Complex Sequential Data Analysis: A Systematic Literature Review of Existing Algorithms

Kudakwashe Dandajena, Isabella M. Venter, Mehrdad Ghaziasgar and Reg Dodds
Department of Computer Science, University of the Western Cape
Bellville, Western Cape
kudadandaz@gmail.com,{iventer,mghaziasgar,rdodds}@uwc.ac.za

ABSTRACT

This paper provides a review of past approaches to the use of deep-learning frameworks for the analysis of discrete irregular-patterned complex sequential datasets. A typical example of such a dataset is financial data where specific events trigger sudden irregular changes in the sequence of the data. Traditional deep-learning methods perform poorly or even fail when trying to analyse these datasets. The results of a systematic literature review reveal the dominance of frameworks based on recurrent neural networks. The performance of deep-learning frameworks was found to be evaluated mainly using mean absolute error and root mean square error accuracy metrics. Underlying challenges that were identified are: lack of performance robustness, non-transparency of the methodology, internal and external architectural design and configuration issues. These challenges provide an opportunity to improve the framework for complex irregular-patterned sequential datasets.

CCS CONCEPTS

• **Computing methodologies** → **Supervised learning by regression**; *Feature selection*; *Artificial intelligence*; **Modeling methodologies**.

KEYWORDS

irregular patterns, time series forecasting, parameter, volatile financial prediction, state-of-the-art, extreme weather forecasting, sequential learning and financial signal processing

ACM Reference Format:
Kudakwashe Dandajena, Isabella M. Venter, Mehrdad Ghaziasgar and Reg Dodds. 2020. Complex Sequential Data Analysis: A Systematic Literature Review of Existing Algorithms. In *Conference of the South African Institute of Computer Scientists and Information Technologists 2020 (SAICSIT '20), September 14–16, 2020, Cape Town, South Africa.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3410886.3410899

SAICSIT '20, September 14–16, 2020, Cape Town, South Africa

ACM ISBN 978-1-4503-8847-4/20/09.
https://doi.org/10.1145/3410886.3410899

1 INTRODUCTION

The art of improving the performance of any deep-learning framework is a process of iterated refinements. Currently, there is no single ideal framework that addresses the discontinuous, impulsive and irregular patterns of behaviour associated with irregular-patterned complex sequential datasets [19, 42]. These extreme datasets can be found in many different domains, including: health care, traffic, finance, such as stock prices, meteorology, such as rainfall data and so forth. An automated artefact, beyond the conventional, suitable for solving prediction and regression problems for such datasets will be useful for engineers and academics [35].

This paper considers the advances made towards the design and application of frameworks for irregular-patterned complex sequential analysis based on recent scholarly work.

It was found that there are many deep-learning frameworks aimed at the improvement of the analysis of sequential datasets but that there is no single ideal framework for the analysis of irregular-patterned complex sequential datasets and that the frameworks that were developed were not extensively evaluated using multidimensional evaluation mechanisms [13]. Frameworks based on recurrent neural network (RNN) architecture centred on long short-term memory (LSTM) [27] have been widely identified as the most suitable approach towards addressing unstable sequential behaviour [35]. It is empirically clear that the architectural designs of most present state-of-the-art sequential frameworks are simple extensions of the original LSTM architecture [11]. Most of these are equipped with gating mechanisms for solving vanishing gradient problems.

An exhaustive literature search of all deep-learning related materials is not possible. The influence of the foundational work by Hochreiter and Schmidhuber in 1997 on the original LSTM architecture can be seen in many of these recent research studies [14]. Sequential forecasting competitions have also contributed to the development of present state-of-the-art benchmarks, methodologies, theories and datasets [20]. Some of the most influential competitions include the M1 to M4 Competitions by Makridakis and Hibon, the Sante Fe competitions by the Santa Fe Institute, the knowledge discovery and data mining (KDD) cup competitions by the Association for Computing Machinery's Special Interest Group on KDD, the Kaggle time series competitions by Goldbloom, the global energy forecasting competitions by Tao Hong and the International Journal of Forecasting [17].

Algorithm performance evaluation plays a critical role in the design of improved frameworks. Many performance evaluation criteria or mechanisms have been suggested. These include: efficiency, accuracy, consistency, reliability, stability, explainability, baseline comparison [32]. Most sequential deep-learning frameworks are evaluated using only accuracy as a criterion based on root mean square error as a metric [21].

