



CONTRIBUTION TO SALES FORECASTING BASED ON RECURRENT NEURAL NETWORK IN THE CONTEXT OF A MOROCCAN COMPANY

Ghita Rguiga, Nabil Mouttaki, Jamal Benhra

► To cite this version:

Ghita Rguiga, Nabil Mouttaki, Jamal Benhra. CONTRIBUTION TO SALES FORECASTING BASED ON RECURRENT NEURAL NETWORK IN THE CONTEXT OF A MOROCCAN COMPANY. 13ème CONFERENCE INTERNATIONALE DE MODELISATION, OPTIMISATION ET SIMULATION (MOSIM2020), 12-14 Nov 2020, AGADIR, Maroc, Nov 2020, AGADIR (virtual), Morocco. <hal-03192817>

HAL Id: hal-03192817

<https://hal.science/hal-03192817v1>

Submitted on 8 Apr 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

CONTRIBUTION TO SALES FORECASTING BASED ON RECURRENT NEURAL NETWORK IN THE CONTEXT OF A MOROCCAN COMPANY.

Ghita RGUGA

LRI Laboratory, OSIL ENSEM,
Hassan II University of Casablanca,
Casablanca, Morocco
Ghita.rguiga@ensem.ac.ma,

Nabil MOUTTAKI.

LRI Laboratory, OSIL ENSEM,
Hassan II University of Casablanca,
Casablanca, Morocco
Nabil.mouttaki@ensem.ac.ma

Jamal BENHRA

LRI Laboratory, OSIL ENSEM,
Hassan II University of Casablanca,
Casablanca, Morocco
jamalbenhra@gmail.com

ABSTRACT:

Forecasting sales is a prerequisite for successful management, this avoids overproduction and therefore costly storage, the objective of this work is to provide a reliable sales forecasting model to help organizations make strategic and operational decisions. Indeed, this work accounts for the profitability of the use of techniques based on artificial intelligence and more precisely the application of recurrent neural networks for a correct prediction of sales. We initially developed the necessary adjustments of this type of algorithm and thereafter we opted for a justified use of the long short term memory (LSTM) techniques of recurrent neural networks for forecasting sales instead of the conventional forecasting methods already used. In this sense, we have proposed a prediction model based on the time series of previous sales which reacts and readjusts the parameters of each sequence.

KEYWORDS:

Forecasting sales model, Artificial intelligence techniques, Machine learning techniques, Production management, Decision support, LSTM.

1- INTRODUCTION

Business growth is nowadays determined by a critical factor which is sales forecasting. It consumes a large portion of managers' time and focus. The majority of large B2B companies engage in forecasting on a weekly basis. forecasts are serving as the basis for planning and executing supply chain activities such as sourcing, making, and distributing products and services to customers. (Huang T. et al,2019)

Over the years, trends such as the globalization of supply chains, the explosion of product variety, the shortening of product life cycles and increasingly competitive markets have made forecasting more complex, yet its role more critical (Tonya B. et al , 2019)

Digital technologies such as advanced Point-of-Sale (POS) systems, the “Internet of Things” (IoT), and cloud computing (Dey N. et al , 2018) have enabled firms to collect vast amounts of data in real time. Then these vast amounts of data that are collected and analyzed in near real-time have the potential to improve our understanding of customer behavior, improve demand prediction and

better execute supply chain management (Tannu et al.,2017)

Data-driven decision-making is the practice of basing decisions on the data analysis, rather than on intuition (Tom F. 2019) In the use of data-driven decision making, companies in the top third of their industry are, on average more productive and more profitable than their competitors.

Since sales forecasting has become a prerequisite for successful management. This work reports the effectiveness of using Recurrent Neural Network LSTMs for sales forecasting instead of classical forecasting methods already employed, use of time series data, forecasting options, and management of forecasting systems.

