Journal of Information and Technology
ISSN Online 2617-3573
Stratford
Peer Reviewed Journals & books

A Comprehensive Explainable Framework for Designing Enhanced Deep Learning Models

Kudakwashe Dandajena, Proffesor Isabella M. Venter (PhD), [3]Dr Mehrdad Ghaziasgar (PhD) & [4]Reg Dodds

ISSN: 2617-3573

Stratford Peer Reviewed Journals and Book Publishing
Journal of Information and Technology
Volume 7||Issue 1||Page 47-57||June||2023|
Email: info@stratfordjournals.org ISSN: 2617-3573

A Comprehensive Explainable Framework for Designing Enhanced Deep Learning Models

***[1]Kudakwashe Dandajena, [2]Proffesor Isabella M. Venter, [3]Dr Mehrdad Ghaziasgar & [4]Reg Dodds**

[1]Department of Computer Science, Faculty of Science, University of the Western Cape, Private Bag X17 Bellville Cape Town 7535, South Africa

[2]Department of Computer Science, Faculty of Science, University of the Western Cape, Private Bag X17 Bellville Cape Town 7535, South Africa

[3]Department of Computer Science, Faculty of Science, University of the Western Cape, Private Bag X17 Bellville Cape Town 7535, South Africa

[4]Department of Computer Science, Faculty of Science, University of the Western Cape, Private Bag X17 Bellville Cape Town 7535, South Africa

***Email of corresponding author: 3986658@myuwc.ac.za**

How to cite this article: **Dandajena, K., Venter, I., M., Ghaziasgar, M. & Dodds, R. (2023). A Comprehensive Explainable Framework for Designing Enhanced Deep Learning Models. *Journal of Information and Technology*, 7(1), 47-57.** https://doi.org/10.53819/81018102t3086

Abstract

Many deep learning models that show improved efficacy over current state-of-the-art models are built using ad-hoc design strategies. In this study, a framework was developed to enhance the explainability of deep learning models. The framework systematically explains each step involved in enhancing existing models so that users can understand, replicate and trust them. A design science research methodology was used to develop the framework to identify ambiguities and knowledge gaps in current approaches. Experimentation enhanced current deep learning models. The results of this study revealed that enhancing state-of-the-art deep learning models for prediction is made possible by using the suggested framework. Furthermore, the steps to achieve this are easy to comprehend. The main contribution of this study is the design of an explainable deep learning framework using a repeatable and understandable strategy that researchers can follow for improving state-of-the-art prediction models.

Keywords: *Artificial intelligence ethics, time series prediction, irregular sequential patterns, machine learning models and deep learning framework.*

1.0 Introduction

According to Alkhatib and Bernstein (2019), explainable artificial intelligence (XAI) leads to artificial intelligence (AI) systems that are trustworthy, transparent, and ethically acceptable (Suresh & Guttag, 2021). They are not the only researchers that believe that innovators should be aware of critical ethical issues such as bias, fairness, explainability, and trustworthiness and that AI systems are needed that can interact ethically with other AI systems, with humans as well as function ethically in society (Romanov et al., 2019). This paper considers explainability an integral part of building ethical AI systems and an explainable framework that enhances models that are better understood and trusted will be discussed (Siau & Wang, 2020). Even as recent as two decades ago, the notion that one would carry around a device with a comprehensive set of sensors coupled with state-of-the-art AI-powered tracking and profiling capabilities was unthinkable and unacceptable. Today, the cellphone—exactly such a device—is an integral part of daily life.

A large amount of information is collected from the cellphone user and analyzed to construct a comprehensive behavioral profile of the individual, including shopping habits, locations visited, time spent at locations, social norms, habits, and screen time. This is just one example of some contexts that demonstrates the need for an ethical framework to ensure the use of ethically grounded AI. The increase in data generated by high-end miniaturized and hyper-connected technologies of the fourth industrial revolution has accelerated developments towards super-intelligent systems such as recent and current complex large language models (LLMs) such as chat generative pre-trained transformer (ChatGPT) by OpenAI. This has led to the financial growth of AI, according to Schwab (2017) and Zhou et al. (2019). AI technology is already worth billions and is anticipated to grow considerably. It is claimed by Grandview (2022) that the size of the global AI market, currently valued at $93.5 billion, will grow at an annual rate of 38% to 45% in the next decade. With the considerable monetary gains that AI can achieve, its ethical implementation of it can perhaps be compromised.

