Trends in Machine Learning Applied to Demand & Sales Forecasting: A Review

Juan Pablo Usuga Cadavid, Samir Lamouri, Bernard Grabot

Trends in Machine Learning Applied to Demand & Sales Forecasting: A Review

Juan Pablo Usuga Cadavid [1], Samir Lamouri [1], Bernard Grabot [2]

[1] LAMIH CNRS, Arts et Métiers ParisTech, Paris, France
[2] LGP, ENIT, Tarbes, France

juan_pablo.usuga_cadavid@ensam.eu, samir.lamouri@ensam.eu, bernard.grabot@enit.fr

Abstract. Supply chain management (SCM) is considered as one of the key elements to leveraging a company's success. Enterprises must adapt their supply chain to cutting-edge technologies and techniques in order to improve their performance, reduce costs and provide better service. Several of these recent breakthroughs have been allowed by Machine Learning (ML), providing solutions to complex problems that were until now difficult to solve, or enhancing the results of former methods. Because of the importance of the field, this paper aims to investigate representative applications of ML in SCM, with an emphasis on the specific area of Demand and Sales Forecasting (D&SF). The purpose of this study is twofold: firstly, to point out the most recent ML trends used in D&SF; and secondly, to explain when companies should invest in novel forecasting techniques over traditional methods.

Keywords: Machine learning, Demand and sales forecasting, Supply chain analytics, Supply Chain Management, Traditional forecasting methods.

1. Introduction

Broadly speaking, the main objective of any Supply Chain (SC) is to "*maximize the total generated value, which is the difference between what the client pays for the final product and the costs incurred by the chain in order to fulfill the client's order*" [1]. Mastering SC provides a huge competitive advantage, as it allows a company to maximize this generated value. This is the reason why the vast majority of companies try to be at the forefront of cutting-edge trends.

As a result of progress over the past two decades in computing power (large data storage, more powerful processors, faster internet connection, etc.) problems that seemed to be extremely complex or expensive to solve are now within our reach. New trends such as Big Data, Cybersecurity, Internet of Things (IoT) and Cryptocurrencies have emerged, ready to challenge the aforementioned technological breakthroughs. For instance, IoT, which aims to use embedded systems including sensors and actuators coupled with the internet to enable control and immediate access to information in real time [2, 3], represents one of these challenges, as there will be approximately 25-50 billion connected devices by 2020 [4]. Moreover, some of these devices will have the capacity to generate data in a range of Exabytes (One million Terabytes) [5], information that may be extremely valuable for a company's sake. Therefore, the SC cannot overlook these new trends; it must adapt and integrate them in order to maintain a competitive advantage.

One of the most recent and popular techniques that aims to tackle those new business challenges is Big Data Analytics (BDA). An accurate definition of BDA is given in [6, based on 5], "*the application of advanced analytic techniques including data mining, statistical analysis, predictive analytics, etc. on big data sets as new business intelligence practice*". In short, it is the coordination of Big Data and Machine Learning (ML) techniques to provide trustworthy insights for decision making. Fig. 1, obtained from [7], represents the general architecture and elements in a BDA model. Two main ideas can be extracted from this model: firstly, ML and big data benefit from each other, as they can be coupled to create more complete models. Secondly, the main purpose of BDA is to transform information into useful knowledge. In this context, the objective of this paper is to provide an insight about the most recent ML techniques applied to D&SF by means of a literature review.

The remainder of this article is organized as follows: section 2 defines Supply Chain Analytics; section 3 analyses ML and its learning types; section 4 describes the used methodology for selecting the analyzed articles, discusses new trends found in literature and compares ML models with traditional forecasting methods. Finally, section 5 concludes and discusses future developments for research.

