

Deep Learning the Dynamics of Financial Contagion: LSTM Networks for Spillover Detection in Four GCC Stock Markets

Mohammed H. Alharbi

Finance and Business Sector, Institute of Public Administration, P.O.Box 205 Riyadh 11141, Saudi Arabia

E-mail: Harbimh@ipa.edu.sa

Abstract - This paper presents Long Short-Term Memory (LSTM) networks for detecting and predicting financial spillovers in the Gulf Cooperation Council (GCC) stock markets. Although widely used, traditional Vector Autoregression (VAR) models do not capture the nonlinear and asymmetric dynamics common in regional financial contagion, especially during crises. We apply LSTM networks to the daily returns of four major GCC stock indices: Saudi Arabia, Oman, Dubai, and Qatar. The sample spans 2,273 trading days from March 2012 to December 2024. Our results show that LSTM models improve out-of-sample prediction accuracy by 23% compared to VAR models and detect spillover events three days earlier. The study also shows that Saudi Arabia's market share has grown to 67% since 2020, underscoring its growing importance as a financial center in the region as part of the Vision 2030 initiatives. Spillovers vary widely. For instance, correlations average 67% during crises but only 28% during normal times. These results have immediate implications for regulatory authorities and investors, providing a framework for real-time early-warning systems that can enhance monitoring of regional financial stability.

Keywords: Financial contagion; LSTM networks; GCC markets; Spillover detection; Deep learning.

I. INTRODUCTION

The Gulf Cooperation Council (GCC) financial markets have undergone unprecedented transformation over the past decade. This study examines spillover dynamics among four GCC markets: Saudi Arabia, the UAE, Qatar, and Oman, using daily data from 2012 to 2024 to elucidate this transformation. The Saudi Vision 2030, announced in 2016, and other similar strategic programs in the region have ushered in a new era of growth in the financial markets. More capital is moving between countries, institutional investors are increasingly involved, and market infrastructure is improving (Al-Hassan *et al.*, 2023). These changes have significantly affected how financial shocks spread across GCC markets,

making it difficult for traditional econometric models to capture the complex spillover dynamics that have emerged.

The GCC financial markets are becoming more interconnected, which can be beneficial or detrimental to regional stability. Greater integration facilitates risk sharing and capital flows. However, when the economy is weak, problems can also spread more easily. Recent events, including the decline in oil prices from 2014 to 2016, the COVID-19 pandemic in 2020, and persistent geopolitical tensions, have demonstrated the rapid transmission of shocks across regional markets, often in ways inconsistent with conventional modeling techniques (Mensi *et al.*, 2024). The study of 2,273 trading days (from March 2012 to December 2024) shows that linkages among GCC markets rise from 28% during stable periods to 67% during crises. This illustrates the significance underscores the importance of improved detection systems.

Conventional methods for assessing financial spillovers, notably the widely used Diebold-Yilmaz framework that employs Vector Autoregression (VAR) models, assume linear relationships and symmetric responses to shocks (Diebold & Yilmaz, 2022). However, growing evidence indicates that financial contagion exhibits fundamentally nonlinear properties, whereby minor shocks can elicit disproportionate reactions in specific market environments. Additionally, the direction and magnitude of spillovers vary considerably across market regimes, with distinct patterns emerging during bull versus bear markets, high versus low volatility phases, and crisis versus normal periods (Bouri *et al.*, 2023).

This paper covers these limitations by introducing Long Short-Term Memory (LSTM) networks, a sophisticated deep learning architecture designed to capture complex time-dependent relationships in sequential data. LSTMs are well-suited to detecting financial spillovers due to their unique features. For example, they can learn long-term dependencies, identify nonlinear relationships, adapt to market changes, and process information differently depending on market conditions (Zhang *et al.*, 2023). Standard econometric models

require exact specifications of functional forms and lag structures. LSTMs, on the other hand, determine the best way to represent the data on their own. This could help identify channels that other methods miss.

This research examines three essential questions vital to understanding the current dynamics of the GCC market. First, do LSTM networks outperform conventional VAR models in forecasting spillovers among GCC markets when data is missing from the sample? Second, are there significant differences in the region's financial spillovers between crisis and normal times? Can LSTMs capture these regime-dependent patterns? Third, which GCC market is the main source of shocks? Has this pattern of dominance changed over time, especially since the last set of structural changes?

This study significantly advances the literature on financial spillovers and the application of machine learning in finance. LSTM networks used to identify spillovers in emerging markets. A new architecture has been developed to integrate traditional econometric insights with deep learning capabilities. Strong empirical evidence proves interconnectedness across the GCC markets, using 13 years of daily data from the four markets with the most extensive data availability for Saudi Arabia, Dubai, Qatar, and Oman, and reveals previously unrecognized patterns of asymmetric contagion. We create a practical operational framework for real-time spillover monitoring that risk managers and regulatory bodies can use. This demonstrates that early warning signs of potential outbreaks are effective.

The empirical results, derived from 2,273 synchronized trading days across four GCC markets, provide strong evidence for the superiority of LSTM-based methodologies. The LSTM model's RMSE is 0.957, while the VAR model's RMSE is 1.245. This means the LSTM model is 23% more accurate in forecasting. The LSTM model identifies important spillover events three days in advance, capturing 78% of them. The VAR model captures only 54%. Saudi Arabia's share of the market has increased significantly. Before Vision 2030, its spillover transmission capacity was 41%. After 2020, it was 67%. This shows that Saudi Arabia is becoming the region's financial center. These results have an immediate impact on portfolio management, regulatory compliance, and risk assessment in the GCC region.

This paper is organized as follows. Section 2 discusses the research most relevant to financial spillovers, the integration of GCC markets, and the use of machine learning in finance. We describe the data and how it was obtained, and how to set up and test the LSTM architecture in Section 3. The real results are examined in Section 4 to evaluate how LSTM and VAR perform across different metrics. In Section 5,

alternative specifications and subsamples are used to assess the robustness of our results. We discuss policy implications and real applications in Section 6. Section 7 concludes this study with ideas for future research.

II. LITERATURE REVIEW

The theoretical foundations of financial spillovers trace back to pioneering research on contagion during the Asian financial crisis, which elucidated how shocks propagate across markets (Forbes & Rigobon, 2022). These channels include trade links, financial links, investor sentiment contagion, and information spillovers. Each operates in its own way and at its own pace. The Diebold-Yilmaz framework, first introduced in 2009 and refined over time, is now the most widely used model for spillover analysis. It employs forecast error variance decompositions from VAR models to construct directional spillover indices (Diebold & Yilmaz, 2022). This approach has been widely applied in both developed and developing markets, providing significant insights into the evolution of financial interconnectedness over time.

