ESG Volatility Forecasting using LSTM Networks and Hyperparameter Optimization Techniques

Syrine Ferjani[1*], Boutheina Jlifi[1,2,4,6], Lamjed Ben Said[1], Marc-Arthur Diaye[3], Khaled Guesmi[5]

***Abstract*— Incorporating Environmental, Social, and Governance (ESG) factors has become increasingly important in modern investment strategies for both individual and institutional investors. However, forecasting ESG market volatility remains challenging due to the nonlinear and dynamic nature of financial time series. This study proposes a deep learning framework for ESG volatility forecasting based on an optimized Long Short-Term Memory (LSTM) model. The proposed approach integrates historical financial data, technical indicators, and macroeconomic variables to capture complex market behavior and volatility dynamics.**

To improve model robustness and automate hyperparameter selection, a Genetic Algorithm (GA) was employed to optimize three key LSTM hyperparameters: number of neurons, batch size, and look-back window. In addition to the proposed ESG-Volatility-3GA-LSTM framework, baseline LSTM, Random Search, and Grid Search optimization strategies were evaluated under identical experimental conditions.

Experimental results demonstrate that the GA-optimized LSTM achieves competitive forecasting performance with improved stability across repeated runs. Statistical analyses, including paired t-tests and Wilcoxon signed-rank tests, further confirm the robustness and reproducibility of the proposed framework. The findings highlight the importance of systematic hyperparameter optimization for ESG volatility forecasting under varying market conditions.

Keywords : ESG index, volatility prediction, Hyperparameter tuning, Artificial Intelligence (AI) techniques, Computational Intelligence, Genetic Algorithm, Deep Learning.

I. INTRODUCTION

A. Motivations and Challenges

Conventional investors primarily emphasize maximizing investment returns by assessing profitability metrics and conduct thorough analyses of financial statements to identify the most lucrative stocks available in the market [1]. According to [2], ESG factors, which correspond to non-financial information, have been viewed in recent years as new drivers for financial analysis and creditworthiness evaluation. Policymakers and investors are paying more and more attention to ESG investing [3]. In order to manage financial risk and encourage sustainable investment strategies, investors are gradually incorporating ESG indicators into their decision-making processes [4]. Accurately forecasting the volatility of ESG indexes has become crucial for risk assessment, portfolio management, and market stability as these indices gain popularity. However, because of the complex interactions between financial markets and macroeconomic variables, forecasting the volatility of ESG indices is still difficult. Econometric models like Generalized Autoregressive Conditional Heteroskedasticity (GARCH) and its variations are frequently used in traditional methods for volatility prediction [5], [6], [7]. While these models offer valuable insights, they often fail to capture complex temporal dependencies and nonlinear relationships in financial data. To increase the accuracy of financial prediction, machine learning (ML) and DL techniques have been introduced more recently [8]. Few studies have concentrated on ESG index volatility prediction using advanced optimization techniques, despite the growing interest in applying DL to financial forecasting. Additionally, the majority of current research relies on fixed model configurations and doesn't thoroughly examine how hyperparameter optimization affects predictive performance [9], [10].

B. Objectives and Contributions

This study aims to improve ESG volatility forecasting using an optimized LSTM-based deep learning framework. The proposed approach integrates historical financial data, technical indicators, and macroeconomic variables to better capture the nonlinear behavior of ESG financial markets.

The main contributions of this work are summarized as follows:

- **GA-Optimized LSTM Framework:** We propose an ESG volatility forecasting framework based on LSTM networks optimized using a GA for tuning the number of neurons, batch size, and look-back window.
- **Comparative Experimental Evaluation:** The proposed ESG-Volatility-3GA-LSTM framework is compared with baseline LSTM, Random Search, and Grid Search optimization strategies under identical experimental conditions.
- **Application to Sustainable Finance:** The study demonstrates the potential of combining deep learning and evolutionary optimization for ESG volatility forecasting.