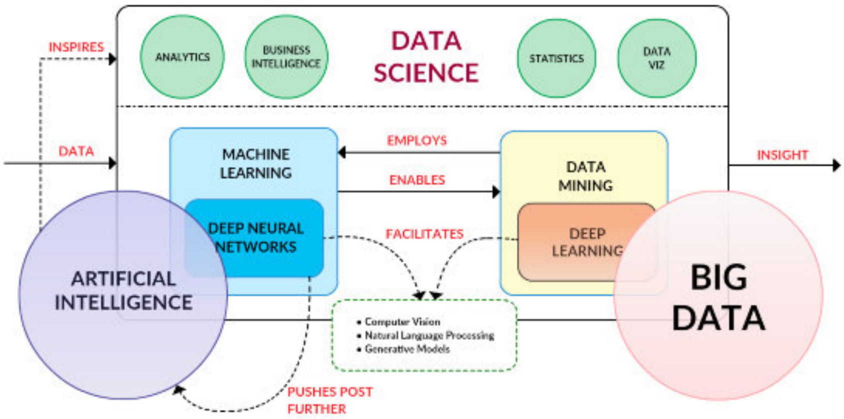

Figure 1: Architecture of a Big Data Analytics model [7]

2. Supply Chain Analytics

In the past few years, BDA has proven to be a true advantage for decision support systems, encouraging SC managers to employ these new techniques in their chain. The use of advanced BDA in a Supply Chain, also called Supply Chain Analytics [5] (SCA), encompasses three main branches: Descriptive analytics, Predictive analytics, and Prescriptive analytics [8]. As this document deals with ML in D&SF, we will mainly focus on predictive analytics. However, the three branches will be explained.

2.1. Descriptive analytics

Descriptive analytics refers to reporting methods and visualization tools. It answers questions about *what is happening* or *what has happened* by providing real time information via technologies such as GPS, RFID, bar codes, etc. [8]. It is an excellent source of real-time information that can be applied in the different business units of a company. For instance, having information about the exact location of the goods in a chain will allow a SC manager to know whether warehouses are under- or over-utilized.

2.2. Predictive analytics

This answers the question, "*what will happen?*" These methods utilize information to create models and perform simulations that will provide insight into future events, letting the most attentive managers foresee strategic actions that will improve the performance of their company. By definition, the results found through these techniques are not 100% accurate, as no method can predict the future. Therefore, a good predictive analytics tool is one that provides the most accurate results in a reasonable amount of time.

One common use of predictive analytics in business is D&SF - this point will be addressed in section 4 of this paper - but there are also several applications in domains such as cost estimation, performance evaluation and strategic sourcing. For instance, [9] applied ML techniques to rapidly estimate the cost of jet engine components, as classical cost estimation methods are pretty accurate but expensive and slow. In performance evaluation, [10] utilized ML for objectively estimating Supply Chain performance based on the 5 Dimensional Balanced Scorecard (5DBSC), with the objective of providing quick results and avoiding biased performance evaluations by managers. Finally, concerning strategic sourcing, [11] employed ML to predict the performance of Third-party logistics (3PL) providers by considering elements like their increment in the market share, experience, location, on-time delivery ratio, etc. The objective was to provide a methodology to choose a 3PL provider given the importance of this decision for the company.

2.3. Prescriptive analytics

This aims to derive recommendations by answering questions about *what should be happening* or *how it can be made to happen*. Prescriptive analytics is supported by the two former analytic branches. Its efficient use demonstrates the most mature state of an enterprise that has effectively applied BDA into its SC. Some authors have proposed examples of how ML could help Prescriptive analytics; for instance, [12] created a ML model for sales forecasting of new books. In their methodology, the authors used Regression Trees, that consists in a series of classifications followed by multiple linear regressions. The advantage of this approach is that it allows the manager to know which are the meaningful variables influencing the new book sales, providing an insight to increase them by quantifying the impact of decisions. The aforementioned study performed by [10] also proposes a prescriptive approach where the manager can change a "performance improvement" parameter, the model pointing which indicators should be improved and by how much (for example, reducing the Supply Chain Time Response from 91 days to 88.75 would enhance the performance by 75%).

Having explained that ML is a core element of BDA, and having shown that it provides indispensable tools for performing SCA, it is worth mentioning some theoretical aspects concerning ML to further clarify the concepts and employed jargon.

