11th International Scientific Conference
"Business and Management 2020"
May 7–8, 2020, Vilnius, LITHUANIA
Section: Business Technologies and Sustainable Entrepreneurship
http://www.bm.vgtu.lt

ISSN 2029-4441 / eISSN 2029-929X
ISBN 978-609-476-231-4
eISBN 978-609-476-230-7
https://doi.org/10.3846/bm.2020.604

STUDY OF DIFFERENT DATA SCIENCE METHODS FOR DEMAND PREDICTION AND REPLENISHMENT FORECASTING AT RETAIL NETWORK

Aleksei Iurasov[1], Giedre Stanelyte[2]

Department of Business Technologies and Entrepreneurship, Faculty of Business Management, Vilnius Gediminas Technical University, Saulėtekio al. 11, LT-10223, Vilnius, Lithuania
E-mails: [1] aleksei.iurasov@vgtu.lt (corresponding author); [2]giedre.stanelyte@stud.vgtu.lt

Received 11 March 2020; accepted 05 May 2020

Abstract. The demand prediction becoming an essential tool to remain or even lead in the competition among the retail businesses. A well-done demand prediction model could help retailer to track the level of inventory, orders and sales in the most effective way in which the best results could be achieved. However, there are many different methods and opinions of how to create a demand prediction model. In this paper, we will analyse the most commonly used methods of Linear regression, Logistic Regression, Probabilistic Neural Network, Bayesian Additive Regression Trees, Random Forest and Fuzzy Logic with their specifications and limitations found in studies of authors. After review performed all methods will be compared according to characteristics selected. Moreover, in order to get more practical results the accuracy of Logistic Regression and Random Forest methods will be compared based on data of milk sales collected from retail network. For constructing of decision support system for retail network, we need to go beyond demand prediction one-step to replenishment forecasting. It was concluded that there is no best method to forecast replenishment and results can differ based on the data and conditions analysing. In every situation authors seeking to select the method with the highest accuracy and the lowest number of errors possible. Limitations of research: limited number of goods and stores included in the modelling.

Keywords: demand prediction, replenishment forecasting, retail network, logistic regression, random forest.

JEL Classification: L81.

1. Introduction

Retail market today is one of the fastest growing markets in the world. This rapid growth of consumption and information technologies provides a number of opportunities for retail companies. A quick reaction and ability to work efficiently in changing business can deliver great results. Unfortunately, not all the organizations are working in the most effective way and modelling of demand prediction and replenishment here could be suggested as solution. A well-done demand prediction and replenishment model could help retailer not even to work more efficiently but also to increase company profit by saving the cost, increasing revenue and customer satisfaction. However even if decision to create a model would be accepted, very often retailers face the problem of lack of the knowledge about the methods those could be applied in modelling demand prediction in retail network. Moreover, as there are many different methods of demand prediction modelling, whether the best method for the most accurate prediction exists.

The aim of this study is to compare the application of different methods and by using the real time, data to evaluate the accuracy of two most commonly mentioned methods in practise. Tasks set for achieving the goal:

- To propose different mathematical methods for prediction and study its application in practice.
- To compare the methods by emphasizing its advantages and disadvantages.
- By forming demand prediction model with help of two different methods evaluate the accuracy of methods used.

Research object is enterprise demand prediction and replenishment modelling. Modelling prepared by using logistic regression and random forest methods. For modelling KNIME Analytics Platform was used. Due to the wide range of the products and business transactions in the analysing market model was formed based on sales of limited number of goods in defined number of stores.

2. Mathematical methods for demand prediction modelling

Demand prediction is the combination of two words. The first one – demand, and second – prediction. Before combining those words into phrase, and further reviewing different methods applicable, it is important to understand economical meaning and value of this phrase. Word demand – could be define as requirement of products and services, while prediction – in general, means making estimation in present for future events, at this case – future demand of products. Demand plays a vital role in the decision making of a business. In competitive market conditions, there is a need to take correct decision and make planning for future events related to business like a sale, production or inventory optimization. The better analysis of different factors are performed – the better decisions based on demand prediction results could be made. As the process by itself is difficult, there are many different methods invented and studies made based on the applications of those methods. In this part – application of the methods of linear regression, Logistic regression, Probabilistic Neural Network, Bayesian Additive Regression Trees, Random Forest and Fuzzy Logic will be reviewed and discussed.