The remainder of this paper is organized as follows. Section II reviews related work on ESG volatility forecast-

[1]SMART Lab, Institut Supérieur de Gestion de Tunis (ISGT), Tunis university, Tunisia.
[2]Ecole Supérieure de Commerce de Tunis (ESCT), Manouba university, Tunisia.
[3]Sorbonne School of Economics , Paris, France.
[4]Centre CRECC, Paris School of Business , Paris, France.
[5] IPAG Business School, Paris, France.
[6]LITIS Laboratory, Le Havre University, Normandie, France.
*Corresponding authors.
E-mail(s):
sirineferjani@esc.u-manouba.tn
boutheina.jlifi@esct.uma.tn
lamjed.bensaid@isg.rnu.tn
Marc-Arthur.Diaye@univ-paris1.fr
kalloda$_fr@yahoo.fr$

ing and deep learning approaches. Section III presents the proposed methodology and optimization framework. Section IV discusses the experimental setup and obtained results. Finally, Section V concludes the paper and outlines future work.

II. LITERATURE REVIEW

The incorporation of AI into sustainable finance has become a crucial innovation for enhancing the evaluation of ESG factors. In the financial sector, AI-powered solutions have shown exceptional abilities in refining decision-making, optimizing portfolio management, and reducing risks [9].

Although there is significant enthusiasm for AI-driven sustainable finance, the empirical evidence supporting AI's effectiveness in ESG prediction is still scattered. Many traditional ESG scoring systems depend on rule-based frameworks, such as ARCH and GARCH models [7], that find it challenging to capture real-time or detailed data, especially in climate-sensitive industries. On the other hand, AI-based models, particularly those utilizing ML, offer the potential for improved predictive capabilities, but their advantages over traditional systems have not been thoroughly investigated [9]. In the context of this study, predicting ESG index volatility is crucial as it captures market dynamics influenced by these non-financial factors, offering investors deeper insights into risk and opportunity within sustainable investment landscapes. In a recent study, a new DL framework called ESG2Risk was implemented to forecast the future volatility of stock prices using ESG news [11].

Most of the studies have focused on price prediction and a few papers tackled the ESG volatility prediction (mentioned in Table I, which is the focus of this study.

TABLE I: Related works of used techniques to forecast the volatility of ESG index

Used Methodology	Used data	Evaluation Metrics	Experiments and Results
DL model architectures like LSTM, gated recurrent unit (GRU), and convolutional neural network (CNN) to forecast ESG index volatility [1]	Fundamental data + Technical indicators Macroeconomic data	RMSE, MAE and R-scores of a test set	The LSTM with 10 neurons showed outperformed other model with: RMSE=0.5849 MAE=0.1425 R=0.9952
Generative Adversarial Networks (GANs) for predicting the volatility of ESG indices. [12]	input company information historical ESG index data.	Accuracy (ACC)	Tested on various number of tasks reaching 100 Achieved ACC= 90.3%
GARCH and LSTM models for predicting the return of ESG volatility [5]	input company information historical ESG index data.	RMSE	GARCH outperformed with RMSE= 0.01135

The mentioned studies in this section lack the analysis on the robustness of the utilized models.

Therefore, our study aimed to address these issues by developing a novel DL-based framework in the context of ESG volatility forecasting trained with the best possible set of influencing factors.

III. PROPOSED METHODOLOGY

A. Choice of the model

Previous studies on ESG volatility forecasting have investigated several deep learning architectures, including CNN, GRU, and LSTM networks [1]. Among these architectures, LSTM demonstrated superior forecasting performance due to its ability to capture long-term temporal dependencies and nonlinear sequential patterns in financial time-series data. Based on these observations and the comparative results reported in prior work, the present study adopts the LSTM architecture as the baseline forecasting model and focuses on improving its hyperparameter configuration through optimization strategies.