3. Machine Learning

Human beings learn through experience: we use a trial and error process in order to discover which actions should be triggered given certain circumstances. This allows us to make abstractions and build knowledge. Machine Learning is somehow similar; it can be seen as algorithms having for objective to improve a performance measure by automatically deriving its own rules and creating its own models based on given information [13, p 2] (for a technical definition, see [14, p. 2]). In general terms, we can identify three types of learning methods: supervised learning, unsupervised learning and reinforcement learning.

3.1. Supervised Learning

This kind of learning is characterized by using data structures that have a collection of features pointing to a result (output). As the desired output is already known, supervised models learn progressively to imitate the required output. In that purpose, the learning system creates its own logic, that may produce outputs to queries on new features [15]. Supervised Learning is often used for classification and regression.

In classification, the objective is to predict the categorical class label of new instances by using rules learnt from past observations [13, p. 3]. Classification may be binary or may concern multiple classes. In binary classification, the model learns a series of rules that will distinguish between two possibilities (i.e. [16] predicted stock trends by classifying whether the price will increase or decrease using Decision Trees and Support Vector Machine); while multiclass classification tries to map a new instance onto one of several possibilities. For instance, differentiating between wine samples coming from three different growers by knowing their intrinsic features such as malic acid, alkalinity of ash, magnesium level, etc. [17].

On the other hand, a regression consists of finding a relationship between predictor variables and a continuous response variable for predicting an outcome [13, p. 5]. We will mainly focus on this kind of problem, as authors have used very powerful ML methods to tackle D&SF issues. Furthermore, even if ML approaches are computationally more expensive, they have proven to deliver better predictive performance than the traditional time-series models [18], proof of the need for research in this field.

3.2. Unsupervised Learning

In contrast to the aforementioned learning type, an unsupervised learning system is not fed with an expected output or explicit feedback to create its rules; instead, the system is supposed to uncover patterns by exploring a data structure and extracting meaningful information [19]. This type of learning is for instance used for data clustering, which is a way of organizing a data set into subgroups in order to recognize patterns or hidden structures. It is also used in dimensionality reduction, which consists in choosing the most meaningful variables for the model. The latter application is convenient when models are complex and computationally expensive, as it reduces the number of features taken into account.

Unsupervised Learning is a helpful tool when forecasting with ML because in practice, identifying hidden trends in data structures provides valuable information (so, clustering is necessary to find patterns). Moreover, data sets often have irrelevant features that only introduce more complexity into the models, so a dimensionality reduction is mandatory. For instance, [20] used K-means for data clustering to reduce the noise in the input data and effectively forecast sales in the printed circuit board industry.

3.3. Reinforcement Learning

Reinforcement learning is related to Supervised Learning, but "*instead of training examples that indicate the correct output for a given input, the training data in reinforcement learning are assume to provide only an indication as to whether an action is correct or not (...)*" [15]. In short, the objective is to maximize the expected reward over time by learning a series of actions that will avoid punishments or penalties. One of the most impressive and most recent application of reinforcement learning is AlphaGo, the Artificial Intelligence program that beat the European champion Fan Hui 5-0 at the Chinese board game Go in October 2015 [21]. A new version of Alpha Go, Alpha Go Zero, was released in 2017. Unlike its earlier versions, Alpha Go Zero only learned by Reinforcement Learning and the results were astonishing: Alpha Go Zero beat Alpha Go 100-0 [22].

Now that ML and SCA have been defined, the role of ML in D&SF will be highlighted. The next section will cover the most recent and most used techniques along with a comparison between "*new*" and classical methods.

4. Machine Learning in Demand & Sales Forecasting

Before addressing the topic, it is important to clarify that Demand Forecasting is different from Sales Forecasting. The latter uses data that have been directly collected from the POS (Point of Sale), so it is subject to the effect of promotions or stock shortages. On the other hand, Demand Forecasting uses data where the effect of promotions or shortages has been corrected, with the objective of reflecting the actual market demand. Despite this subtle yet important difference, this paper will cover both concepts under a single term: Demand & Sales Forecasting (D&SF).