1.1 Statement of the Problem

A wide range of designs and arrangements of deep learning models have been suggested to date. Most models with higher performance or enriched efficacy—that improve on state-of-the-art models—are typically developed through random or ad-hoc design strategies (Dandajena et al., 2020). These trial-and-error methods have led to high levels of model design bias and poor transparency (Chang et al., 2018), poor interpretability (Kuleshov et al .2018), literature inadequacy, contradiction, and inconsistency with one another (Cerqueira et al., 2018). No systematic strategy or process with explainable steps that can be used to improve the design of state-of-the-art models ethically could be found in the current scientific literature (Kearns et al., 2019).

1.2 Research objective

To determine a comprehensive explainable framework for designing enhanced deep learning models.

2.0 Theoretical and Empirical Literature

AI is a system that can read external input data accurately, learn from it, and use what it has learned to achieve specific goals (Sepp¨al¨ et al., 2021). Daily life mainly depends on these tools and systems in various fields, such as health, education, business, and socio-economic services (Romanov et al., 2019). It is worth noting that the underlying data-driven technologies—algorithms, and models—are created by designers. In most cases, the designer's interest does not necessarily align with those of the users. Furthermore, users at different levels of sophistication, indicating the need for an ethical approach (Whittlestone et al., 2019), are using AI systems, in the form of machine learning models, in society. Ethical AI should be used for the greater good—it should not inflict harm to any being—human or animal (Whittlestone et al., 2019). AI moral standards or ethics are complex issues that may cripple the AI industry's innovation, adoption, and development (Siau & Wang, 2020).

Several approaches have been suggested for improving ethics within the domain of AI (Thamik & Wu, 2022). Ethical guidelines for AI were introduced in many countries guided by national, economic, development, and competitive interests. This includes the Japanese Society for Artificial Intelligence's Ethical Guidelines (2017), the Canadian Montreal Declaration on Responsible AI (2017), the United Kingdom House of Lords' five AI principles for a cross-sector AI code of conduct (2018), Google's AI ethics principles (2019), Australia's Ethics Framework (Dawson et al., 2019). European Commission's High-Level Expert Group on AI guidelines (2019), the Chinese Development Plan for New Generation of AI of July 2017 (Xu et al., 2019), and the Institute of Electrical and Electronics Engineers (IEEE) general AI principles (Whittlestone et al., 2019). The sphere of artificial intelligence ethics was altered in April 2017 by a $50 million investment by the United States of America's Defense Advanced Research Projects Agency (Xu et al., 2019). This demonstrates that states, institutions, and individuals accept the strategic importance of AI ethics by establishing guidelines for societal, security, and economic good (Hagendorff, 2020).

While earlier research has been done to improve the performance of AI models based on accuracy and efficiency, explainability is rarely considered part of the multidimensional performance evaluation criteria used to assess such models' performance (Dandajena et al., 2020). The explainability of a machine-learning model is often inversely proportional to its prediction accuracy, i.e., the better the prediction accuracy is, the lower the model explainability (Xu et al., 2019). Understanding machine-learning models has become complicated since many existing articles contradict one another (Dandajena et al., 2021). The bulk of existing deep learning approaches can be classified as black boxes since they do not provide enough explanation of how these models function or are derived (Aivodji, 2019). This raises issues of bias, suitability, and transparency that reduce the chances of attaining human trust in AI systems (Suresh & Guttag, 2021).

3.0 Methodology

Using design science research (DSR) as a methodology focused on producing an artifact, such a framework was developed. The DSR research strategy comes with standard guidelines to describe and justify the methodologies and tools chosen for this research. It has a robust loop-back function to accommodate changes at various stages during the execution cycle. Knowledge creation and contribution are overall. This makes it an ideal research technique to improve AI systems explain ability, reproducibility, and trustworthiness characteristics

(Venable et al., 2017). As a result, this study views explain ability as a value chain for AI comprising a set of those characteristics. The framework was deployed to enhance current deep-learning models for predicting time series.