GAs were selected for hyperparameter optimization due to their effectiveness in exploring complex and nonlinear search spaces without requiring gradient information. Previous studies deep learning optimization have demonstrated that GA-based tuning can improve model robustness and automate the selection of suitable hyperparameter configurations [13],[14]. Compared with manual tuning approaches, GA provides a systematic search mechanism capable of balancing exploration and exploitation during optimization.

The GA optimization process focused on three influential LSTM hyperparameters: the number of neurons, batch size, and look-back window. These parameters were selected because they directly affect model complexity, temporal dependency learning, and training stability in time-series forecasting tasks. Previous studies have shown that optimizing these hyperparameters can significantly influence LSTM forecasting performance in financial applications [15]. Restricting the optimization process to these parameters also helps maintain a manageable search space and computational cost during GA optimization.

B. Algorithm Design

This framework combines a GA with a LSTM network to optimize hyperparameters and improve prediction accuracy. The optimized LSTM has been previously implemented in few studies across different contexts, demonstrating its versatility and effectiveness [13], [14]. This methodology consists of a primary algorithm and a number of core procedures, where each step addresses a specific challenge in the forecasting process as shown in figure 1.

Step 1: Data preprocessing and normalization The raw dataset is first cleaned to remove inconsistencies and missing values. The data is then divided into training, validation, and test sets. A Min–Max normalization is also applied to both the input variables and the target variable.

Step 2: Sequence construction Since LSTM networks require sequential inputs, the time-series data is transformed into supervised learning sequences using a sliding window mechanism. The parameter look_back determines the number of past observations used to predict the next time step.

Step 3: Hyperparameter optimization using GA To enhance model performance, a GA is used to automatically search for the optimal hyperparameters of the LSTM network. The algorithm initializes a population of candidate solutions representing different combinations of neurons, batch size, and look-back window. Through iterative processes including fitness evaluation, selection, crossover, and mutation, the GA identifies the hyperparameter configuration that minimizes the validation error.

Step 4: Optimized LSTM training and forecasting Using the optimal hyperparameters obtained from the GA, the LSTM model is trained with early stopping to prevent overfitting. To ensure robustness, the model is trained through multiple independent runs. The trained model then generates predictions on the test dataset.

Step 5: Performance evaluation and statistical validation The forecasting performance of the proposed ESG-Volatility-3GA-LSTM model is evaluated using several metrics, including Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and the determination coefficient (R^2). To verify the statistical significance of the improvements, the ESG-Volatility-3GA-LSTM model is compared with a baseline LSTM model using statistical tests such as the paired t-test, Wilcoxon signed-rank test, and the Diebold–Mariano test.

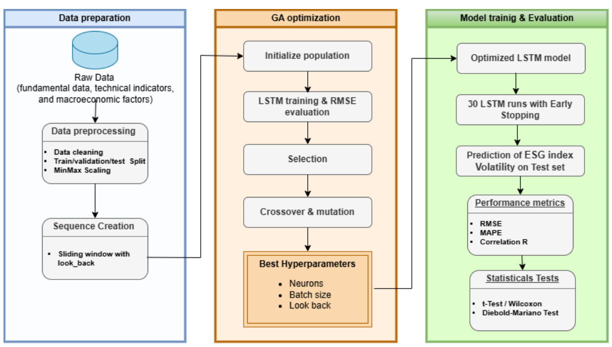

Fig. 1: The Proposed ESG-Volatility-3GA-LSTM Framework

This algorithm describes the main procedure of the proposed ESG-Volatility-3GA-LSTM forecasting framework. It takes a raw time-series dataset as input and processes it through several stages including preprocessing, hyperparameter optimization using a GA, and final model training for forecasting. Each stage contributes to improving the prediction accuracy of the volatility time series.

The algorithm requires several important inputs:

D **: Raw dataset** The original time-series dataset containing historical observations used for volatility prediction.