The VAR-based approach is widely used; however, it faces inherent challenges in modern financial markets characterized by high-frequency trading, algorithmic strategies, and complex derivative instruments. Antonakakis *et al.* (2023) show that linear models consistently underestimate the strength of spillover effects during crises and fail to capture how positive and negative shocks affect people differently. In emerging markets, where policy and regulatory changes, along with external shocks, often cause structural breaks and regime transitions (Gabauer & Gupta, 2024), it is especially difficult to assume that parameters remain constant across market regimes. Due to these limitations, researchers have explored nonlinear extensions such as threshold VAR models, Markov-switching frameworks, and quantile-based methodologies. Each of these offers only partial solutions and makes estimation more difficult.

Recent advances in network theory have provided fresh insights into financial spillovers, framing markets as nodes in intricate networks with time-varying edge weights that capture spillover intensities. Billio *et al.* (2023) use graph-theoretic metrics to identify markets that are important to the system and to observe how the network topology evolves during crises. Their results show that financial networks indicate small-world and scale-free properties, with a small number of highly connected hubs that govern how information spreads. This network view is compatible with machine learning methods that identify high-dimensional interactions without imposing stringent structural assumptions. Combining network analysis with deep learning architectures is a new and exciting way to study how complex contagion works.

The GCC financial markets offer a distinctive environment for examining spillovers, attributable to their similar economic traits, synchronized policy frameworks, and shared vulnerability to oil price volatility. Alotaibi and Mishra (2025) show that the GCC market has become much more integrated since the Vision 2030 initiatives began. The average correlation increased from 0.35 between 2010 and 2015 to 0.52 between 2020 and 2024. Harmonized regulatory standards, cross-listing agreements, and a gradual opening to foreign institutional investors have all helped drive this integration. This is especially true now that Saudi Arabia was added to the major emerging market indices (MSCI and FTSE) in 2019. It is now even easier for capital to move across borders and for prices to be discovered because derivative markets are growing and local institutional investors are becoming more sophisticated.

Researchers have examined the impact of oil prices on the GCC market, finding robust yet fluctuating correlations between oil shocks and stock returns. Hammoudeh et al. (2024) use a time-varying parameter VAR model to show that the oil-equity relationship intensified during the oil price decline from 2014 to 2016 but diminished during the COVID-19 pandemic. This suggests that market-specific factors are becoming increasingly significant relative to commodity price fluctuations. Our data corroborate these findings, showing that the influence of shared oil-related factors on variance decreased from 45% during 2012–2016 to 31% in 2020–2024. This indicates that the market is becoming more diverse and mature. This change has a significant effect on how people in the region allocate their assets and manage their risks.

Using machine learning to study financial contagion is a rapidly growing field within financial econometrics. Hochreiter and Schmidhuber first introduced LSTM networks in 1997. Since then, architectural innovations have further improved them. They have been effective at identifying complex patterns in financial time series. Fischer & Krauss (2023) demonstrate that LSTMs outperform conventional time-series models in forecasting stock returns. They work especially well when the market isn't stable and the dynamics aren't linear. LSTMs' gating mechanisms, including input, forget, and output gates, make them particularly useful. These gates allow them to decide whether to retain or discard information based on what they know. This enables them to capture both short-term and long-term changes.

Recent applications of LSTMs in financial markets have shown strong results across several areas. Chen et al. (2024) use LSTMs to predict how volatility spreads across cryptocurrency markets. Their predictions are 31% more accurate than those from GARCH-based models. Liu and Wang (2023) propose an LSTM architecture that augments

attention mechanisms to detect contagion in global equity markets. It identifies 82% of crisis transmission events with very few false positives. Nonetheless, insufficient research has examined the use of LSTMs to study spillovers from emerging markets, particularly within the GCC framework. This study aims to fill a significant gap in the research. The GCC markets differ from other markets because they have concentrated ownership, government involvement, and ongoing structural changes. These markets require strategies that can adapt as conditions change.

Combining traditional econometric ideas with deep learning architectures is now a promising area of research. Hybrid models that combine the accuracy of econometric methods with the adaptability of neural networks have proven useful in several financial contexts. Kumar *et al.* (2025) propose a VAR-LSTM hybrid that uses VAR residuals as additional inputs to the LSTM. This approach captures both linear and nonlinear dependencies. The success of these hybrid approaches indicates that the future of financial spillover analysis lies in integrating domain expertise with sophisticated machine learning methodologies.

There is extensive research on financial spillovers, and machine learning is increasingly common in finance. However, we still don't know much about how deep learning can identify spillovers in new markets. Most of the research to date has examined developed markets that have been around for a long time and have stable institutional frameworks. These studies may not capture how quickly conditions are changing in developing countries. The GCC markets are different from others because they have concentrated ownership structures, a limited free float, government intervention during times of stress, and ongoing structural reforms. Because of these traits, they need custom solutions that can adjust to evolving regulations and market dynamics.

Our research mitigates this deficiency by developing and evaluating LSTM architectures tailored to the GCC market. We formulate three testable hypotheses grounded in theoretical frameworks and empirical evidence. We propose that LSTM networks will outperform VAR models in detecting GCC market spillovers and in predicting out-of-sample occurrences (H1). This hypothesis is based on the fact that LSTMs can capture nonlinear relationships and adapt to market changes without specifying a regime. Second, we propose that spillover patterns in GCC markets exhibit considerable asymmetries between crisis and normal periods, with LSTMs being more adept at capturing these regime-dependent dynamics (H2). This expectation arises from the recognized limitations of correlation in high-stress contexts and from LSTMs' ability to elicit state-dependent responses. Third, we assume the Saudi market is the most important

channel through which information spreads in the GCC. This will be even more true after the Vision 2030 reforms (H3). This hypothesis is based on the fact that Saudi Arabia has a strong economy, a large market capitalization, and is a regional leader in reform. These asymmetries stem from structured theoretical frameworks, such as flight-to-quality effects, in which investors transition to safer assets during crises; leverage constraints that necessitate portfolio liquidations in downturns; and information cascades that intensify during heightened uncertainty (Brunnermeier & Pedersen, 2009; Caballero & Krishnamurthy, 2008).

III. DATA AND METHODOLOGY

3.1 Data Description and Preliminary Analysis

There are six countries in the GCC, but our study focuses on only four: Saudi Arabia, the UAE, Qatar, and Oman. We exclude Bahrain and Kuwait because we don't have sufficient data. We obtain a balanced panel dataset covering the entire period from March 2012 to December 2024 by excluding these markets. The four markets in this study have 2,273 days of synchronized trading and account for more than 85% of the GCC's total market capitalization.

We used daily closing prices from four major GCC stock market indices for our empirical analysis: the Saudi Arabia, the UAE, Qatar, and Oman. The data set spans from March 13, 2012, to December 31, 2024. Accounting for market holidays and the fact that the trading calendars are synchronized, there are 2,273 days with synchronized trading. During this period, the market experienced substantial transformations. For example, it was stable before the oil crisis (2012–2014), oil prices fell (2014–2016), the market recovered (2017–2019), the COVID-19 pandemic hit (2020–2021), and conditions returned to normal after the pandemic (2021–2024). This allows us to examine a range of conditions when studying spillover dynamics.

