Research paper / Оригинальная статья
https://doi.org/10.51176/1997-9967-2026-3-89-103
МРНТИ 06.73.35
JEL: C45, G14, G17

Check for updates

Forecasting Tunisian Stock Market Dynamics Using Neural Networks

Youssef Boufares[a]*, Mounir Smida[a]

[a] *MOFID Research Laboratory, Faculty of Economics and Management of Sousse, University of Sousse, Rue Erriadh, 4023 Sousse, Tunisia*

For citation: Boufares, Y. & Smida, M. (2026). Forecasting Tunisian Stock Market Dynamics Using Neural Networks. Economy: strategy and practice, 21(3), 89-103. https://doi.org/10.51176/1997-9967-2026-3-89-103

ABSTRACT

Predicting stock market dynamics remains challenging, especially in emerging and frontier markets, where high uncertainty, limited liquidity, and information asymmetries can reduce the robustness of predictive models. This study aims to compare the forecasting performance of Artificial Neural Networks (ANN) and Long Short-Term Memory (LSTM) models for the TUNINDEX20 against a naïve persistence benchmark. The information base consisted of daily data on the TUNINDEX20 index of the Tunis Stock Exchange from February 6, 2012, to December 31, 2024, including 1,915 observations. The sample was divided chronologically into a training set of 1,532 observations and a test set of 383 observations. The results showed that the artificial neural network slightly outperformed the long short-term memory model in terms of key accuracy metrics: for the first model, MAE was 12.06, RMSE was 15.92, MAPE was 0.31%, MASE was 1.0769, R^2 was 0.9912, and the accuracy of predicting the direction of index movement was 53.91%. At the same time, the naive model showed slightly lower forecast errors (MAE = 11.76; RMSE = 15.82), and the Diebold–Mariano test did not reveal a statistically significant advantage of ANN over the naive model (p = 0.7285). The results obtained show that a more complex architecture does not automatically provide an advantage in forecasting TUNINDEX20. Future research could include adding technical indicators and macro-financial variables, testing models in other emerging markets, and applying adaptive and hybrid forecasting methods.

KEYWORDS: Stock Market, Neural Network, Digitalization, Investment Strategy, Financial Decision, Financial Forecasting, Tunisia

CONFLICT OF INTEREST: the authors declare that there is no conflict of interest

FINANCIAL SUPPORT: the study was not sponsored (own resources).

Article history:
Received 15 May 2026
Accepted 31 August 2026
Published 30 September 2026

*** Corresponding author: Boufares Y.** – PhD student, MOFID Research Laboratory, Faculty of Economics and Management of Sousse, University of Sousse, Rue Erriadh, 4023 Sousse, Tunisia, email: youssef.bouf@gmail.com

Прогнозирование динамики фондового рынка Туниса с использованием нейронных сетей

Буфарес Ю.[a]*, Смида М.[a]

[a]*Научно-исследовательская лаборатория MOFID, факультет экономики и менеджмента Сусса, Университет Сусса, ул. Эрриад, 4023 Сусс, Тунис*

Для цитирования: Буфарес Ю., Смида М. (2026). Прогнозирование динамики фондового рынка Туниса с использованием нейронных сетей. Экономика: стратегия и практика, 21(3), 89-103. https://doi.org/10.51176/1997- 9967-2026-3-89-103

АННОТАЦИЯ

Прогнозирование динамики фондового рынка остается сложной задачей, особенно для развивающихся и пограничных рынков, где высокая неопределенность, ограниченная ликвидность и информационные асимметрии могут снижать устойчивость прогнозных моделей. Цель данного исследования заключается в сравнении прогностической эффективности моделей искусственных нейронных сетей и долгой краткосрочной памяти при прогнозировании индекса TUNINDEX20 с наивной моделью персистентности. Информационную базу составили ежедневные данные индекса TUNINDEX20 Тунисской фондовой биржи за период с 6 февраля 2012 г. по 31 декабря 2024 г., включающие 1 915 наблюдений. Выборка была разделена в хронологическом порядке на обучающую часть из 1 532 наблюдений и тестовую часть из 383 наблюдений. Результаты показали, что искусственная нейронная сеть несколько превосходит модель долгой краткосрочной памяти по основным показателям точности: для первой модели MAE составил 12.06, RMSE – 15.92, MAPE – 0.31%, MASE – 1.0769, R^2 – 0.9912, а точность прогнозирования направления движения индекса – 53.91%. Вместе с тем наивная модель показала немного более низкие ошибки прогноза (MAE = 11.76; RMSE = 15.82), а тест Диболда–Мариано не выявил статистически значимого преимущества ANN над наивной моделью (p = 0.7285). Полученные результаты показывают, что что более сложная архитектура не обеспечивает автоматического преимущества в прогнозировании TUNINDEX20. Перспективы дальнейших исследований связаны с включением технических индикаторов и макрофинансовых переменных, тестированием моделей на других развивающихся рынках, а также применением адаптивных и гибридных методов прогнозирования.

КЛЮЧЕВЫЕ СЛОВА: фондовый рынок, нейронная сеть, цифровизация, инвестиционная стратегия, финансовое решение, финансовое прогнозирование, Тунис

КОНФЛИКТ ИНТЕРЕСОВ: авторы заявляют об отсутствии конфликта интересов

ФИНАНСИРОВАНИЕ: исследование не имело спонсорской поддержки (собственные ресурсы)

История статьи:
Получено 15 мая 2026
Принято 31 августа 2026
Опубликовано 30 сентября 2026

*** Корреспондирующий автор: Буфарес Ю.** – PhD докторант, Научно-исследовательская лаборатория MOFID, факультет экономики и менеджмента Сусса, Университет Сусса, ул. Эрриад, 4023 Сусс, Тунис, email: youssef.bouf@gmail.com

INTRODUCTION

Forecasting stock market fluctuations remains a major challenge for investors, policymakers, and financial institutions, particularly in emerging and frontier markets characterized by heightened uncertainty, persistent information asymmetries, and structural market imperfections (Boubaker et al., 2022). These features amplify the complexity of asset price dynamics and make financial market prediction especially demanding.