D&SF is one of the most important elements of any SC. It aims to coordinate all the SC partners by reducing the information lag between the final client and the n-tier supplier. This provides several advantages, such as mitigation of the bullwhip effect, better resources and aggregate capacity planning, and reduction of stock shortages or oversupply costs [23]. The most widely known statistical techniques meant to forecast are also called *traditional forecasting methods*.

Traditional forecasting methods are based on time series. This means they are applied under the hypothesis that the past demand can statistically estimate the future demand. Normally, these methods are easy to apply and present good performance in markets whose demand is mostly stable [1]. Unfortunately, this is not often the case: the demand also depends on exogenous factors that are not effectively represented by past values. For instance, on-demand ride services like UBER, Lyft or Didi Chuxing cannot estimate their demand by only relying on time series, they must take into account other elements like weather conditions (such as humidity, temperature, etc.), time of the day or day of the week [18].

To satisfy this need, other kinds of forecasting, known as causal modeling, proposes methods including exogenous elements such as macro-economic variables, weather conditions, marketing strategies, etc. [1]. These techniques allow to tackle the limits encountered in time series models. In that vein, ML itself could be considered as providing causal modeling because it can handle time series, categorical variables, fuzzy variables, text analysis, images, and other elements.

4.1. New trends in literature

ML has been applied to D&SF since 1980 through methods like Artificial Neural Networks (ANN) [24]. During the past two decades, these methods delivered interesting results and demonstrated some potential, but several ANN research applications lacked validity because of validation or implementation issues [25]. This problem was probably caused by the scarcity of data for effectively training ANNs, as this technique has a good generalization capacity but needs a lot of data, data well distributed and time for training. Previous limitations, such as limited storage capacity, low computing power and slow internet connections, could have influenced the reluctance to use ML in D&SF.

Nowadays, ML benefits from a good reputation, probably because most of the aforementioned constraints have been overcome. Thanks to this, new applications have been published leading to new trends and techniques. In order to identify these new trends and meet our research objective, a literature study was performed by surveying three databases (ScienceDirect, Scopus and IEEE) with the following sets of keywords in *titles, abstracts,* and *keywords*: ‘Machine learning’ and ‘Supply Chain’, ‘Machine learning’ and ‘Supply Chain Management’, ‘Neural Networks’ and ‘Forecasting’ and ‘Supply Chain’, ‘Big Data’ and ‘Machine Learning’ and ‘Forecasting’, ‘Machine Learning’ and ‘Sales Forecasting’, ‘Machine learning’ and ‘Demand Forecasting’. Afterwards, only papers addressing Demand or Sales forecasting topics and published between 2009 and 2017 were preselected. Finally, from the preselected papers, a short list was built by choosing the most pertinent articles for analysis: recent papers (published after 2015) were only analyzed if they were cited at least once, while papers preceding 2015 were considered if they were frequently cited.

The short list was composed of 10 publications. For each study, the application, the type of features found in the dataset, the method utilized for data preprocessing, and the used ML techniques for D&SF are presented. Table 1 presents the nomenclature used while the results of our literature study are given in Table 2, together with some meaningful theoretical references.

Table 1. Nomenclature.

T	Time series variables (Past demand or sales values)
En	Endogenous variables of the model (price, number of POS, weeks on sale, etc.)
Ex	Exogenous variables of the model (Weather, time, day of the week, spatial location, etc.)
ELM	Extreme Learning Machine [32]
ANN	Artificial Neural Network (with Back-propagation learning) [33]
DT	Decision Trees [34]
KNN	K-nearest-neighbor [35]
RF	Random forests [36]. This is an Ensemble Learning technique, more information in [37]
SVM	Support Vector Machine (In the regression form) [38]
CNN	Convolutional Neural Network [39]
LSTM	Long Short-Term Memory [40]
SRA	Stepwise regression analysis
WES	Winter's Exponential Smoothing
SOM	Self-Organizing Maps [41]
DBN	Deep Belief Network

Table 2: Some applications of D&SF with ML from 2009 to 2017.
*Compares ML with traditional methods, ** Contains a literature comparison between ML and traditional methods