It seems as if design and configuration issues of internal-, external- and hyper-parameters hinder the optimal performance of existing frameworks for extreme datasets. This points to the need for the development of newer and more transparent frameworks that reveal performance robustness in these environments. It is important to appreciate performance strengths and weaknesses of existing novel frameworks on known datasets, before suggesting any new design or architectural improvements to existing frameworks [26].

Using a systematic literature review, recent research publications were identified based on the following specific inclusion criteria: keywords, publication timelines, algorithms or framework relevance in terms of complex datasets, accreditation and citation quality of the journal.

The survey explored the following questions:
(1) How efficient are existing deep-learning algorithms for analysing complex sequential datasets?
(2) How should current deep-learning algorithms be adapted to deal with irregular-patterned complex sequential datasets?

The primary aims of this systematic literature survey were to:
(1) Identify well-known state-of-the-art deep-learning frameworks for complex sequential analysis.
(2) Identify the challenges in current methods of analysing irregular-patterned complex sequential datasets.

This paper highlights the current state of affairs of deep-learning frameworks for the transparent analysis of irregular-patterned complex sequential datasets and their challenges. Transparency in this instance refers to explainable frameworks or models that can be easily and independently replicated at any given time through experiments.

2 MATERIALS AND METHODS

A systematic literature review, shown in Figure 1, was the preferred methodology to demonstrate the breadth and depth of the existing body of knowledge of deep-learning frameworks but also to identify inconsistencies and gaps in this body of knowledge [39].

According to Peter et al. [23], applying a systematic literature review strategy offers a comprehensive, focused, reliable, repeatable and thorough literature overview [1, 24, 37].

Referring to Figure 1, the systematic literature review methodology used in this research had the following phases:

Stage 1—Identification of specific keywords to search for irregular-patterned complex sequential analysis, as well as identifying appropriate online research database platforms, for example Google Scholar. The identified keywords were used to search for published accredited peer reviewed journal articles on online research database platforms [24]. The title of the journal article was used to identify which articles to consider and these were then stored for later evaluation.

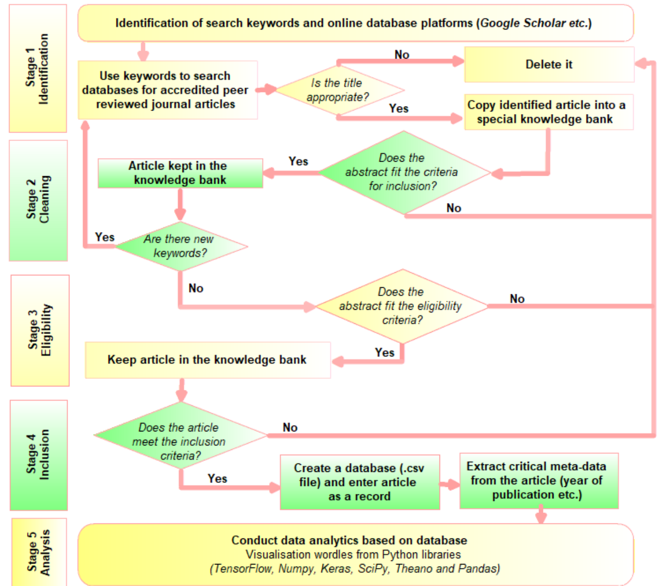

Figure 1: Systematic literature review methodology

Stage 2—Cleaning of the comprehensive list of the identified journal articles that were collected during Stage 1. This stage entailed a qualitative evaluation of the fitness of each article based on the abstract of the article. The articles that did not pass the criteria for inclusion were then deleted. If new keywords were found in the abstracts, these were added to the initial list of keywords and Stage 1 was repeated for these new keywords, in a *grounded theory fashion* [34].

Stage 3—Eligibility of the articles was identified in Stage 2. A detailed qualitative screening exercise was applied to the output of Stage 2, based on specific eligibility criteria [31]. This stage produced a comprehensive list of articles—those that did not adhere to the eligibility criteria were discarded.

Stage 4—Inclusion of the state-of-the-art articles was based on inclusion criteria. Articles identified in this stage were recorded in a database. For this stage the whole article was considered. Critical meta-data was qualitatively identified and recorded in fields, for example the year of publication, each with an appropriate column header. The articles not included in this database were now discarded. The output of this stage was a `.csv` file.