This paper is presented in several sections. The following section presents a review of previous works on the use of data analysis techniques for sales forecasting. The case study and the analyzed data are described in the third section. The Methodology section provides a description of the methodology followed and the performance

evaluation criteria applied. The fifth section presents all the results obtained as well as a discussion focused on these results, and finally, the conclusions of the work are presented.

2- LITERATURE REVIEW

To explore the issue of sales forecasting, several projects have been developed by researchers over the years, however, to the best of knowledge, the application of deep learning algorithms in sales forecasting is still incipient with a very limited number of published works that we are going to cite in this paper.

A study of (Bohdan P. et al, 2018) Machine-Learning Models for Sales Time Series Forecasting was conducted with a main goal considering approaches and case studies of using machine learning for sales forecasting. The effect of machine-learning generalization has also been considered. This effect can be used to make sales predictions when there is a small amount of historical data for specific sales time series in the case when a new product or store is launched. A stacking approach for building regression ensemble of single models has been studied. The results show that the performance of predictive models for sales time series forecasting can be improved by using stacking techniques.

Furthermore, a research on (Canton et al, 2019) Sales Prediction through Neural Networks for a Small Dataset has shown the applicability of neural networks in a case where a company reports a short time-series given the changes in its warehouse structure. Given the neural networks independence from statistical, it uses a multilayer-perceptron to get the sales forecasting of this enterprise. thus, learning rates variations do not significantly increase the computing time, and the validation fails with an error minor to five percent which is a high standard.

Then a work was conducted on (Kalaivasan T R et al, 2019) Sales Forecasting using RNN which explains the concept of the system architecture of sales forecasting using recurrent neural networks with variant features, the foremost part is based on sales propelled by the sales administrator of a company. Managing the information on the stock and updating the stock value. The generality of RNN is the continuous form of the working part. It is discreetly used for the sales updating norms, amount of product that has been produced, and the wide

continuity of product sales in the previous year's till on to the date. the accuracy of these sales forecasting enables companies to take important business decisions and can improve in their performance for the profitable sales, the LSTM (Long Short Term Memory) in this work has been used for classifying, processing and making predictions along with the time series, by which the products are been checked for the sales forecasting and reported by checking up with the time series of previous year's sales of the already existing products, this is the continuous range of checking up the product that is sold that this work has proposed.

In our knowledge, these are the works already lead on prediction of sales, therefore these works are still limited since the first study is issued by using a small dataset which makes the data non representative, also as mentioned on the second study, checks are based on the same period of the previous year which doesn't tolerate evolution and changesets, the third study performs on new products which means that there is no history data, also the use of simple regression limits its performance, in this work, forecasting is based on a large amount of records which is necessary for training data using the recurrent neural network LSTMs rather than convolutional networks.

3- CONTEXTUALISATION OF THE STUDY

Since Deep learning is a subset of the field of machine learning, which is a subfield of artificial intelligence (AI), we will discuss in this section the impact of each technology on sales forecasting

3.1- Artificial Intelligence and sales forecasting:

Good sales forecasting helps businesses to grow, But, for many years, forecasting has relied on the human element, emotions and hunches that can make or break a quarter. Just as big data and artificial intelligence (AI) pervade many aspects of how we work, it is doing the same for forecasting. AI equips businesses with improved lead scoring capabilities, enabling them to determine the fate of sales opportunities more accurately, and, ultimately, produce more accurate sales forecasts. AI can sift through large volumes of real-time and historical data and identify the most lucrative sales leads. Using intricate algorithms, AI can fine-tune lead scores by accounting for a multitude of different

factors. Also, AI is driven by machine learning, which means the software trains on data and becomes better over time.

since neural networks have been well proven in the field of prediction, we have chosen to explain their role in sales prediction (Canton,2019).