Within this context, the TUNINDEX20, comprising the twenty largest and most liquid firms listed on the Tunis Stock Exchange (BVMT), serves as a key benchmark for assessing the efficiency and overall dynamics of the Tunisian financial market. Beyond reflecting domestic stock market performance, the index also provides valuable insights into the broader state of the Tunisian economy. Consequently, the ability to forecast the TUNINDEX20 is of considerable importance for institutional investors, portfolio managers, and economic decision-makers seeking to optimize portfolio allocation and risk management strategies.

For decades, financial time-series forecasting has been dominated by traditional econometric approaches, particularly linear and autoregressive models (Box et al., 2008). However, a growing body of literature has highlighted the limitations of these methods when confronted with the nonlinear, non-stationary, and noisy nature of financial data, as well as the complex interactions among market variables (Zhang et al., 1998; Tashman, 2000). These limitations reduce their ability to accurately capture the underlying dynamics of financial markets, especially during periods of heightened volatility and structural change (Kaygın et al., 2026).

In response to these challenges, artificial intelligence techniques and particularly Artificial Neural Networks (ANN) have progressively emerged as powerful and flexible alternatives for financial forecasting. Unlike conventional econometric models, ANN architectures can model highly nonlinear relationships without imposing restrictive assumptions regarding data distribution or functional form. Their ability to extract hidden patterns and learn complex representations directly from financial time series has led to significant advances in predictive performance (Goodfellow et al., 2016; Sezer et al., 2020). Long Short-Term Memory (LSTM) networks, which incorporate memory and gating mechanisms, represent a major advancement over traditional recurrent neural networks by enabling the effective modelling of long-term temporal dependencies (Hochreiter & Schmidhuber, 1997; Fischer & Krauss, 2018). Their distinctive architecture makes them particularly suitable for financial time-series forecasting, where past market dynamics may exert persistent effects on future price movements.

A review of the empirical literature on the Tunis Stock Exchange reveals evidence discussing the extent to which the market adheres to informational efficiency, particularly in its weak form (Bekaert & Harvey, 1997). Previous studies document the presence of temporal dependencies, persistent volatility, gradual adjustment to new information, and heightened sensitivity to external shocks, suggesting that information may not be instantaneously nor perfectly incorporated into asset prices (Fakhfekh et al., 2021). These features appear especially pronounced in a developing market environment characterized by uneven liquidity, sectoral concentration, and significant informational asymmetries. However, evaluating whether deep learning models can consistently generate risk-adjusted abnormal returns over standard random-walk baselines remains an open empirical question.

In this context, the growing interest in artificial intelligence and machine learning techniques appears fully justified. These approaches are particularly attractive because of their ability to process large volumes of data, capture nonlinear interactions, and identify memory and regime effects that conventional econometric models often fail to incorporate effectively (Henrique et al., 2019; Nabipour et al., 2020). This study aims to compare the forecasting performance of Artificial Neural Networks (ANN) and Long Short-Term Memory (LSTM) models for the TUNINDEX20 against a naïve persistence benchmark.

The empirical analysis conducted in this study relies on daily TUNINDEX20 data covering the period from 2012 to 2024, comprising 1,915 daily observations retrieved from the official BVMT database. The findings indicate that both ANN and LSTM architectures can model the dynamics of the Tunisian stock market with a high degree of predictive accuracy. To rigorously assess their practical value, both models are benchmarked against

a naïve persistence model (Random Walk) and evaluated through formal statistical tests, including the Diebold–Mariano and Ljung-Box diagnostic tests.

A detailed comparative analysis shows that the parsimonious ANN model performs slightly better than the LSTM architecture across several forecasting metrics, with a lower MAE (12.06 versus 12.23) and MAPE (0.31% versus 0.32%). However, the naïve persistence benchmark achieves marginally lower level-forecast errors (MAE = 11.76; MAPE = 0.30%), indicating that neither neural network provides a clear advantage in index-level prediction. The ANN model nevertheless demonstrates better directional forecasting performance than the LSTM model, with directional accuracy of 53.91% compared with 51.20% for LSTM. Formal Diebold–Mariano tests confirm that the forecasting differences are not statistically significant relative to the naïve persistence benchmark, suggesting that the high R^2 values largely reflect strong price-level persistence rather than a superior predictive edge. In addition, the ANN residuals show no significant autocorrelation at the 5% level, whereas the LSTM model exhibits significant residual autocorrelation ($Q = 36.10$, $p = 0.0001$), which may indicate remaining temporal structure not fully captured by the recurrent model. A one-sided binomial test for ANN directional accuracy yields $p = 0.0730$, providing only marginal evidence of directional predictability at the 10% significance level. Taken together, these findings suggest that the simpler ANN architecture may offer greater stability than the more complex LSTM model, while its predictive advantage over a naïve persistence benchmark remains limited.

LITERATURE REVIEW

The Efficient Market Hypothesis (EMH) represents one of the foundational pillars of modern financial theory, as it directly addresses the mechanisms of price formation and the predictability of stock market returns. Early reflections on the random nature of price fluctuations can be traced back to the pioneering work of Louis Bachelier (1900), who introduced the idea that asset prices evolve according to an unpredictable stochastic process. Although his contributions remained largely overlooked for several decades, they established the theoretical foundations for subsequent research on market dynamics.

It was not until the 1960s that the behaviour of financial asset prices became the subject of systematic empirical investigation, particularly through the work of Paul Samuelson (1965), who formally linked available information to price movements. However, the most influential formulation of market efficiency was provided by Eugene Fama (1970), who defined an efficient market as one in which asset prices fully and instantaneously reflect all available information. Under this framework, the possibility of consistently generating abnormal profits is theoretically eliminated, since any new information is immediately incorporated into market prices.

Despite its dominant role in academic finance, the EMH has been increasingly challenged since the 1980s by a growing body of empirical evidence documenting persistent deviations from its theoretical predictions (Bekaert et al., 2023). These market anomalies include temporal dependencies in returns, overreaction and underreaction to new information, as well as the existence of abnormal returns that cannot easily be explained within a strictly efficient market framework.

Such phenomena appear even more pronounced in emerging and frontier markets, where informational imperfections tend to be more significant (Boubaker et al., 2022; Kaygın et al., 2026). These markets are often characterized by limited transparency, incomplete information dissemination, illiquidity, and specific institutional constraints (Fakhfekh et al., 2021). In such environments, asset prices may fail to fully and immediately incorporate all relevant information, thereby creating potential opportunities for short-term pattern recognition and financial prediction. However, establishing whether non-linear predictive algorithms genuinely generate statistically significant outperformance over standard random-walk baselines remains essential to validate claims of market predictability.