Stage 5—Analysis of the database using a quantitative method. This empirical stage employed data analytic operations using Python libraries such as TensorFlow, Numpy, Keras, SciPy, Theano and Pandas, within the Jupyter Notebook environment. This stage mapped the relationship between key features of the records across identified fields. During this stage the relationship of records was visualised as *wordles*, or word clouds. Finally, the used codes, tools

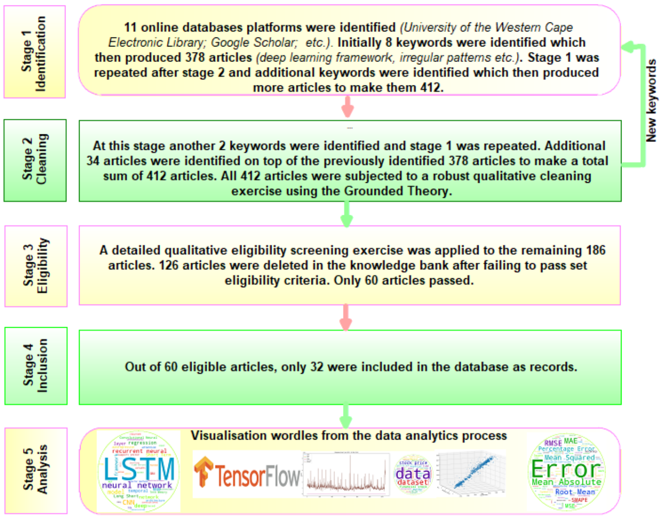

Figure 2: Processing the results of the systematic review

and the database of identified literature was uploaded on an online platform to allow free access to other researchers.[1]

3 RESULTS

The results are segmented into two sections. The first section highlights the methodological implementation results which feeds into our last main results section which provides the overall research findings of the systematic literature review.

3.1 Methodological implementation results

Figure 2 summarises the phases of the systematic literature review to provide the results of the implemented methodology.

Stage 1—Initially eight search keywords were identified. These were: deep-learning framework, sequential algorithm optimisation, irregular patterns, time series forecasting, parameter, volatile financial prediction and extreme weather forecasting. After the cleaning process of Stage 2, two more keywords were identified, namely sequential learning and financial signal processing. Thus 10 keywords in total were used for identifying articles. The following 11 online research database platforms were used for the search: University of the Western Cape Electronic Library, Google Scholar, CiteSeerx, GetCITED, Microsoft Academic Research, Bioline International Directory of Open Access Journals, PLOS ONE, Papers with Codes, BioOne, Science and Technology of Advanced Materials, New Journal of Physics, ScienceDirect and NIPS. This stage produced a total of 412 search results in the form of a comprehensive list of peer reviewed articles. We focused on the results not exceeding the first 10 pages of each online database platform. These articles were stored for further processing.

Stage 2—Abstracts of 378 articles identified in Stage 1 were subjected to a robust qualitative cleaning exercise using the *grounded theory approach* [34] which revealed two further keywords that were put through Stage 1 again to identify more possible articles. This produced 34 more articles. Eventually 226 of the 412 articles were found not to be the right fit for the study and were removed—186 articles were kept.

Stage 3—The remaining 186 articles were evaluated based on eligibility criteria: identified keywords, publication timelines from 2016 to date, algorithms or framework relevance in terms of complex datasets, accreditation and quality of the journal, i.e., its citation index [32]. Abstracts of each article were read in detail. A further 126 articles were trimmed off after failing to meet the eligibility standards—60 articles were kept.

Stage 4—All 60 articles in the folder were now scrutinised and read in depth. An excel `.xlsx` database was created to capture a record of each article that was deemed to be eligible.[2] Critical meta-data from the article were recorded in identified fields such as: journal source web link, journal name, journal title, authors, pages, timelines (day, month, and year), editor, volume, issue number, city, country, continent, standard number, day accessed, month accessed, year accessed, data set, data set type, data set sources, dataset description, research problem, research objective, implementation framework, architecture properties, baseline models, best models, methodology, evaluation mechanism or criteria, evaluation metric, results, novelty, future recommendations and gaps comment. Tabulated data fields were critically identified to suit empirical data analytics operations for visualisation which would be implemented in Stage 5. A total of 28 ineligible articles were deleted in Stage 4 of the process since they were not a direct fit to the aim of the literature review study. Only 32 articles were recorded in the database and the rest were deleted. The `.xlsx` database file was further cleaned and converted to a `.csv` file.[3]

Stage 5—A program was implemented to analyse the database—the cleaned `.csv` result of Stage 4. The output of this program was the visualisation of the analysis in the form of graphs, wordles and schematic diagrams. Finally, the code, tools and the literature search database were uploaded to GitHub to allow open access of all experimental study material and code to other researchers. The literature research database can be freely accessed from the GitHub website. Schematic diagrams were recorded as annexures.