3.2- Neural Networks and sales forecasting:

However classical models have a significant dependence on statistical assumptions, which diminish their capacity to predict a behavior that was not previously registered. To deal with such problem, Artificial Neural Networks (ANNs) appear as a suitable alternative. These techniques try to mimic neural cells behavior, which makes them non-linear models by nature. Because ANNs are data-driven, they have a self-adaptive feature. This adaptation capacity makes them suitable to analyze problems where data is not complete, or when its documentation was not appropriately gathered (i.e., data with noise). Given that ANNs allow the approximation of almost all non-linear continuous function, they are suitable for sales forecasting. However, in the classical literature, there is not a unique approach to sales forecasting because ANNs may have different structures according to the problem. For example, electricity demand forecasting considers a hybrid structure since data may present seasonal effects

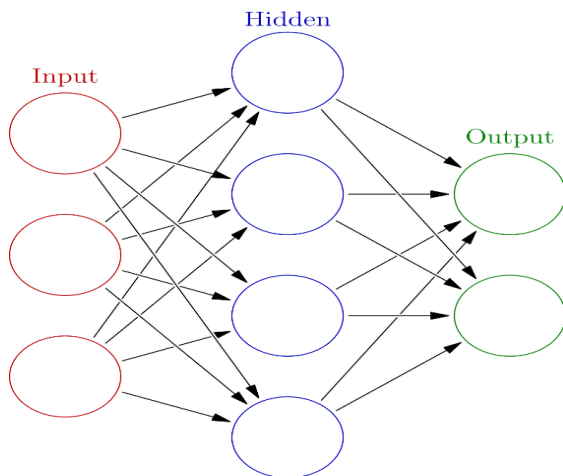


Figure 1 simple neural network graphic

Graphically, an ANN is a graph (figure 1) with neurons (nodes) and their synapses (edges) (Bohdan M.et al, 2019). These models may learn and identify consumers' behavior patterns according to the input data and the architecture of the ANN. The structure

of an ANN summarizes the interconnection between neurons, which are organized in layers. The most uncomplicated architecture considers the input layer, the output layer, and the hidden layer, whose objective is the identification and communication of pattern to get the demand forecasting. ANN with only one hidden layer. The hidden layer serves to adjust the importance of the input data in the determination of future behavior. Therefore, it is possible to add as many hidden layers as required.

Although no theoretical tool indicates the optimal number of hidden layers, experiments demonstrate that more than one hidden layer does not significantly improve the prediction's quality. Also, there is a positive relationship between the number of hidden layers and the computing time. Thus, one hidden layer is enough to get a proper prediction, mainly when the database is small.

yet, it will take both the human and technological sides working together to achieve true success.

3.3- Recurrent Neural Networks and sales forecasting:

A Recurrent Neural Network is basically a standard neural network that has been extended across time by having edges which feed into the next time step instead of into the next layer in the same time step. RNN is designed to recognize sequences, for example, a speech signal or a text. It has cycles inside that implies the presence of short memory in the net

One of the issues of the recurrent neural network is related to sequence as a direct cycle is formed with their connections. The next state can be retained using the current output as input for the next step resembling as running a program with standard inputs and selected internal variables (figure 2) (Sharma V,2019)

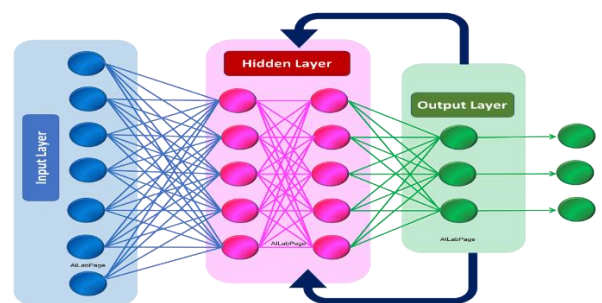


Figure 2 Recurrent Neural Network graphic

Long Short-Term Memory usually just called “LSTMs” (figure 3) (sunny P.,2020). - are a special kind of RNN, capable of learning long-term dependencies. They were introduced by Hoch Reiter & Schmid Huber (1997) (Oikina ,2015), and were refined and popularized by many people in following work. They are widely used on a large set of issued and designed specifically to eliminate long-term dependency. The default behavior LSTM can capture and store information for a longer span of time. The main component of LSTM is the cell state, a horizontal line that keeps running all through. The cell state always runs down the entire chain with less frequent interactions. It is very easy for information to just flow unchanged along with it. LSTM has all the rights to include or exclude the information to the cell state, that are let through by structures called gates. It is a way to decide whether to let the information through or not. It consists of a sigmoid network layer and performs multiplication operation. The output value of the sigmoid layer ranges between zero and one indicating the individual component to decide whether to go through or not. A value zero indicates “let nothing through” and one indicates “let everything through”.