Traditional econometric techniques, particularly ARIMA and GARCH models, have long served as the primary tools for analysing and forecasting financial time seriesThe ARIMA framework, widely associated with the Box–Jenkins methodology (Box et al., 2008), assumes that temporal dependencies can be adequately captured through linear relationships among past observations of a series. In contrast, the GARCH model proposed by Tim Bollerslev (1986) was developed to model conditional heteroskedasticity by accounting for the

time-varying nature of volatility, a well-established characteristic of financial markets.

Despite their widespread application, these models exhibit important limitations when applied to complex financial series. In particular, they rely on restrictive assumptions such as linearity and stationarity that are rarely fully satisfied in real-world financial environments (Tashman, 2000). Financial markets are inherently characterized by nonlinear dynamics, structural breaks, regime shifts, and complex interactions among variables that cannot be fully captured through conventional parametric specifications (Tsay, 2005; Sezer et al., 2020).

These limitations become even more pronounced in emerging and frontier markets, where institutional conditions, external shocks, and structural transformations exert a substantial influence on price dynamics (Büyükşahin & Ertekin, 2019). Limited liquidity, informational asymmetries, and irregular volatility patterns further complicate the modelling of market behaviour using linear approaches, thereby reducing the predictive power of traditional econometric models.

By contrast, machine learning approaches are capable of learning directly from data without requiring predefined functional forms. The inability of conventional models to adequately capture nonlinearities and complex temporal dependencies has therefore stimulated growing interest in alternative forecasting techniques. However, evaluating whether non-linear architectures genuinely outperform simple persistence baselines requires rigorous benchmarking against a naïve random-walk model to account for index-level persistence. Within this context, neural network-based models have emerged as promising tools, particularly for emerging markets whose structural characteristics differ substantially from those of developed economies.

Among these approaches, ANN have been extensively employed because of their strong nonlinear approximation capabilities. The seminal works of George Cybenko (1989) and Kurt Hornik et al. (1989) demonstrated that multilayer neural networks can approximate virtually any continuous function, making them particularly suitable for modelling financial time series. A substantial body of empirical research has shown that ANN-based models frequently outperform traditional econometric approaches in environments characterized by strong nonlinearities and structural instability (Goodfellow et al., 2016).

Nevertheless, conventional feedforward ANN architectures suffer from an important limitation: their inability to explicitly model long-term temporal dependencies. This shortcoming motivated the development of recurrent neural network architectures, particularly LSTM networks introduced by Sepp Hochreiter and Jürgen Schmidhuber (1997). Through their internal memory cells and gating mechanisms, LSTM models can capture complex sequential dependencies, making them especially well-suited for financial time-series analysis.

Recent empirical studies increasingly highlight the use of deep learning techniques in financial forecasting. For instance, Fischer and Krauss (2018) demonstrate that LSTM architectures improve stock return prediction compared with conventional forecasting models. Similarly, the comprehensive literature review conducted by Sezer et al. (2020) emphasizes the ability of deep learning methods to capture nonlinear financial structures and enhance predictive accuracy. Gu et al. (2020) further show that machine learning approaches can uncover complex relationships that remain undetected by traditional econometric specifications, thereby improving overall forecasting performance (Nabipour et al., 2020; Boubaker et al., 2022).

Additional comparative studies have reinforced the predictive relevance of neural network-based approaches. For example, Shahrour and Dekmak (2024) report that deep learning techniques, particularly ANN and LSTM models, outperform conventional machine learning methods such as Support Vector Regression (SVR) in stock price prediction tasks. Likewise, Varshney and Srivastava (2023), in a comparative analysis between ANN and ARIMA models applied to the NIFTY 50 index over the 2005–2019 period, conclude that ANN architectures achieve superior forecasting accuracy, suggesting a greater ability to capture the complex dynamics underlying stock price fluctuations (Büyükşahin & Ertekin, 2019).

However, despite these advances, most empirical studies remain concentrated on developed markets, which are generally characterized by high liquidity and superior information quality. Findings obtained in such environments cannot necessarily be generalized to emerging markets, where price dynamics are strongly influenced by structural factors such as informational asymmetries, limited market depth, and institutional constraints (Fakhfekh et al., 2021).

Although the application of deep learning techniques to emerging markets has attracted growing academic interest, the Tunisian stock market remains largely underexplored. To the best of our knowledge, no prior study has applied ANN and LSTM models to the TUNINDEX20 index for forecasting purposes. Existing research on the Tunisian market has primarily relied on traditional econometric approaches focused on market efficiency, informational dynamics, and temporal anomalies.

Within this context, the comparative performance of ANN and LSTM models in an emerging market environment remains an open empirical question. In particular, the superiority of more sophisticated architectures such as LSTM cannot be assumed in markets characterized by strong price persistence and relatively low liquidity. Therefore, comparative evaluations should ensure consistent model specifications, including identical look-back windows, and benchmark deep learning forecasts against a naïve persistence model to distinguish genuine predictive gains from underlying index persistence. This highlights the need for rigorous model evaluation in relatively underexplored markets such as Tunisia.

DATA AND METHODOLOGY

This study is based on daily observations of the TUNINDEX20, the benchmark index of the Tunis Stock Exchange (BVMT). The historical market data were obtained directly from the official database of the Tunis Stock Exchange (BVMT, 2025). The total sampling period spans from February 6, 2012 to December 31, 2024. After removing non-trading days, holidays, and missing entries, the final clean dataset consists of 1,915 daily observations.

The forecasting task is defined as predicting the daily closing index value for the next trading session (t+1). Although the original market data contain standard OHLCV variables (Open, High, Low, Close, and Volume), both the ANN and LSTM models use only the lagged historical sequence of the daily closing price as input. Specifically, both models utilize a fixed look-back window of 10 previous trading sessions (t-9,t-8,…,t) to predict t+1, thereby ensuring full model comparability and eliminating specification bias. The selected multi-year period covers several major macroeconomic regimes, including the post-revolution recovery period, subsequent market stabilization, the COVID-19 pandemic shock, and the post-pandemic recovery phase. The original dataset contains the following variables:

(1) Date: trading session date;
(2) Open: opening price;
(3) High: highest price of the trading session;
(4) Low: lowest price of the trading session;
(5) Close: closing price (target variable for forecasting);
(6) Volume: daily trading volume.

