Anomaly Detection with Machine Learning Technique to Support Smart Logistics

Nittaya Kerdprasop[1,2(✉)], Kacha Chansilp[2], Kittisak Kerdprasop[1,2], and Paradee Chuaybamroong[3]

[1] Data and Knowledge Engineering Research Unit, Nakhon Ratchasima, Thailand
[2] School of Computer Engineering, Suranaree University of Technology, Nakhon Ratchasima, Thailand
{nittaya, kacha, kerdpras}@sut.ac.th
[3] Department of Environmental Science, Thammasat University, Pathumthani, Thailand
paradee@tu.ac.th

Abstract. Accurate planning and cost-effective management on product delivery are among key factors leading to the success of most manufacturing sectors in the current era of the fourth industrial revolution, or Industry 4.0. In this work, we focus our study on the development of a model based on machine learning technique to support smart logistics by detecting anomaly events on the big stream of electronic orders obtained from ubiquitous customers. We use the data-driven approach to build the order-anomaly model and present the built model as the classification and regression tree. Our model in a tree formalism is to be used as an automatic detector for unusual events inherent in the customers' order stream. The anomaly events should invoke special care in the smart logistics environment that delivery tasks are performed in an automatic manner. Early detection of anomaly ordering events is expected to improve accuracy on delivery planning.

Keywords: Anomaly detection · Order anomaly model · Smart logistics · Classification and regression tree · Machine learning

1 Introduction

Traditional logistics process deals with the delivery of products to the right place at the right time. The main processes in logistics include the three main steps [1] of order processing and transport planning, order shipping, and order billing. For smart logistics, digital data from ubiquitous sources and state-of-the-art technologies play their crucial roles on turning the regular shipping process into the intelligent transport system (ITS) by means of the radio frequency identification (RFID) tagging with goods to be delivered. The current capability of ITS has been argued [2] to be extended to fulfill the need of being an efficient linkage between the smart city, the customer side, and the smart industry, the producer side.

© Springer Nature Switzerland AG 2019
S. Misra et al. (Eds.): ICCSA 2019, LNCS 11619, pp. 461–472, 2019.
https://doi.org/10.1007/978-3-030-24289-3_34

The high-quality ITS can be considered an important supporter for the manufacturing movement toward the fourth stage of industrialization [3, 4]. We, therefore, propose in this work the application of data-driven modeling through the machine learning approach. Our modeling process is based on the classification and regression algorithm to learn anomaly events from product ordering stream. The learned model is designed to be an intelligent part of the supply chain management module for monitoring unusual events.

After presenting literature review in Sect. 2, we explain our modeling method in Sect. 3. The model evaluation results are illustrated in Sect. 4. We conclude our work in Sect. 5.

2 Literature Review

In logistics and supply chain management, an accurate forecasting of customers' order is significant for not only a cost-effective manufacturing planning, but also a timely material and product distribution [5] to fulfill the demand of customers. The needs for an accurate forecasting technique can be found in so many literature across a variety of health-related domains ranging from drinking water distribution network [6, 7], to electricity and home energy demands [8–10], outpatient space in the hospital [11], amount of food donations [12], and retail pharmacies [13].

Forecasting has long been an important tool in the business domain to project customers' demand trends. Retail industry is the business section that has heavily applied forecasting for efficient business planning and managing. Computational forecasting method has been used to predict the fashion demand [14], to project the number of tourists [15], to estimate the growth in on-line retailing [16], to identify an optimal retail location [17], to plan replenishment in the discount retail chain [18], to approximately quantify the need for high-end luxury products [19], and to estimate the demand for retailing through automated vending system [20].

Business processes have indeed been strongly linked to industrial movements since the first industrial revolution with mass mechanization to the fourth one with full digitization through the cyber-physical systems. In the current movement of the Industry 4.0, numerous machine learning techniques and soft computing approaches have been extensively applied in many smart manufacturing industries. Intelligent analysis has been applied within a number of supply chain management processes [21–26].