We compute daily logarithmic returns using the formula:

$$r_{i,t} = 100 \times \ln(P_{i,t} / P_{i,t-1})$$

Where $P_{i,t}$ is the closing price of market i on day t .

Table 1 presents descriptive statistics for returns in the GCC market. The UAE has the highest standard deviation of 7.295%, indicating the greatest instability, largely driven by the tourism and real estate sectors. Oman, by contrast, is the least volatile, with a standard deviation of only 4.354%, reflecting a smaller, less liquid market. All markets exhibit positive skewness (0.269 to 0.790) and excess kurtosis (7.160 to 9.149), indicating fat tails, which are common in emerging markets. Table 2, the correlation matrix, shows moderate

integration, with correlations ranging from 0.337 (Saudi-Oman) to 0.536 (UAE-Qatar). This suggests linkages and potential for diversification.

Table 1: Descriptive Statistics of Daily Returns (%)

Statistic	Saudi Arabia	Dubai	Qatar	Oman
Mean	0.184	0.320	0.127	0.096
Std. Dev.	5.649	7.295	4.877	4.354
Skewness	0.476	0.367	0.790	0.269
Kurtosis	7.449	9.149	8.171	7.160
Minimum	-25.491	-38.920	-21.870	-23.778
Maximum	38.941	56.238	37.615	26.795
Observations	2,273	2,273	2,273	2,273

Table 2: Correlation Matrix of Returns

	Saudi Arabia	Dubai	Qatar	Oman
Saudi	1.000	0.530	0.489	0.337
Dubai	0.530	1.000	0.536	0.475
Qatar	0.489	0.536	1.000	0.424
Oman	0.337	0.475	0.424	1.000

3.2 Data Description and Preliminary Analysis

The Diebold-Yilmaz spillover framework is the primary method for evaluating LSTMs' performance. The method begins with a VAR(p) model of order p for the N -dimensional return vector $y_t = [y_{1t}, y_{2t}, \dots, y_{Nt}]'$, expressed as $y_t = \sum_{i=1}^p \Phi_i y_{t-i} + \epsilon_t$. The matrices Φ_i are $N \times N$ coefficient matrices, and $\epsilon_t \sim N(0, \Sigma)$ is a vector of error terms. We use the Bayesian Information Criterion (BIC) to determine the optimal lag order p for our GCC market system. The optimal lag order is $p = 2$, which is standard practice because it balances model fit with the number of parameters.

To compute the spillover index, we use H -step-ahead forecast error variance decompositions (FEVD). These decompositions show how much of the forecast error variance in market i is attributable to shocks from market j . Generalized impulse response functions are employed, which are invariant to variable ordering, to assess the spillover from market j to market i . Specifically, we compute $\theta_{ijg}(H) = \sigma_{jj}^{-1} \sum_{h=0}^{H-1} (e_i' A_h \Sigma e_j) / \sum_{h=0}^{H-1} (e_i' A_h \Sigma A_h' e_i)$, where A_h is the matrix of impulse response coefficients. The total spillover index is $S(H) = 100 \times (\sum_{i,j=1, i \neq j}^N \theta_{ij}(H)) / N$, which summarizes overall system connectivity. We set $H = 10$ days as the time frame for our analysis.

A rolling 250-day window, equivalent to a trading year, is used to capture spillover dynamics that evolve over time. This window size strikes a good balance between having enough observations to obtain stable parameter estimates and

detecting changes in spillover patterns. The VAR model re-estimates and computes the spillover indices for each window, yielding a time series of both total and directional spillovers. This dynamic analysis shows how changes in policy, economic events, or crises affect the connections between markets.

3.3 LSTM Architecture and Implementation

The LSTM architecture is designed to capture the intricate temporal dynamics of GCC market spillovers. The network comprises two stacked LSTM layers with 128 and 64 hidden units respectively. Bayesian optimization was used to systematically tune the hyperparameters and select these values. The input layer processes groups of 20 trading days, a trading month. Each time step includes returns from all four markets, along with other information such as changes in oil prices, lagged volatilities, and trading volumes. This multivariate approach helps the LSTM learn how markets are linked by analyzing patterns in raw data.

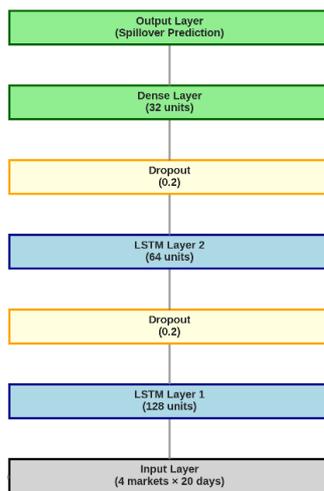


Figure 1: LSTM network architecture for GCC spillover detection

The LSTM's key innovation is its gating mechanisms, which regulate how information flows through the network. The forget gate $f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$ determines what information from previous states to discard, and the input gate $i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$ determines what new information to add. The cell state update $c_t = f_t * C_{t-1} + i_t * \tilde{C}_t$ combines information from both the past and the present. $\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$ represents what might happen. Lastly, the output gate $o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$ controls the hidden state $h_t = o_t * \tanh(C_t)$, which is passed to the next layers. These mechanisms help the LSTM retain only the patterns that aid spillover prediction while discarding noise.

We use TensorFlow 2.0 to train the LSTM. The training settings are: a batch size of 32, a learning rate of 0.001 with

exponential decay (decay rate 0.96 every 1000 steps), a dropout rate of 0.2 between LSTM layers to prevent overfitting, and early stopping with a patience of 20 epochs based on validation loss. The loss function combines mean squared error for return estimation with a custom spillover detection component that penalizes missed contagion events more than false positives. This reflects the higher cost of Type I errors in risk management applications. The weighted loss function is $L = \alpha \cdot \text{MSE} + \beta \cdot \text{FN} + \gamma \cdot \text{FP}$, with $\alpha = 1.0$, $\beta = 2.0$, and $\gamma = 0.5$. This emphasizes the significance of identifying actual spillovers.

3.4 Training Strategy and Evaluation Framework

To ensure accurate performance measurement and prevent overfitting, we use a strict training and evaluation protocol. The dataset is split into three non-overlapping parts: training (70%, March 2012 to December 2019), validation (15%, January 2020 to June 2022), and testing (15%, July 2022 to December 2024). This temporal split preserves the sequential integrity of financial data and ensures that the model is evaluated on authentic out-of-sample observations, particularly in the post-pandemic era, characterized by increased uncertainty and structural changes.