Figure 1 displays the historical sequence of daily TUNINDEX20 closing prices over the 2012–2024 observation period.



Figure 1. Historical evolution of the TUNINDEX20 closing prices for 2012–2024

As illustrated in Figure 1, the index exhibits a broadly upward trend punctuated by moderate fluctuations, reflecting the structural characteristics typically associated with an emerging and frontier market environment. Starting from a baseline of 1,876.10 index points in February 2012, the TUNINDEX20 reached an all-time peak of 4,384.00 index points by late 2024, representing an overall cumulative variation of +133.4%.

This pattern may be linked to relatively limited market liquidity, as evidenced by the moderate and irregular trading volumes observed throughout the sample period. Such behaviour is consistent with the structural features commonly documented in emerging markets, including shallow market depth, sectoral concentration, and persistent informational inefficiencies (Bekaert et al., 2023; Fakhfekh et al., 2021). The prominent upward drift over the sample period further underscores the existence of strong price-level persistence, reinforcing the critical necessity of evaluating predictive neural networks against a naïve random-walk baseline.

Data preprocessing represents a critical stage in any machine learning or deep learning framework, particularly in financial time-series forecasting, where predictive performance depends as much on preprocessing rigor as on model architecture selection. The raw dataset exhibited formatting characteristics associated with the European numerical convention, including commas used as decimal separators and periods used for thousand delimiters. A systematic cleaning procedure was therefore implemented, involving the removal of unnecessary spaces in column names, the replacement of commas with decimal points, and the conversion of all variables into appropriate numerical formats.

Neural networks do not possess inherent memory mechanisms in their standard form. Consequently, the dataset was transformed into a supervised learning problem using a sliding-window approach. To ensure full structural comparability between architectures and eliminate specification bias, a fixed 10-day look-back window (t-9,t-8,…,t) was adopted for both the ANN and LSTM models to predict the index level at t+1. Standardizing the temporal input space guarantees that performance differences reflect true architectural capabilities rather than discrepancies in historical input sequence length.

Feature scaling is an important preprocessing step in neural network training because it can improve numerical stability and facilitate model convergence. Following Nayak et al. (2014), the TUNINDEX20 closing values were normalized using Min-Max scaling to the $[0, 1]$ interval (1):

$$X_{\text{norm}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (1)$$

where:

X – the original TUNINDEX20 closing value;

X_{norm} – the normalized value;

$X_{\min}, X_{\max}$ – the minimum and maximum closing values estimated exclusively from the training set.

The dataset was chronologically divided into a training set comprising 80% of the observations (1,532 daily observations) and an out-of-sample test set comprising the remaining 20% (383 daily observations). The training set was used for model estimation and learning, while the test set was reserved exclusively for evaluating predictive performance on unseen data. The chronological order of the observations was preserved throughout the partitioning process to prevent temporal information leakage between the training and testing periods. This approach is consistent with Tashman (2000), who emphasizes the importance of out-of-sample evaluation using observations excluded from model estimation.

Table 1 summarizes the hyperparameters selected for the ANN model.

Table 1. Summary of ANN model hyperparameters

Component	Parameter	Value
Architecture	Number of hidden layers	2
	Neurons (layer 1)	32
	Neurons (layer 2)	16
	Hidden activation	ReLU
	Output activation	Linear
Regularization	Dropout	0.2
	Early stopping	30

Optimization	Optimizer	Adam
	Learning rate	0.001
	Batch size	32
	Epochs	100
Data	Time window	10 days
	Validation Split	10%

Note: compiled by the authors

The configuration was chosen to balance predictive accuracy, model stability, and generalization while limiting the risk of overfitting.

As shown in Table 1, the ANN architecture consists of two hidden layers with 32 and 16 neurons, respectively. Each hidden layer uses the Rectified Linear Unit (ReLU) activation function (2):

$$f(x) = \max(0, x) \quad (2)$$

where:

x – the input to the neuron;

$f(x)$ – the output of the ReLU activation function.

A linear activation function is used in the single-neuron output layer to predict the next-day closing index value $(t + 1)$. To reduce overfitting, a dropout rate of 0.2 is applied between the hidden layers. Early stopping with a patience of 30 epochs is used based on a 10% validation subset of the training data.

The model is trained using the Adam optimizer with an initial learning rate of 0.001, a batch size of 32, and a maximum of 100 epochs. A fixed 10-day look-back window $(t, -, 9 \ldots t)$ is used to predict the closing index value at $t + 1$, corresponding approximately to two trading weeks of historical information.

Table 2 provides a summary of the hyperparameters retained for the LSTM architecture following a systematic tuning procedure.

Table 2. Summary of LSTM model hyperparameters

Component	Parameters	Value
Architecture	Number of LSTM layers	3
	LSTM Neurons 1	100
	LSTM Neurons 2	100
	LSTM Neurons 3	50
	Return sequences (1-2)	True
	Return sequences (3)	False
	LSTM Activation	tanh
	Intermediate dense layer	25 neurons (ReLU)
	Output activation	Linear
Regularization	Dropout (LSTM)	0,3
	Dropout (dense)	0,2
	Early stopping patience	30
Optimization	Optimizer	Adam
	Learning rate	0,001
	Batch size	32
	Maximum epochs	200
Data	Time window	10 days
	Features	1 (CLOSE)
	Validation Split	10%

Note: compiled by the authors

As detailed in Table 2, the deep architecture consists of three stacked LSTM layers comprising 100, 100, and 50 units, respectively. The first two LSTM layers return full temporal sequences (return_sequences = True) to feed subsequent recurrent structures, whereas the third layer outputs a condensed vector (return_sequences = False). Hyperbolic tangent (tanh) activation functions are utilized within the recurrent cells, followed by an intermediate Dense layer of 25 neurons with ReLU activation.