In the digital world of the twenty-first century, which generates vast amounts of data, artificial intelligence (AI) technologies are being integrated into various decision-making procedures. Many facets of society, including sensitive ones, are deploying and using this AI technology. Both creators and users of AI technology need to comprehend how AI functions, how it performs tasks, and why a given decision was made. This will give the critical AI technology community the knowledge to review the technology's operation and conclusions. Explain ability is an essential subject of AI technology that fosters the creation of acceptable, responsible, transparent, repeatable, inclusive, and dependable systems. A trade-off exists between achieving algorithm explain ability and maintaining performance robustness (Xu et al., 2019). More straightforward and less robust algorithms are easier to explain. Research by Dignum (2017) indicated that understanding the ethical implications of AI systems should be a priority of any deep learning development (Gonen & Goldberg, 2019). This work argues that a state-of-the-art framework can only be said to be optimal with a comprehensive explanation. The path to explainable models begins with a simplified description. The framework developed in this study provides a transparent and traceable localized explanation of enhancing a model for predicting time series. This ensures trustworthiness when such systems are deployed (Arrieta et al., 2020). The proposed framework illustrated in Figure 1 is a tool that can be used to analyze and manage risks related to constructing deep learning models. It tailors explain ability at the framework's various steps, increasing the understanding of the models designed. This section demonstrates how each framework step contributes to creating an enhanced explainable deep learning model (Avodji et al., 2019). Furthermore, the framework will be applied to a specific domain, and the subsequent results will be provided.

Step 1—Explainable identification

The first step of the framework, illustrated in Figure 1, is a systematic literature review blended with grounded theory. Relevant recent peer-review papers should be listed during this literature review, considering the most current articles. This process should be governed by structured identification, cleansing, eligibility, and inclusion criteria (Dandajena et al. 2020). This step will produce a baseline data bank for further exploration.

Step 2—Explainable exploration, evaluation, and selection

This step involves a comprehensive quantitative and qualitative analysis of the data collected in Step 1. It assists with the selection of appropriate datasets as well as models, and evaluation methods from articles identified in Step 1. It is important not to be biased when choosing datasets and models that could be used in the next step.

Step 3—Explainable implementation

An implementation algorithm with discrete steps should be designed for applying it to the datasets, models and evaluation methods identified in the previous step. The aim of this algorithm should be to recreate baseline models. These baseline models will be used to evaluate and validate the development of an enhanced model. Analytical information generated from core research articles and the selection of baseline models should provide ground truth insights. Algorithm 1 in Section 5 of Annexures provides procedural guidance of this step.