The first method – linear regression is the basic and often used type of predictive analysis. Linear regression uses only one independent variable as a predictor, which has an effect to dependent variable (outcome). The main idea of regression analysis is to determinate the strength of predictor, to forecast an effect, and to use the model formed in forecasting the further results. However, the method when there is only one predictor is not clear enough to predict possible subsequent development of the analyzing process, therefore the authors commonly choose a *Multiple Linear regression.* This means that linear regression can be extended and instead of one independent variable, we have a many different variables (Anghelache, 2015). As the method is treated as clearly understandable, it is widely used in various studies and researches to predict outcomes selected by different authors. Author (Anghelache, 2015) used multiple linear regression to analyze the final consumption and gross investment influence to Romania GDP. Based on the data collected authors formed an equation and based on statistical tests evaluate the accuracy of the model. Authors conclude that higher number of factors in regression model allows the researcher to draw results that are more conclusive in macroeconomic analysis. Authors (Aghdaei et al., 2017) also used a regression in building of energy simulation model. The aim of the study was to predict the space heating and cooling requirements in different cities of Australia. Findings of the research show that the linear regression with simple independent variables can predict the requirements for space heating and cooling of the residential buildings in the specific climates within acceptable errors. It can even be applied in the studies where the relation between the financial news as an independence variable and the stock price of financial market is evaluating. In this case the author of the work named regression as machine learning-based approach (Ihlayyel et al., 2018). However, as method is quite simple, for better results it has to be used in combination with some rules, methods or algorithms. In the researches mentioned above to reduce the modelling cost of the parametric analysis authors (Aghdaei et al., 2017) use methods of Taguchi and Analysis of variance (ANOVA). Other author Ihlayyel mentioned above applied Enhanced ELR-BoW (Bag of Words) algorithm. Only after simplification of the data regression models were formed. The research of the authors (Cankurt & Subasi, 2015) also tried to compare the linear regression with neural network and agree with the idea above. Formed forecast shows that neural network present higher accuracy than the linear regression when there is no combination with linear regression and other methods.

When it comes to regression it is always important to mention – logistic regression. Unlike traditional linear regression – logistic regression is appropriate for modelling a binary variable. In wider perspective it means that results can procedure two outcomes (1 or 0), those could be considered as "positive" and "negative". Such results are useful in the practice however can't be received by using simple linear regression. This is mostly due to two main reasons. Firstly, a simple linear regression can only predict values outside the acceptable range. Secondly, as the dichotomous experiments can only have one of two possible outcomes for every experiment, the residuals will not be normally allocated about the predicted line. Performing analytics with logistic regression includes three main goals: prediction that the outcome or response variable equals to 1, categorization of outcomes and predictions and finally, access to the odds or risk associated with model predictors (Grömping, 2016). Logistic regression is considered as very important statistical procedure in predictive analytics in areas of health-care, medical analysis, social statistics and economy. The authors Joubert, Verster, and

Raubenheimer (2019) included logistic regression as commonly used method to predict probability components and loss severity in study of Loss Given Default (LGD) evaluation in banks. In authors study probability components were modelled by making use of logistic regression binary outcomes (write-offs or not write-offs). Moreover in study logistic regression was used in combination with method of survival analysis, what let to increase the model's predictive power and accuracy of results.