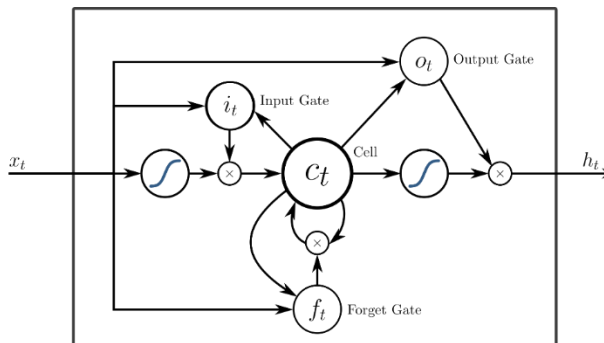


Figure 3 The structure of an LSTM unit

Finally, human side is no more needed to achieve the performance needed since

4- PROBLEM DEFINITION

In this work we are going to present a method for sales forecasting using Recurrent Neural Networks LSTMs based on a data set recorded from a CRM ,and compare the results with prediction made by classical neural network and a recurrent neural network , in this section we are going to present some analysis of the data chosen for experimentation , then

the results given by each method of prediction , a comparative study is leaded as well

4.1- Data preparation

Our dataset represents a two years record of a B2C organization's sales for a specific product, the references have been anonymized to protect personal data, privacy, and the legitimate interest of legal persons (figure 4)

	B	D	E	F	G
	ID_C	Base (dh)	Turnover (dh)	DATE	TYPE
1					
2	2235445	300	320	02/11/2019	4
3	3402737	345	1120	30/10/2019	1
4	3960265	615	1120	30/10/2019	4
5	1933126	1065	1920	30/10/2019	3
6	3596948	1290	2320	30/10/2019	2
7	3643653	80	320	30/10/2019	4
8	2962291	780	1120	30/10/2019	2
9	796964	990	1420	30/10/2019	2
10	3776992	1690	2420	30/10/2019	2
11	925504	2040	2920	30/10/2019	3
12	2773856	1020	1120	30/10/2019	2
13	1368140	1290	1420	27/10/2019	5
14	1333243	2190	2420	27/10/2019	3
15	3959279	2640	2920	27/10/2019	3
16	3864872	48	40	27/10/2019	1
17	1427062	269.58	286.2	19/10/2019	2
18	3162625	210	220	19/10/2019	2
19	3230395	228	240	19/10/2019	4
20	2285745	500	720	19/10/2019	3
21	170822	675	1150	18/10/2019	1
22	1356878	801	1150	18/10/2019	1
23	1745265	1375	1970	18/10/2019	2
24	281904	1662	2380	18/10/2019	3
25	1286761	45	70	18/10/2019	3
26	1975950	9.37	165.1	18/10/2019	1
27	211000	440	630	18/10/2019	4

Figure 4: Data Recorded

The principal component analysis helped finding our target data to get a time series based on the total amount of daily sale (figure 5)

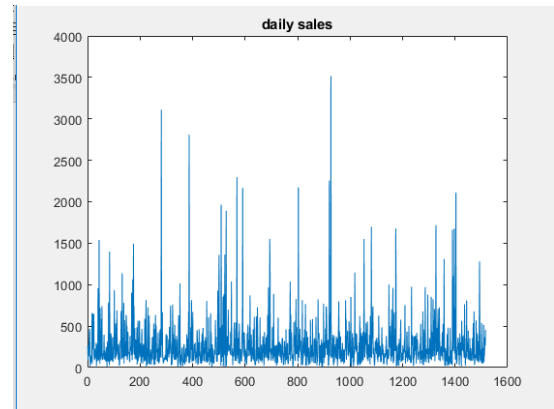


Figure 5: Data Visualization plot

4.2- Training methods for Deep Learning Networks used

There is an abundant amount of great knowledge about deep learning, here are specified the powerful deep learning methods that are applied to our forecasting problem.