P **: GA parameters** A set of configuration parameters controlling the behavior of the GA, including the population size, number of generations, and mutation rate. These parameters guide the evolutionary search for optimal LSTM hyperparameters.

H **: Hyperparameter search space** A predefined set of candidate values for the LSTM hyperparameters optimized by the GA. It includes:

neurons : Number of hidden units in the LSTM layer, controlling the model's learning capacity.

batch_size : Number of training samples processed before updating the network weights.

look_back : Length of the sliding window used to transform the time series into input sequences.

R **: Number of training replicates** Number of independent model training runs performed using the optimal hyperparameters to ensure robustness and reduce randomness from neural network initialization.

M_{best} **: Optimized ESG-Volatility-3GA-LSTM model** The final LSTM forecasting model trained using the best hyperparameter configuration identified by the GA.

Y_{pred} **: Forecasted volatility values** Predicted volatility values produced by the optimized ESG-Volatility-3GA-LSTM model on the test dataset.

Algorithm 1 ESG-Volatility-3GA-LSTM Hyperparameter Optimization for Volatility Forecasting

Data: X, Y : Input features and target volatility
$PopSize$: Population size
N_g : Number of GA generations
C : Crossover probability
M : Mutation probability
H : Hyperparameter ranges {neurons, batch_size, look_back}
Result: Optimal hyperparameters H_{best}
$Fitness\_Set \leftarrow \emptyset$ // Initialize fitness values
$Pop \leftarrow Initial\_Population(PopSize, H)$
for $t \leftarrow 1$ to $PopSize$ **do**
    $neurons_t \leftarrow Pop[t].neurons$
    $batch_t \leftarrow Pop[t].batch\_size$
    $lookback_t \leftarrow Pop[t].look\_back$
    $(X_{train}, X_{val}, Y_{train}, Y_{val}) \leftarrow Split\_Data(X, Y)$
    $Model \leftarrow Build\_LSTM(neurons_t)$
    $Model.Train(X_{train}, Y_{train}, batch_t)$
    $Y_{pred} \leftarrow Model.Predict(X_{val})$
    $fitness \leftarrow RMSE(Y_{val}, Y_{pred})$
    $Pop[t].fitness \leftarrow fitness$
    $Fitness\_Set \leftarrow Fitness\_Set \cup \{fitness\}$
end for
for $g \leftarrow 1$ to N_g **do**
    $(indiv_1, indiv_2) \leftarrow RouletteSelection(Pop, Fitness\_Set)$
    $(child_1, child_2) \leftarrow Crossover(C, indiv_1, indiv_2)$
    $(child'_1, child'_2) \leftarrow Mutation(M, child_1, child_2)$
    $Pop \leftarrow UpdatePopulation(Pop, child'_1, child'_2)$
end for
$Best \leftarrow SelectBest(Pop)$
$neurons_{best} \leftarrow Best.neurons$
$batch_{best} \leftarrow Best.batch\_size$
$lookback_{best} \leftarrow Best.look\_back$
return $(neurons_{best}, batch_{best}, lookback_{best})$

IV. Experiments and Results

A. Dataset Description

The dataset used in this study is a combination of fundamental, macroeconomic, and technical variables, originally collected and processed in prior research [1].

Fundamental data: This category includes historical financial variables that provide basic information regarding index performance. The closing price of the index is used as the primary variable representing daily market performance.

Macroeconomic data: Macroeconomic indicators are included to capture the overall health of the financial market. The selected variables are CBOE Volatility Index (VIX), Effective Federal Funds Rate (EFFR), Civilian Unemployment Rate (UNRATE), Consumer Sentiment Index (UMCSENT), US Dollar Index (USDX).

Technical indicators: Technical indicators are commonly used by active traders to analyze short-term market dynamics. The included features are Volatility (used both as an input and the target variable), Moving Average Convergence Divergence (MACD), Relative Strength Index (RSI), Sharpe Ratio (SR).