Article	Application	Data set	Preprocessing technique	ML D&SF technique
[26]*	Fashion products sales forecasting	T, En	-	ELM, ANN
[23]*	Clothing industry sales forecasting	T, En, Ex	-	Fuzzy-ANN, K-means + DT
[12]*	New books sales forecasting	T, En	Correlation feature selection, Regressional Relief	ELM, kNN, DT, ANN, RF, SVM
[27]*	Hospital Length of stay forecasting	T, En, Ex	Cubist DT	DT, Cubist DT, ANN, SVM
[18]*	On-demand ride services forecasting	T, En, Ex	RF	CNN, ANN, LSTM, CNN + LSTM
[20]*	Printed circuit board sales forecasting	T, En, Ex	SRA, WES, K-means, SOM	K-means + Fuzzy-ANN, ANN
[28]	Off -season longan supply forecasting	T, En, Ex	Rank by weights	Fuzzy-ANN, SVR
[29]**	Supermarket demand forecasting	T	-	ANN
[30]	Load demand forecasting	T, En, Ex	EMD	SVR, ANN, DBN, RF
[31]	Load demand forecasting	T, Ex	-	ANN

By analyzing the previous trends presented in Table 2, we can conclude the following:

- First of all, it is clear that ML can be applied to D&SF in a broad range of different industries: from estimating the demand by hour, day and location of on-demand ride services like UBER [18], to forecasting the sales of new books [12].

- Even if there are more recent methods than the classical ANN with back-propagation, this tool is still an attractive model because of its generalization capacity and excellent performance for representing non-linear relations. Furthermore, it can be effectively mixed with Fuzzy methodologies and unsupervised ML techniques like K-means [20]. On the other hand, ANN are black boxes that can hardly help to structure new knowledge.

- [27] and [12] highlighted that methods such as RF and DT provide a peerless level of interpretability, along with good accuracy and decent computing time.

- RF and DT can be considered as both Predictive analytics and Prescriptive analytics tools, because they allow the manager to know the reality behind the model and understand how its decisions will impact the result.

- Fuzzy approaches are convenient when incorporating uncertain or imprecise information in D&SF, like the weather [18, 28].

- Most of the studies include endogenous and exogenous inputs in their models, which shows the good flexibility of ML for dealing with a wide range of inputs. Moreover, [12] tackled the problem of implementing D&SF in new products for which no historical sales data is available (sales of new books). For achieving that, historical data of *other* previously published books was used along with endogenous variables like the *number of weeks on sale,* the *retail price* or the *main subject*.

- [12, 27, 18, 20, 28, 30] effectively applied Data Preprocessing techniques, which resulted in simpler models, lower computational expense and good accuracy. As we are in an age of generation of a high amount of data that also contains noisy and meaningless variables, this provides a valuable insight: analyzing the most relevant variables in real world applications is mandatory, as this can tell the difference between a useful and useless model.

In next section, ML is compared to traditional D&SF methods.

4.2. Comparison between Machine Learning and Classical forecasting methods

Traditional models offer huge advantages in terms of simplicity and accuracy, as they can perform D&SF in a matter of seconds for several SKUs (Stock Keeping Units) [26]. Nevertheless, they need to be thoroughly designed by an expert who can fit the precise needs of the enterprise to a specific model. Moreover, they fail to include exogenous variables. ML models partially solve this problem because they can include other types of data like endogenous and exogenous variables, allowing for a better representation of reality. What is more, ML techniques that are properly implemented outperform most of the traditional forecasting methods [26]. In order to analyze whether this statement is also valid in our literature study, depicted in Table 2, Table 3 mentions what were the traditional and ML models applied by the authors. The table also shows which method provides the best performance (retained model).

Table 3. Studies comparing ML with traditional methods in D&SF.