4.2.1- Back-Propagation

Back-prop is simply a method to compute the partial derivatives (or gradient) of a function, which has the form as a function composition (as in Neural Nets). When you solve an optimization problem using a gradient-based method (gradient descent is just one of them), you want to compute the function gradient at each iteration.

For a Neural Nets, the objective function has the form of a composition. There are 2 Common ways to compute the gradient:

- (i) Analytic differentiation. You know the form of the function. You just compute the derivatives using the chain rule (basic calculus).
- (ii) Approximate differentiation using finite difference. This method is computationally expensive because the number of function evaluation is $O(N)$, where N is the number of parameters. This is expensive, compared to analytic differentiation. Finite difference, however, is commonly used to validate a back-prop implementation when debugging.

4.2.2- Levenberg–Marquardt Algorithm

The Levenberg-Marquardt algorithm, also known as the damped least-squares method, has been designed to work specifically with loss functions which take the form of a sum of squared errors. It works without computing the exact Hessian matrix. Instead, it works with the gradient vector and the Jacobian matrix. Consider a loss function which can be expressed as a sum of squared errors of the form.

Batch Normalization

Naturally, neural networks including deep networks require careful tuning of weight initialization and learning parameters. Batch normalization helps relaxing them a little.

4.2.3- Long Short-Term Memory:

The LSTM network has the following three aspects that differentiate it from a usual neuron in a recurrent neural network:

It has control on deciding when to let the input enter the neuron.

It has control on deciding when to remember what was computed in the previous time step.

It has control on deciding when to let the output pass on to the next time stamp.

The strength of the LSTM is that it decides all this based on the current input itself.

4.3- Data processing

At first we experimented prediction with a simple neural network with one hidden layer, a backpropagation network performing on the Levenberg Marquart training method (figure 6), and since we are based on a time series dataset, a closed loop network (figure 7) has also been experimented considering sequences relationships, last method we used is a recurrent neural network with two LSTM layers

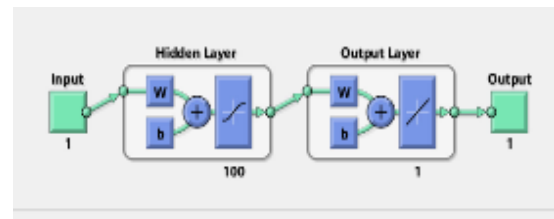


Figure 6: Simple Neural Network

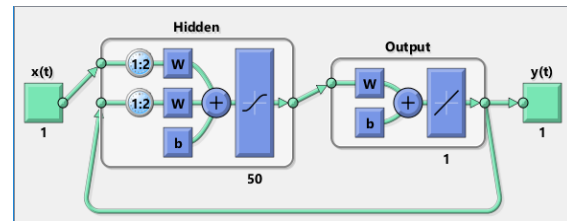


Figure 7 Closed Loop Neural Network

5- RESULTS

5.1-algorithms tuning

we configured the algorithms according to several parameters of which we quote, the number of hidden layers, the number of neurons in each layer and the training methods, then we have retained the best result of root-mean-square error shown in (figure 9), (figure 11) and (figure 13)

Algorithm	Hidden layer 1	Hidden layer 2	Training method	Performance	Data concatenation
ANN	100	0	Levenberg-Marquart	++	true
ANN	50	30	Backpropagation	--	false
RNN	100	0	Closed loop	--	false
RNN	50	0	Closed loop	--	true
LSTM	80	100		++	false
LSTM	200	150		++	false
LSTM	100	200		--	false

5.2- Data prediction:

First, we have separated the data for test, training, and prediction proportionally as 70% ,15% and 15%, then we have proceeded to many tunings of the algorithms to get more accuracy on the predicted data.