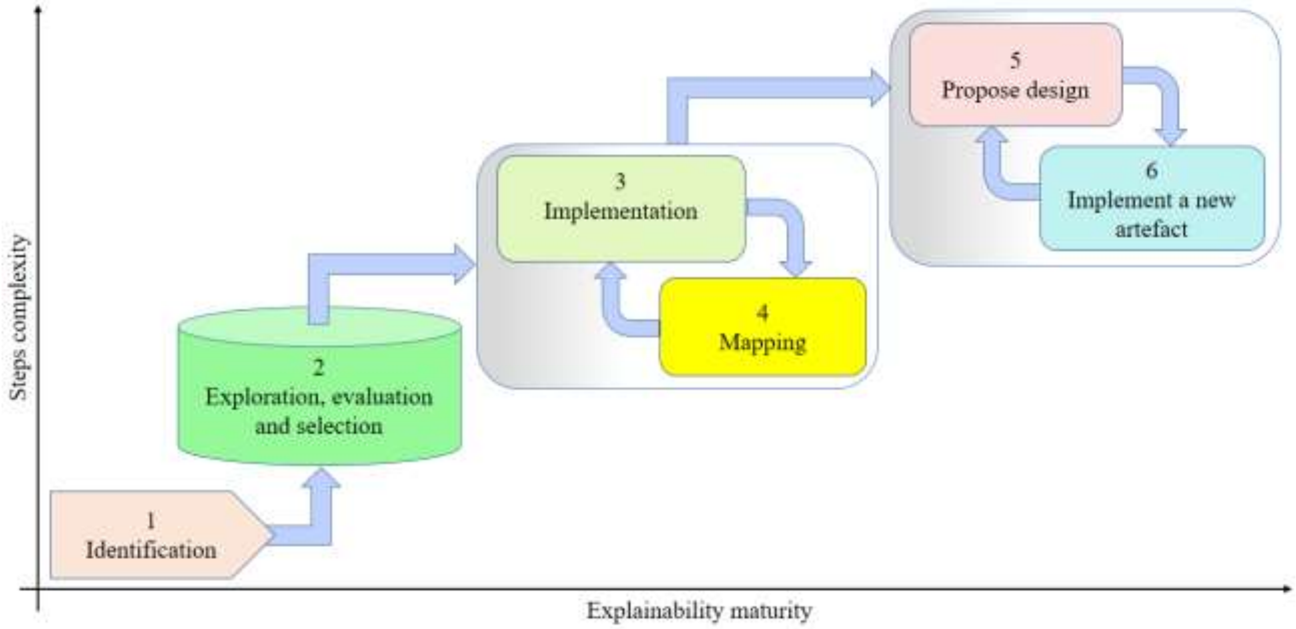


Figure 1: Maturity of explainability in six steps of the framework

Step 4—Explainable mapping

This step entails compiling and tabulating all variables and intrinsic elements connected with the different ways in which the architectures of the baseline models might be used to build an improved model. These variables should be considered as candidate components, to derive an enhanced model by leveraging existing models. Steps 3 and 4 should be applied iteratively.

Steps 5 and 6—Propose, design, and implement a new explainable model

Using the algorithm in Step 3 together with a table of variables produced in Step 4, Step 5 entails the creation of an enhanced model. Step 6 implements the model created in Step 5. As can be seen in Figure 1, Steps 5 and 6 form a loop-back mechanism which tracks every adjustment made during the enhancement process. In Step 6 the enhanced model is deployed and compared with the selected state-of-the-art baseline models to determine whether there is an improvement. The process continues until a satisfactory outcome has been obtained or a predefined maximum number of iterations has been reached.

4.0 Findings and Discussion

To demonstrate the usefulness of the proposed framework, it was applied to the domain of financial currency prediction. The systematic literature review produced 412 recent peer-reviewed articles using 11 online databases. Of these, 32 articles were deemed appropriate for the research. The box and whisker plot and Billauer's algorithm were used to identify forex time series datasets with the most irregular patterns (Alhatib et al., 2019). The daily currency exchange rate data of the GB pound versus the US dollar from 1990 to 2016 had the highest number of irregular patterns and was thus selected as the training dataset. The daily currency exchange rate data of the Japanese yen versus the US dollar was identified as having the second most irregular patterns and thus was used as an unseen validation dataset for testing model performance (Chang et al., 2018). From the 32 articles, 69 models with 34 deep learning architectures were identified—for more detail, see the Github link in Section 5. Of these, a pool of the 12 best-performing baseline models was identified for implementation—see Table 1.

Table 1: List of baseline deep learning models identified in Step 2 of the proposed framework

Model	Architecture	Remarks
1.	LSTM(32) + Dropout (0.2) + Dense (1)	Derived from Azlan et al. (2019) and Mihaita et al. (2019).
2.	LSTM(32) + LSTM(64) + Dropout(0.2) + LSTM(128) + Dropout(0.5) + Dense(1)	Influenced by Glenski et al. (2019) and Chalvatzisa et al. (2019).
3.	Bi(LSTM(50)) + Dense(10) + Dense(10) + Dense(1)	A gated LSTM suggested by Sardelicha and Manandhara (2018).
4.	Bi(GRU(50)) + Dense(10) + Dense(10) + Dense(1)	A gated GRU mentioned by Sardelicha and Manandhara (2018).
5.	LSTM(100) + Dropout(100) + Attention(SeqSelfAttention(32)) + LSTM(16) + Dense(10) + Dense(10) + Dense(1))	Derived from experiments by Huang (2019).
6.	LSTM(32) + Conv1D(32) + Dropout(0.2) + Conv1D(16) + Conv1DTr(16) + Dropout(16) + Conv1D(32) + Conv1D (16) + Attention(SeqSelfAttention(1)) + LSTM(16) + Dropout(0.2) + Dense (1)	As indicated by Makinen et al. (2018) and Huang (2019).
7.	LSTM(32) + Dropout(100) + Attention(SeqSelfAttention(32)) + LSTM(16) + Dense(10) + Dense(10) + Dense(1)	As implemented by Liu (2018).
8.	LSTM(32) + Dropout(0.2) + Attention(SeqSelfAttention(32)) + Bi(LSTM(32)) + Bi(LSTM(32)) + Dense(10) + Dense(1)	Demonstrated by Sardelicha and Manandhara (2018).
9.	LSTM(32) + Conv1D(32) + Dropout(0.2) + Conv1D(16) + Conv1DTranspose(16) + Dropout(0.2) + Conv1DTranspose(32) + Conv1DTranspose(1) + GRU(32) + Dropout(0.5) + Dense(1)	Suggested by Maggiolo and Spanakis (2019).
10.	GRU(32) + GRU(64) + Dropout(0.2) + GRU(128) + Dense(1)	Designed by GRU by Qin (2019).
11.	LSTM(32) + LSTM(64) + RepeatVector(64) + LSTM(64) + TimeDist(1) + LSTM(128) + Dropout(128) + Dense(1)	Suggested by Qin (2019).
12.	LSTM(50) + Dropout + LSTM(100) + Dropout(0.5) + GRU(100) + LSTM(100) + Dropout(0.5) + LSTM(100) + Dropout (0.5) + Dense(100) + Dense(10) + Dense(10) + Dense(1)	Implemented by Bai (2019).