3.2 Systematic literature review results

The most important finding from this study is the fact that current algorithms cannot yet optimally analyse complex datasets. In most of the recent state-of-the-art papers, Chinese publications dominate deep learning framework research. The novelty of existing deep-learning frameworks is largely centred on their application mechanism. From these publications it appears that researchers normally design an architecture and then uniquely apply it to a particular problem

[1] All the experimental code is given in the Jupyter Notebook files on the GitHub website at: https://github.com/Dandajena/SDA/.

[2] Accessible at: https://github.com/Dandajena/SDA/blob/master/Database.xlsx.

[3] Accessible at: https://github.com/Dandajena/SDA/blob/master/Database.csv.

together with well-known architectures, such as generic recurrent neural networks and long short-term memory to determine which of the architectures is the most optimal for the specific application. We next discuss the results in terms of identified challenges, frameworks, datasets and their evaluation.

3.2.1 Complex sequential analysis challenges. The analysis of complex sequential datasets—characterised with spontaneous or volatile behaviour—is far from attaining a maturity status. According to the literature, most modern deep-learning frameworks cannot address many of the following challenges:

- Complexity in modelling and capturing extremely long-term sequential patterns using traditional deep-learning algorithms such as RNN [15]. These algorithms lack transparency and explainability within the implementation of deep-learning models [5].
- The analysis of highly variable, noisy and volatile datasets (the problematic aspects of sequential datasets) Ma et al. (2019) [19], Zhang et al. (2018) [43] leads to performance disadvantages such as consistency or inconsistency, sensitivity to outliers, extreme values and computing inefficiency [3, 4].
- Most state-of-the-art deep-learning frameworks [9] experience efficiency performance problems when exposed to different sequential datasets [36].
- Model reliability and generalisation, is a problem when predictive neural frameworks are used. These associated problems are caused by the stochasticity of stock features in financial stock price datasets [10].
- Precision challenges when forecasting within financial environments. This is associated with extreme sequential financial market datasets [26].
- Lack of accurate, reliable, and interpretable modern deep-learning models for uncertainty estimation over continuous variables [18].
- Lack of a comprehensive comparison analysis of existing deep-learning models for sequential learning. Recent research focused on one-step forecasting, based on smaller datasets [27].
- Accurate sequential forecasting is a challenge when using existing uncertainty estimation models. This is a problem when dealing with its probabilistic formulation which is difficult to tune, scale and it adds exogenous variables, i.e., other variables outside the existing variables [44].
- The simultaneous forecasting of the inflow and outflow of crowds in regions of a city is complex because of spatial dependencies, temporal dependencies and external influence factors [42].
- The complexity of sequential stock price datasets which require extensive analysis resources [40].
- The deficiency of traditional models towards the capturing of complex nonlinear or dynamic dependencies between time steps and between multiple time series [16].
- Predictive performance challenges associated with existing models and the lack of well-established and explainable literature for sequential predictive machine learning methods [38].
- Dealing with robust and accuracy challenges which are associated with existing sequential forecasting [22]
- The ever-growing requirement of computing power, time and resources of sequential forecasting modelling. This is particularly associated with unstable extreme weather patterns [6].
- Performance deficiency of sequential neural network models [13] when forecasting in multivariate dataset environments [7].

3.2.2 Identified sequential deep-learning frameworks. To address some of the aspects of these research challenges and gaps, various researchers have applied different techniques and methods to find solutions. The resources needed to deploy deep-learning frameworks with good performance, is a problem as well as poorly configured internal and external parameters and hyperparameters for such analysis. Existing research work focuses on the need to further explore and advance existing deep-learning models. Some of these are: long short-term memory, adversarial LSTM, CapsNets, LSTM-convolutional neural networks, gated recurrent unit, and attention, bidirectional and temporal convolutional neural networks [9].

There is a need to explicitly combine the aspect of sequential dataset complexity as an optimisation technique in the design of more advanced deep-learning predictive algorithms or models or frameworks. These frameworks require a multi-dimensional evaluation mechanism which considers existing complex datasets [33].

In terms of on-going research, most researchers are citing the need to develop state-of-the-art optimised algorithms that address challenges associated with complex sequential environments. There is a wide range of sequential analysis frameworks as illustrated in Annexure 1 which were designed to resolve these challenges.

