computer sciences and mathematics forum

Proceeding Paper

What are Recurrent Expansion Algorithms? Exploring a Deeper Space than Deep Learning †

Tarek Berghout [1,*] and Mohamed Benbouzid [2,3]

[1] Laboratory of Automation and Manufacturing Engineering, University of Batna 2, Batna 05000, Algeria
[2] UMR CNRS 6027 IRDL, University of Brest, 29238 Brest, France; mohamed.benbouzid@univ-brest.fr
[3] Logistics Engineering College, Shanghai Maritime University, Shanghai 201306, China
* Correspondence: t.berghout@univ-batna2.dz
† Presented at the 1st International Online Conference on Mathematics and Applications; Available online: https://iocma2023.sciforum.net/.

Abstract: Machine-learning applications nowadays usually become a subject of data unavailability, complexity, and drift resulting from massive and rapid changes in data Volume, Velocity, and Variety (3V). Recent advances in deep learning have brought many improvements to the field providing, generative modeling, nonlinear abstractions, and adaptive learning to address these challenges respectively. In fact, deep learning aims to learn from representations that provide a consistent abstraction of the original feature space, which makes it more meaningful and less complex. However, data complexity related to different distortions such as higher levels of noise, for instance, remains difficult to overcome and challenging. In this context, recurrent expansion (RE) algorithms are recently unleashed to explore deeper representations than ordinary deep networks, providing further improvement in feature mapping. Unlike traditional deep learning, which extracts meaningful representations through inputs abstraction, RE allows entire deep networks to be merged into another one consecutively allowing exploration of Inputs, Maps, and estimated Targets (IMTs) as primary sources of learning; a total of three sources of information to provide additional knowledge about their interaction in a deep network. Besides, RE makes it possible to study IMTs of several networks and learn significant features, improving its accuracy with each round. In this context, this paper presents a general overview of RE, its main learning rules, advantages, disadvantages, and its future opportunities, while reviewing an important state-of-the-art and some illustrative examples.

Keywords: recurrent expansion; deep learning; machine learning

Citation: Berghout, T.; Benbouzid, M. What are Recurrent Expansion Algorithms? Exploring a Deeper Space than Deep Learning. *Comput. Sci. Math. Forum* **2023**, *2*, x. https://doi.org/10.3390/xxxxx

Academic Editor(s):

Published: 28 April 2023

1. Introduction

Machine learning models generally aim to build prediction models using a set of training samples to be able to generalize them on the coming unseen samples related to the same phenomena [1]. Basically, a well-trained machine learning model should have resilience and adaptability against rapidly varying 3V of big-data [2]. As a result, the current growth in 3V of big-data has led to the emergence of deep learning philosophy in a variety of applications [3]. To better understand the difference between machine learning and deep learning models, we bring this simple mathematical description. If we consider a machine learning model as a function f that approximates a set of training samples x by estimating their targets $\tilde{y}$, then Formula (1) will look very much like a definition of machine learning. Thus, traditional machine learning in this case does not consider the effect of feature representations on the training process, while only focusing on hyperparameters optimization as the main problem of approximation and generalization. Indeed, deep learning comes by emphasizing the consideration of improving feature representations by introducing a well-defined intermediate feature mapping taking into account the 3V issues of big-data. Mathematically, if we assume that a well-structured feature map is defined as $\varphi(x)$, the mathematical representations of the deep network can be demonstrated as in Formula (2).

$$\tilde{y} = f(x) \tag{1}$$

$$\tilde{y} = f(\varphi(x)) \tag{2}$$

$\varphi(x)$ in this case can be any type of deep learning network feature maps, including but not limited to autoencoder (AE) hidden layers of all types [4], recurrent neural network (RNN) hidden layers and its variants [5], convolutional neural network (CNN) feature maps [6], deep belief network (DBN) hidden layers [7].

In fact, deep learning has undergone a massive evolution targeting the current rapid change of 3V of big-data by introducing a well-sophisticated complex nonlinear feature mapping process, adaptive learning features and generative modeling targeting all problems of complexity, dynamism and availability of data.