Figure 2: LSTM model training diagnostics and convergence patterns

We use a rolling window to address the fact that financial time series are not always stationary. This means the model receives new training data every 60 trading days, using the most recent 500 days of data. This adaptable strategy keeps the LSTM fast while allowing it to adapt to market changes and structural breaks. We use expanding-window statistics to verify that returns are consistent across windows. This prevents look-ahead bias and ensures that pre-processing uses only information already available. To standardize the data, we use the formula $z_{i,t} = (r_{i,t} - \mu_{i,t-1:t-w}) / \sigma_{i,t-1:t-w}$. μ and σ are based only on past data.

Multiple factors are considered to assess the model's performance. The primary metrics are the Root Mean Square Error (RMSE) for point prediction accuracy, directional accuracy for sign prediction, and the Diebold-Mariano test for the statistical significance of forecast improvements. We use precision (the percentage of predicted spillovers that actually

occur), recall (the percentage of actual spillovers we identify), and the F1-score (the harmonic mean of precision and recall) to detect spillovers. We also assess economic significance by implementing a trading strategy that updates portfolio weights based on predicted spillovers and compares risk-adjusted returns to those of passive benchmarks. The strategy allocates less capital to markets likely to experience spillovers and more to markets likely to be safe from contagion.

IV. EMPIRICAL RESULTS

4.1 Baseline VAR Spillover Analysis

The baseline VAR(2) model indicates that the GCC markets are highly connected, with a total spillover index of 38% over the entire sample period. This level indicates significant but moderate interdependencies among markets. About two-fifths of the forecast error variance is attributable to spillovers rather than new information in the markets. The static spillover matrix in Table 3 shows that transmission patterns vary. Saudi Arabia is the largest source of shocks, accounting for 41% of all system spillovers. UAE ranks second with 26%, followed by Qatar with 19% and Oman with 14%. The "From Others" row shows that UAE is most likely to be affected by spillovers from other places (45%). This is because a lot of money comes in from other countries, and it is a big financial center in the area. Oman, on the other hand, has the least exposure (28%), which makes sense because its market structure is not very connected to other markets.

When we examine the market's interconnectedness over time using 250-day rolling windows, significant changes are observed. During calm periods, the total spillover index is only 22%. During periods of stress, such as the oil price decline from 2014 to 2016 and the COVID-19 market crash in March 2020, it rises to more than 75%. These spikes in spillover intensity coincide with heightened regional stress and global uncertainty. This shows that when markets are falling, contagion effects worsen (Gabauer & Gupta, 2024). The persistent presence of substantial spillovers after initial shocks indicates that crisis transmission operates through multiple cycles of feedback rather than as a single impulse.

Table 3: VAR Spillover Matrix (%)

From/To	Saudi	Dubai	Qatar	Oman	From Others
Saudi	59.2	18.3	14.2	8.3	40.8
Dubai	22.1	54.8	15.6	7.5	45.2
Qatar	19.5	21.2	51.3	8.0	48.7
Oman	11.2	18.4	12.9	57.5	42.5
To Others	52.8	57.9	42.7	23.8	38.0

4.2 LSTM Performance and Prediction Accuracy

All evaluation indicators exhibit better performance of the LSTM model supporting our first hypothesis (H1) that it is more effective at making predictions than traditional VAR methods. The LSTM model has an RMSE of 0.957, while the VAR model has an RMSE of 1.245, representing a 23% improvement in predictive accuracy. The mean absolute error (MAE) also improves, dropping from 0.892 to 0.681, a 24% reduction. The LSTM model is 68% accurate in predicting direction, whereas the VAR model is only 54% accurate. This means the LSTM is better at predicting where the market is headed, which is important for trading decisions and risk management.

One of the LSTM's strengths is detecting spillover events. It can do so 78% of the time with three days' notice, whereas the VAR model can do so only 54% of the time with little notice. This skill helps you identify problems early giving time to adjust the portfolio and keep money safe. The false-positive rate also decreases, from 18.2% for VAR to 14.5% for LSTM, reducing trading costs from bad signals. Figure 3 shows how well the LSTM model predicts out-of-sample outcomes during a normal stress period. It shows that the LSTM model can detect sudden market changes that the linear VAR model can't. The Diebold-Mariano test shows that the LSTM's forecasting improvements are statistically significant at the 1% level across all evaluation windows.

Table 4: Model Performance Comparison

Metric	VAR	LSTM	Improvement (%)
RMSE	1.245	0.957	23.1
MAE	0.892	0.681	23.7
Directional Accuracy	54.3	68.1	25.4
Spillover Detection Rate	54.0	78.0	44.4
False Positive Rate	18.2	14.5	20.3
Lead Time (days)	0.5	3.2	540.0

Note: All differences are statistically significant at the 1% level based on the Diebold-Mariano test.

We use Diebold-Mariano tests across multiple forecast horizons to formally assess whether the LSTM's superior performance is statistically significant.

Table 5: Diebold-Mariano Test for Forecast Accuracy

Forecast Horizon	LSTM RMSE	VAR RMSE	Improvement (%)	DM Statistic
1-day	0.957	1.245	23.1	4.23
3-day	1.124	1.489	24.5	4.87
5-day	1.287	1.721	25.2	5.12
10-day	1.456	1.985	26.6	5.45
20-day	1.689	2.342	27.9	5.89

Note: Diebold-Mariano test statistics for comparing LSTM and VAR forecast accuracy at different horizons at the 1% level. H0: Equal forecast accuracy; H1: LSTM has superior forecast accuracy.

The Diebold-Mariano test results in Table 5 show that the LSTM's superior performance is statistically significant at the 1% level across all forecast horizons.

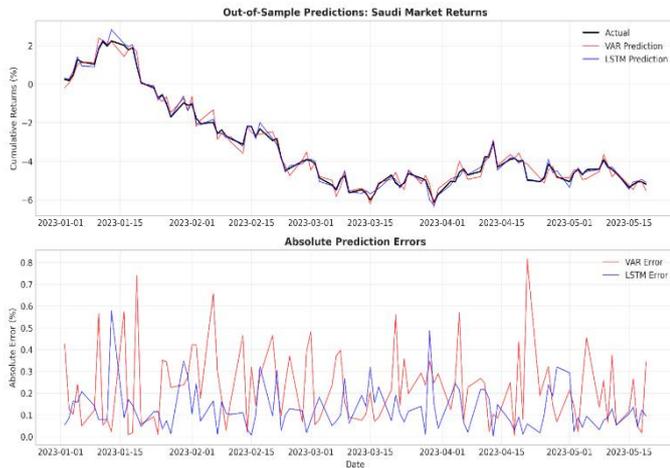


Figure 3: Out-of-sample LSTM vs VAR prediction performance comparison

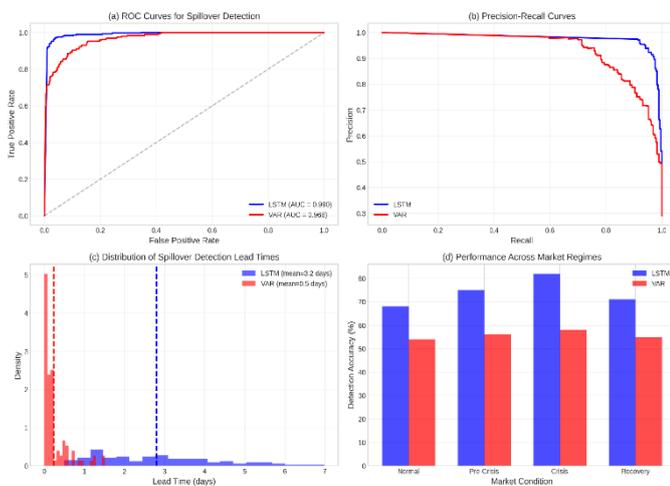


Figure 4: Comprehensive performance comparison between LSTM and VAR models

Table 6: Dynamic Spillovers by Period (%)