Figure 2 points to the fact that deep-learning frameworks, based on the recurrent neural network architecture, long short-term memory by Hochreiter and Schmidhuber [14] dominate the field of sequential forecasting. These gated architectures address exploding and vanishing gradient problems associated with neural networks. Gated architectures are made up of input, forget and output gates or modules to provide them with sequential learning capabilities which decide which critical information to keep or discard during sequential modelling. In sequential modelling, Tang et al. [35] indicated that LSTM networks are excellent for capturing features with longer sequences or time span unlike the gated recurrent unit of Cho et al. [8].

According to Zhang et al. [43] there are other competing modern frameworks that can deal with this problem effectively. These are: deep convolutional neural networks (CNNs), capsule neural network, dilated RNN, dilated CNN, ensemble CNN-LSTM-attention, DEQ-trellisNet and DEQ-transformer, attention-based mechanisms by Young, et al. [41], bidirectional recurrent neural networks by Schuster and

Figure 3: Results of state-of-the-art deep-learning framework for sequential challenges

Paliwal [29], temporal convolutional networks by Bai et al.(2018) [2], deep Bayesian neural network, deep sequential spatio-temporal residual neural network, dual self-attention network and memory-based ordinal regression deep neural networks.

The attention mechanism of any deep-learning framework has the ability to select hidden states and patterns within a particular dataset which makes them more attractive for modelling complex situations. This mechanism may even produce an unexpected performance when combined with unidirectional, bidirectional or multidirectional mechanisms. The sequential robustness of any framework is correlated with the nature of the problem set being resolved, the selected dataset and the evaluation criteria.

It is clear that the art of identification, selection, configuration, deployment and evaluation of any deep-learning framework has not been exhausted. Furthermore many of these frameworks have not been explained adequately and thus the experimental work is unclear.

3.2.3 Complex sequential datasets. The majority of existing sequential data types listed in Annexure 2 were either univariate or multivariate or both and were sourced from different domains such as: traffic, financial stock markets, meteorological weather and climate information, energy consumption, natural language sentiment processing, telecommunications, astronomy, etc.

Figure 4 shows experimental data analysis results showing irregular sequential patterns of the financial dataset from NASDAQ stock market (2 January 2012–26 December 2016) by [26] and the wordle analysis of sequential datasets.

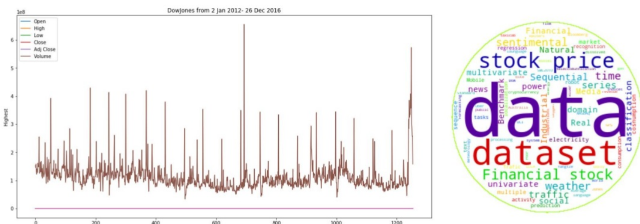

Figure 4: Irregular patterns of the financial dataset

3.2.4 Evaluation. Performance evaluation of most modern deep-learning frameworks is based on one of the following mechanisms: prediction accuracy, efficiency, correlation, baseline, consistency, visualization sharpness, robustness and computational complexity. There is no single framework that has been evaluated in terms of its multidimensional performance in which critical issues such as efficiency, accuracy, consistency, reliability, stability and transparency were considered.

In terms of performance metrics, there are over 40 metrics listed in Annexure 3 which were applied in the evaluation processes of the different frameworks. The systematic literature review process revealed that several of the accuracy metrics were most often used. The majority of these accuracy metrics are categorised and measured based on absolute, squared, relative, symmetric and percentage error types Bal et al. [3]. Mean absolute error (MAE), mean square error (MSE) and root mean squared error (RMSE) dominated this space since they are easy to compute [28]. RMSE and MSE in particular have sensitivity challenges associated with the stability of the selected frameworks [35]. Any results that yield the best (lowest) RMSE value, demonstrate the stability of such an algorithm or framework as they are less sensitive to outliers.

Selection of the correct performance evaluation metrics is central to the identification of a proper framework for any given problem. A particular framework can be deemed sub-optimal because the wrong evaluation metrics were chosen [12]. When comparing the performance of state-of-the-art baseline frameworks it may be necessary to consider applying different combinations of evaluation criteria.

3.3 Existing recommendations

Future work recommendations from current studies point to the need to improve on: the interpretability of modern deep-learning models, resolving prediction problems associated with volatile time series data, and lowering performance computing costs. They also suggest the exploration of modern deep-learning frameworks such as: adversarial networks, temporal convolution networks, transformer networks, and CapsNets, by combining them with other architectures such as the dilated CNNs. Furthermore model optimisation should be explored through its manipulation of both internal and external architectural properties, deploying future frameworks on complex univariate and multivariate cases with exogenous environments.