Machine learning has the ability to learn from sales data, and then report in an automatic manner about consumer behavior and preferences. Various machine learning techniques have been applied for such tasks. These techniques include clustering [27, 28], artificial neural network [29], fuzzy models [30], support vector machine [31], and many others [32, 33].

Several recent studies have reported that advanced machine learning techniques such as support vector machine, genetic algorithm, artificial neural network, and deep learning outperform traditional statistical-based forecasting. We also apply machine learning technique using the classification and regression tree algorithm to learn

anomaly events from the historical product delivery data. Details of our methodology are explained in the following section.

3 Modeling Method for Anomaly Detection

3.1 Product Order Data Characteristics

The focus of this work is to develop a model to detect anomaly customer order events aiming to support smart supply chain management system, especially in product delivery planning. Our model building method is data-driven in the sense that the historical data are used in the training process of the automatic model creation step. The product order data used in our work are public-domain contributed by researchers from Brazil [34].

The original dataset contains 60 records of the daily delivery parcel service company in Brazil. Each record contains information related to the product distribution, as summarized and shown in Table 1. Pair of attributes showing very strong correlation (that is, correlation > ±0.70) are highlighted with bold red font as shown in Table 2.

Table 1. Data attribute characteristics.

Name	Meaning	Type
A1	Banking orders (1)	Integer
A2	Banking orders (2)	Integer
A3	Banking orders (3)	Integer
A4	Day of the week (Monday to Friday)	Category
A5	Fiscal sector orders	Real
A6	Non-urgent orders	Real
A7	Order type A	Real
A8	Order type B	Real
A9	Order type C	Real
A10	Orders from the traffic controller sector	Integer
A11	Total orders	Real
A12	Urgent orders	Real
A13	Week of the month	Category

3.2 Modeling Process

The purpose of this work is to derive anomaly events from customer orders. But the original dataset contains only 60 records and such a dataset is too small to capture anomaly events. We thus over-sampling the dataset to contain 100,000 records and make a series of steps as shown in Fig. 1.

The anomaly detection modeling process starts from computing anomaly index value, which is a measure of dissimilarity of a data record as compared to other data records in the same group [35]. All data records are then ranked based on this index

Table 2. Correlation among data attributes.

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
A2	0.26											
A3	0.22	-0.11										
A4	-0.05	-0.58	-0.01									
A5	0.004	-0.06	0.30	-0.13								
A6	0.73	0.79	0.13	-0.42	-0.06							
A7	0.68	0.29	0.23	-0.07	0.07	0.56						
A8	0.59	0.71	0.07	-0.38	-0.12	0.83	0.44					
A9	0.33	0.72	0.03	-0.45	0.01	0.75	0.22	0.52				
A10	-0.16	0.24	0.23	-0.34	0.20	0.25	-0.15	0.13	0.44			
A11	0.63	0.80	0.10	-0.44	-0.05	0.94	0.56	0.90	0.81	0.25		
A12	0.23	0.66	0.03	-0.52	-0.01	0.57	0.41	0.51	0.77	0.24	0.73	
A13	0.39	0.15	-0.16	-0.21	0.001	0.24	0.25	0.31	-0.04	-0.19	0.21	0.12

value from the highest to the lowest. Data record with the highest anomaly index value reflect the unusualness of that record. We set the minimum threshold of anomaly index as 2.0. Data records having anomaly index lower than this threshold are considered normal cases. Otherwise, the data will be labeled as anomaly events.

After identifying anomaly data records (steps 1–3), we select normal cases from the least anomalous group of with the same amount as the abnormal cases (step 4). The purpose of this step is to generate a training set with balancing proportion between the two groups of data (normal cases and anomalous cases). The data set is then used to train the classification and regression learning algorithm [36] to build the tree model called classification and regression tree, or CART, in step 5.

The CART model is finally assessed its classifying performance on separating normal customer orders from the abnormal ones. On evaluating model performance (step 6), we use the out-of-sample method, that is, the hold-out test data (around 30% of the 100,000 data records) are unseen by the tree-learning algorithm. The final product of the proposed methodology is the CART model that has the capability of automatic differentiating normal orders from the anomalous ones.