Despite these massive improvements in deep learning, the continuous massive growth and rapid change of 3V of big-data has always made it difficult to demand new, up-to-date features to make learning systems more adaptive and resilient. The main research gaps in this case are related to the complexity and the dynamism resulting from this massive growth of 3V of big-data.

Accordingly, as an insight to solve these challenges, a recent algorithm called recurrent expansion with multiple repeats (REMR) is introduced in the literature. The REMR algorithm aims to introduce deeper space then deep learning investigating and experimenting new, more meaningful feature representations leading to better performance.

As a main contribution of this paper with respect to the previous research gap analysis and since there is now a literature review introducing this new tool, this paper is introduced to provide a clear state-of-the-art of this new method, its learning rules, while going through various illustrative examples, and finally to illustrate future opportunities.

This paper is organized as follows. Section 2 is devoted to presenting REMR algorithm and its main learning rules through some explanations, illustrations of flowcharts and some important pseudo-codes while comparing the representations of the general REMR formula to the previous ones (1–2). Second, some examples will be brought in Section 3 based on previous work and clarified to highlight the advantages and limitations of REMR algorithms. Section 4 will conclude this paper by discussing some future opportunities.

2. REMR Background

REMR is a representation learning algorithm built on a philosophy that aims to explore and profit from pre-trained deep learning models and their behaviors towards unseen sample predictions when building new models. REMR is previously used to solve learning problems related to imbalanced classification of complex and dynamic data, with missing values [8], as well as for time series predictions in the near future [9,10]. REMR has proven its ability to improve both deep learning and small-scale machine learning algorithms.

Accordingly, REMR algorithms are designed to fully merge deep learning networks into others consecutively in different rounds to build more complex architecture and experiencing new, more meaningful feature representations. Accordingly, and unlike deep learning algorithms, which use inputs x as the main source of information, REMR exploits three main sources, namely inputs x, maps $\varphi(x)$ and estimated targets $\tilde{y}$ previously called IMTs as main training features. Moreover, REMR actually uses IMTs of several trained deep networks, studies their behaviors and builds a new robust and more accurate model. Mathematically, following the previous illustrations in Formulas (1) and (2), REMR can be expressed similarly as in Formula (3). m is the maximum number of training rounds.

Since IMTs, which are referred to as x_{k+1} are generally expected to be massive, especially under multiple layers of deep learning nonlinear mappings, a well-defined data processing ρ as in Formula (4) is required before feeding the next deep network.

$$\tilde{y}_{k+1} = f(x_{k+1})|k = 1 \rightarrow m \quad (3)$$

$$x_{k+1} = \rho([x, \varphi_k(x_k), \tilde{y}_k])|k = 1 \rightarrow m \quad (4)$$

Considering the previous representation of Formula (3) and compared to Formulas (1) and (2), the REMR appears more complex than any other deep network witnessed so far referring to deeper nonlinear abstractions.

The flowchart introduced in Figure 1 is dedicated to displaying a clear illustration of REMR learning rules. This flowchart shows the positions of IMTs and demonstrates the process of collecting, processing, and feeding back IMTs to upcoming deep learning networks.

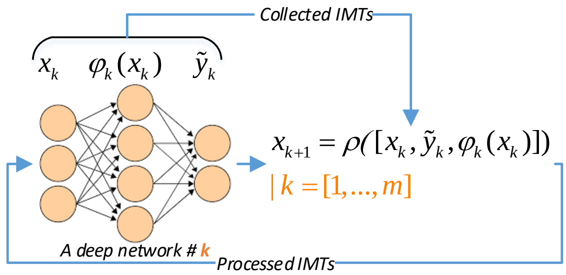

Figure 1. IMTs flows in REMR algorithms.