4 DISCUSSION

Existing work on deep-learning frameworks for sequential analysis has produced a large pool of publications covering wide range of issues such as methods, experimental design, optimisation techniques, input signal issues in the form of datasets, application areas and theories and many others, however, there are important systematic gaps, limitations and inconsistencies in these studies. It is clear that researchers have different objectives when designing and analysing deep-learning frameworks. This creates an unnecessary discord in identification of a systematic evaluation methodology.

The deep-learning algorithms listed in Annexure 1 are suboptimal and not efficient in analysing complex sequential environments. They lack transparency, interpretability, and their performance evaluation is not exhaustive. These issues have not yet been adequately documented and publications contradict one another. For example how can any experiment conclude that the designed state-of-the-art framework architecture is the best when it only uses performance accuracy based on one criterion?

Historical sequential datasets from financial stock markets have produced understandable volatile, non-linear and chaotic characteristics. The financial stock market domain is very sensitive to changes such as the Covid-19 pandemic. It is highly correlated to these financial events. This makes the domain an interesting environment for developing an enhanced deep-learning framework for accurate analysis of sequential irregular patterns.

It is possible that a deep-learning framework that has been trained on complex datasets, when exposed to normal or ordinary environments, will outperform other frameworks in terms of robustness and efficiency.

There is room for exploring better optimised sequential frameworks that have the potential of producing a better sequential analysis performance. Current deep-learning algorithms can be adapted to analyse irregular-patterned complex sequential datasets as a means of improvement. These new approaches need to focus on a more detailed cross cutting approach that address efficiency, accuracy, consistency and reliability issues.

5 ANNEXURES

The annexures can be found on the GitHub website. https://github.com/Dandajena/SDA/blob/master/Annexures.pdf

6 CONCLUSION AND FUTURE RECOMMENDATIONS

The goal of this research was to determine which deep-learning frameworks are currently being used to analyse irregular-patterned complex sequential datasets. Using a systematic literature research methodology the issues associated with the performance of well-known state-of-the-art deep learning frameworks for irregular-patterned complex sequential analysis and their respective challenges were identified.

It was found from existing literature that several researchers feel that deep-learning algorithms are suboptimal and not efficient in analysing complex sequential datasets [25]. It is thus clear that there is a need to improve the existing performance evaluation methods into a unified multidimensional evaluation method. It cannot be claimed that a state-of-the-art framework is optimal without applying an extensive, transparent and traceable performance evaluation procedure on the results of such a framework.

To address this deficiency it is worth considering a combination of: (1) selecting a proper algorithm architecture and redesigning it; (2) sensitively tuning it; as well as (3) evaluating its performance in a multidimensional mode based on complex irregular-patterned complex sequential datasets.

This will provide a potential approach towards the development of a new breed of robust deep-learning frameworks which are efficient, accurate, consistent, reliable, stable and transparent. This can only be achieved by selecting an existing sequential dataset on which research has already been done as a way of minimising the experimental variables and parameters. It will allow comparison.

The systematic literature review results show that the financial stock market domain—particularly from well-established financial markets—provides irregular-patterned sequential datasets. Furthermore, it is important to rank their volatility before selecting and adopting these datasets in complex sequential modelling experiments. Improving the design of algorithms based on such extreme complex sequential datasets provides much needed theoretical, methodological and experimental contributions on the performance exploration of deep-learning frameworks. Extreme scenarios are associated with big data and it has engulfed our lives. Faster analysis through state-of-the-art frameworks may offer the much needed scientific, engineering, academic and business solutions to make the world a better place [30].

Finally, to enrich a future experimental setup, it is advisable to consider some best properties of the sequential recurrent neural networks proposed by Zhang et al. [43] as baseline architecture. This process could be coupled with other novel architectures such as temporal convolution networks, transformer networks, CapsNets, unidirectional as well as attention, bidirectional or multidirectional mechanisms.

ACKNOWLEDGMENTS

This work is supported by Research Committee of the University of the Western Cape and the Telkom/Aria Technology Africa Centre of Excellence.