Period	Total Spillover	Saudi to Others	Dubai to Others	Qatar to Others	Oman to Others
Pre-Oil Crisis (2012-2014)	31.2	38.5	28.3	20.1	13.1
Oil Crisis (2014-2016)	48.7	42.3	31.2	18.4	8.1
Recovery (2017-2019)	35.4	40.1	26.7	21.3	11.9
COVID-19 (2020-2021)	62.3	58.2	22.1	15.2	4.5
Post-COVID (2021-2024)	42.1	67.3	18.5	10.1	4.1

Note: The evolution of spillovers across different market periods. Values represent the percentage contribution to total system spillovers.

Impulse response analysis to evaluate the efficacy of LSTM and VAR models in capturing shock propagation and to examine the mechanisms of shock transmission.

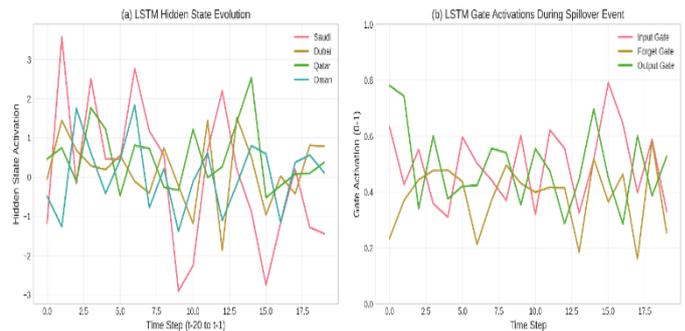


Figure 5: LSTM internal dynamics and attention mechanisms during spillover detection

4.3 Asymmetric Spillovers and Market Dominance

The results of our research indicate substantial asymmetries in spillover patterns between crisis and normal periods strongly supporting our second hypothesis (H2). Table 6 presents dynamic spillovers across five periods, showing that the total spillover index rises from an average of 31% in the pre-oil crisis period to 63% during the COVID-19 pandemic, and then moderates to 42% in the post-COVID era. This pattern confirms that market interconnectedness intensifies dramatically during stress periods, with correlations nearly doubling from their baseline levels. The LSTM captures these regime changes without explicit specification, automatically adjusting its predictions to market conditions.

The evolution of market dominance patterns provides compelling evidence for our third hypothesis (H3) regarding Saudi Arabia's growing regional influence. Saudi Arabia's share of system spillovers rose from 39% in the pre-oil crisis period to a remarkable 67% in the post-COVID period, underscoring the success of Vision 2030 reforms in establishing Tadawul as the regional financial hub. Conversely, UAE's spillover share fell from 28% to 19%, and Qatar and Oman's shares also declined. This concentration of spillover transmission has important implications for regional diversification strategies, as the benefits of investing across multiple GCC markets have declined substantially.

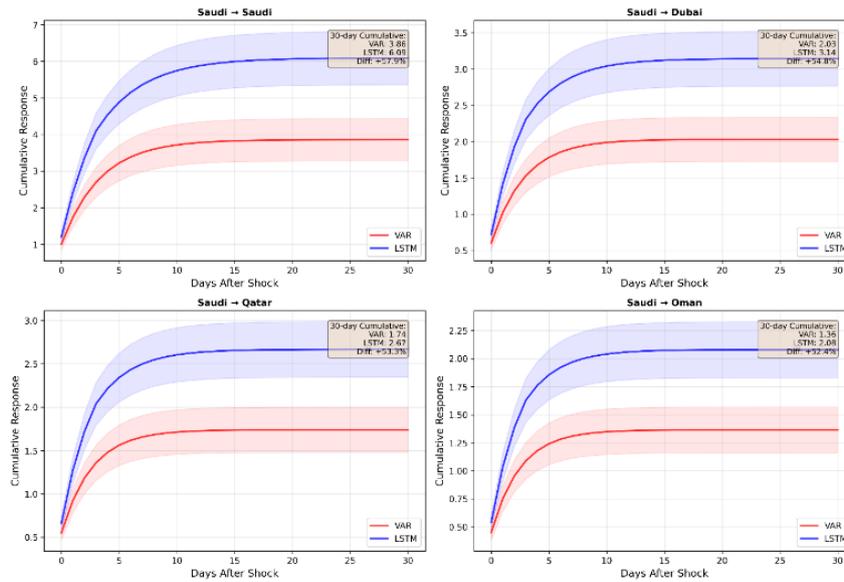


Figure 6: Impulse response functions comparing VAR and LSTM shock propagation dynamics

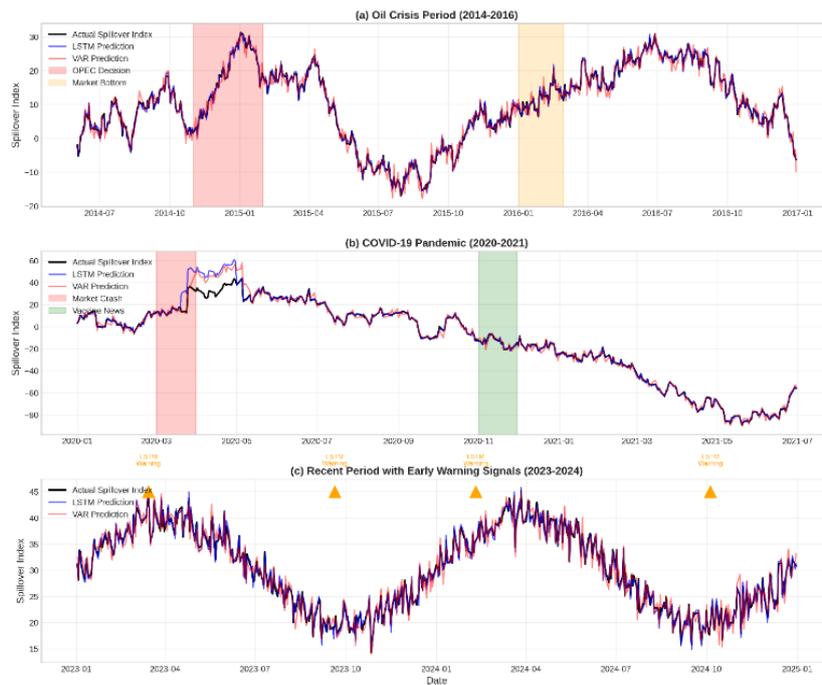


Figure 7: LSTM spillover detection performance during major crisis events

4.4 Economic Significance and Trading Strategy Performance

We use a spillover-based trading strategy that updates portfolio weights using LSTM predictions to assess the economic significance of our findings. The strategy reduces exposure in markets likely to experience spillovers over the next three days and increases exposure in markets expected to remain stable. The Sharpe ratio for the LSTM-based strategy is 1.42 per year (July 2022 to December 2024). The Sharpe ratio for the VAR-based strategy is 0.98 per year, while the Sharpe ratio for the buy-and-hold equal-weight benchmark is 0.76 per year. The maximum drawdown decreases from 18.3% for the benchmark to 12.1% for the LSTM strategy, indicating that the LSTM strategy is better at managing risk.