REMR does not really show an approximation process by hyperparameter optimization which is usually done via gradient descent algorithms whereas it's more like submerging of deep networks that does so [11]. Therefore, the traditional formulas of the loss function will not be useful in evaluating its convergence behavior at every round. In this context, REMR uses a special metric developed on the basis of the behavior of the expected loss function at each round. The REMR philosophy expects that with each round and assuming the performance of the learning process improves, the loss function should exhibit better, smoother and faster loss behavior of each deep network at each next round. These improved loss function characteristics could result in less area under the loss curve (AULC). Therefore, the AULC parameter addressed in Formula (5) is elected as the main metric for evaluating the convergence of the learning process in REMR. ite and l_k represent respectively the maximum number of iterations in each round and the loss function.

$$AULC_k = \int_0^{ite(k)} l_k(x_k)\, dx_k \quad (5)$$

To better understand the REMR algorithm and its main steps, its pseudo-codes are introduced in Algorithm 1.

Algorithm 1. REMR algorithm.

Inputs: x_k, y_k, m, ite_k
Outputs: $\tilde{y}_m$
% Start training process
For $k = 1 \rightarrow m$
% Train the deep network and collect IMTs
$x_{k+1} = \rho_k([x_k, \varphi_k(x_k), \tilde{y}_k]);$
% Evaluate the AULC
$\mathrm{AULC}_k = \int_0^{ite(k)} l_k(x_k)\, dx_k;$

% Reinitialize parameters
$x_k = x_{k+1};$
End(For)

3. Some Illustrative Examples

This section presents two main examples obtained from previous applications of REMR. The first comes from a classification problem [8] the second is related to regression problem [12]. These sections are not really dedicated to display numerical evaluations because some of these results are already revealed in previous works [8–10]. Accordingly, the main objective of this section is to illustrate the advantages of REMR from the point of view of data representations by showing some important data visualizations.

3.1. Classification Problem

The classification problem addressed in this case is closely related to imbalanced complex and dynamic data suffering from higher cardinality level under missing values. These data were subject of REMR training using LSTM network as the main element of recurrent expansion. Information on LSTM hyperparameters can be found at [8]. Figure 2 is devoted to show some important results at this stage. The REMR was unwound for 20 rounds. First, the results in Figure 2a show clear convergences in AULC behavior as expected during the training process. This confirms the first REMR learning hypothesis explained earlier in Section 2. Second, Figure 2b,c are related to processed IMTs of first and last rounds respectively. The purpose of this second illustration is to show the effect of REMR mappings on representations. REMR in this case not only shows feature separation ability, but also shows agglomeration and outlier correction ability.

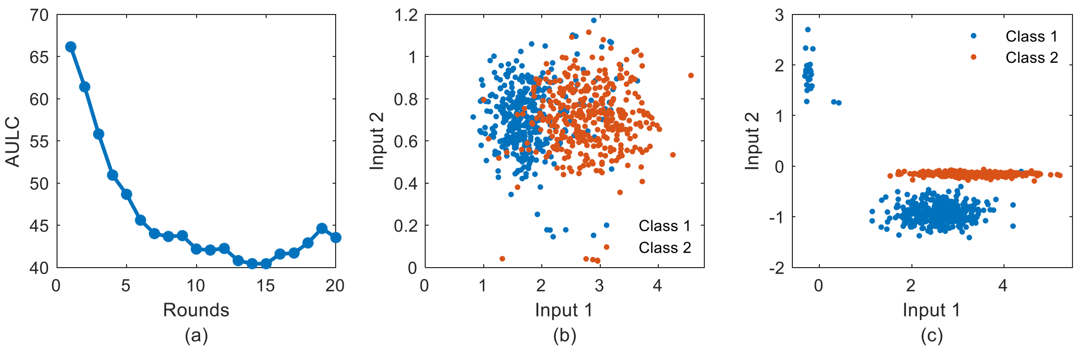

Figure 2. REMR performances for a complex classification problem: (**a**) Behavior of AULC function; (**b**) Data scatters at round 1; (**c**) Data scatters at round 20 (i.e., final round).