4 Model Creation and Evaluation Results

4.1 Anomaly Detection Model

After computing anomaly index of each data record in step 1, we found that there are 1,142 records showing the anomaly index values higher than 2.0. We then select other 1,142 data records from the group of normal data that have the lowest anomaly index value. The justification of this selection is for creating a balance data set having the same amount of normal and abnormal cases. Moreover, the selection of data group having lowest anomaly index is for the high contrast to the anomaly cases. We investigate characteristics of both normal (illustrated in Fig. 2) and anomalous cases (Fig. 3) with the web graph tool to explore the dominant attributes.

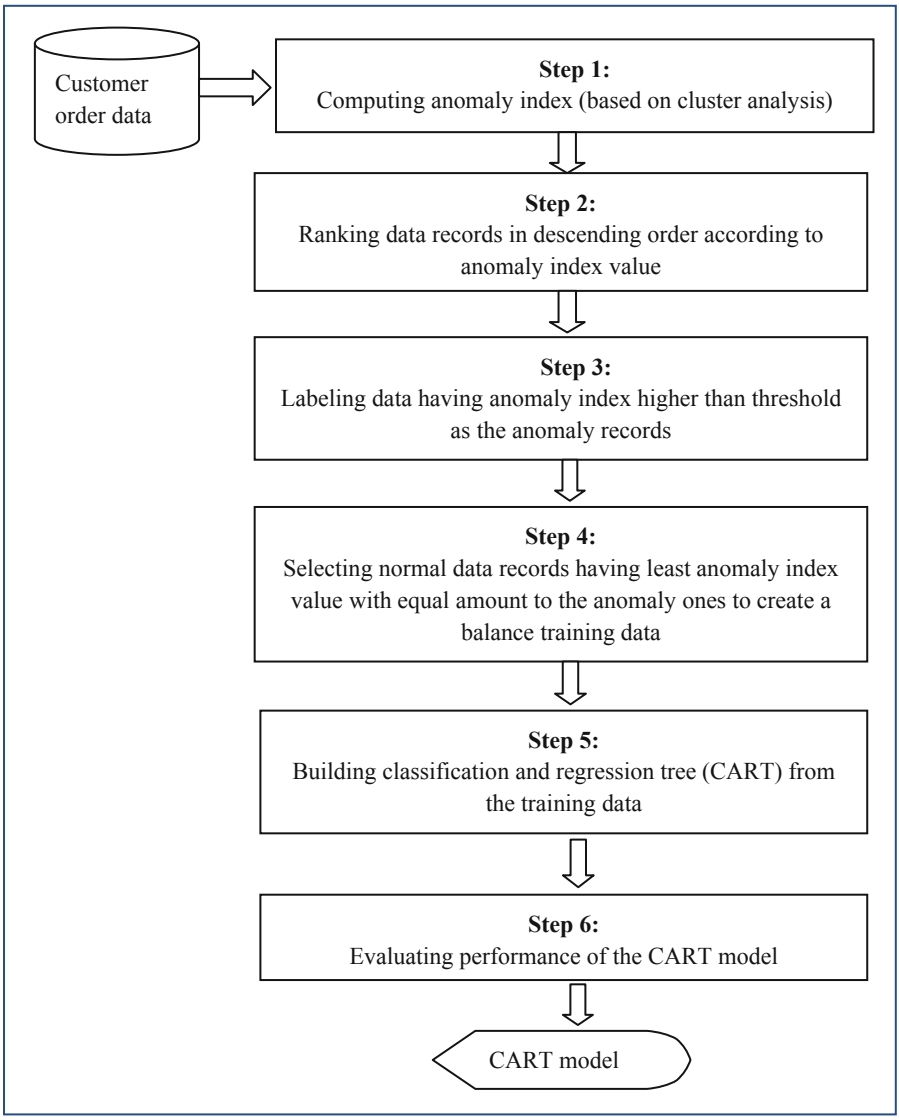


Fig. 1. Proposed method to learn normal and anomaly events from the historical logistics data.