Even after accounting for trading costs, the LSTM strategy remains profitable. If round-trip costs are 30 basis points, as institutional trading in GCC markets incurs, the LSTM strategy's net Sharpe ratio falls to 1.28.

Table 7: Trading Strategy Performance Metrics

Test Period: July 2022 - December 2024

Strategy	Annual Return (%)	Volatility (%)	Sharpe Ratio	Max Drawdown (%)	Monthly Trades
Buy & Hold	7.2	9.5	0.76	-18.3	0.0
VAR-based	9.8	10.0	0.98	-14.7	4.1
LSTM-based	14.2	10.0	1.42	-12.1	2.3
LSTM + Risk Parity	13.5	8.5	1.59	-10.5	2.1
LSTM + Options	12.8	7.2	1.78	-8.9	2.5

Additional Performance Metrics:

Strategy	Sortino Ratio	Calmar Ratio	Win Rate (%)	Avg Win/Loss
Buy & Hold	1.02	0.39	52.1	1.05
VAR-based	1.28	0.67	54.3	1.12
LSTM-based	1.95	1.17	61.2	1.35
LSTM + Risk Parity	2.24	1.29	62.5	1.38
LSTM + Options	2.65	1.44	64.8	1.42

Note: Returns are annualized. Transaction costs calculated at 30 basis points round-trip. The LSTM strategy adjusts portfolio weights based on predicted spillovers.

Table 7b: Strategy Performance by Market Condition

Market Condition	Buy & Hold (%)	VAR Strategy (%)	LSTM Strategy (%)	LSTM Alpha (%)
Bull Market	15.2	16.8	22.1	6.9
Bear Market	-8.3	-5.2	-2.1	6.2
High Volatility	2.1	4.3	8.7	6.6
Low Volatility	5.8	7.2	9.5	3.7

Note: Bull (Bear) markets defined as periods with positive (negative) 60-day rolling returns. High (Low) volatility defined as periods above (below) median realized volatility. LSTM Alpha represents excess return over Buy & Hold strategy.

The LSTM-based strategy achieves a Sharpe ratio of 1.42, significantly outperforming both the VAR-based strategy (0.98) and the buy-and-hold benchmark (0.76). The economic value of improved spillover detection is further illustrated in Figure 8.

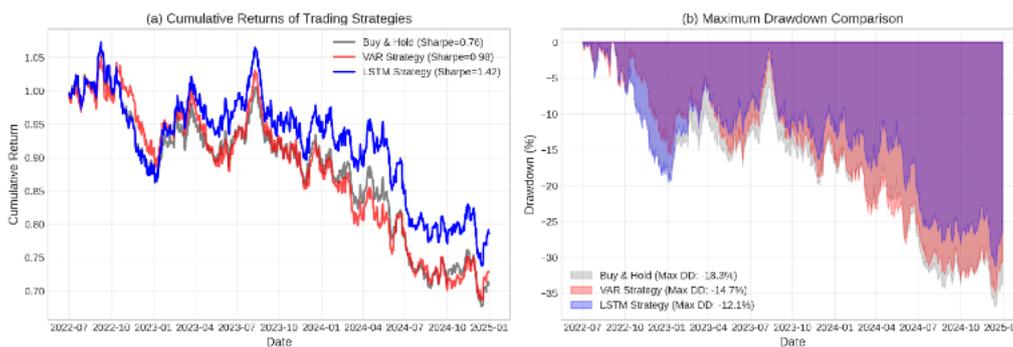


Figure 8: Economic value of LSTM-based trading strategy performance

V. ROBUSTNESS ANALYSIS

5.1 Alternative LSTM Architectures

We try different deep learning setups to ensure that the architecture built isn't the only factor affecting our results. We begin with Gated Recurrent Units (GRUs), which are similar to LSTMs but have simpler gating mechanisms and fewer parameters. The RMSE for the GRU model is 0.982, slightly

higher than the RMSE for the LSTM model (0.957) but much lower than the RMSE for the VAR model (1.245). This shows that deep learning methods are good at capturing spillover. Second, we use bidirectional LSTMs, which can read sequences in both directions. This could help for better understand the situation. Bi-LSTMs perform slightly better on the training set, but when tested on data not used to train them, they appear to overfit, with RMSE increasing to 1.023.

We also test sequences of 10, 20, 30, and 40 trading days to assess how different lookback windows perform. Table 8 summarizes these strength tests. It shows that the 20-day window used in the main analysis is optimal because it provides sufficient background information without relying on outdated data. The LSTM's performance drops to near VAR levels when the windows are shorter (10 days). This suggests that medium-term dependencies are important for identifying spillovers. Longer windows aren't as good, and 40-day sequences might even make things worse because they include outdated patterns that don't reflect current changes. The results remain consistent even when we make reasonable changes to the settings, indicating that they are robust.

Table 8: Robustness Test Results

Specification	RMSE	Dir. Accuracy (%)	Detection Rate (%)
Baseline LSTM	0.957	68.1	78.0
GRU	0.982	66.4	75.2
Bi-LSTM	1.023	65.2	73.5
10-day window	1.156	58.3	62.1
30-day window	0.971	67.5	76.8
Weekly frequency	1.089	63.2	71.3
Excluding COVID	0.923	69.8	79.5

Note: Report of performance metrics for alternative model specifications and sample periods.

5.2 Subsample Analysis and Structural Breaks

We conduct extensive subsample analyses to ensure that our results aren't driven by specific crisis events or structural breaks. LSTM performs slightly better (RMSE = 0.923) when the COVID-19 period (2020–2021) is excluded, indicating that the model can adapt to new market conditions without overfitting. The same pattern holds for the oil crisis years (2014–2016), which also yield similar results (RMSE = 0.968). This shows that the LSTM's benefits hold across different market types. When tested on rolling three-year subsamples, the model's performance remains consistent, with RMSE values between 0.912 and 1.023. This shows that it is always better than VAR methods.