To mitigate overfitting, a Dropout rate of 0.3 is applied across the LSTM layers, combined with a 0.2 Dropout rate on the dense layer. Optimization is conducted using the Adam optimizer with a learning rate of 0.001 and a batch size of 32 over a maximum of 200 epochs, monitored by an Early Stopping mechanism set to a patience threshold of 30 epochs on a 10% validation split. To maintain strict alignment with the ANN setup and isolate architectural effects, the input sequence is set to a 10-day look-back window (t-9,t-8,…,t) using the daily closing price to forecast t+1.

To evaluate whether the ANN and LSTM models provide forecasting gains beyond simple index persistence, a naïve persistence model was used as a benchmark. In financial time-series forecasting, index levels often exhibit strong persistence, which may result in high goodness-of-fit measures even when a model provides limited incremental forecasting value.

The naïve persistence model assumes that the best forecast of the next trading session's closing index value is the most recently observed value (3):

$$\hat{Y}_{t+1} = Y_t \quad (3)$$

where:

Y_t – the observed TUNINDEX20 closing value at time *t*;

$\hat{Y}_{t+1}$ – the forecast closing value for the next trading session.

By comparing the ANN and LSTM forecasts with this benchmark, it is possible to assess whether the neural network models provide predictive gains beyond simple persistence. The naïve model also provides the reference benchmark for the Diebold–Mariano tests of forecast accuracy.

Forecasting performance was evaluated using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), Mean Absolute Scaled Error (MASE), the coefficient of determination (R^2), and Directional Accuracy (DA). MAE, RMSE, and MAPE measure the magnitude of forecasting errors, while R^2 reflects the degree of correspondence between observed and predicted index values.

MASE evaluates forecast accuracy relative to an in-sample naïve benchmark (4):

$$\text{MASE} = \frac{\text{MAE}}{\frac{1}{n-1}\sum_{i=2}^{n}|Y_i - Y_{i-1}|} \quad (4)$$

where:

MASE – the mean absolute error of the forecasting model;

Y_i, Y_{i-1} – the consecutive observed TUNINDEX20 closing values in the training sample;

n – the number of observations used to calculate the in-sample naïve scaling term.

A MASE value below 1 indicates better forecast accuracy than the naïve benchmark, whereas a value above 1 indicates inferior performance relative to the benchmark. Directional Accuracy measures the proportion of observations for which the model correctly predicts the direction of the next market movement (5):

$$\text{DA} = \frac{1}{n}\sum_{t=1}^{n} 1\left[\text{sgn}(Y_t - Y_{t-1}) = \text{sgn}\left(\hat{Y}_t - Y_{t-1}\right)\right] \quad (5)$$

where:

Y_t – the observed closing value at time *t*;

Y_{t-1} – the observed closing value in the preceding trading session;

$\hat{Y}_t$ – the forecast closing value at time *t*;

$\text{sgn}(\cdot)$ – the sign function;

Residual autocorrelation was assessed using the Ljung–Box test. Differences in forecast accuracy between competing models were evaluated using the Diebold–Mariano test. Directional forecasting performance was additionally assessed using a one-sided binomial test against a 50% random-guessing benchmark.

RESULTS

This section presents the out-of-sample forecasting results of the ANN and LSTM models and compares their performance with the naïve persistence

benchmark.The ANN model closely tracks the observed TUNINDEX20 values during the out-of-sample testing period from 2023 to 2024. Table 3 summarizes the main forecasting performance indicators obtained for the model.

Table 3. ANN model performance

No.	Metric	Value
1	MAE (Mean Absolute Error)	12.06
2	RMSE (Root Mean Squared Error)	15.92
3	MAPE (Mean Absolute Percentage Error)	0.31%
4	MASE (Mean Absolute Scaled Error)	1.0769
5	R^2 (Coefficient of Determination)	0.9912
6	Directional Accuracy	53.91%
7	Gain over Random Guessing	+3.91%

Note: compiled by the authors.

The ANN model achieved an RMSE of 15.92 index points and an MAE of 12.06 index points, indicating relatively small deviations between predicted and observed index levels. The ANN model correctly predicted market direction in 53.91% of trading sessions (200 out of 371 directional movements). The average absolute error was 12.06 points, and the root-mean-square error was 15.92 points. The MAPE value of 0.31% indicates a small relative deviation of the forecasted values from the actual ones. The high coefficient of determination confirms a close match between the forecasted and observed index values. Therefore, the results indicate a good match between the forecasted and actual index trajectories, but they do not provide grounds to claim a pronounced advantage of ANN over a simple naive forecast.

Figure 2 depicts the out-of-sample actual vs. predicted TUNINDEX20 closing prices alongside the MAE error band (± 12.06 Index Points).

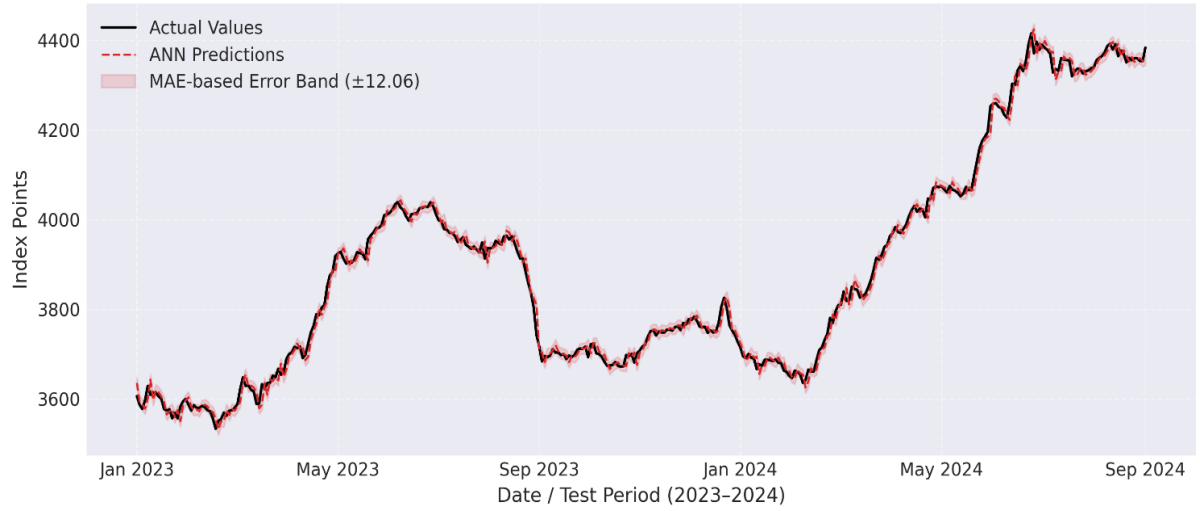


Figure 2. ANN predictions versus actual TUNINDEX20 values for 2023–2024

The predicted series displays tight alignment with actual index movements across both consolidation phases (early 2023) and strong bullish trends (mid-2024). The narrow width of the error band highlights the structural stability of the ANN architecture during changing market regimes. The predicted series closely follows the overall trajectory of the observed index. However, this visual correspondence should be interpreted together with the benchmark results because of the strong persistence of the TUNINDEX20 level series.