The thickness of lines in the web graph represents the weight of association that has been counted from number of data records supporting that relationship. The thick line shows strong association, while the thin line illustrates less significant relationship. It can be seen from the web graph in Fig. 2 that for the group of normal product order, the most outstanding attribute related to normal order is *week of the month*. The other three attributes expecting useful for identifying the normal orders are *orders from the traffic controller sector*, *day of the week*, and *banking orders (3)*. These three attributes are, however, less significant than the *week of the month* attribute.

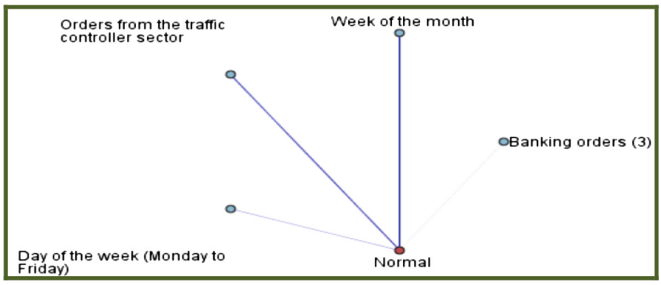


Fig. 2. Associations among attributes in the normal group of product orders.

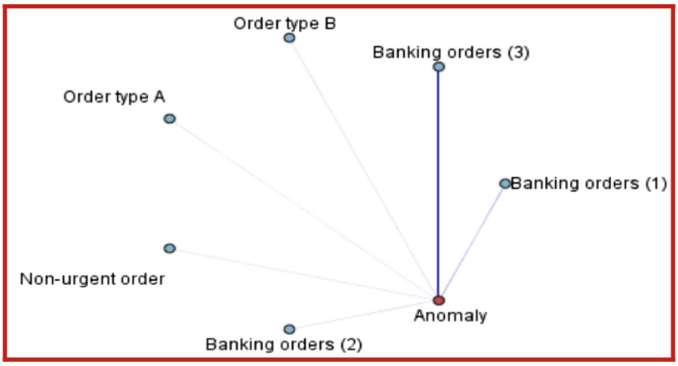


Fig. 3. Associations among attributes in the anomaly order group.

For the case of anomalous order shown in Fig. 3, it can be clearly observed that *banking orders (3)* is the most prominent attribute. Other relationships in Fig. 3 shown with very thin lines that hardly observable represent minimal contribution to pinpoint the anomaly cases.

After combining data from the two groups to build the CART model, the outcome is the one shown in Fig. 4. The meaning of this model is as follows:

Normal-order cases have the following characteristics:

1: *banking orders (3)* $\leq$ 34496.50 AND
total orders $\leq$ 333.334 AND
orders from the traffic controller sector > 26033.50

Anomaly-order cases have the following characteristics:

1: *banking orders (3)* $\leq$ 34496.50 AND
total orders $\leq$ 333.334 AND
orders from the traffic controller sector $\leq$ 26033.50
2: *banking orders (3)* $\leq$ 34496.50 AND
total orders > *333.334*
3: *banking orders (3)* > 34496.50