3.2. Regression Problems

The regression problem adopted in this work is related to an approximation of a linear deterioration functions based on a set of nonlinear and complex features detailed in [12]. Similarly, LSTM was also involved in this experiment while the hyperparameters are tuned similarly to LSTM network introduced in [12] for 30 rounds. Figure 3 presents results obtained on REMR performance. The AULC function of Figure 3a is still able to carry out a great convergence of REMR algorithm. This also confirms the ability of REMR to fit both regression and classification even if the loss function is of different type. Figure 3b,c are devoted to show an example of deterioration trajectory curve fitting using REMR at the first and last rounds respectively. A 99% confidence interval is used to provide insight into predictions uncertainties. The results show that REMR algorithm attempts at each round to approach desired responses. This reflects the benefits of learning from different IMTs produced from different deep networks to teach the model to behave appropriately. We believe that the reason for this large approximation is closely related to new sources of information (i.e., mainly estimated targets) providing additional knowledge to the system. Also, learning consecutively from different IMTs will be helpful in understanding the effect of changing inputs, maps, and targets respectively on the approximation process leading to producing better results each time based on better understanding of these behaviors.

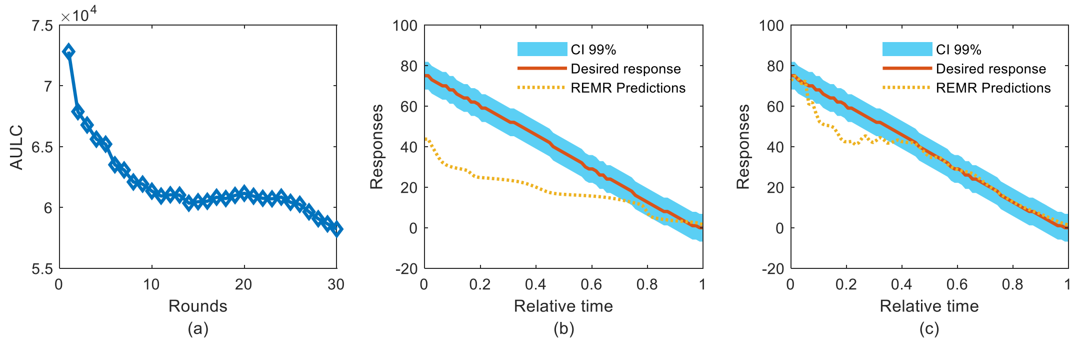

Figure 3. REMR performances for a complex regression problem: (**a**) Behavior of AULC function; (**b**) REMR predictions at round 1; (**c**) REMR predictions at round 30 (i.e., final round).

3.3. Advantages of REMR

To summarize; From previous studies done so far using the REMR algorithm [8–10], REMR shows its ability to improve the learning performance of deep networks through the following aspects:

- Improve feature representations through understanding the different IMTs of different deep networks;
- Providing a new source of information, such as estimated targets for example, helps introduce additional knowledge into the system through a kind of transductive transfer learning;
- The REMR pseudocode in Algorithm 1 shows that building an REMR algorithm does not require much intervention and is simple to design with only a few hyperparameters (i.e., k, m, and ite);
- REMR not only shows the ability to improve feature representations, but also the correction of outliers.

3.4. Disadvantages of REMR

Since it is undeniable that REMR was able to achieve excellent prediction results from both classification and regression, it also experiences some burdens during model reconstructions, including the following:

- When building an REMR model by merging complex deep networks, it means that the computational complexity will be increased and will require more computational power than ordinary deep learning. This means that this architecture is computationally expensive;
- Convergence of the AULC function strongly depends on IMTs initialization IMTs. This means that stacking in the inappropriate IMTs will cause the AULC function to diverge, which could lead to a worse outcome at each round;
- So far, IMTs initialization is evaluated by the basis of errors and trials, which means that it takes a lot of intervention when rebuilding the model at this stage.