As it is mention above, one more method often included into studies is – Probabilistic Neural Network (PNN). PNN is the method of artificial intelligence that allow to form a complex nonlinear relationship between response variables and explanatory ones. The network was introduced in late 20ths – in 1990 by Specht. The main characters after presentation of network were – easy to use and possibility to interpret the network's structure in the form of a probability density function which is simple to understand. For those reasons the method is used in various sectors to analyze. Authors Penpece and Elma (2014) used neural networks for the purpose to forecast Sales Revenue in Grocery Retailing Industry. Based on results received authors stated that neural network method is more organic and predicting results better than the other methods. Revenue forecasts calculated based on neural networks were very close to actual data of sales revenue. However, the model also faced some problems of estimation of probability density function and high space complexity of PNN pattern layer. Moreover, model by itself has only one parameter of training and the smoothing parameter (σ), which must be optimized in order to make the network achieve the highest prediction ability. Based on different scientists' network is composed of 3 either 4 layers: an input layer, a pattern layer, a summation layer, and an output layer. Some scientists (Sun et al., 2017) does not count last layer and named third layer as Classes. The neurons in the input layer are simply the features of input vectors. The pattern layer consists of as many neurons as training examples. In the summation layer, the number of neurons is equal to the cardinality of classes in the data set. Finally, the output layer consists of a single neuron that provides the classification result. As the structure of the network is considered as a complex and probabilistic neural network is a frequently exploited model in the field of data mining in different researches of scientists, certain PNN reduction techniques have been established. These techniques include dynamic decay adjustment algorithm (DDA) (Berthold & Diamond, 1998) backpropagation mechanism (Sun et al., 2017), dimensionality reduction (Kusy, 2015) and also some other, presented earlier (in 1991–1994) – learning vector quantization (Burrascano, 1991) or maximum-likelihood algorithm techniques. All the techniques have the own specific parameters and the influence on the network. Further mainly parameters and the practical value of techniques are reviewed:

- Dynamic adjustment algorithm (DDA). Operation of algorithm required two phases – training and classification. New neurons are added if necessary. Less than five epochs are needed to complete training. The algorithm can be proven to terminate when a finite number of training examples is used. And finally, only two thresholds are required to be adjusted manually (Berthold & Diamond, 1998).
- Dimensionality reduction. For the reduction creation model by author (Kusy, 2015) two main steps – feature selection (methods of single decision tree (SDT) and random forest (RF)) and feature extraction (method of principal component analysis (PCA)) need to be follow. The main idea of the SDT is treated method as a predictive model – it maps the input data into desired targets. If the desired targets take the form of groups to which the data belong, SDT is treated as classification tree. After this process RF utilizes the collection of independent decision trees formed by SDT. Within the training process, the trees grow in parallel, not interacting until all of them have been grow. Once the training is completed, the need to move to the next phase appears and predictions of single trees are combined to make the overall prediction of RF. For the further step – principal component analysis is used. PCA is one of most popular feature extraction method to use. The methods combine the statistical technique which converts a set of input features into a set of new values by means of linear transformation. The results are names principal components and are linearly uncorrelated. With the help of method patterns of similarities and differences in data can be identified. Once these patterns are determined, the data can be compressed by decreasing the number of dimensions without significance loss of information. According to

the author Kusy (2015), who applied this variation of method for medical data classification tasks, the results showed an increase of prediction ability and decrease in computational time needed to complete the task in every single case.

- Backpropagation mechanism (BP). The authors (Parry et al., 2011) define the method as feed-forward neural network and included it into the test of prediction accuracy in the first-time adoption of DVD players. The authors forecasted performance of the logit model and the three neural network models. At the end it was found that the PNN algorithm significantly outperforms the logit model and the two remaining neural network algorithms. BP was one of those methods. However in 2017 scientists (Sun et al., 2017) present the idea of new PNN model in combination with backpropagation algorithm. New BP-PNN model has two phases – learning phase, where the idea is to receive the initialized value of the variable weights and training phase – where the error function is propagated back. Based on the analyses performed by the authors comparing with PNN, BP-PNN has less components in the pattern layer, which helps to reduce the space complexity of the model. Moreover, comparing with PNN, BP-PNN is designed with the much clearer structure and ability to identify the importance of indicators. Nevertheless, there are still some limitation of model emphasized by the authors – the model required further studies as the number of parameters trained is higher than in other models and thus requires a long time for calculations in the model.