Monthly volatility is computed as the rolling standard deviation of monthly returns[1], and is then annualized by multiplying by $\sqrt{12}$:

$$\text{Volatility}_{annual} = \text{Volatility}_{monthly} \times \sqrt{12} \tag{1}$$

The dataset combines these variables to capture both short-term market movements and longer-term economic trends, forming a comprehensive input for the proposed ESG-Volatility-3GA-LSTM volatility forecasting framework.

The target variable corresponds to the annualized ESG volatility index *volat_ann*, modeled as a continuous real-valued regression output. In the considered dataset, the target values range approximately from 0.0077 to 0.2131. Prior to training, MinMax normalization is applied using parameters estimated exclusively from the training set, while all reported metrics are computed after inverse transformation to the original volatility scale.

B. Experimental Setup

All experiments were implemented in Python using the TensorFlow/Keras deep learning framework. The experimental environment was executed on Google Colab using an NVIDIA T4 GPU accelerator. The TensorFlow version used during experimentation was TensorFlow 2.20.0.

C. Dataset Preparation and Preprocessing

The proposed ESG-Volatility-3GA-LSTM framework was evaluated using the dataset described in Section IV-A. The data were split into training, validation, and test sets based on the following:

TABLE II: Chronological Dataset Partitioning

Subset	Period
Training	2013-01-02 – 2018-05-28
Validation	2018-05-29 – 2020-03-16
Testing	2020-03-17 – 2021-11-30

The training subset was used for model learning, while the validation subset was exclusively reserved for hyperparameter optimization during the GA search process. The test subset remained completely unseen until the final evaluation stage. This chronological separation ensures that future observations are never used during training or hyperparameter optimization, thereby preventing temporal leakage.

All input features were normalized using Min-Max scaling. Importantly, scaling parameters were estimated exclusively from the training data and subsequently applied to the validation and testing subsets. All reported evaluation metrics were computed after inverse transformation to the original target scale.

Since LSTM models require sequential temporal input, sliding-window sequences were generated using a look-back mechanism. For each prediction step, the model receives a fixed number of previous observations and predicts the corresponding volatility value.

The GA was employed to optimize three hyperparameters: number of LSTM neurons, batch size, and look-back window length. To ensure robustness, all experiments were repeated for $R = 30$ independent replicates. Performance metrics include RMSE, MAE, and R^2.

D. Hyperparameter Optimization

The GA was used to search the hyperparameter space $\mathcal{H}$ for optimal values. The fitness function was defined as the RMSE on the validation set:

$$\varphi_i = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(y_j - \hat{y}_j)^2} \tag{2}$$

The GA was used to optimize three LSTM hyperparameters:

- Number of LSTM neurons,
- Batch size,
- Look-back window size.

The search ranges used in the GA optimization process are summarized in Table III.

TABLE III: GA Hyperparameter Search Space

Hyperparameter	**Search Range**
Number of neurons	[10, 200]
Batch size	[4, 32]
Look-back window	[5, 60]

The GA was configured using a population size of 10 individuals, 5 generations, and a mutation rate of 0.2. Selection, crossover, and mutation operators were iteratively applied to evolve candidate solutions toward improved forecasting performance on the validation set.

The optimization process with GA converged relatively quickly across generations, showing that the evolutionary search was able to efficiently identify suitable LSTM hyperparameters. The best solution was obtained at Generation 4, where the lowest validation RMSE value reached approximately 0.0644.

The best hyperparameters found by the GA are summarized below:

- Number of neurons: 149
- Batch size: 7
- Look-back window: 6

These results indicate that the model achieved better forecasting performance using a relatively short historical window together with a larger number of neurons to capture complex temporal patterns in ESG volatility data.

The optimized hyperparameters obtained from the GA were then used for the final ESG-Volatility-3GA-LSTM model training and evaluation on the unseen test set.

