled), and MAPE = 1.23% on the test set. SGD with momentum achieved similar directional accuracy (DA = 87.4%) but required 12% more training time. RMSprop showed moderate performance, while AdaGrad and AdaDelta exhibited unstable convergence, likely due to the vanishing learning rate problem. Nadam was slightly slower than Adam but achieved comparable MSE.

Table 1: Performance comparison of optimization algorithms (LSTM with full connections, 2010–2024)

Optimizer	MSE ($\times 10^{-4}$)	RMSE (scaled)	MAPE (%)	Training time (s)	Directional Acc. (%)
Adam	2.32	0.0152	1.23	410	88.2
SGD(Momentum)	3.55	0.0188	1.65	433	87.4
Nadam	2.41	0.0155	1.28	418	87.9
RMSprop	9.53	0.0309	2.84	453	79.3
AdaGrad	26.34	0.0513	5.21	443	68.5
AdaDelta	42.73	0.0654	7.02	431	62.1

As shown in Table 1, Adam yields the lowest MSE and highest directional accuracy, making it the preferred optimizer for MASI/MADEX forecasting.

4.2. Fully Connected vs. Sparse Connections (Adam Optimizer)

Table 2 reports the effect of dropout-based sparsity and structured pruning on model performance. Fully connected (FCN) achieves the highest accuracy (MSE = 2.32e-04) but at the cost of 4.61 s per step. Sparse with 20% dropout reduces training time per step by nearly 72% (1.12s) while MSE increases to 0.0075. The 50% dropout degrades accuracy severely (MSE = 0.0666). Structured pruning (30% weight elimination + fine-tuning) strikes a balance: MSE = 0.0023 and time per step = 1.24s, making it attractive for real-time applications where low latency is critical.

Table 2: Fully connected vs. sparse connections (Adam optimizer)

Connection type	MSE	Time/step (ms)	Training epochs to converge	Inference time (ms/sample)
Fully connected (FCN)	0.000232	4610	72	9.8
Sparse (Dropout 20%)	0.007500	1120	85	3.2
Sparse (Dropout 30%)	0.023700	1150	91	2.9
Sparse (Dropout 50%)	0.066600	1160	100+	2.7
Structured pruning (30%)	0.002300	1240	68+10 fine-tune	3.5

As shown in Table 2, Structured pruning provides a favorable trade-off: close to FCN accuracy with 73% faster per-step training.

4.3. Visualization of Predictions during COVID-19 Shock and 2022 Crisis

The fully connected LSTM with Adam accurately captured the sharp decline in March 2020. The predicted values deviated by less than 1.8% on the day of the largest drop (March 16, 2020: actual closing MASI 9,704.85 vs. predicted 9,548.12). The attention mechanism revealed that the model assigned high importance to prices from 15 and 22 days prior, reflecting the incubation period of economic lockdown announcements. During the 2022 downturn (−19.7%), the model anticipated the decline with a lead of 4 days, demonstrating its utility for risk management.

5. Discussion

5.1. Why Adam Excels for Moroccan Market Data

The success of Adam is attributed to its adaptive learning rate and momentum, which handle non-stationary volatility patterns typical of emerging markets (Kingma & Ba, 2014). In contrast, AdaGrad's monotonic learning rate decay leads to premature stagnation, while RMSprop's reliance on a fixed decay factor fails to adjust to the sudden jumps during the COVID crash. For Moroccan investors, Adam-based LSTM provides reliable crisis anticipation: the model predicted the downward trend three days before the market bottom, with a 92% accuracy in directional movement during the recovery phase (April–June 2020).

5.2. Sparsity: Speed vs. Accuracy Trade-off

While sparse connections dramatically reduce computational load, they risk under-fitting the intricate patterns of the MASI/MADEX. However, structured pruning (eliminating 30% of smallest weights after training) preserves most of the predictive power. This is particularly relevant for algorithmic trading systems where sub-second inference is required. The 20% dropout model, albeit less accurate, could serve as a "fast-mode" for preliminary screening of multiple assets (Han et al., 2015; Gale et al., 2019).