Most of these models were based on recurrent neural network variants, dilated convolutional neural networks, attention mechanisms, bi-directional mechanisms, and other architectures. Multidimensional performance evaluation criteria were adopted to assess the performance of the enhanced model concerning baseline models.

All variables and intrinsic aspects related to how the baseline models might be used were compiled and tabulated. Creating a knowledge bank of the variables involved in deep learning applications provided a technique for controlling and trimming the number of variables to be evaluated and optimized.

An enhanced prediction model based on bidirectional, gated recurrent units, self-attention mechanism, and long-short-term memory was developed for the currency exchange rate datasets—see Table 2. The framework assisted in designing this enhanced model, which outperformed the state-of-the-art baseline models. The supercomputing resources from the South African Centre for High-Performance Computing were used for these experiments.

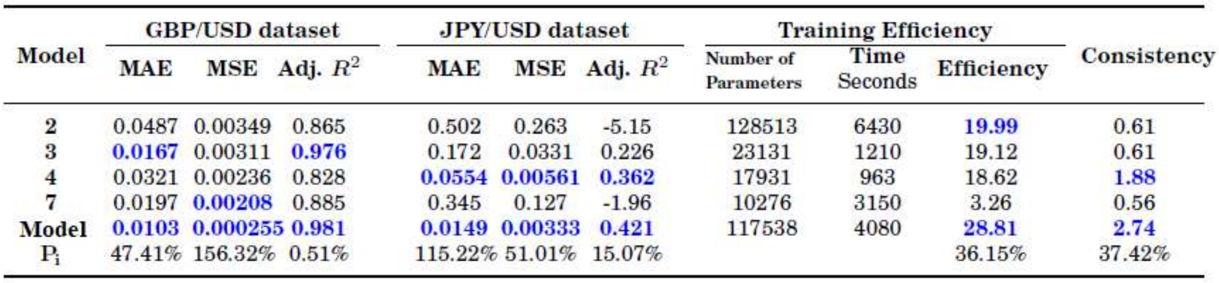
When considering the data in Table 3, the proposed framework produced an enhanced model, which outperforms the highest-performing baseline model for each metric. These results quantified and proved that the produced model best on the test set, where the enhanced model outperforms the best baseline models by 47% in mean absolute error (MAE) and 156% in mean squared error (MSE) while providing a slight increase in adjusted r-squared (R2).

Table 2: Architectural design of an enhanced model produced by the proposed framework.

Model	Architecture	Remarks
Model	SeLFISA BiD(GRU(32)) + SeqSelfAtt(att width=30) + Dropout(0.2) + BiD(LSTM(32)) + BiD(GRU(32)) + BiD(LSTM(32)) + BiD (GRU(32)) + LSTM (32) + GRU(32) + Dense(1)	Proposed enhanced model.

The gain in MSE is noteworthy given that the MSE is sensitive to outliers, and an increase in MSE shows that the enhanced model is significantly more robust to outliers, i.e. irregular patterns. The same model outperforms the baseline models on the unseen currency exchange rate testing dataset set of JPY/US. For this data set, there was a 115% improvement in MAE, 51% in MSE, and 15% in R2. This is supported by the best consistency performance value of 2.74, about 1.5 times greater than the most consistent baseline model, Model 4, with a performance consistency of 1.88. This promising outcome indicates that the proposed framework has created a model adaptable to an unknown dataset in the same format as the training set.

Table 3: Results of top-performing baseline models compared with enhanced model.