For comparative analysis, additional experiments using Grid Search and Random Search optimization strategies were also conducted under the same preprocessing pipeline, evaluation protocol, and testing conditions to ensure fair comparison and reproducibility.

E. Model Performance

Table IV reports the mean and standard deviation of RMSE, MAE, and R² over 30 replicates for the ESG-Volatility-3GA-LSTM model and a baseline LSTM model. The baseline LSTM model was independently reimplemented and evaluated under the same preprocessing pipeline, dataset partitioning strategy, and experimental protocol as the proposed ESG-Volatility-3GA-LSTM framework. To evaluate the effectiveness of the proposed GA-based optimization approach, Random Search and Grid Search strategies were also implemented for LSTM hyperparameter tuning. The comparison was conducted using the same dataset splits, preprocessing pipeline, scaling procedure, evaluation metrics, and testing protocol to ensure fairness.

TABLE IV: Performance comparison of forecasting models over 30 runs.

Model	RMSE	MAE	R^2
Baseline LSTM	0.007437 ± 0.001928	0.005703 ± 0.002140	0.839635 ± 0.115822
ESG-Volatility-3GA-LSTM	0.008223 ± 0.001314	0.005363 ± 0.000826	0.938192 ± 0.019075
Random Search LSTM	0.008149 ± 0.002092	0.005106 ± 0.001205	0.963508 ± 0.012253
Grid Search LSTM	0.009384 ± 0.002930	0.005357 ± 0.001048	0.951482 ± 0.013078

The experiments demonstrate that hyperparameter optimization techniques such as GA, Random Search, and Grid Search do not necessarily guarantee superior forecasting accuracy compared with a carefully configured baseline LSTM model.

F. Statistical Analysis

To evaluate the statistical significance of forecasting differences between models, paired t-tests, Wilcoxon signed-rank tests, Cohen's d effect size analysis, and the Friedman test were conducted on the RMSE values obtained across 30 independent runs.

TABLE V: Statistical comparison between the baseline LSTM and optimized models.

Comparison	t-test p	Wilcoxon p	Cohen's d
Baseline vs ESG-Volatility-3GA-LSTM	0.0547	0.0549	-0.3718
Baseline vs Random Search	0.1809	0.1579	-0.2546
Baseline vs Grid Search	0.0050	0.0043	-0.5645

The statistical results indicate that the differences between the baseline LSTM and both ESG-Volatility-3GA-LSTM and Random Search LSTM were not statistically significant at the 5% significance level. However, Grid Search LSTM produced significantly higher forecasting errors compared with the baseline model.

Additionally, the Friedman test yielded a p-value of 0.004467, confirming the existence of statistically significant differences among the evaluated forecasting approaches overall.

The boxplot representation in Figure 2 further illustrates the distribution and variability of RMSE values across the 30 runs for each model.

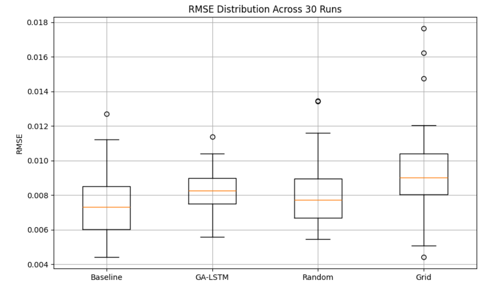

Fig. 2: Boxplot of the models

To further analyze prediction errors, the distribution of residuals is presented in Figure 3 showing that most errors are concentrated near zero.

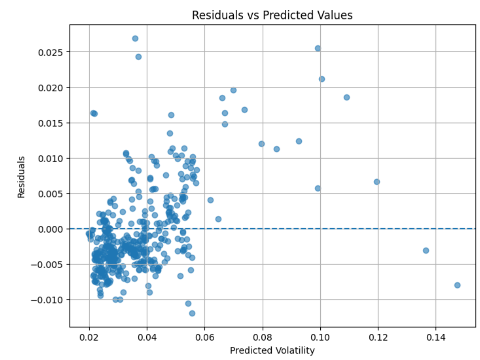

Fig. 3: Error Distribution (ESG-Volatility-3GA-LSTM)

Figure 4 shows that ESG-Volatility-3GA-LSTM achieved the lowest RMSE variability across 30 runs. This result highlights the ability of GA-based hyperparameter optimization to improve the stability and reproducibility of LSTM forecasting, even when average prediction accuracy remains comparable to other optimization strategies.

