Temporal Action Analysis in Metaheuristics: A Machine Learning Approach

Panagiotis Kalatzantonakis[0000−0002−0729−6583], Angelo Sifaleras[0000−0002−5696−7021], and Nikolaos Samaras[0000−0001−8201−7081]

Department of Applied Informatics, School of Information Sciences, University of Macedonia, 156 Egnatia Str., Thessaloniki 54636, Greece
pkalatzantonakis@uom.edu.gr, sifalera@uom.gr, samaras@uom.gr

Abstract. This study explores the use of Autoregressive Integrated Moving Average (ARIMA) and Long Short-Term Memory (LSTM) machine learning models in metaheuristic algorithms, with a focus on a modified General Variable Neighborhood Search (GVNS) for the Capacitated Vehicle Routing Problem (CVRP). We analyze the historical chain of actions in GVNS to demonstrate the predictive potential of these models for guiding future heuristic applications or parameter settings in metaheuristics such as Genetic Algorithms (GA) or Simulated Annealing (SA). This "optimizing the optimizer" approach reveals that, the history of actions in metaheuristics provides valuable insights for predicting and enhancing heuristic selections. Our preliminary findings suggest that machine learning models, using historical data, offer a pathway to more intelligent and data-driven optimization strategies in complex scenarios, marking a significant advancement in the field of combinatorial optimization.

Keywords: Intelligent Heuristic Decision-Making · Data-Driven Metaheuristic Strategies · Machine Learning Enhanced Combinatorial Optimization · Offline Metaheuristic Algorithm Configuration

1 Introduction

1.1 Metaheuristics in Combinatorial Optimization

Metaheuristic algorithms have evolved significantly to address complex and NP-hard challenges in combinatorial optimization [12, 13], such as the CVRP [4]. Historically, these algorithms have evolved from simple solution-seeking methods to sophisticated adaptive frameworks capable of intelligently navigating complex solution spaces. This evolution reflects a continuous effort to enhance efficiency and effectiveness in finding near-optimal solutions, especially in computationally demanding scenarios.

1.2 Machine Learning Integration in Metaheuristics

The integration of machine learning into metaheuristics marks the latest advancement in this field, representing a significant leap in computational intelligence. Building upon historical progress, our research incorporates ARIMA [5] and LSTM models [7] into the GVNS for the CVRP, aiming to capture and leverage temporal dynamics in optimization processes. This integration is not just an innovation, but a response to the growing need for more precise predictive capabilities in dynamic environments. Based on recent studies in parallel execution [9, 8, 1], learning-based neighborhood search [10, 14], and large neighborhood search adaptations [6], our approach seeks to harness the potential of machine learning to further refine and improve metaheuristic strategies.

2 Methodological and Experimental Setup

In this section, we outline our comprehensive methodological framework, which begins with the innovative reconfiguration of the GVNS [2, 11] for the CVRP. This approach is crucial for generating a robust dataset, essential for the subsequent training and optimization of ARIMA and LSTM models.

2.1 GVNS-Driven Data Collection and Analysis

The proposed modification of the GVNS metaheuristic consists of a different neighborhood selection step, and it is tailored towards generating unbiased data across CVRP instances. Through multithreaded data collection and extensive preprocessing, including normalization and structuring, we prepare the data set for pattern analysis using ARIMA and LSTM models. This strategic approach seeks to evolve traditional metaheuristic algorithms into intelligent, adaptive systems.

Data collection during GVNS iterations involves tracking each heuristic's application and outcome, quantified as binary values (success or failure) and continuous values (degree of solution improvement). This detailed data collection is crucial for a robust analysis. The data set is then preprocessed for analysis. ARIMA models are used for regression analysis to identify linear trends, while LSTM networks address classification issues, adapt in processing sequential data. This combination allows for a comprehensive analysis of patterns in decision making. We used CVRP instances from CVRPLib [3] (sets A, B, and X), which offer various complexities, to validate our methodology in a structured environment.

2.2 Model and Parameter Optimization

In our study, both ARIMA and LSTM models underwent meticulous optimization processes to enhance their predictive accuracy for the GVNS algorithm. ARIMA's role was to forecast the "reward" value, with our analysis revealing some seasonality potentially influenced by GVNS's cyclic phases. The Augmented Dickey-Fuller and Kwiatkowski–Phillips–Schmidt–Shin (KPSS) tests, combined with Fourier Transform and Seasonal Decomposition, confirmed the time series' stationarity, leading us to favor ARIMA over SARIMA. We explored a range of parameters, evaluated the models on Akaike Information Criterion

(AIC) and Mean Squared Error (MSE), and settled on the ARIMA (5, 1, 5) model for its optimal balance of AIC and high PRAUC, indicating its effectiveness in predicting heuristic improvements. In optimizing the ARIMA model, specific parameters are pivotal: “p” representing the order of autoregression, “d” the degree of differencing, and “q” the moving average window, that together define the model's structure. The performance of ARIMA (5, 1, 5) was evaluated using the Mean Squared Error (MSE) and Akaike Information Criterion (AIC), with lower values in both indicating better model fit. AIC was particularly crucial for comparing the quality of different models. Additionally, the Precision-Recall Area Under Curve (PRAUC) metric was utilized to assess the binary classification effectiveness of ARIMA, an important aspect given the imbalanced nature of our dataset.