REFERENCES

[1] D. Andreini and C. Bettinelli. 2017. *Business Model Innovation: From Systematic Literature Review to Future Research Directions.* Springer, Cham, Switzerland.
[2] S. Bai, J.Z. Kolter, and V. Koltun. 2018. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv:1803.01271v2* [cs.LG] (2018), 1–14.
[3] C. Bal, S. Demir, and C.H. Aladag. 2016. A comparison of different model selection criteria for forecasting EURO/USD exchange rates by feed forward neural network. *Int. Journal of Computing, Communications and Instrumentation Engineering* 3, 2 (2016), 271–275.
[4] C. Chalvatzisa and D. Hristu-Varsakelis. 2019. High-performance stock index trading: making effective use of a deep long short-term

memory network. *arXiv:1902.03125v2* [q-fin.ST] (2019), 1–30.
[5] Y.-Y. Chang, F.-Y. Sun, Y.-H. Wu, and S.-D. Lin. 2018. A memory-network based solution for multivariate time-series forecasting. *arXiv:1809.02105v1* [cs.LG] (2018), 1–8.
[6] A. Chattopadhyay, E. Nabizadeh, and P. Hassanzadeh. 2020. Analog forecasting of extreme-causing weather patterns using deep learning. *Journal of Advances in Modelling Earth Systems* 12, 2 (2020), 1–14.
[7] Y. Chen, Y. Kang, Y. Chen, and Z. Wang. 2019. Probabilistic forecasting with temporal convolutional neural network. *arXiv:1906.04397v3* [stat.ML] (2019), 1–27.
[8] Kyunghyun Cho, Bart van Merriënboer, Dzmitry Bahdanau, and Yoshua Bengio. 2014. On the properties of neural machine translation: Encoder-decoder approaches. *arXiv:1409.1259v2* [cs.CL] (2014), 1–9.
[9] K. Dashtipour, M. Gogate, A. Adeel, C. Ieracitano, H. Larijani, and A. Hussain. 2018. Exploiting deep learning for Persian sentiment analysis. *arXiv:1808.05077v1* [cs.CL] (2018), 1–9.
[10] F. Feng, H. Chen, J. He, X.and Ding, M. Sun, and T.-S. Chua. 2019. Enhancing stock movement prediction with adversarial training. *arXiv:1810.09936v2* [q-fin.TR] (2019), 1–7.
[11] K. Greff, R. Srivastava, J. Koutnik, B. Steunebrink, and J Schmidhuber. 2017. LSTM: A search space odyssey. *Transactions on Neural Networks and Learning Systems* 2 (2017), 1–12.
[12] A. Gunawardana and G. Shani. 2009. A survey of accuracy evaluation metrics of recommendation tasks. *Journal of Machine Learning Research* 10 (2009), 2935–2962.
[13] H. Hewamalage, C. Bergmeir, and K. Bandara. 2019. Recurrent neural networks for time series forecasting: Current status and future directions. *arXiv:1909.00590* [cs.LG] (2019), 1–51.
[14] S. Hochreiter and J. Schmidhuber. 1997. Long short-term memory. *Neural Computation* 9, 10 (1997), 1735–1780.
[15] Y. Hua, Z. Zhao, R. Li, X. Chen, Z. Liu, and H. Zhang. 2018. Deep learning with long short-term memory for time series prediction. *arXiv:1810.10161* [cs.NE] (2018), 1–9.
[16] S. Huang, D. Wang, X. Wu, and A. Tang. 2019. DSANet: Dual self-attention network for multivariate time series forecasting. In *CIKM '19: Proceedings of the 28th ACM International Conference on Information and Knowledge Management*. ACM, Beijing, 2129–2132.
[17] R. Hyndman. 2020. A brief history of forecasting competitions. *International Journal of Forecasting* 36 (2020), 7–14.
[18] V. Kuleshov, N. Fenner, and S. Ermon. 2018. Accurate uncertainties for deep learning using calibrated regression. *arXiv:1807.00263* [cs.LG] (2018), 1–9.
[19] X. Ma, P. Karkus, D. Hsu, and W. Lee. 2019. Particle filter recurrent neural networks. *arXiv:1905.12885v1* [cs.LG] (2019), 1–16.
[20] S. Makridakis, E. Spiliotis, and V. Assimakopoulos. 2020. The M4 Competition: 100,000 time series and 61 forecasting methods. *International Journal of Forecasting* 36 (2020), 54–74.
[21] Nijat Mehdiyev, David Enke, Peter Fettke, and Peter Loos. 2016. Evaluating forecasting methods by considering different accuracy measures. *Procedia Computer Science* 95 (2016), 264–271.
[22] B.P. Orozco, G. Abbati, and S.J. Roberts. 2018. MOrdReD: Memory-based ordinal regression deep neural networks for time series forecasting. *arXiv:1803.09704v4* [stat.ML (2018), 1–30.