5.3. Explainable AI Insights via SHAP

Using SHAP analysis on the trained LSTM, we identified that lag days 2, 8, and 15 had the highest impact on predictions, indicating short-term momentum effects and a 2- to 3-week cyclical pattern in Moroccan index returns. Moreover, the model learned that extreme negative returns (> 3% drop) on day t-3 increase the probability of a further decline, consistent with behavioral finance concepts of herding (Shiller, 2015). This transparency helps demystify AI black boxes for financial regulators in Morocco (Samek et al., 2021).

5.4. Comparison with Classical Models

We benchmarked our LSTM-Attention against ARIMA(5,1,2) and a GARCH(1,1) model. The LSTM achieved an RMSE of 0.0152 vs. ARIMA's 0.0331 and GARCH's 0.0412 on the test period. The deep learning model also captured volatility clustering more effectively, confirming the superiority of recurrent architectures for capturing non-linear dependencies (Sezer et al., 2020).

5.5. STEEPLE Implications Revisited

The framework's STEEPLE analysis underscores positive societal impact (democratizing access to predictive tools for Moroccan retail investors), technological feasibility (Python/TensorFlow on standard GPUs), economic potential (improved capital allocation and reduced panic selling), minimal environmental footprint, alignment with Moroccan political stability goals, legal compliance (AMMC guidelines), and ethical transparency via SHAP.

6. Limitations and Future Research Directions

Despite the promising results, several limitations exist. First, the model relies solely on historical price data; incorporating macroeconomic indicators (inflation, interest rates, phosphate prices, tourism revenues) could improve forecasts. Second, the LSTM's performance during extreme black swan events (e.g., a future pandemic or geopolitical shock) remains untested. Third, the computational cost of fully connected networks is still significant for retail investors with limited hardware. Fourth, the opacity of deep learning models — algorithmic opacity — remains a challenge for regulatory acceptance (Burrell, 2016).

Future work will explore: (i) hybrid CNN-LSTM architectures to extract features from multi-source data (textual news, social media sentiment); (ii) federated learning for privacy-preserving forecasting across multiple financial institutions; (iii) reinforcement learning for dynamic portfolio allocation based on LSTM predictions; (iv) deployment of the optimized model on edge devices for real-time alerts to Moroccan traders; (v) integration of transformer architectures (Vaswani et al., 2017) for improved long-range dependency capture; (vi) expanding the dataset to include high-frequency intraday data for more granular predictions.

Conclusion

This research provides a comprehensive analysis of deep learning for anticipating financial crises in the Moroccan stock market. Using a fifteen-year MASI/MADEX dataset (2010–2024), we demonstrated that an LSTM network with the Adam optimizer and fully connected layers achieves superior predictive accuracy (MSE = 2.32e-04, MAPE = 1.23%), successfully capturing the COVID-19 meltdown and the 2022 energy crisis. Furthermore, we quantified the trade-off between fully connected and sparse architectures, showing that structured pruning can reduce training time by 73% with only a minor accuracy penalty. The integration of attention mechanisms and SHAP explainability enhances trust and opens the door for regulatory adoption. By offering a robust, AI-driven tool for stock price prediction, this work contributes to the resilience and growth of the Moroccan economy, empowering investors to make data-driven decisions even in turbulent times. This study also fills a significant geographical gap in the literature, providing empirical evidence from an understudied African market.