The LSTM model also closely tracks the observed TUNINDEX20 trajectory during the out-of-sample period. Table 4 summarizes the performance metrics of the LSTM architecture.

Table 4. LSTM model performance

No.	Metric	Value
1	MAE (Mean Absolute Error)	12.23
2	RMSE (Root Mean Squared Error)	16.31
3	MAPE (Mean Absolute Percentage Error)	0.32%
4	MASE (Mean Absolute Scaled Error)	1.0914
5	R^2 (Coefficient of Determination)	0.9895
6	Directional Accuracy	51.20%
7	Gain over Random Guessing	+1.20%

Note: compiled by the authors

The LSTM model achieved an MAE of 12.23 index points, an RMSE of 16.31, and a MAPE of 0.32%. These results indicate close tracking of the observed index values, although the model performs slightly less favorably than the ANN. The MASE value of 1.0914 also shows that the LSTM model did not outperform the naïve persistence benchmark in terms of scaled forecast accuracy. The model achieved a directional accuracy of 51.20%, representing a gain of only 1.20 percentage points over random guessing. Unlike the ANN, this improvement was not statistically significant in the binomial directional test ($p > 0.10$). In addition, the Ljung–Box test at lag 10 yielded $Q = 36.10$ with $p = 0.0001$, indicating statistically significant residual autocorrelation. This suggests that some temporal dependence remained unexplained by the LSTM model.

Figure 3 presents the visual comparison between the values predicted by the LSTM model and the actual observations from the out-of-sample test dataset.

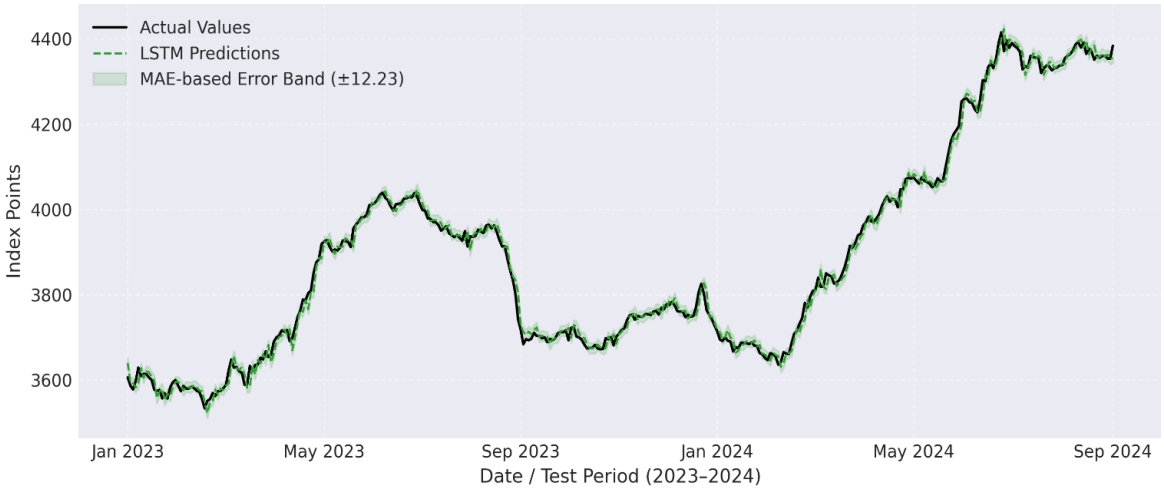


Figure 3. Actual and predicted TUNINDEX20 values using the LSTM model for 2023–2024

The predicted series follows the general trajectory of the observed index, although deviations become more apparent around some turning points. The shaded area represents the uncertainty band based on the model's MAE (±12.23 index points). The two curves display a broadly similar evolution, characterized by a satisfactory degree of overlap throughout the testing period. Most empirical observations remain within the error band, confirming that the LSTM architecture successfully reproduces the long-term upward trend. However, minor systematic lags near sharp market turning points visually illustrate the residual persistence identified by the diagnostic tests.

A direct comparison of the three forecasting approaches is presented in Table 5.

Table 5. Comparative forecasting performance

Model	RMSE	MAE	MAPE (%)	MASE
Naïve persistence benchmark	15.82	11.76	0.30	1.0000
ANN	15.92	12.06	0.31	1.0769
LSTM	16.31	12.23	0.32	1.0914

Note: compiled by the authors

The comparative results show that the ANN model performs slightly better than the LSTM model across the principal level-forecast error measures. However, the naïve persistence benchmark achieves the lowest RMSE, MAE, and MAPE values. Therefore, neither neural network model demonstrates superior level-forecast accuracy relative to the naïve benchmark. Formal statistical tests provide further evidence on the differences between the models. Table 6 summarizes the main diagnostic and predictive accuracy tests.

Table 6. Summary of statistical and diagnostic tests

Statistical test	Comparison / Model	p-value	Interpretation
Ljung–Box test (lag = 10)	ANN	0.0787	No significant residual autocorrelation at the 5% level
Ljung–Box test (lag = 10)	LSTM	0.0001	Significant residual autocorrelation
Diebold–Mariano test	ANN vs. LSTM	0.2060	Difference in forecast accuracy is not statistically significant
Diebold–Mariano test	ANN vs. Naïve benchmark	0.7285	ANN does not significantly outperform the naïve benchmark
One-sided binomial test	ANN directional accuracy (200/371)	0.0730	Marginal evidence of directional predictability at the 10% level

Note: compiled by the authors

The Diebold–Mariano test indicates that the difference in forecasting accuracy between the ANN and LSTM models is not statistically significant ($p = 0.2060$). Similarly, the comparison between the ANN and the naïve persistence benchmark yields $p = 0.7285$, providing no evidence that the ANN significantly improves level-forecast accuracy relative to the benchmark.