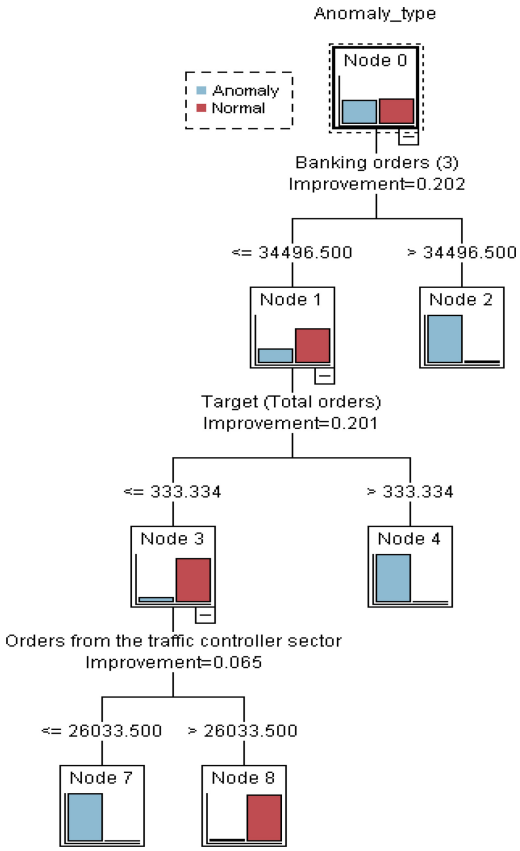


Fig. 4. A CART model for classifying normal product orders from the anomaly ones system.

The CART model captures *banking orders (3)*, *total orders*, and *orders from the traffic controller sector* as the three most discriminative attributes to differentiate normal orders from the abnormal ones. Importance of other attributes can be seen from Fig. 5.

On the full scale of 1.0 to reflect the highest importance on predicting the target outcome of normal/anomaly order, the attribute *banking orders (3)* is the most prominent one with the importance value of 0.44. The attribute *total orders* comes in second with the importance value of 0.38. The attribute *orders from the traffic controller sector* is the third one with 0.08 importance. The attributes *banking orders (1)*, *order type A*, *urgent order*, and *non-urgent order* have equal importance value of 0.01.

The final step is the accuracy evaluation of the CART model to differentiate anomaly orders from the normal ones. We perform the hold-out method by splitting data into a training data set (containing 1,566 records) and a test data set (with the remaining 718 records). The CART model can classify correctly 696 records (out of the 718 records). This comprises of the 96.94% of the accuracy rate, and the error rate is as low as 3.06%. We consider it a very satisfied model for detecting anomaly events.

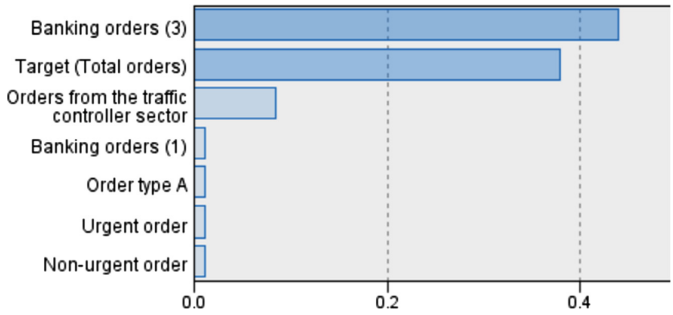


Fig. 5. Importance analysis of attributes to identify the total orders as either normal or anomaly.

4.2 Model Performance Enhancement

To further improve the performance of the CART model, we investigate the possibility of applying boosting technique. The main advantage of boosting is that instead of generating a single model for predicting the target attribute, we rather use ten models to predict the target. We generate ten CART models, then vote the outcome of the target attribute as either anomaly or normal product order based on the majority voting of the ten CART models.

On generating each CART model, the training data are randomly chosen to comprise of 1,566 records. Therefore, the detail in each model can be different depending on the random samples used as a training dataset. Moreover, from all available 13 predictors, the model may use only a subset of all available predictors. The top-five frequently used attributes on building a CART model is shown in Fig. 6. It can be seen from the figure that the five attributes (*week of the month*, *urgent order*, *total orders*, *orders from the traffic controller sector*, *order type C*) are used in every of the CART ensemble.

The performances of all ten models are also summarized in Table 3. The predictive accuracy of a single model is varied from 77.50% to 98.50%. However, to use boosting technique on predicting the anomaly/normal product order, all models are used as an ensemble. The final predicting outcome is obtained through the voting scheme. Therefore, the performance of the ensemble is much better than a single model. The ensemble quality in terms of the accuracy from voting scheme is displayed in Fig. 7.