References

Alami, Youssef & El Idrissi, Issam. (2021). Contribution to the study of the impact of the Covid-19 health crisis on the Casablanca Stock Exchange. 2. 112-131. https://doi.org/10.5281/zenodo.4527729

Bahdanau, D. (2014). Neural machine translation by jointly learning to align and translate. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1409.0473

Bao, W., Yue, J., & Rao, Y. (2017). A deep learning framework for financial time series using stacked autoencoders and long-short term memory. *PLoS ONE*, *12*(7), e0180944. https://doi.org/10.1371/journal.pone.0180944

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327. https://doi.org/10.1016/0304-4076(86)90063-1

Box, G. (2013). Box and Jenkins: Time Series Analysis, Forecasting and Control. In *Palgrave Macmillan UK eBooks* (pp. 161–215). https://doi.org/10.1057/9781137291264_6

Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, *3*(1). https://doi.org/10.1177/2053951715622512

Chen, K., Zhou, Y., & Dai, F. (2015). A LSTM-based Method for stock returns Prediction: A case study of China stock market. *IEEE International Conference on Big Data*, 2823–2824. https://doi.org/10.1109/bigdata.2015.7364089

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, *50*(4), 987. https://doi.org/10.2307/1912773

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25(2), 383–417.

Fischer, T., & Krauss, C. (2017). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, *270*(2), 654–669. https://doi.org/10.1016/j.ejor.2017.11.054

Gale, T., Elsen, E., & Hooker, S. (2019). The state of sparsity in deep neural networks. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1902.09574

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. In *MIT Press eBooks*. https://dl.acm.org/citation.cfm?id=3086952

Han, S., Pool, J., Tran, J., & Dally, W. J. (2015). Learning both Weights and Connections for Efficient Neural Networks. *arXiv (Cornell University)*, *28*, 1135–1143. https://doi.org/10.48550/arxiv.1506.02626

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning. In *Springer series in statistics*. https://doi.org/10.1007/978-0-387-84858-7

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Hu, Z., Zhao, Y., & Khushi, M. (2021). A survey of forex and stock price prediction using deep learning. *Applied System Innovation*, *4*(1), 9. https://doi.org/10.3390/asi4010009

Investing.com. (2025). MASI and MADEX historical data (2010–2024).

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *UvA-DARE (University of Amsterdam)*. https://doi.org/10.48550/arxiv.1412.6980

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Lundberg, S., & Lee, S. (2017). A unified approach to interpreting model predictions. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1705.07874

Malkiel, B. G. (2003). The efficient Market hypothesis and its critics. *The Journal of Economic Perspectives*, *17*(1), 59–82. https://doi.org/10.1257/089533003321164958

Nahil, A., & Lyhyaoui, A. (2018). Short-term stock price forecasting using kernel principal component analysis and support vector machines: the case of Casablanca stock exchange. *Procedia Computer Science*, *127*, 161–169. https://doi.org/10.1016/j.procs.2018.01.111

Rao, P. S., Srinivas, K., & Mohan, A. K. (2020). A survey on stock market prediction using machine learning techniques. In *Lecture notes in electrical engineering* (pp. 923–931). https://doi.org/10.1007/978-981-15-1420-3_101

Ruder, S. (2016). An overview of gradient descent optimization algorithms. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1609.04747

Sahibi, Y., & Amine, A. (2010). Impact de l'adoption des indices MASI et MADEX sur la capitalisation boursière de la bourse de Casablanca. *Mémoire Online*.

Samek, W., Montavon, G., Lapuschkin, S., Anders, C. J., & Müller, K. (2020). Explaining Deep Neural Networks and Beyond: A review of Methods and applications. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2003.07631

Sezer, O. B., Gudelek, M. U., & Ozbayoglu, A. M. (2019). Financial Time Series Forecasting with Deep Learning : A Systematic\n Literature Review: 2005-2019. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1911.13288

Shiller, R. J. (2015). Irrational Exuberance Revised and Expanded third Edition. Princeton, NJ: Princeton University Press. https://doi.org/10.2307/j.ctt1287kz5

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). *Attention is all you need*. arXiv. https://arxiv.org/abs/1706.03762

Zhu, W., Lan, C., Xing, J., Zeng, W., Li, Y., Shen, L., & Xie, X. (2016). Co-Occurrence feature learning for skeleton based action recognition using regularized deep LSTM networks. *Proceedings of the AAAI Conference on Artificial Intelligence*, *30*(1). https://doi.org/10.1609/aaai.v30i1.10451