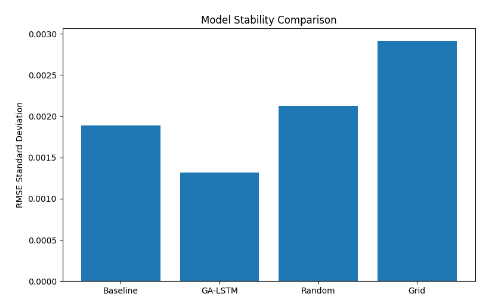

Fig. 4: Comparison of RMSE standard deviation across 30 independent runs

G. Robustness Under Different Market Conditions

To evaluate generalization under varying market regimes, the proposed framework was assessed during periods characterized by elevated financial volatility, including the COVID-19 market disruption period. Experimental observations indicate that the model maintains stable predictive behavior despite increased market uncertainty, suggesting robustness under changing volatility conditions.

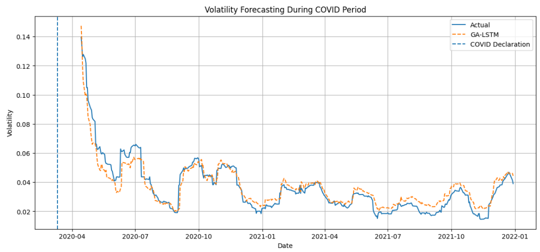

Fig. 5: Error Distribution Comparison

H. Discussion

The experimental results show that the proposed ESG-Volatility-3GA-LSTM framework achieves competitive and stable forecasting performance for ESG volatility prediction. Compared with the baseline LSTM, the GA-optimized model obtained slightly improved MAE and R^2 values while maintaining comparable RMSE performance across repeated runs. The lower variability observed in several experiments also indicates improved prediction stability and reproducibility.

The results further confirm the effectiveness of LSTM networks for financial time-series forecasting due to their ability to capture temporal dependencies and nonlinear market behavior. In addition, the comparative analyses with Random Search and Grid Search highlight the importance of systematic hyperparameter optimization for improving training consistency.

Statistical analyses based on paired t-tests and Wilcoxon signed-rank tests indicate that the observed improvements remain moderate, suggesting that GA optimization mainly enhances model stability rather than producing substantial forecasting gains.

Despite these advantages, GA-based optimization introduces additional computational cost because multiple neural network models must be trained across several generations and candidate solutions. Consequently, a trade-off exists between optimization quality and computational efficiency.

V. CONCLUSION AND FUTURE WORK

This study proposed a ESG-Volatility-3GA-LSTM framework for ESG volatility forecasting by integrating historical financial data, macroeconomic variables, and technical indicators within a deep learning architecture. A GA was employed to optimize key LSTM hyperparameters, namely the number of neurons, batch size, and look-back window. To improve robustness and reproducibility, multiple independent runs and early stopping were adopted.

Experimental results demonstrated competitive forecasting performance with stable RMSE, MAE, and R^2 values across repeated runs. Additional comparisons with baseline LSTM, Random Search, and Grid Search were conducted under identical experimental conditions. Statistical analyses, including paired t-tests and Wilcoxon signed-rank tests, were also performed to evaluate the robustness of the obtained results.

The main contribution of this work lies in establishing a rigorous and reproducible evaluation framework for ESG volatility forecasting using systematically optimized LSTM architectures. The results suggest that hyperparameter optimization mainly improves model stability and reproducibility rather than producing dramatic predictive gains.