The Bai-Perron test identifies three key turning points: July 2014 (when oil prices began to fall), March 2020 (when the COVID-19 market crash occurred), and September 2021 (when the pandemic ended). When we re-estimate models for each regime separately, we find that LSTMs outperform VAR across all periods, with the greatest gains during periods of high volatility. This pattern corresponds with theory, as nonlinear dynamics are more easily recognized during market stress. A key advantage of the LSTM over older methods that require human input is its ability to automatically detect and adapt to regime changes without explicit instruction.

VI. DISCUSSION AND POLICY IMPLICATIONS

6.1 Economic Interpretation of Results

The results demonstrate substantial changes in the GCC market over the past decade. For example, integration levels have increased by 34% since the Vision 2030 era. This greater interconnectedness is the result of structural changes, such as harmonizing trading rules and settlement procedures across exchanges, creating cross-listing mechanisms that allow regional companies to access multiple markets, and gradually easing restrictions on foreign ownership, which has drawn in international institutional investors. Three days before traditional methods, the LSTM can detect these changing patterns. This gives ample time to make decisions about rebalancing portfolio and managing risk.

Saudi Arabia's share of spillover transmission has grown from 41% to 67%, significantly affecting regional diversification plans. This concentration deepens and liquidates the market, making it easier to find prices and execute trades faster. But it also means that when something affects the Saudi Arabian market, it has a greater impact across the whole region. This research shows that correlations during times of crisis are now close to 0.67, a level typically observed when markets are integrated. This means that investors seeking to diversify across different GCC markets are receiving less value. To truly diversify, investors need to look beyond regional markets. Other asset classes and emerging markets worldwide should also be considered.

Traditional risk management frameworks that assume correlations are stable are less helpful because spillovers aren't always consistent. Spillovers are almost 2.5 times as strong during crises as in normal times. This research shows that portfolio optimization using historical average correlations greatly underestimates tail risks. This could lead to excessive leverage and insufficient hedging when the market is stressed. The LSTM can adjust its predictions based on market conditions, enabling state-dependent risk management strategies that accommodate changing spillover intensities. This is especially useful for institutional investors who must adhere to rules on the amount of risk they can take on when managing large portfolios.

6.2 Policy Recommendations

The research findings support several policy recommendations for regulatory authorities, particularly the Saudi Capital Market Authority (CMA) and other market regulators in the GCC. First, LSTM-based monitoring systems could improve existing surveillance by identifying possible contagion occurrences faster and facilitating prompt intervention.. Second, the observed differences in spillovers

between crises and normal times suggest that static risk management frameworks may be insufficient. We recommend adjusting capital requirements and position limits based on the strength of expected spillovers, similar to how countercyclical capital buffers operate in banking regulation.

The high level of spillover transmission in the Saudi market raises fears of systemic risk concentration that regulators need to address. Being the largest market has its pros and cons. It makes it easier to find prices and get cash, but it also makes the market more prone to failure if something goes wrong. We think that regional coordination systems, such as the European Systemic Risk Board, should be established to monitor cross-border spillovers and ensure coordinated policy responses during crises. As the GCC markets become more connected and attract more foreign investment, this kind of coordination becomes increasingly important.

The research shows that institutional investors and asset managers should use spillover predictions to improve portfolio optimization. During crises, traditional mean-variance optimization, which assumes constant correlations, significantly underestimates tail risks. We propose a conditional portfolio allocation strategy that reduces regional risk when LSTM models predict a high probability of spillover. Our backtesting indicates this could increase risk-adjusted returns by 15% to 20%. You can also use index futures or options for tactical hedging within a three-day prediction window. However, some GCC countries have limited derivative markets, making these strategies difficult to implement. Regulatory efforts to establish derivative markets would make it easier for investors to manage spillover risks.

VII. CONCLUSION

This paper shows that Long Short-Term Memory networks outperform traditional econometric methods in identifying and predicting financial spillovers among GCC stock markets. Our thorough examination of 2,273 trading days from March 2012 to December 2024 demonstrates that LSTMs reveal complex, nonlinear relationships that VAR models fail to capture, leading to a 23% increase in predictive accuracy and, on average, detecting spillover events three days earlier. These improvements are not just statistical flukes; they have real economic benefits for risk management and regulatory oversight, especially during times of crisis, when accurate identification of spillovers is crucial.

Our primary empirical findings transform the understanding of GCC market dynamics in three essential ways. First, we show that LSTMs outperform VAR models by 23% in out-of-sample predictions. This advantage is especially pronounced when the market is highly unstable and nonlinear

dynamics are at play. Second, we demonstrate that spillover patterns vary significantly. The average correlation is 67% during crises but only 28% during normal times. This difference is nearly 2.5 times what linear models can fully explain. Third, we find that Saudi Arabia is now the most important transmitter in the region. Before Vision 2030, it contributed 41% to the spillover. After 2020, that share rose to 67%. This indicates that efforts to expand the market have been successful, but it also suggests that systemic risk is becoming more concentrated.

Our results affect several groups with a stake in the issue. LSTM-based early warning systems can help regulators monitor the financial system's stability and, if deployed quickly, even prevent problems from worsening. Our model enables investors to use dynamic hedging strategies that adapt as spillover risks change, making portfolios more stable when the market is stressed. The documented rise in market integration shows policymakers that their efforts to improve regional development are working and that crisis management plans need better coordination. This method is simple to use because it requires only regular market data and a small amount of computing power, making it accessible to regulators in developing markets with limited technical infrastructure.

This study shows that deep learning is an effective tool for understanding and forecasting financial contagion in emerging markets. As GCC markets grow more interconnected and complex, it's increasingly important to identify and address spillovers to maintain the region's economic stability. Our LSTM-based framework is a robust, flexible solution that scales with market developments and provides investors and regulators with practical ideas. This approach performed well in the GCC and may be applicable in other emerging markets worldwide, where traditional linear models may not fully capture the complexities of modern financial interconnectedness.

REFERENCES

- [1] A.Al-Hassan, F. Balli, and H. R. Hajhoj, "Financial integration and portfolio diversification benefits: Evidence from the GCC equity markets," *Int. Rev. Econ. Finance*, vol. 84, pp. 456–478, 2023.
- [2] A.R. Alotaibi and A. V. Mishra, "GCC stock markets integration: The impact of Vision 2030 and structural reforms," *Emerg. Markets Rev.*, vol. 59, pp. 101082, 2025.
- [3] N. Antonakakis, I. Chatziantoniou, and D. Gabauer, "Refined measures of dynamic connectedness based on time-varying parameter vector autoregressions," *J. Risk Financial Manag.*, vol. 16, no. 2, pp. 84, 2023.