Article	Application	Traditional models	Best ML model	Retained model
[26]	Fashion products sales forecasting	Polynomial regression	ELM	ELM + Polynomial regression
[23]	Clothing industry sales forecasting	Average profile, flat profile	K-means + DT	K-means + DT
[12]	New books sales forecasting	Multiple linear regression	DT, RF	DT, RF
[27]	Hospital Length of stay forecasting	Poisson simple regression	Cubist DT	Cubist DT
[18]	On-demand ride services forecasting	Historical average, ARIMA, moving average	CNN + LSTM	CNN + LSTM
[20]	Printed circuit board sales forecasting	WES	K-means + Fuzzy ANN	K-means + Fuzzy ANN

The results show that ML truly outperforms traditional models in D&SF. It is worth mentioning that [23] also recognized the mitigation of the Bullwhip Effect, thanks to ML. In addition, [26] proposed a hybrid method with ELM and Polynomial regression that aims to reduce the computing time of the forecast when dealing with numerous SKUs. [29] also performed a literature study comparing ML with traditional methods and found similar results.

The fact that ML outperforms traditional models at D&SF does not necessarily mean that companies should change of their forecasting tools. This leads us to consider an important question: when should enterprises invest in ML for D&SF?

D&SF must be flexible and reactive, especially when dealing with short-term forecasting, which is necessary in sectors such as the clothing industry. The calculation of the forecast must be agile, because most companies have hundreds, thousands or even more SKUs, and all of them may need an immediate estimation. If not applied appropriately, ML models like ANN could take hours or even days to train [42]. Moreover, after the training phase, these methods are still time-consuming, making them less suitable for D&SF. For instance, [12] applied both SVM and the classical Linear Regression using all the features. The former took around 6000 seconds to predict sales, whereas the latter only took around 980 seconds. In this case, SVM would not be convenient for applications dealing with several case scenarios, as the relation between accuracy and computing time is not attractive.

Even if Data Preprocessing techniques could substantially reduce computing time, they are still complex and therefore expensive to establish in a company, as they require qualified people with proper equipment. This difficulty is especially present for SMEs, where available resources are limited and employees seldom have advanced knowledge on the topic. In that case, if an enterprise is positioned in a stable market and if the historical demand is sufficient for achieving good accuracy in D&SF with traditional methods, this company should delay the migration towards ML techniques until it really identifies an added value in using them.

On the other hand, if a company sells products in a market subject to constant evolution, where it is mandatory to be at the vanguard of trends for being competitive, the ML models for D&SF will be a valuable asset. Nevertheless, it is not an easy change, as firms must secure three valuable aspects: data storage capacity, data processing capacity, and employees' qualifications. Data storage is critical, as ML needs Big Data to perform well, and the latter often comes in complex data sets that can surpass the storage capacity of several PCs; the data processing capacity can vary from models that can be trained in a single PC to models needing more onerous configurations like 20 CPU cores, 251 GB RAM and one GPU, as in [18]; the company must consider its goals versus the required equipment. Finally, employees' qualifications are vital to build the model, test it and keep it running effectively. This could be expensive, as companies tend to use consulting services for this work.

In short, a company should favor ML rather than traditional forecasting methods for D&SF when its economic environment truly requires a digital transformation, and when the enterprise can gather the resources required to take on the challenge of Supply Chain Analytics.

5. Conclusions and further research

Dealing with a high volume of data is one of the more common challenges imposed by modern business trends. Since data has become one of the most valuable resources, SC managers are eager to extract relevant information that will lead to a competitive advantage. This paper used a literature study to investigate the application of ML in D&SF as a method to achieving this advantage. Ten recent research papers applying ML to D&SF were selected and analyzed in order to identify new trends. One of the findings was that ML broadens the reach of D&SF, as it is able to handle complex variables. More precisely, Fuzzy-ANN approaches showed excellent performance when dealing with imprecise data like weather variables, whereas DT and RF offered invaluable interpretation capacity. Furthermore, Data Preprocessing techniques proved to substantially reduce the complexity of the models, enabling both good accuracy and reasonable computing time.

At this stage, it can be noticed that no study using data mining techniques on D&SF was found, which is a bit surprising. Indeed, techniques such as rule mining may allow to analyze large amounts of data and to identify correlations between items that can provide explicit knowledge, able to improve the skills of the decision makers. This may be of specific interest for better formalizing the role of the exogenous variables, and this facility can be used in combination with traditional predictive methods.