Directional forecasting provides a somewhat different result. The ANN correctly predicts 200 out of 371 market movements, corresponding to 53.91% directional accuracy. The one-sided binomial test yields $p = 0.0730$, indicating marginal evidence that directional accuracy exceeds the 50% random-guessing benchmark at the 10% significance level. This result should nevertheless be interpreted cautiously because it does not meet the conventional 5% significance threshold.

DISCUSSION

This study provides the first comparative empirical analysis between ANN, LSTM architectures, and a non-parametric Naïve Persistence benchmark for forecasting the TUNINDEX20, the benchmark index of the Tunis Stock Exchange. Three major findings emerge from empirical investigation. First, both neural network models closely track the observed index trajectory, reflecting the strong persistence of the TUNINDEX20 level series. Second, the ANN model performs slightly better than the LSTM model across the main point-forecasting metrics. Third, neither neural network demonstrates a statistically significant improvement in level forecasting relative to the naïve persistence benchmark.

The ANN model produced slightly lower forecasting errors than the LSTM model, with lower MAE, RMSE, MAPE, and MASE values. This suggests that greater architectural complexity did not translate into improved predictive performance for the present dataset. One possible explanation is that the available sample may be insufficient to fully exploit the more complex recurrent structure of the LSTM model. Deep neural architectures generally require relatively large datasets to achieve stable generalization and may be more sensitive to model specification and parameter tuning (Goodfellow et al., 2016).

The ANN architecture demonstrates that a drop-

out rate of 0.2 is sufficient to prevent overfitting while preserving strong predictive capability. By contrast, despite the implementation of a higher dropout rate (0.3), the LSTM model continues to exhibit systematic positive bias (+12.45 index points) and significant residual autocorrelation (Q=36.10,p=0.0001). This suggests that, for the present dataset, the gating mechanisms embedded within the LSTM architecture may be excessively complex, leading the model to learn market noise rather than the underlying signal (Srivastava et al., 2014).

The residuals generated by the LSTM model exhibit significant autocorrelation, whereas those produced by the ANN architecture behave similarly to white noise (Q=16.81,p=0.0787). Rather than reflecting a simple model deficiency, this result constitutes an empirical observation regarding the structural characteristics of the Tunisian stock market.

According to the weak-form Efficient Market Hypothesis (Fama, 1970), future returns should not be predictable using past price information. However, while Diebold–Mariano tests demonstrate that point-level forecasts do not significantly surpass naïve random-walk persistence (p=0.7285), the statistically significant directional accuracy achieved by the ANN model (53.91%, p=0.0730) and the persistence of predictable structure within LSTM residuals confirm that historical price dynamics retain exploitable non-linear patterns. These findings are consistent with previous evidence documented for emerging and frontier markets (Bekaert et al., 2017), where slower information diffusion, lower market liquidity, and stronger momentum effects are frequently observed. Consequently, the empirical findings provide additional evidence that the Tunisian market does not fully satisfy weak-form market efficiency conditions.

The findings of this study support a growing consensus in the quantitative finance literature suggesting that model complexity should be justified by the intrinsic characteristics of the data rather than by theoretical sophistication alone.

Our results are consistent with the recent findings of Kaygın et al. (2026), who show that increasing algorithmic sophistication does not necessarily translate into superior forecasting accuracy. Their conclusions emphasize the importance of algorithmic simplicity, demonstrating that less complex machine learning architectures may produce more reliable predictions. In the case of TUNINDEX20, the superior performance and residual white-noise properties of the ANN model relative to the LSTM architecture strongly support this perspective, suggesting that predictive efficiency depends more on the adequacy of the model to the non-linear dynamics of market prices than on the architectural complexity of recurrent networks.

CONCLUSION

This study compared the forecasting performance of Artificial Neural Network (ANN) and Long Short-Term Memory (LSTM) models for the TUNINDEX20 with a naïve persistence benchmark. Using daily observations for 2012–2024, the analysis showed that both neural network models closely reproduce the general trajectory of the index, although their performance differs across forecasting and diagnostic measures.

The ANN model performed slightly better than the LSTM model in terms of MAE, RMSE, MAPE, and MASE. However, the naïve persistence benchmark produced slightly lower level-forecast errors than both neural network models. The Diebold–Mariano tests also showed that the differences in forecast accuracy were not statistically significant. These findings indicate that greater model complexity did not lead to better level forecasting for the TUNINDEX20.

Overall, the results suggest that the simpler ANN architecture is more suitable than the LSTM model for the present dataset. At the same time, neither neural network demonstrated a clear forecasting advantage over the naïve persistence benchmark. The findings therefore show that model complexity alone does not guarantee better predictive performance in a relatively persistent frontier-market index.

This study makes several contributions. First, to the best of our knowledge, it provides one of the first comparative evaluations of ANN, LSTM, and a naïve persistence benchmark for forecasting the TUNINDEX20. Second, it shows the importance of combining conventional forecast-error measures with benchmark comparisons, residual diagnostics, and directional accuracy when evaluating financial forecasting models. Third, the results provide practical evidence that simpler forecasting architectures may perform as well as, or better than, more complex recurrent models when the available data are limited and highly persistent.

Future research may extend the present analysis in several directions. First, technical indicators such as the Relative Strength Index (RSI), Moving Average Convergence Divergence (MACD), and Bollinger Bands could be incorporated into the forecasting framework to assess whether additional market information improves predictive performance. Second, macro-financial variables, including money supply, interest rates, exchange rates, foreign portfolio flows, and news sentiment, could be added to examine whether broader information improves forecasts beyond past index values alone. Third, the models could be tested across other frontier and emerging equity markets to assess the generalizability of the findings. Finally, future studies could examine adaptive and hybrid forecasting approaches, including model re-estimation under changing market conditions and reinforcement-learning-based strategies.