Model	GBP/USD dataset			JPY/USD dataset			Training Efficiency			Consistency
	MAE	MSE	Adj. R^2	MAE	MSE	Adj. R^2	Number of Parameters	Time Seconds	Efficiency	
2	0.0487	0.00349	0.865	0.502	0.263	-5.15	128513	6430	**19.99**	0.61
3	**0.0167**	0.00311	**0.976**	0.172	0.0331	0.226	23131	1210	19.12	0.61
4	0.0321	0.00236	0.828	**0.0554**	**0.00561**	**0.362**	17931	963	18.62	**1.88**
7	0.0197	**0.00208**	0.885	0.345	0.127	-1.96	10276	3150	3.26	0.56
Model	**0.0103**	**0.000255**	**0.981**	**0.0149**	**0.00333**	**0.421**	117538	4080	**28.81**	**2.74**
P_i	47.41%	156.32%	0.51%	115.22%	51.01%	15.07%			36.15%	37.42%

The blue numbers in Table 3, Model 4, show that it is the best-performing model for predicting the JPY/USD dataset. Figure 2 is the performance of Model 4's prediction, whereas Figure 3 shows the performance of the enhanced model for predicting the JPY/USD dataset.

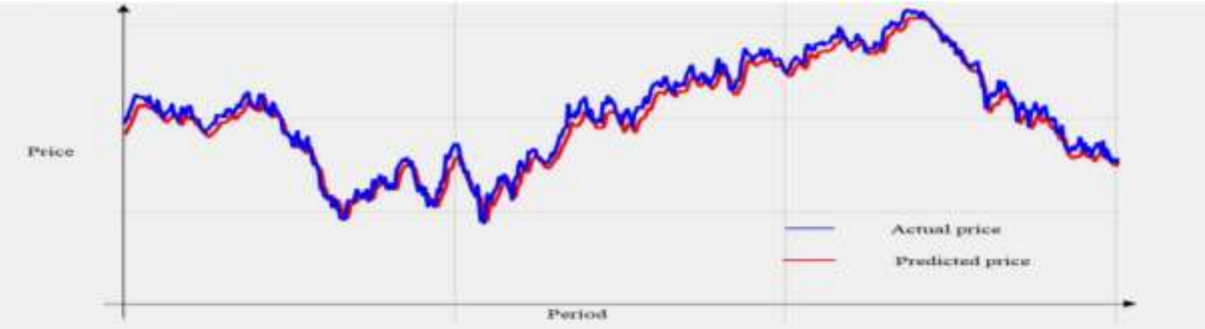


Figure 2: Model 4—Bi (GRU (50)) + Dense (10) + Dense (10) + Dense (1) by Sardelicha and Manandhara (2018)

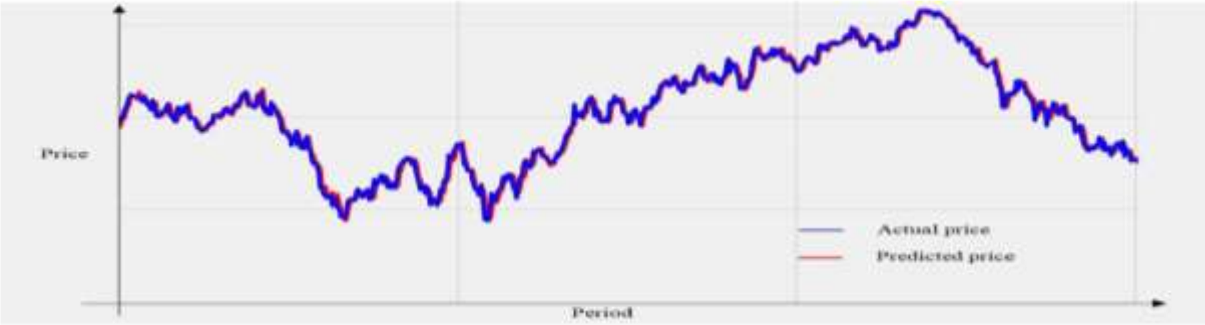


Figure 3: Enhanced Model—Bi (GRU (32)) + SeqSelfAtt(30) + Dropout(0.2) + Bi(LSTM(32)) + Bi(GRU(32)) + Bi(LSTM(32)) + Bi(GRU(32)) + LSTM(32) + GRU(32) + Dense(1)

5.0 Conclusion

In conclusion, the study emphasized the significance of explainability in deep learning models and presented a framework that enhances transparency and interpretability. Predicting the framework for financial currency resulted in an enhanced model that demonstrated design consistency, superior performance, and improved trustworthiness. The framework offers researchers a reliable and trusted tool for developing effective deep-learning models by addressing the need for clear principles and explanations. Further, this work contributes to the ongoing discussion on AI ethics, emphasizing the importance of establishing comprehensive guidelines to ensure the ethical use of AI technologies for the common good.