As adapting a company to new technologies often comes with doubts raised by uncertainty, a comparison between ML and traditional forecasting methods was made with the objective of providing insight to managers who are willing to implement ML in their processes. The results of this study show that ML is a more suitable technique than traditional forecasting methods in terms of accuracy, particularly when models contain exogenous and endogenous variables. Furthermore, it enables the identification of hidden patterns in demand that can be used as a baseline to identify new market trends.

For further research, a similar study should be performed on Data Preprocessing techniques as they offer significant advantages in terms of reduction of computational costs when applying ML models. Finally, addressing these new trends to SMEs is vital, as they are numerous but often lack the financial means or experience to implement cutting-edge technologies.

6. References

1. Chopra, S., and Meindl, P.: Supply Chain Management. Pearson Education, Inc, 5th Edition, 2013.
2. Atzori, L., Iera, A., and Morabito, G.: The internet of things: A survey. Computer Networks, pages 2787–2805, 2010.
3. Cecchinel, C., Jimenez, M., Riveill, M., and Mosser, S.: An architecture to support the collection of big data in the internet of things. IEEE World Congress on Services, pages 442–449, 2014.
4. Mahdavinejad, M.S., Rezvan, M., Barekatain, M., Adibi, P., Barnaghi, P., and Sheth, A.P.: Machine learning for internet of things data analysis: A survey. (in press, 2018).
5. Tiwari, S., Wee, H.M., and Daryanto, Y.: Big data analytics in supply chain management between 2010 and 2016: Insights to industries. Computers & Industrial Engineering, 115:319–330, 2018.
6. Russom, P.: Big data analytics, tdwi best practices report, fourth quarter. Technical report, tdwi.org, 2011.
7. Mayo, M.: The data science puzzle, explained, 2016. Available at www.kdnuggets.com/2016/03/data-science-puzzle-explained.html/2.
8. Souza, G.C.: Supply chain analytics. Business Horizons, 57(5):595–605, 2014.
9. Loyer, J.L., Henriques, E., Fontul, M., and Wiseall, S.: Comparison of machine learning methods applied to the estimation of manufacturing cost of jet engine components. International Journal of Production Economics, 178:109–119, 2016.
10. Fan, X., Zhang, S., Wang, L., Yang, Y., and Hapeshi, K.: An evaluation model of supply chain performances using 5dbsc and lmbp neural network algorithm. Journal of Bionic Engineering, 10:383–395, 2013.
11. Vahdani, B., Razavi, F., and Mousavi, S.M.: A high performing meta-heuristic for training support vector regression in performance forecasting of supply chain. Neural Computing and Applications, 27:2441–2451, 2016.
12. Castillo, P.A., et al: Applying computational intelligence methods for predicting the sales of newly published books in a real editorial business management environment. Knowledge-Based Systems, 115:133–151, 2017.
13. Raschka, S., and Mirjalili, V.: Python Machine Learning, 2nd Ed. Packt Publishing, Birmingham, UK, 2 edition, 2017.
14. Mitchell, T.M.: Machine Learning. McGraw Hill, 1997.
15. Jordan, M.I., and Mitchell, T.M.: Machine learning: Trends perspectives and prospects. Science, 349:255–260, 2015.
16. Wang, D., Liu, X., and Wang, M.: A dt-svm strategy for stock futures prediction with big data. IEEE 16th International Conference on Computational Science and Engineering, 2013.
17. Lichman, M.: UCI Machine Learning Repository, 2013. Available at https://archive.ics.uci.edu/ml/datasets/wine.
18. Ke, J., Zheng, H., Yang, H., and Chen, V. (Michael): Short-term forecasting of passenger demand under on-demand ride services: A spatio-temporal deep learning approach. Transportation Research Part C, 85:591–608, 2017.
19. Zhou, L., Pan, S., Wang, J., and Vasilakos, A.V.: Machine learning on big data: Opportunities and challenges. Neurocomputing, 237:350–361, 2017.
20. Chang, P.C., Liu, C.H., and Fan, C.Y.: Data clustering and fuzzy neural network for sales forecasting: A case study in printed circuit board industry. Knowledge-Based Systems, 22:344– 355, 2009.
21. Silver, D., et al.: Mastering the game of go with deep neural networks and tree search. Nature, 529:484–489, 2016.
22. Silver, D., et al.: Mastering the game of go without human knowledge. Nature, 550:354–359, 2017.
23. Thomassey, S.: Sales forecasts in clothing industry: The key success factor of the supply chain management. International Journal of Production Economics, 128:470–483, 2010.
24. Hippert, H.S., Pedreira, C.E., and Souza, R.C.: Neural networks for short-term load forecasting: a review and evaluation. IEEE Transactions on Power Systems, 16(1):44–55, Feb 2001.
25. Adya, M. and Collopy, F.: How effective are neural networks at forecasting and prediction? a review and evaluation. Journal of Forecasting, 17(5-6):481–495, 1998.
26. Yu, Y., Choi, T.M., and Hui, C.L.: An intelligent fast sales forecasting model for fashion products. Expert Systems with Applications, 38(6):7373 – 7379, 2011.
27. Turgeman, L., May, J.H., and Sciulli, R.: Insights from a machine learning model for predicting the hospital length of stay (los) at the time of admission. Expert Systems with Applications, 78(Supplement C):376 – 385, 2017.
28. Leksakul, K., Holimchayachotikul, P., and Sopadang, A.: Forecast of off-season longan supply using fuzzy support vector regression and fuzzy artificial neural network. Computers and Electronics in Agriculture, 118(Supplement C):259 – 269, 2015.
29. Slimani, I., Farissi, I.E., and Al-Qualsadi. S.A.: Configuration of daily demand predicting system based on neural networks. In 2016 3rd International Conference on Logistics Operations Management (GOL), pages 1–5, May 2016.