AUTHOR CONTRIBUTIONS

Conceptualization and theory: YB and MS; research design: YB; data collection: YB; analysis and interpretation: YB and MS; writing draft preparation: YB and MS; supervision: MS; correction of article: YB; proofread and final approval of article: YB and MS. All authors have read and agreed to the published version of the manuscript.

REFERENCES

Bachelier, L. (1900). Théorie de la spéculation. *Annales Scientifiques de l'École Normale Supérieure, 17,* 21–86. https://doi.org/10.24033/asens.476

Bekaert, G., & Harvey, C. R. (1997). Emerging equity market volatility. *Journal of Financial Economics, 43*(1), 29–77. https://doi.org/10.1016/S0304-405X(96)00889-6

Bekaert, G., Harvey, C. R., & Mondino, T. (2023). Emerging equity markets in a globalised world. *Emerging Markets Review, 56*,101034. https://doi.org/10.1016/j.ememar.2023.101034

BVMT. (2025). *Tunis Stock Exchange*. Retrieved June 20, 2025 from https://tunis-stockexchange.com/

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

Boubaker, H., Saidane, B., & Ben Saad Zorgati, M. (2022). Modelling the dynamics of stock market in the Gulf Cooperation Council countries: Evidence on persistence to shocks. *Financial Innovation, 8*, Article 46. https://doi.org/10.1186/s40854-022-00348-3

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2008). *Time series analysis: Forecasting and control (4th ed.).* John Wiley & Sons. https://doi.org/10.1002/9781118619193

Büyükşahin, Ü.Ç., & Ertekin, Ş. (2019). Improving forecasting accuracy of time series data using a new ARIMA-ANN hybrid method and empirical mode decomposition. *Neurocomputing, 361*, 151-163. https://doi.org/10.1016/j.neucom.2019.05.099

Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems, 2*(4), 303–314. https://doi.org/10.1007/BF02551274

Fakhfekh, M., Jeribi, A., & Ben Salem, M. (2021). Volatility dynamics of the Tunisian stock market before and during the COVID-19 outbreak: Evidence from the GARCH family models. *International Journal of Finance & Economics, 28*(2), 1653-1666. https://doi.org/10.1002/ijfe.2499

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance, 25*(2), 383–417. https://doi.org/10.2307/2325486

Fischer, T., & Krauss, C. (2018). 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

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning.* MIT Press.

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *Review of Financial Studies, 33*(5), 2223–2273. https://doi.org/10.1093/rfs/hhaa009

Henrique, B. M., Sobreiro, V. A., & Kimura, H. (2019). Literature review: Machine learning techniques applied to financial market prediction. *Expert Systems with Applications, 124*, 226–251. https://doi.org/10.1016/j.eswa.2019.01.012

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

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks, 2*(5), 359–366. https://doi.org/10.1016/0893-6080(89)90020-8

Kaygın, C. Y., Gün, M., Akarsu, O. N., Bağcı, H., & Yanık, A. (2026). Algorithmic stability in turbulent markets: Unveiling the superiority of shallow learning over deep architectures in cryptocurrency forecasting. *Mathematics, 14*(6),989. https://doi.org/10.3390/math14060989

Mehtab, S., & Sen, J. (2020). Stock price prediction using CNN and LSTM-based deep learning models. In Proceedings of the 2020 International Conference on Decision Aid Sciences and Application (DASA) (pp. 448–453). IEEE. https://doi.org/10.1109/DASA51403.2020.9317207

Nayak, R. K., Misra, D., & Behera, H. S. (2014). Impact of data normalization on stock index forecasting. International *Journal of Computer Information Systems and Industrial Management Applications, 6,* 257–269.

Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review, 6*(2), 41–49.

Sezer, O. B., Gudelek, M. U., & Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning: A systematic literature review (2005–2019). *Applied Soft Computing, 90*, 106181. https://doi.org/10.1016/j.asoc.2020.106181

Shahrour, M., & Dekmak, M. (2024). Intelligent stock prediction: A neural network approach. *International Journal of Financial Engineering, 11*(1), 2250016. https://doi.org/10.1142/S2424786322500165

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research, 15*(56), 1929–1958. https://doi.org/10.5555/2627435.2670313

Tashman, L. J. (2000). Out-of-sample tests of forecasting accuracy: An analysis and review. *International Journal of Forecasting, 16*(4), 437–450. https://doi.org/10.1016/S0169-2070(00)00065-0

Tsay, R. S. (2005). *Analysis of financial time series (2nd ed.).* John Wiley & Sons. https://doi.org/10.1002/0471746193

Varshney, S., & Srivastava, P. (2023). A comparative study of future stock price prediction through artificial neural network and ARIMA modelling. *NMIMS Management Review, 31*(4), 229–239. https://doi.org/10.1177/09711023241230367

Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting, 14*(1), 35–62. https://doi.org/10.1016/S0169-2070(97)00044-7

Information about the authors

***Youssef Boufares** – PhD student, Researcher, MOFID Research Laboratory, Faculty of Economics and Management, University of Sousse, Sousse, Tunisia, email: youssef.bouf@gmail.com, ORCID ID: https://orcid.org/0009-0009-9913-8571

Mounir Smida – Professor, PhD, Director of Research Laboratory, Faculty of Economics and Management, University of Sousse, Sousse, Tunisia, email: mounirsmida2006@gmail.com

Авторлар туралы мәліметтер

***Буфарес Ю.** – PhD докторант, MOFID ғылыми-зерттеу зертханасы, Сусс экономика және менеджмент факультеті, Сусс университеті, Сусс, Тунис, email: youssef.bouf@gmail.com, ORCID ID: https://orcid.org/0009-0009-9913-8571

Смида М. – профессор, PhD, зерттеу зертханасының жетекшісі, Сусс экономика және менеджмент факультеті, Сусс университеті, Сусс, Тунис, email: mounirsmida2006@gmail.com

Сведения об авторах

***Буфарес Ю.** – PhD докторант, Научно-исследовательская лаборатория MOFID, факультет экономики и менеджмента Сусса, Университет Сусса, Сусс, Тунис, email: youssef.bouf@gmail.com, ORCID ID: https://orcid.org/0009-0009-9913-8571

Смида М. – профессор, PhD, руководитель исследовательской лаборатории, факультет экономики и менеджмента Сусса, Университет Сусса, Сусс, Тунис, email: mounirsmida2006@gmail.com