[23] M. Peter, T. Diekötter, and K. Kremer. 2019. Participant outcomes of biodiversity citizen science projects: A systematic literature review. *Sustainability* 11, 10 (2019), 1–18.
[24] C. Pickering, J. Grignon, R. Steven, D. Guitart, and J. Byrne. 2015. Publishing not perishing: How research students transition from novice to knowledgeable using systematic quantitative literature reviews. *Studies in Higher Education* 40, 10 (2015), 1756–1769.
[25] S. Pouyanfar, S. Sadiq, Y. Yan, H. Tian, Y. Tao, M. Reyes, M.-L. Shyu, S.-C. Chen, and S. Iyengar. 2018. A survey on deep learning: Algorithms, techniques, and applications. *Comput. Surveys* 51, 5 (2018), 1–36.
[26] X.-Y. Qian. 2017. Financial series prediction: Comparison between precision of time series models and machine learning methods. *arXiv:1706.00948v3* [cs.LG] (2017), 1–9.
[27] H. Qin. 2019. Comparison of deep learning models on time series forecasting: A case study of dissolved oxygen prediction. *arXiv:1911.08414v2* [eess.SP] (2019), 1–16.
[28] S. Rasp, P.D. Dueben, S. Scher, J.A. Weyn, S. Mouatadid, and N. Thuerey. 2020. WeatherBench: A benchmark dataset for data-driven weather forecasting. *arXiv:2002.00469v1* [physics.ao-ph] (2020), 1–13.
[29] Mike Schuster and Kuldip K. Paliwal. 1997. Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing* 45, 11 (1997), 2673–2681.
[30] K. Schwab. 2016. *The Fourth Industrial Revolution*. World Economic Forum, Geneva, Switzerland.
[31] K. Sokol and P. Flach. 2020. Explainability fact sheets: a framework for systematic assessment of explainable approaches. In *Fairness, Accountability, and Transparency (FAT20)*. ACM, Barcelona, Spain, 1–22.
[32] K. Sokol and P. Flach. 2020. Explainability fact sheets: A framework for systematic assessment of explainable approaches, fairness, accountability, and transparency. *arXiv:1912.05100* [cs.LG] (2020), 1–22.
[33] Y.-G. Songa, Y.-L. Zhoub, and R.-J. Hanc. 2018. Neural networks for stock price prediction. *Journal of Difference Equations and Applications* 1902.3125v2 [q-fin.ST] (2018), 1–13.
[34] A. Strauss and J. M. Corbin. 1997. *Grounded Theory in Practice*. Sage Publications Inc., Thousand Oaks, Ca.
[35] Q. Tang, M. Yang, and Y. Yang. 2019. ST-LSTM: A deep learning approach combined spatio-temporal features for short-term forecast in rail transit. *Journal of Advanced Transportation* 2019 (2019), 1–8.
[36] V. Umayaparvathi and K. Iyakutti. 2017. Automated feature selection and churn prediction using deep learning models. *International Research Journal of Engineering and Technology* 4 (2017), 1846–1854.
[37] E. van Laar, A. van Deursen, J. van Dijk, and J. de Haan. 2017. The relationship between 21st century skills and digital skills: A systematic literature review. *Computers in Human Behaviour* 72 (2017), 577–588.
[38] C. Vitor, T. Luís, and S. Carlos. 2019. Machine learning vs statistical methods for time series forecasting: Size matters. *arXiv:1909.13316* [stat.ML] (2019), 1–9.
[39] Y. Xiao and M. Watson. 2019. Guidance on conducting a systematic literature review. *Journal of Planning Education and Research* 39, 1 (2019), 93–112.
[40] L. Xinyi, L. Yinchuan, Y. Hongyang, Y. Liuqing, and L. Yang. 2019. DP-LSTM: Differential privacy-inspired LSTM for stock prediction using financial news. In *33rd Conference on Neural Information Processing Systems*. ACM, Vancouver, Canada, 1–9.
[41] T. Young, D. Hazarika, S. Poria, and E. Cambria. 2018. Recent trends in deep learning based natural language processing. *arXiv:1708.02709* [cs.CL] (2018), 1–32.
[42] J. Zhang, Y. Zheng, and D. Qi. 2017. Deep spatio-temporal residual networks for citywide crowd flows prediction. *arXiv:1610.00081* [cs.AI] (2017), 1–8.
[43] Q. Zhang, R. Luo, Y. Yang, and Y. Liu. 2018. Benchmarking deep sequential models on volatility predictions for financial time series. *arXiv:1811.03711* [cs.LG] (2018), 1–14.
[44] L. Zhu and N. Laptev. 2017. Deep and confident prediction for time series at Uber. *arXiv:1709.01907v1* [stat.ML] (2017), 1–8.