30. Qiu, X., Ren, Y., Suganthan, P.N., and Amaratunga, G.A.J.: Empirical mode decomposition based ensemble deep learning for load demand time series forecasting. Applied Soft Computing, 54(Supplement C):246 – 255, 2017.
31. Rahman, M.N., Esmailpour, A., and Zhao, J.: Machine learning with big data an efficient electricity generation forecasting system. Big Data Research, 5(Supplement C):9 – 15, 2016. Big data analytics and applications.
32. Huang, G., Huang, G.B., Song, S., and You, K.: Trends in extreme learning machines: A review. Neural Networks, 61(Supplement C):32 – 48, 2015.
33. Rumelhart, D.E., Hinton, G.E., and Williams, R.J.: Learning representations by backpropagation errors. Letters to Nature, 323:533–536, 1986.
34. Hunt, E.B., Marin, J., and Stone, P.J.: Experiments in induction. Academic Press, 1966.
35. Fix, E. and Hodges, J.L.: Discriminatory analysis, nonparametric discrimination: Consistency properties. Technical report, Randolph Field, 1951.
36. Tin Kam Ho. Random decision forests. In Proceedings of the Third International Conference on Document Analysis and Recognition (Volume 1) - Volume 1, ICDAR '95, pages 278–, Washington, DC, USA, 1995. IEEE Computer Society.
37. Dietterich, T.G.: Ensemble Methods in Machine Learning, pages 1–15. Springer Berlin Heidelberg, Berlin, Heidelberg, 2000.
38. Boser, B.E., Guyon, I.M., and Vapnik. V.: A training algorithm for optimal margin classifiers. In Proceedings of the Fifth Annual Workshop on Computational Learning Theory, COLT '92, pages 144–152, New York, NY, USA, 1992. ACM.
39. LeCun, Y., Haffner, P., Bottou, L., and Bengio, Y.: Object Recognition with Gradient Based Learning, pages 319–345. Springer Berlin Heidelberg, Berlin, Heidelberg, 1999.
40. Hochreiter, S. and Schmidhuber, J.: Long short-term memory. Neural computation, 9:1735–80, 12 1997.
41. Kohonen, T.: Self-organized formation of topologically correct feature maps. Biological Cybernetics, 43(1):59–69, Jan 1982.
42. Huang, G.B., Zhu, Q.Y., and Siew, C.K.: Extreme learning machine: Theory and applications. Neurocomputing, 70(1):489 – 501, 2006. Neural Networks.