The accuracy rate shown in Fig. 7 has been evaluated from the training dataset. To fairly compare predictive performance of the ensemble model based on boosting technique against a single CART model, we apply a separate test set. The test set is applied to both the ensemble model and the single CART model scheme. We found that on testing with the out-of-sample data, the ensemble based on boosting technique yields the accuracy rate as high as 98.89%, which is almost two percents better that a single CART model.

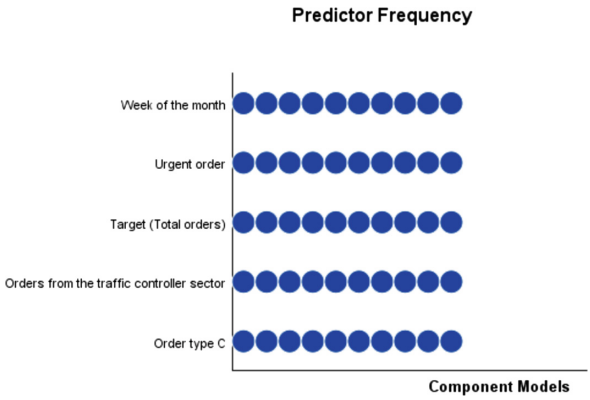


Fig. 6. Frequently used attributes for predicting anomaly/normal product order of the CART ensemble built from ten models.

Table 3. Performances of the ten CART models.

Model	Number of predictors	Model size (#nodes)	Accuracy (%)
1	12	7	98.30
2	13	11	96.70
3	13	15	92.40
4	12	15	94.30
5	12	9	94.10
6	13	13	95.70
7	13	19	77.50
8	12	15	90.90
9	11	11	98.50
10	12	11	85.40

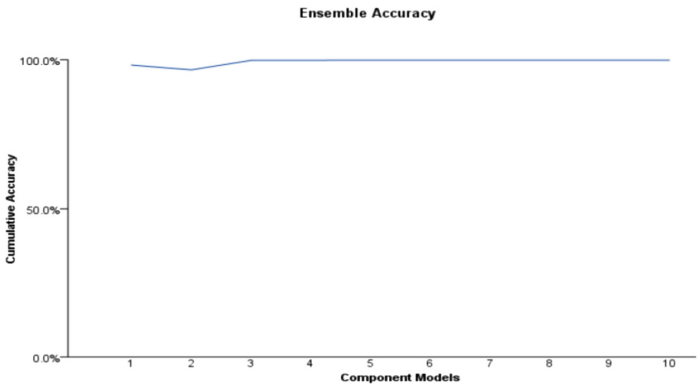


Fig. 7. Ensemble accuracy from a combination of CART models.

5 Conclusion

In the current era of the fourth industrial revolution, or the so-called Industry 4.0, manufacturing, logistics, supply chain management, and many business sectors are facing challenging movement toward smart operations and intelligent production. In this work, we propose the idea of employing a self-learning methodology to support smart logistics. In this work, we devise a method that can learn anomaly detection model automatically from the product order data collected from various customer groups. The proposed model induction method employs the power of machine learning technique on analyzing attribute importance on classifying normal customer order events from the anomaly cases.

In our design, we use the classification and regression tree induction technique as a main algorithm for learning automatically characteristics of anomaly orders that can be differentiated from the normal orders. The selection of tree formalism is due to the fact that the tree structure can facilitate reasoning on tracing back the pre-condition leading to the predicted outcome. On the data preparation step, we use anomaly index computed from the cluster analysis to justify whether the data record is normal or abnormal cases. The anomalous records are extracted and incorporated with the same amount of normal records to train the tree induction model. The evaluation result of normal/anomaly event detection of the tree model shows the accuracy rate as high as 96.94%. This result reveals that our proposed method yield an accurate discriminative model that is appropriate for embedded in the smart logistics and supply chain management system.

On evaluating performance of the tree model, we also apply the ensemble method by using a boosting technique to build ten models. Then applying all ten models to predict the outcome of the target attribute. All ten predicting results from the ensemble are combined based on the voting scheme. We found from the experimental result that the ensemble of tree models can improve predicting accuracy from 96.94% to 98.89%. The ensemble is thus a promising technique to build a tree model to detect anomaly event. We thus plan to further investigate other ensemble schemes on this particular domain.