REFERENCES

A¨ıvodji, U., Arai, H., Fortineau, O., Gambs,S., Hara, S., and Tapp, A., (2019). *Fairwashing: the risk of rationalization*, Proceedings of the 36th International Conference Machine Learning Research, vol. 97, 161–170.

Alkhatib, A. and Bernstein, M. (2019). *Street-level algorithms: A theory at the gaps between policy and decisions*, Proceedings of Conference on Human Factors in Computing Systems, Glasgow, UK, 1-13.

Arrieta, A. B., Rodr´ıguez, N. D., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garc´ıa, S. , Gil-Lopez, S., Molina, D., Benjamins, R. , Chatila, R.. and Herrera, F.(2020).

Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI, Information Fusion, vol. 58, 82–115.

Azlan, A., Yusof, Y., and Mohsin, M. F. M., (2019). *Determining the impact of window length on time series forecasting using deep learning*, International Journal of Advanced Computer Research, vol. 9, no. 44, pp. 260– 267.

Bai S., and Koltun, J. Z. K. V. (2019), *Deep equilibrium models*," 33rd Conference on Neural Information Processing Systems (NeurIPS 2019), (Vancouver), pp. 1–16.

Cerqueira, V., Torgo, L., and Soares, C., (2019). *Machine learning vs statistical methods for time series forecasting*: Size matters, arXiv:1909.13316, vol. [stat.ML], 1–9.

Chalvatzisa, C., and Hristu-Varsakelis, D. (2018). *High-performance stock index trading: making effective use of a deep long short-term memory network*, arXiv:1902.03125, vol. [q-fin.ST], pp. 1–30.

Chang, Y.-Y., Sun, F.-Y., Wu, Y.-H., and Lin, S.-D., (2018). *A memory-network based solution for multivariate time-series forecasting*, Association for the Advancement of Artificial Intelligence, vol. [cs.LG], 1–8.

Dandajena, K., Venter, I. M., Ghaziasgar, M., and Dodds, R, (2020). *Complex sequential data analysis: A systematic literature review of existing algorithms*, Conference of the South African Institute of Computer Scientists and Information Technologists 2020, (Cape Town), pp. 44–50, 2020.

Dandajena, K., Venter, I. M., Ghaziasgar, M., and Dodds, R., (2021). *Selecting datasets for evaluating an enhanced deep learning framework*, SATNAC 2021, (Drakensberg, Natal), pp. 1–7, 2021.

Dawson, D., Schleiger, E., Horton, J., McLaughlin, J., Robinson, C., Quezada, G., Scowcroft, J. and Hajkowicz, S., (2019). *Artificial intelligence*: Australia's ethics framework, CSIRO, Australia, Tech. Rep.

Dignum, V. (2018). *Ethics in artificial intelligence: introduction to the special issue*, Ethics and Information Technology, vol. 20, 1–3.

Dignum,V. (2017). *Responsible autonomy*, Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI-17), Macao, 4698–4704.

Glenski, M., Weninger, T., and Volkova, V., (2019). *Improved forecasting of cryptocurrency price using social signals*, arXiv:1907.00558, vol. [q-fin.ST], pp. 1–11.

Gonen, H. and Goldberg, Y. (2019). *Lipstick on a pig: Debiasing methods cover up systematic gender biases in word embeddings but do not remove them*, Conference of the North American Chapter of the Association for Computer Linguistics, vol. 1, Minneapolis, MN, 609–614.

Grand View Research, (2022). *Market Analysis Report*, Grand View Research, Tech. Rep.

Hagendorff, T. (2020). *The ethics of AI ethics—an evaluation of guidelines*, Minds and

Machines, vol. 30, no. 1, 99–120.

Huang, S., Wang, D., Wu, X. and Tang, A. (2019). DSANet: *Dual self-attention network for multivariate time series forecasting*, CIKM '19: Proceedings of the 28th ACM International Conference on Information and Knowledge Management, (Beijing), pp. 2129–2132, ACM.