Review emphasized that the method of PNN PNN needs to be combine with one of the algorithms reviewed for the purpose to receive the better results.

One more popular method – Bayesian Additive Regression Trees (BART) by authors Pratola et al. (2014) were described as nonparametrically method. Logan, Sparapani, McCulloch, and Laud (2019) define BART method as fully nonparametric and flexible model of prediction which can deal with complex functional forms as well as interactions among wide range of variables The authors Ajidarma and Irianto (2019) in their study agree with the mentioned definition by adding the idea that BART regression method harnesses dimensionally adaptive random basis elements. The method consisting from a prior and likelihood and is a function of an ensemble of trees. As method can combine and compare a number of different factors it is widely used in the researches of different authors. Method can use different algorithms with whose help more precise search of the model space and variation across the algorithm draws can be organized. However there are some problems as publicly available version of algorithm in R package can process only the limited number of observations (Pratola et al., 2014). Moreover the authors Linero and Yang (2018) performed a study for one more problem of decision threes to analyze. According to the authors in those methods – a high possibility of deficiency in ensembles is possible and could be caused by lack of smoothness and vulnerability to the curse of dimensionality. The idea of soft decision trees was suggested and implemented on the BART method. The authors demonstrate that their methods can have meaningful improvement over existing methods. Yet as there are still a lot of limitation, further studies of idea is performing. Ajidarma and Irianto (2019) have included the BART in the analysis and prediction the growth of electric automobile industry in different states of United States. BART method was used to analyze the relationship between several factors chosen and the sales of electric vehicles. With help of the algorithms (on this case Markov chain Monte Carlo (MCMC) algorithm was chosen) four BART models were generated and fitted into the data. The models identified the top predictors those correlation with the sales of electric vehicles were confirmed in the further steps. Authors of the work concluded that to use the method was beneficial – as the method enables a full assessment of prediction uncertainty while remaining highly competitive in terms of prediction accuracy.

Moving forward to Random forest method – it is one more decision trees method which consist of a chosen number of decision trees, which are used for classification and regression analysis (Feng & Wang, 2017). Authors Gupta, Rawat, Jain, Arora, and Dhami (2017) in their study of decision methods described this random forest method as tool that form the ability of multiple varied analyses, organization strategies, supply and demand prediction modelling, perceptive variables and importance ranking on the record-by-record basis for deep data understanding. Instead of tool authors Yin, Lee, and Wong (2012) define random forest method as the algorithm with an ensemble random method. Author Ghatasheh (2014) agree with the definitions

mentioned and emphasized some more positive attributes, such as an immunity to overfit, good estimation of internal errors, and high accuracy comparing with other learning algorithms. Because of these reasons' method is included in various calculations and studies performed. In 2017 authors Feng & Wang made a study, where the demand of the bicycle rental was evaluating. Two methods – multiple linear regression analysis and random forest were included into calculation. However, research shows that accuracy of linear regression model to forecast is too low even though the normal distribution of factors and good relationship between them was identified. Authors identified the high possibility of error due to specific characteristics of factors. The method was changed to random forest method and this improve the accuracy of result to 82%.

The last method reviewed in this article – Fuzzy method. The authors Agápito et al. (2019) characterize Fuzzy method as logic which with a help of specific set of rules can make an associations between linguistic and numeric data of a database. Rules can be provided by two groups – artificial intelligence algorithms, which are the target of researches in nowadays and by group of experts. The author Syahputra (2016) in her study agree with the definition about the logic, but additionally emphasized that this logical function recognizing only two parameters, either „Yes" or „No" („1" or „0"). The logic mostly used to create an expert systems and knowledge – based control settlements. However there are some limitations, those are important to know before analysing the method in more detail. Firstly, the method is case-dependent – every time the changing scenario can have a different influencing factors, where each of them need to be evaluated as important. Secondly, contribution of domain experts is of significant importance in process of forming control settlements (Yadav et al., 2018). Moreover there also could be the problem of too many rules. The problem usually arise among the rule-based models, when rules are created for every single factors and the outliers, those were not identified in the beginning of modelling are detected only in the process (Berthold, 2003). Depside it, the fuzzy logic method is applying in studies, yet usually in combination with other methods. The author Syahputra (2016) performed a study to predict a vehicle fuel consumption by using the combination of artificial neural networks and fuzzy logic (ANFIS). Prediction was made for different models of cars based on two criteria – weight and age. Results show that with an increase of weight of the motor vehicle, an amount of fuel needed to travel the same distance is increasing. Moreover car age also affects fuel efficiency. For the younger car, the higher efficiency of fuel consumption is calculated. The same combination of methods were also used by author Ridwan (2019) who used adaptive network-based fuzzy inference system (ANFIS) in the study to predict the price of good – lamp.