Despite its effectiveness, the proposed framework remains computationally expensive due to repeated model training during optimization. In addition, only a limited subset of hyperparameters and LSTM-based architectures was considered.

Future work may investigate alternative optimization methods such as PSO or Bayesian Optimization, as well as more advanced architectures including stacked LSTM, bidirectional LSTM, attention-based models, and transformers. Larger and more recent ESG datasets may also be explored to further assess generalization and scalability.

REFERENCES

[1] H. N. Bhandari, N. R. Pokhrel, R. Rimal, K. R. Dahal, and B. Rimal, "Implementation of deep learning models in predicting esg index volatility," *Financial Innovation*, vol. 10, no. 1, p. 75, 2024.

[2] E. Pineau, P. Le, and R. Estran, "Importance of esg factors in sovereign credit ratings," *Finance Research Letters*, vol. 49, p. 102966, 2022.

[3] B. Jlifi, M.-A. Diaye, K. Guesmi, S. Ferjani, and N. Hamdi, "Reconciling divergence among esg scores: A bi-level artificial intelligence based methodology for corporate governance controversies prediction," *Computational Economics*, pp. 1–43, 2026.

[4] S. L. Gillan, A. Koch, and L. T. Starks, "Firms and social responsibility: A review of esg and csr research in corporate finance," *Journal of corporate finance*, vol. 66, p. 101889, 2021.

[5] A. K. Mishra, R. Kumar, and D. P. Bal, "Esg volatility prediction using garch and lstm models," *Financial Internet Quarterly*, vol. 19, no. 4, pp. 97–114, 2023.

[6] R. Chen, "Forecasting the volatility of esg stock indexes by using the arma-garch model," in *International Conference on Economic Management and Green Development*. Springer, 2024, pp. 459–471.

[7] M. S. Naik and N. S. Wagle, "Relationship between volatility and sustainability index returns in the indian stock market: A garch analysis," *IUP Journal of Accounting Research & Audit Practices*, vol. 24, no. 1, pp. 41–53, 2025.

[8] P. Borkowski and M. Z. Abdullazade, "Esg indicators and sectoral volatility: Garch vs. hybrid machine learning," *Finance Research Letters*, p. 108103, 2025.

[9] A. M. Elhady and S. Shohieb, "Ai-driven sustainable finance: computational tools, esg metrics, and global implementation," *Future Business Journal*, vol. 11, no. 1, p. 209, 2025.

[10] A. A. Davidescu, I. Bîrlan, E. M. Manta, and C. M. Geambaşu, "Artificial intelligence in esg and sustainable finance: A bibliometric analysis of research trends," in *Proceedings of the International Conference on Business Excellence*, vol. 19, no. 1. Bucharest University of Economic Studies, 2025, pp. 1506–1517.

[11] T. Guo, N. Jamet, V. Betrix, L.-A. Piquet, and E. Hauptmann, "Esg2risk: A deep learning framework from esg news to stock volatility prediction," *arXiv preprint arXiv:2005.02527*, 2020.

[12] J. Shi and J. Li, "Application of deep learning in predicting esg index volatility," in *Proceedings of the 2024 International Conference on Big Data Mining and Information Processing*, 2024, pp. 199–203.

[13] B. Jlifi, M. Medini, and C. Duvallet, "A guided genetic algorithm-based ensemble voting of polynomial regression and lstm (gga-polreg-lstm) for congestion prediction using iot and air quality data in sustainable cities," *The Journal of Supercomputing*, vol. 80, no. 13, pp. 18 797–18 837, 2024.

[14] B. Jlifi, S. Ferjani, and C. Duvallet, "A genetic algorithm based three hyperparameter optimization of deep long short term memory (ga3p-dlstm) for predicting electric vehicles energy consumption," *Computers and Electrical Engineering*, vol. 123, p. 110185, 2025.

[15] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.