Acknowledgment. This work was financially supported by grants from the National Research Council of Thailand and Suranaree University of Technology through the funding of the Data and Knowledge Engineering Research Unit.

References

1. Singh, P.M., van Sinderen, M.J., Wieringa, R.J.: Smart logistics: an enterprise architecture perspective. In: 29th International Conference on Advanced Information Systems Engineering (CAiSE 2017), pp. 9–16 (2017)
2. Oonk, M.: Smart logistics corridors and the benefits of intelligent transport system. In: Blanquart, C., Clausen, U., Jacob, B. (eds.) Towards Innovative Freight and Logistics. Wiley, New Jersey (2016)

3. Stock, T., Seliger, G.: Opportunities of sustainable manufacturing in industry 4.0. Proc. CIRP **40**, 536–541 (2016)
4. Delfmann, W., ten Hompel, M., Kersten, W., Schmidt, T., Stolzle, W.: Logistics as a science – central research questions in the era of the fourth industrial revolution. Logist. Res. **11**, 1–13 (2017)
5. Syntetos, A.A., Babai, Z., Boylan, J.E., Kolassa, S., Nikolopoulos, K.: Supply chain forecasting: theory, practice, their gap and the future. Eur. J. Oper. Res. **252**(1), 1–26 (2016)
6. Rangel, H.R., Puig, V., Farias, R.L., Flores, J.J.: Short-term demand forecast using a bank of neural network models trained using genetic algorithms for the optimal management of drinking water networks. J. Hydroinform. **19**(1), 1–16 (2017)
7. Lopez Farias, R., Puig, V., Rodriguez Rangel, H., Flores, J.: Multi-model prediction for demand forecast in water distribution networks. Energies **11**(3), 660 (2018). 1–12
8. AL-Musaylh, M.S., Deo, R.C., Li, Y., Adamowski, J.F.: Two-phase particle swarm optimized-support vector regression hybrid model integrated with improved empirical mode decomposition with adaptive noise for multiple-horizon electricity demand forecasting. Appl. Energy **2017**, 422–439 (2018)
9. Koolen, D., Sadat-Razavi, N., Ketter, W.: Machine learning for identifying demand patterns of home energy management systems with dynamic electricity pricing. Appl. Sci. **7**(11), 1160 (2017)
10. Wang, Y., Chen, Q., Gan, D., Yang, J., Kirschen, D.S., Kang, C.: Deep learning-based socio-demographic information identification from smart meter data. IEEE Trans. Smart Grid (2018). https://doi.org/10.1109/tsg.2018.2805723
11. Jiang, S., Chin, K.S., Wang, L., Qu, G., Tsui, K.L.: Modified genetic algorithm-based feature selection combined with pre-trained deep neural network for demand forecasting in outpatient department. Expert Syst. Appl. **82**, 216–230 (2017)
12. Pugh, N., Davis, L.B.: Forecast and analysis of food donations using support vector regression. In: 2017 IEEE International Conference on Big Data (Big Data), pp. 3261–3267 (2017)
13. Papanagnou, C.I., Matthews-Amune, O.: Coping with demand volatility in retail pharmacies with the aid of big data exploration. Comput. Oper. Res. **98**, 343–354 (2018)
14. Ren, S., Chan, H.L., Ram, P.: A comparative study on fashion demand forecasting models with multiple sources of uncertainty. Ann. Oper. Res. **257**(1–2), 335–355 (2017)
15. Chang, Y.W., Tsai, C.Y.: Apply deep learning neural network to forecast number of tourists. In: 31st International Conference on Advanced Information Networking and Applications Workshops (WAINA), pp. 259–264 (2017)
16. Qin, Z., Bowman, J., Bewli, J.: A Bayesian framework for large-scale geo-demand estimation in on-line retailing. Ann. Oper. Res. **263**(1–2), 231–245 (2018)
17. Fisher, M., Glaeser, C.K., Su, X.: Optimal retail location: empirical methodology and application to practice. SSRN Electron. J. (2016). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2842064
18. Akyuz, A.O., Uysal, M., Bulbul, B.A., Uysal, M.O.: Ensemble approach for time series analysis in demand forecasting: Ensemble learning. In: 2017 IEEE International Conference on INnovations in Intelligent SysTems and Applications (INISTA), pp. 7–12 (2017)
19. Qu, T., Zhang, J.H., Chan, F.T.S., Srivastava, R.S., Tiwari, M.K., Park, W.Y.: Demand prediction and price optimization for semi-luxury supermarket segment. Comput. Ind. Eng. **113**, 91–102 (2017)
20. Semenov, V.P., Chernokulsky, V.V., Razmochaeva, N.V.: Research of artificial intelligence in the retail management problems. In: 2017 IEEE II International Conference on Control in Technical Systems (CTS), pp. 333–336 (2017)