Kearns, M. J., Roth, A., and Sharifi-Malvajerdi, S., (2019). *Average individual fairness: Algorithms, generalization and experiments*, arXiv:1905.10607, vol. [CoRR], 1–46.

Kuleshov,V., Fenner, N., and Ermon, S. (2018). *Accurate uncertainties for deep learning using calibrated regression*, Proceedings of the 35th International Conference Machine Learning Research, Stockholm, Sweden, vol. 80, 1–9.

Liu, J., Zhang, T., Han, G., and Gou, Y., (2018). *TD-LSTM: Temporal dependencebased LSTM networks for marine temperature prediction*, Sensors, vol. 18, no. 3697, pp. 1–13.

Maggiolo, M., and Spanakis, G., (2019). *Autoregressive convolutional recurrent neural network for univariate and multivariate time series prediction*, arXiv:1903.02540, vol. [cs.LG], pp. 1–8.

Makinen,M., Kanniainen, J. Gabbouj, M., and Iosifidis, A., (2018), *Forecasting of jump arrivals in stock prices: New attention-based network architecture using limit order book data*," arXiv:1810.10845, vol. [q-fin.TR], pp. 1–29.

Mihaita, A. S., He, H. Li, Z., and Rizoiu, M.-A., (2018). *Motorway traffic flow prediction using advanced deep learning*, arXiv:1907.06356, vol. [cs.LG], pp. 1–10.

Philippe, E., and Carlos, A., (2012). Time-series data mining, ACM Computing Surveys, vol. 45, 18–26.

Qin,H. (2019). *Comparison of deep learning models on time series forecasting: a case study of dissolved oxygen prediction*, arXiv:1911.08414, vol. [eess.SP], pp. 1–16.

Romanov, A., De-Arteaga, M., Wallach, H. M., Chayes, J. T., Borgs, C., Chouldechova, A. , Geyik, S. C., Kenthapadi, K., Rumshisky, A., and Kalai, A. T. (2019). *What's in a name? Reducing bias in bios without access to protected attributes*, Conference of the North American Chapter of the Association for Computer Linguistics, Mineapolis, MN, 1–10.

Salujaa, R. Malhi, A., Knapiˇc, S., Fr¨amling, K., and Cavdar, C. (2019). *Towards a rigorous evaluation of explainability for multivariate.*

Sardelicha, M., and Manandhar, S., (2018). *Multimodal deep learning for short term stock volatility prediction*. arXiv:1812.10479, vol. [q-fin.ST], pp. 1– 40.

Schwab, K. (2017). *The Fourth Industrial Revolution*. New York, NY: Crown Business, 2017.

Sepp¨al¨a, A. Birkstedt, T., and M¨antym¨aki, M. (2021). From ethical AI principles to governed AI, 42nd International Conference on Information Systems, Austin, TX, 1–18.

Siau, K. and Wang, W. (2020). *Artificial intelligence (AI) ethics: Ethics of AI and ethical*, Journal of Database Management, vol. 31, no. 2, 74–87.

Suresh, J. and Guttag, V. J. (2021). *A framework for understanding sources of harm throughout the machine learning life cycle*, Equity and Access in Algorithms, Mechanisms, and Optimization '21, NY, USA, 1-9.

Thamik, H., and Wu, J., (2022). *The impact of artificial intelligence on sustainable development in electronic markets*, Sustainability, vol. 14, no. 7, 1–20.

Venable, J.R, Pries-Heje,J and Baskerville, R.L, (2017). *Choosing a design science research methodology*, Australasian Conference on Information Systems (ACIS) 2017 Proceedings, pp. 1–12.

Whittlestone, J., Nyrup, J., Alexandrova, A., and Cave, S. (2019). *The role and limits of principles in ai ethics: Towards a focus on tensions*, Proceedings of the Conference on AI, Ethics, and Society (AIES '19), Honolulu, HI, 195–200.

Xu, F., Uszkoreit, H., Du, Y., Fan, W., Zhao, D., and Zhu, J. (2019). *Explainable AI: A brief survey on history, research areas, approaches and challenges*, Natural Language Processing and Chinese Computing, Dalian, China, 563–57.

Zhou, F., Zhou, H. M., Yang, Z. and Yang, L. (2019). EMD2FNN: *A strategy combining empirical mode decomposition and factorization machine based neural network for stock market trend prediction*, Expert Systems with Applications, vol. 115.