We configured the algorithms according to several parameters of which we quote, the number of hidden layers, the number of neurons in each layer and the training methods, then we have retained the best result of root-mean-square error shown in (figure 9) , (figure 11) and (figure 13).

After training the models we plotted the original targets , the network predictions and the expected data for the last two months recorded to compare the results , therefore , the training of the simple network (figure 8) and the recurrent (figure 10) one has performed well but not reliable on some periods, instead, the result of the LSTM prediction (figure 12) is stable and nearer to the expected values

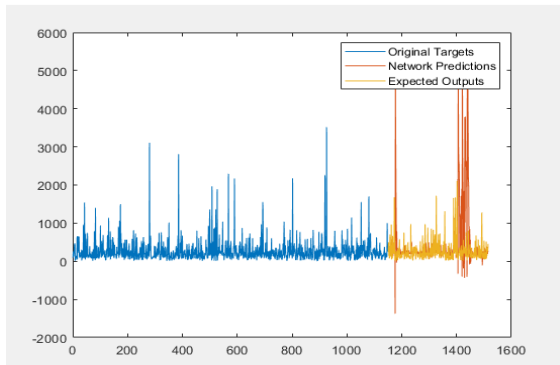


Figure 8 Simple Neural Network Prediction

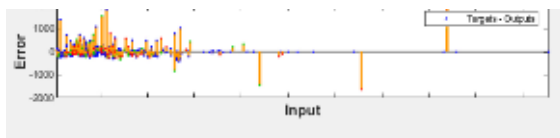


Figure 9: ANN prediction error

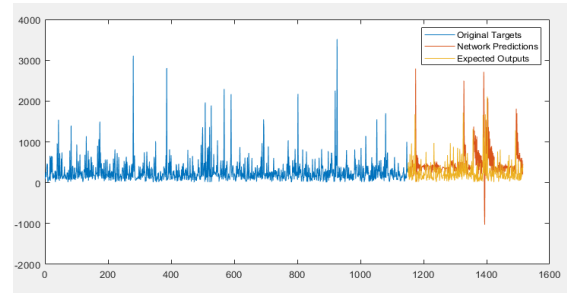


Figure 10: Closed Loop Network Prediction

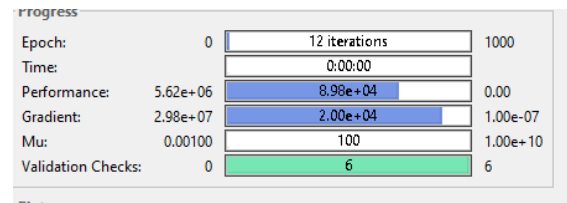


Figure 11: RNN prediction error

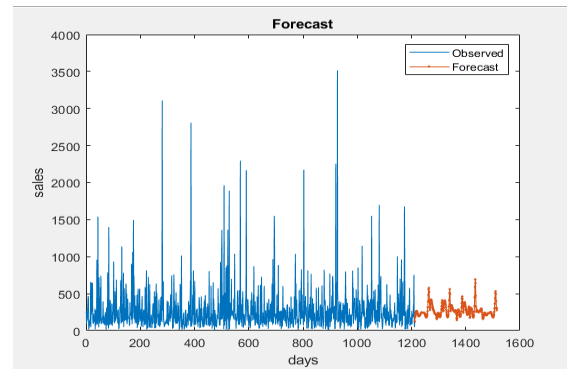


Figure 12: LSTM prediction

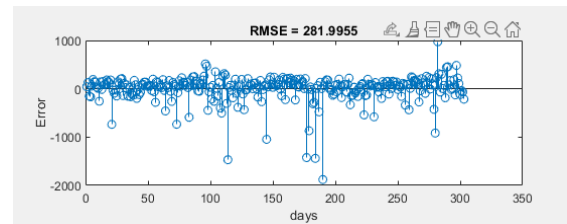


Figure13: LSTM prediction error

5.3- Discussion of the results:

As justified on the previous sequences, the comparative analysis of the different methods used as demonstrated on section 5.1 shows that the LSTMs perform the best on our sales data , and build a reliable model to predict sales for the specific product for the coming months , updating our history with new data.