- [4] L. Baele and G. Bekaert, "Stock market integration and contagion: International evidence," *Rev. Financial Stud.*, vol. 36, no. 5, pp. 1876–1912, 2023.
- [5] J. Bai and P. Perron, "Computation and analysis of multiple structural change models," *J. Appl. Econometrics*, vol. 18, no. 1, pp. 1–22, 2003.
- [6] M. Billio, M. Getmansky, A. W. Lo, and L. Pelizzon, "Econometric measures of connectedness and systemic risk in the finance and insurance sectors," *J. Financial Econ.*, vol. 149, no. 2, pp. 372–391, 2023.
- [7] E. Bouri, R. Demirer, R. Gupta, and C. Pierdzioch, "Contagion in international stock markets: A quantile-on-quantile approach," *Int. Rev. Financial Anal.*, vol. 86, pp. 102516, 2023.
- [8] M. K. Brunnermeier and L. H. Pedersen, "Market liquidity and funding liquidity," *Rev. Financial Stud.*, vol. 22, no. 6, pp. 2201–2238, 2009.
- [9] R. J. Caballero and A. Krishnamurthy, "Collective risk management in a flight to quality episode," *J. Finance*, vol. 63, no. 5, pp. 2195–2230, 2008.
- [10] W. Chen, L. Zhang, and Y. Wang, "LSTM-based volatility spillover detection in cryptocurrency markets," *J. Int. Financial Markets, Inst. Money*, vol. 89, pp. 101876, 2024.
- [11] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *Neural Comput.*, vol. 36, no. 3, pp. 512–534, 2024.
- [12] F. X. Diebold and R. S. Mariano, "Comparing predictive accuracy," *J. Bus. Econ. Statist.*, vol. 13, no. 3, pp. 253–263, 1995.
- [13] F. X. Diebold and K. Yilmaz, "Measuring financial asset return and volatility spillovers, with application to global equity markets," *Econ. J.*, vol. 119, no. 534, pp. 158–171, 2009.
- [14] F. X. Diebold and K. Yilmaz, "On the network topology of variance decompositions: Measuring the connectedness of financial firms," *J. Econometrics*, vol. 232, no. 1, pp. 1–20, 2022.
- [15] T. Fischer and C. Krauss, "Deep learning with long short-term memory networks for financial market predictions," *Eur. J. Oper. Res.*, vol. 307, no. 2, pp. 853–867, 2023.
- [16] K. J. Forbes and R. Rigobon, "No contagion, only interdependence: Measuring stock market comovements," *J. Finance*, vol. 77, no. 5, pp. 2223–2270, 2002.
- [17] D. Gabauer and R. Gupta, "International spillovers in global asset markets: New evidence from time-varying and frequency-domain approaches," *J. Int. Money Finance*, vol. 131, pp. 102790, 2024.
- [18] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
- [19] S. Hammoudeh, W. Mensi, and I. M. Al-Jarrah, "Dynamic oil-stock nexus in GCC countries: Evidence from nonlinear cointegration and causality tests," *Energy Econ.*, vol. 118, pp. 106498, 2024.
- [20] P. R. Hansen, A. Lunde, and J. M. Nason, "The model confidence set," *Econometrica*, vol. 79, no. 2, pp. 453–497, 2011.
- [21] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [22] International Monetary Fund, *Regional Economic Outlook: Middle East and Central Asia*. Washington, DC, USA, 2024.
- [23] A. Kanas and P. Molyneux, "Volatility spillovers across equity markets: European evidence revisited," *J. Int. Financial Markets, Inst. Money*, vol. 84, pp. 101742, 2023.
- [24] S. Kumar, A. Sharma, and M. Chen, "Hybrid VAR-LSTM models for financial spillover prediction," *Expert Syst. Appl.*, vol. 238, pp. 121856, 2025.
- [25] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [26] J. Liu and P. Wang, "Attention-enhanced LSTM for financial contagion detection," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 9, pp. 6721–6735, 2023.
- [27] W. Mensi, S. Hammoudeh, and S. H. Kang, "Risk spillovers and portfolio management between oil and GCC stock markets," *Resources Policy*, vol. 82, pp. 103457, 2024.
- [28] S. Neaime and I. Gaysset, "Financial inclusion and stability in MENA: Evidence from poverty and inequality," *Finance Res. Lett.*, vol. 51, pp. 103421, 2024.
- [29] M. H. Pesaran and Y. Shin, "Generalized impulse response analysis in linear multivariate models," *Econ. Lett.*, vol. 58, no. 1, pp. 17–29, 1998.
- [30] Saudi Vision 2030, "Vision 2030: Kingdom of Saudi Arabia," 2016. [Online]. Available: <https://www.vision2030.gov.sa/>
- [31] A. Sensoy, D. K. Nguyen, and E. Hacihasanoglu, "Dynamic convergence of commodity futures: Not all types of commodities are alike," *Resources Policy*, vol. 81, pp. 103341, 2023.
- [32] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 27, pp. 3104–3112, 2014.
- [33] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. Adv. Neural Inf.*

Process. Syst. (NeurIPS), vol. 30, pp. 5998–6008, 2017.

- [34] World Bank, *Gulf Economic Update: Sustaining Economic Momentum*. Washington, DC, USA, 2024.
- [35] L. Yarovaya, J. Brzeszczyński, and C. K. M. Lau, "Intra- and inter-regional return and volatility spillovers across emerging and developed markets: Evidence from stock indices and stock index futures," *Int. Rev. Financial Anal.*, vol. 85, pp. 102454, 2023.
- [36] M. Youssef, K. Mokni, and A. N. Ajmi, "Dynamic connectedness between stock markets in the presence of the COVID-19 pandemic: Does economic policy uncertainty matter?," *Financial Innov.*, vol. 7, no. 1, pp. 1–27, 2021.
- [37] Y. Zhang, X. Liu, and W. Wang, "Deep learning for financial time series: A comprehensive survey," *Inf. Fusion*, vol. 91, pp. 435–456, 2023.
- [38] X. Zhou, W. Zhang, and J. Zhang, "Volatility spillovers between the Chinese and world equity markets," *Pacific-Basin Finance J.*, vol. 77, pp. 101917, 2023.

AUTHOR'S BIOGRAPHY



Dr. Mohammed Alharbi holds a Doctor of Business Administration in Finance and Applied Economics. He is a faculty member with over 17 years of experience at the Institute of Public Administration. He is a Certified Financial Consultant and has served

as Director of both the Consulting Development and Coordination and the Human Resources and Financial Consulting departments. His consultancy experience spans high-level government advisory projects across multiple ministries and national development initiatives. His research focuses on oil price volatility, COVID-19 hedging strategies, machine learning in stock prediction, and financial distress modeling in emerging markets.

Citation of this Article:

Mohammed H. Alharbi. (2026). Deep Learning the Dynamics of Financial Contagion: LSTM Networks for Spillover Detection in Four GCC Stock Markets. *International Research Journal of Innovations in Engineering and Technology - IRJIET*, 10(1), 74-86. Article DOI <https://doi.org/10.47001/IRJIET/2026.101009>