As all methods reviewed are different, for better understanding it is important to summarize what are the advantages and disadvantages of every method and what are the areas the method could be applied. The comparison of methods shows that there are no one perfectly suitable method to use (see Table 1). All the methods have their advantages and also areas to improve. For further calculations two methods: Logistic Regression (as the most popular for easy interpretation and implementation) and Random Forest (by studies emphasized as one of the most accurate learning algorithm) were selected for demand modelling and replenishment forecasting.

3. Methodology of Logistic regression and Random forest

Data processing for the practical part of this study is conducted by using KNIME Analytics Platform. KNIME (Konstanz Information Miner) is an open-source Big Data Analytics Platform, which used for data analytics and reporting. These processes organized in KNIME in form of workflows. The main principals of workflow are – visualization, modularity and easy extensibility (Berthold et al., 2009). The workflow consist of many nodes, those are processing the data and transporting results via connections between the nodes. The work in this workflow is organized based on the structure of 4 main stages: a) data collection; b) pre-processing c) model development and training and d) prediction and review of the results (scores) (Ranji et al., 2019). Included methods – logistic regression and random forest.

Logistic regression, as mentioned above models the probabilities for classification problems with two possible outcomes (1 & 0, e.g. "No action"&"Replenish"). The method is widely used in various fields of the practise. Logistic regression curve is constructed using the natural logarithm of the "odds" of the target variable. The formula of logistic regression:

$$P = \frac{1}{1 + e^{-(a+bX)}}, \tag{1}$$

where: P – probability of 1; e – the base of natural logarithm; a, b – parameters of the model.

Table 1. Advantages, disadvantages and application of mathematical methods proposed