21. Aqil, D.S.M., Ali, S.M., Burney, S.: A survey of soft computing applications for decision making in supply chain management. In: 2017 IEEE 3rd International Conference on Engineering Technologies and Social Sciences (ICETSS), pp. 1–6 (2017)
22. Martínez-López, F.J., Casillas, J.: Artificial intelligence-based systems applied in industrial marketing: an historical overview, current and future insights. Ind. Mark. Manage. **42**(4), 489–495 (2013)
23. Milgrom, P., Tadelis, S.: How artificial intelligence and machine learning can impact market design. National Bureau of Economic Research, Cambridge, MA, w24282 (2018)
24. Simchi-Levi, D.: The new frontier of price optimization. MIT Sloan Manag. Rev. **59**(1), 21–26 (2017)
25. Simchi-Levi, D., Wu, M.X.: Powering retailers' digitization through analytics and automation. Int. J. Prod. Res. **56**(1–2), 809–816 (2018)
26. Syam, N., Sharma, A.: Waiting for a sales renaissance in the fourth industrial revolution: machine learning and artificial intelligence in sales research and practice. Ind. Mark. Manag. **69**, 135–146 (2018)
27. Chen, I.F., Lu, C.J.: Sales forecasting by combining clustering and machine-learning techniques for computer retailing. Neural Comput. Appl. **28**(9), 2633–2647 (2017)
28. Murray, P.W., Agard, B., Barajas, M.A.: Forecasting supply chain demand by clustering customers. IFAC-PapersOnLine **48**(3), 1834–1839 (2015)
29. Shakya, S., Kern, M., Owusu, G., Chin, C.M.: Neural network demand models and evolutionary optimisers for dynamic pricing. Knowl.-Based Syst. **29**, 44–53 (2012)
30. Efendigil, T., Önüt, S., Kahraman, C.: A decision support system for demand forecasting with artificial neural networks and neuro-fuzzy models: a comparative analysis. Expert Syst. Appl. **36**(3), 6697–6707 (2009)
31. Carbonneau, R., Laframboise, K., Vahidov, R.: Application of machine learning techniques for supply chain demand forecasting. Eur. J. Oper. Res. **184**(3), 1140–1154 (2008)
32. Feng, Q., Shanthikumar, J.G.: How research in production and operations management may evolve in the era of big data. Prod. Oper. Manag. **27**(9), 1670–1684 (2018)
33. Böse, J.H., et al.: Probabilistic demand forecasting at scale. VLDB Endow. **10**(12), 1694–1705 (2017)
34. Pinto Ferreira, R., Martiniano, A., Ferreira, A., Ferreira, A., Jose Sassi, R.: Study on daily demand forecasting orders using artificial neural network. IEEE Latin Am. Trans. **14**(3), 1519–1525 (2016)
35. McCormick, K., Abbott, D., Brown, M.S., Khabaza, T., Mutchler, S.R.: IBM SPSS Modeler Cookbook. Packt Publishing, Birmingham (2013)
36. Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J.: Classification and Regression Tree. Wadsworth & Brooks/Cole. Advanced Books & Software (1984)