3.5. Future Opportunities of REMR

As results of the aforementioned advantages of REMR algorithms, future research opportunities of REMR algorithms revolve around the following:

- Target IMTs initialization problems by performing experiments on IMT selection and optimization in a few primary rounds. This will be useful in deciding whether we go

with this model or not without consuming a lot of computing resources throughout the training rounds;
- Explore available REMR architecture inspired by ensemble learning and parallel architectures to help deliver even better initial IMTs.

4. Conclusions

This paper reviews REMR algorithms through important cutting-edge work. It shows the main learning rules with respect to some illustrative flowcharts and pseudo-codes. Moreover, it illustrates the performance of REMR by going through some important illustrations for both classification and regression. It also has significant pros, cons, and significant future opportunities to work on at this point. The most important conclusion, in this case is, that REMR behaves appropriately as expected if and only if the IMTs are better initialized initially. This means that future opportunities will revolve around it to better improve its performance and application.

Author Contributions: Conceptualization, T.B.; methodology, T.B.; software, T.B.; validation, T.B. and M.B.; formal analysis, T.B.; investigation, T.B.; resources, T.B.; data curation, T.B.; writing—original draft preparation, T.B.; writing—review and editing, T.B. and M.B.; visualization, T.B.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Jordan, M.I.; Mitchell, T.M. Machine learning: Trends, perspectives, and prospects. *Science* **2015**, *349*, 255–260. https://doi.org/10.1126/science.aaa8415.
2. Ivanov, T.; Korfiatis, N.; Zicari, R.V. On the inequality of the 3V's of Big Data Architectural Paradigms: A case for heterogeneity. *arXiv* **2013**, arXiv:1311.0805
3. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.
4. Bank, D.; Koenigstein, N.; Giryes, R. Autoencoders. *arXiv* **2020**, arXiv:2003.05991v2.
5. Yu, Y.; Si, X.; Hu, C.; Zhang, J. A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures. *Neural Comput.* **2019**, *31*, 1235–1270. https://doi.org/10.1162/neco_a_01199.
6. Li, Z.; Liu, F.; Yang, W.; Peng, S.; Zhou, J. A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects. *IEEE Trans. Neural Networks Learn. Syst.* **2022**, *33*, 6999–7019. https://doi.org/10.1109/TNNLS.2021.3084827.
7. Hinton, G. Deep belief networks. *Scholarpedia* **2009**, *4*, 5947. https://doi.org/10.4249/scholarpedia.5947.
8. Berghout, T.; Benbouzid, M.; Ferrag, M.A. Deep Learning with Recurrent Expansion for Electricity Theft Detection in Smart Grids. In Proceedings of the IECON 2022—48th Annual Conference of the IEEE Industrial Electronics Society, Brussels, Belgium, 17–20 October 2022; pp. 1–6.
9. Berghout, T.; Benbouzid, M.; Amirat, Y. Improving Small-scale Machine Learning with Recurrent Expansion for Fuel Cells Time Series Prognosis. In Proceedings of the IECON 2022—48th Annual Conference of the IEEE Industrial Electronics Society, Brussels, Belgium, 17–20 October 2022; pp. 1–5.
10. Berghout, T.; Benbouzid, M.; Bentrcia, T.; Amirat, Y.; Mouss, L. Exposing Deep Representations to a Recurrent Expansion with Multiple Repeats for Fuel Cells Time Series Prognosis. *Entropy* **2022**, *24*, 1009. https://doi.org/10.3390/e24071009.
11. Tian, Y.; Zhang, Y.; Zhang, H. Recent Advances in Stochastic Gradient Descent in Deep Learning. *Mathematics* **2023**, *11*, 682. https://doi.org/10.3390/math11030682.
12. Berghout, T.; Mouss, M.-D.; Mouss, L.; Benbouzid, M. ProgNet: A Transferable Deep Network for Aircraft Engine Damage Propagation Prognosis under Real Flight Conditions. *Aerospace* **2022**, *10*, 10. https://doi.org/10.3390/aerospace10010010.