Hyperparameter tuning of the LSTM model was conducted using the Hyperband method, targeting key parameters such as LSTM units, dropout rate, and learning rate, which unveiled a preference for a BiLSTM structure to better capture temporal dependencies. The optimization involved systematic exploration of critical hyperparameters. LSTM units affect the model's complexity and its ability to discern data patterns, with higher units offering greater complexity at the cost of computational resources. Dropout rate mitigates overfitting by omitting units during training, while the learning rate is vital for effective model training, avoiding minima overshoots. The choice of loss function (MSE, MAE, Binary Cross-Entropy) influences error quantification, and activation functions (sigmoid, ReLU, tanh) affect data signal processing, crucial for learning. BiLSTM's bidirectional approach improves predictive accuracy by utilizing past and future data. Batch size and epochs set the training sample size and cycles, and the optimizer (SGD, RMSprop, Adam) impacts learning speed and efficiency. The attention mechanism further refines the model by concentrating on particular input sequence segments, boosting performance on complex time-series tasks.

Table 1 details the models that perform the best. Also, the final hyperparameter configuration for the BiLSTM model, as detailed in Table 2, was strategically chosen to strike a balance between computational resources and predictive accuracy. This resulted in an optimized BiLSTM model. Both the ARIMA and LSTM models were meticulously fine-tuned to complement each other, thus providing comprehensive predictive insights within the GVNS framework.

Table 1. Top 3 ARIMA Models

Rank	p	d	q	AIC	MSE
Top 3 Models Based on AIC					
1	5	0	5	470.7783	2.0643
2	5	1	5	471.7483	1.8388
3	5	1	6	472.6345	1.8477
Top 3 MSE-based models					
1	3	1	6	495.4785	1.8168
2	5	1	2	487.4079	1.8381
3	6	1	6	475.7171	1.8388

Table 2. LSTM Model Hyperparameters

Hyperparameter	Range	Best Value
LSTM Units	32 to 512 (step: 32)	256
Dropout Rate	0.0 to 0.5 (step: 0.1)	0.1
Learning Rate	1e-4 to 1e-2 (sampling="log")	0.0079
Loss Functions	MSE, MAE, Binary Cross-Entropy	Mean Squared Error
Activation Functions	sigmoid, relu, tanh	tanh
Bidirectional setting	True/False	True
Batch Size	32 to 512	256
Epochs	10 to 100	100
Optimizers	SGD, RMSprop, Adam	Adam
Attention Mechanism	True/False	True

3 Results and Analysis

Our study showcases the potential of machine learning, particularly ARIMA and LSTM, in interpreting the sequence of actions in metaheuristic algorithms such as GVNS, GA and SA. By analyzing historical data from heuristic applications, we demonstrate how these models can predict and influence future heuristic choices, thus optimizing the decision-making process within these algorithms.

The selection of ARIMA and LSTM models in our study illustrates the complexity of decision-making in forecasting actions. ARIMA effectively predicts continuous outcomes such as "reward", providing linear insights, while LSTM excels in binary classification, crucial for different decision-making scenarios. The ARIMA(5, 0, 5) model, with its high AIC, accurately predicts "reward" values. This suggests that maintaining a focus on recent historical actions, up to five steps back, could be crucial to accurately forecasting outcomes in metaheuristic processes. Despite its limitations in accuracy and PRAUC, its ability to capture short-term historical trends is notable.

Conversely, the BiLSTM model significantly surpasses ARIMA in both accuracy and PRAUC, demonstrating its superior capability in binary classification and effective handling of sequential data. This highlights its potential as a robust tool for guiding heuristic decisions in metaheuristic algorithms. For a detailed comparison of the performance of the models, particularly highlighting their respective strengths in predictive accuracy, readers are encouraged to refer to Table 3, which presents a comprehensive overview of the performance metrics of the top models.