6- CONCLUSION

This research sheds new light on sales forecasting, by proposing a new prediction model using the recurrent neural network LSTMs, a sequence-based method that reacts and readapt parameters each sequence. This paper includes works on sales forecasting already elaborated by scientists, Regarding these researches, this paper discusses their limits and proposes a model to cope with that ,the objective of this work is to provide a reliable model for sales prediction to help organizations on decision making on both strategic and operational decisions , The model was applied to a recent dataset of recorded sales of a specific product. However, the proposed model can be generalized to be implemented to other products.

Despite the challenges, forecasting is seen as the issue to provide a clearer picture of customer behavior; and can generate accuracy that will help not only to improve the efficiency of the supply chains but also to enhance the company's revenues. we are looking forward to increase interest in prediction technologies and to develop new methods and tools for turning data into more accurate forecasts for use in making actionable operational decisions.

7- PERSPECTIVES

The perspective of this work consists in training the model on other datasets to improve its performance, also to tune the model according to all parameters and get optimal results.

REFERENCES

- Bohdan Pavlyshenko. Using Stacking Approaches for Machine Learning Models. In Proceedings of the 2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP), Lviv, Ukraine, conference paper, August 2018.
- Rosa María Cantón, Croda Damian, Omar Caballero, sales prediction through Neural Network for a small dataset, international journal of interactive multimedia and artificial intelligence ,2019
- Foster Provost and Tom Fawcett, Data Science for Business: What you need to know about data mining and data-analytic thinking Paperback – 19 Aug. 2013
- Foster Provost and Tom Fawcett
Data Science for Business: What You Need to Know about Data Mining and Data-Analytic ASIN: 1449361323
- Huang tao, Robert Fildes, Didier Soopramanien, Forecasting retailer product sales in the presence of structural change, Elsevier, December 2019
- Kalaiairasan T R, Anandkumar V, Ratheesh KumarA
International Journal of Innovative Technology and Exploring Engineering (IJITEE), Sales Forecasting using RNN
- Nilanjan Dey, Aboul Ella Hassanien, Chintan Bhatt, Amira S. Ashour and Suresh Chandra Satapathy
Internet of Things and Big Data Analytics Toward Next-Generation Intelligence,2018
- Oinkina and Hakyll, Understanding LSTM Networks Posted on August 27, 2015
- Shaudhary S. What is a simple explanation of a neural network to someone with limited background in machine learning and AI, 2018
- Tannu , Indian Journal of Computer Science and Engineering (IJCSE) ,2017
- Shaudhary S. What is a simple explanation of a neural network to someone with limited background in machine learning and AI, 2018
- Said benkachcha, Jamal Benhra, Hicham El Hassani, Ahmed Haroun Sabry, Amélioration de la gestion des stocks par la fiabilisation de la prévision de la demande Un modèle de prévision basé sur les réseaux de neurones artificiels, conference paper, May 2015
- Said Benkachcha, Jamal Benhra, Hicham El Hassani, Seasonal Time Series Forecasting Models based on Artificial Neural Network,2015

Said Benkachcha, Hicham El Hassani, Causal Method and Time Series Forecasting model based on Artificial Neural Network, August 2013 International Journal of Computer Applications

Tonya Boone, Ram Ganeshan, Aditya Jain, Nada R. Sanders, Forecasting sales in the supply chain: Consumer analytics in the big data era, 2018 International Institute of Forecasters. Published by Elsevier B.V.

V.Sharma, Deep Learning – Introduction to Recurrent Neural Networks, Posted on January 8, 2019