In conclusion, the findings of this study have far-reaching implications for the broader field of optimization and algorithm design. The successful integration of ARIMA and LSTM models within metaheuristics such as GVNS, GA, and SA demonstrates a promising path toward more intelligent, data-driven decision-making processes. This approach can be extended to other complex optimization

Table 3. Top Models Performance (ARIMA & LSTM)

Metric	Best Model based on AIC (5, 0, 5)	Best Model based on MSE (3, 1, 6)	LSTM Model
MSE	2.06431	1.81686	-
AIC	470.7783	495.4785	-
Accuracy	0.52658	0.53291	0.788956
PRAUC	0.57666	0.48160	0.673414

scenarios, opening up new avenues for research in algorithm efficiency and effectiveness. Future studies might explore the integration of different machine learning models or delve into real-time data adaptation, further advancing the field of combinatorial optimization. By leveraging historical data to inform heuristic choices, this research contributes to the ongoing evolution of metaheuristic algorithms, moving them toward more adaptive, predictive, and efficient frameworks.

4 Exploring the Future of Machine Learning in Metaheuristics

Our study, focused on analyzing data generated by a modified VNS approach for CVRP, indicates the potential of machine learning models like ARIMA and LSTM in enhancing metaheuristic algorithms. While our research is specific to VNS, the principle can be extended to other metaheuristics such as Genetic Algorithms and Simulated Annealing. For example, in GA, the history of genetic operations could be analyzed to predict their effectiveness, while in SA, the sequence of temperature adjustments and their outcomes could inform future adjustments. Integrating ML into these algorithms involves challenges such as adapting to unique operational frameworks, ensuring data quality, and managing computational demands. Future research should explore the broad application of machine learning models in various optimization contexts, integrate real-time data for adaptive strategies, and investigate advanced machine learning methodologies. This trajectory aims to significantly enhance the problem-solving capabilities of metaheuristics, leading to more optimized solutions in diverse and complex optimization scenarios.

5 Conclusions

This study represents a pioneering effort to blend machine learning with metaheuristics, specifically through the lens of time-series analysis. Integrating ARIMA and LSTM models into the VNS framework for the CVRP demonstrated the potential to significantly enhance the algorithm's decision-making process. Our preliminary findings pave the way for future research in this direction, promising more efficient and effective solutions in combinatorial optimization's vast and challenging domain. The generalization of this approach to other metaheuristics holds substantial promise, heralding a new era in the development of optimization strategies.

References

1. Abdelhafez, A., Luque, G., Alba, E.: Parallel execution combinatorics with metaheuristics: Comparative study. Swarm and Evolutionary Computation **55**, 100692 (2020)
2. Brimberg, J., Salhi, S., Todosijević, R., Urošević, D.: Variable neighborhood search: The power of change and simplicity. Computers & Operations Research **155**, 106221 (2023)
3. CVRPLIB - all instances. http://vrp.atd-lab.inf.puc-rio.br/index.php/en, (Accessed on 01/02/2024)
4. Dantzig, G.B., Ramser, J.H.: The truck dispatching problem. Management science **6**(1), 80–91 (1959)
5. Dickey, D.A., Fuller, W.A.: Distribution of the estimators for autoregressive time series with a unit root. Journal of the American statistical association **74**(366a), 427–431 (1979)
6. Hendel, G.: Adaptive large neighborhood search for mixed integer programming. Mathematical Programming Computation pp. 1–37 (2022)
7. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural computation **9**(8), 1735–1780 (1997)
8. Kalatzantonakis, P., Sifaleras, A., Samaras, N.: Cooperative versus non-cooperative parallel variable neighborhood search strategies: a case study on the capacitated vehicle routing problem. Journal of Global Optimization **78**(2), 327–348 (2020)
9. Kalatzantonakis, P., Sifaleras, A., Samaras, N.: On a cooperative VNS parallelization strategy for the capacitated vehicle routing problem. In: Matsatsinis, N., Marinakis, Y., Pardalos, P.M. (eds.) Learning and Intelligent Optimization (LION 13). vol. 11968, pp. 231–239. Springer (2020)
10. Kalatzantonakis, P., Sifaleras, A., Samaras, N.: A reinforcement learning-variable neighborhood search method for the capacitated vehicle routing problem. Expert Systems with Applications **213**, 118812 (2023)
11. Mladenović, N., Hansen, P.: Variable neighborhood search. Computers & Operations Research **24**(11), 1097–1100 (1997)
12. Monteiro, A.C.B., França, R.P., Arthur, R., Iano, Y.: The fundamentals and potential of heuristics and metaheuristics for multiobjective combinatorial optimization problems and solution methods. In: Multi-Objective Combinatorial Optimization Problems and Solution Methods, pp. 9–29. Academic Press (2022)
13. Talbi, E.G.: Machine learning into metaheuristics: A survey and taxonomy. ACM Computing Surveys (CSUR) **54**(6) (Jul 2021)
14. Thevenin, S., Zufferey, N.: Learning variable neighborhood search for a scheduling problem with time windows and rejections. Discrete Applied Mathematics **261**, 344–353 (2019)