Characteristics	Linear Regression	Logistic regression	PNN	BART	Random forest	Fuzzy logic
Basic principles of method	Determination of relationship between independent and dependent variables.	Prediction and determination of relationship between variables when the dependent variable is binary.	Classification of patterns based on learning from examples.	Creation of sum-of-trees model and regularization prior on the parameters of that model.	Combination of tree predictors where each tree depends on the values of a random vector sampled independently with the same distribution for all trees in the forest.	Form of many-valued logic in which the truth values of variables may be any real number between 0 and 1 both inclusive.
Authors	Anghelache; Aghdaei; Kokogiannakis, Daly & McCarthy; Ihlayyel, Sharef, Nazri & Bakar; Aghdaei et al.; Cankurt, Subasi.	Grömping; Joubert, Verster, & Raubenheimer.	Penpece & Elma; Sun,Wu&Li; Berthold & Diamond; Kusy;Sun; Burrascano;Parry, Cao& Song.	Pratola; Logan, Sparapani, McCulloch, & Laud; Ajidarma & Irianto; Pratola; Linero & Yang.	Feng & Wang; Gupta, Rawat, Jain, Arora, & Dhami; Yin, Lee, & Wong; Ghatasheh.	Agápito; Syahputra; Yadav, Kumar, Kumar, & Yadav; Berthold.
Advantages	– Simple estimation procedure; – Easy to understand interpretation on a modular level (i.e. the weights).	– Gives not only a measure of how relevant a predictor is, but also its direction of association; – Is easy to implement, efficient to train.	– Less time consumed to train virtually; – Relatively sensitive to outliers; – Can calculate probability scores.	– Provides a flexible approach to fitting a variety of regression models and algorithms while avoiding strong parametric assumptions.; – Able to represent interactions; – Can handle missing values.	– One of the most accurate learning algorithms; – Runs efficiently on large database; – Provide a reliable feature estimate; – Offer estimates of the test error and missing data.	– Using simple mathematics for non linear, integrated and complex systems; – Is based on linguistic model; – Has rapid operations; –Can handle trouble with inaccurate data.
Disadvantages	– Can oversimplifies the problems, and make a linear relationship where there is no such; –Is sensitive to outliers; –Linearity between variables. Hard to achieve in real world; – The data must be independent.	– Dependent variable is restricted; – If number of observations < number of features, lead to possible overfit.	– Required a lot of memory for data; – Usually long testing time; – Possible large computational cost.	– In publicly available version of algorithm in R package can process only the limited number of observations (Pratola et al., 2014); – A high possibility of deficiency due to lack of smoothness and vulnerability to the curse of dimensionality. (Linero & Yang, 2018).	– An ensemble model is inherently less interpretable than an individual decision tree; – Training a high number of trees could lead to large computational cost and can use a lot of memory.	– Case-dependent; – Contribution of domain experts is of significant importance in process of forming control settlements (Yadav et al., 2018) – There also could be the problem of too many rules; – Based on the size, high memory and cost required.
Business processes were methods are applied	Inventory, sales prediction, evaluation of marketing effectiveness, promotions.	Supply chain management, inventory, sales forecasting.	Selection of retail stores location, prediction of sale revenue, promotions.	Supply chain management, demand and orders prediction modelling, correlated outcomes.	Prediction of retail demand, optimal retail location, pricing.	Market trend analysis, retail service quality evaluation, consumption and demand predictions.

In the logistic regression the parameter *a* moves the curve left and right and parameter *b* – defined the steepness of the curve. To obtain model coefficients, those best relate predictors to target logistic regression uses maximum likelihood estimation method.

Random forest method – adding to the definitions above, this method also could be defined as an algorithm consisting of a collection of tree structured classifiers that created for independently and identically distributed random vectors. Each tree casts a unit vote for the most popular class for each input when the response is binary. For regression problems, the random forest forecast is the average prediction from the regression trees (Weng et al., 2018). Main steps of the method are showed in the Figure 1:

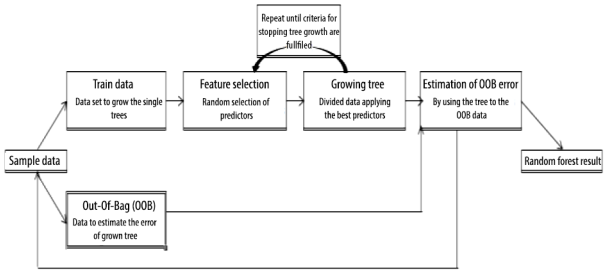

Figure 1. Random forest algorithm
(source: Boulesteix, Janitza, and Kruppa (2012))

Firstly, random samples from a given dataset are selected. Then when the data is selected, the best predictors is found and applied on the data analyzing. This step has to be repeated until criteria for stopping tree growth are fulfilled. Moreover, it is important to estimate and include into calculations the Out of Bag (OOB) error. Finally, when all the requirements are fulfilled, final prediction result is formed.

4. Demand prediction and replenishment modelling on behalf of Logistic regression and Random forest methods

In the practical part the KNIME workflow was created, the logistic regression and random forest methods based demand prediction and replenishment models were generated and assessed (see Figure 2).

As it is mentioned in the methodology part, there are some main steps to do to prepare the models. In the first step data is collected. Data selected to use in this practical case:

- Sales of one product – milk (dependent variable). Sales data collected from 10 different shops (product number, product name, quantity sold, etc.).
- Time information: year; day, day of month, day of week.
- Other additional information, as an independent variables – product delivery information (frequency, delivery time), number of customers per day, average basket information, discounts information, holidays, weather information.

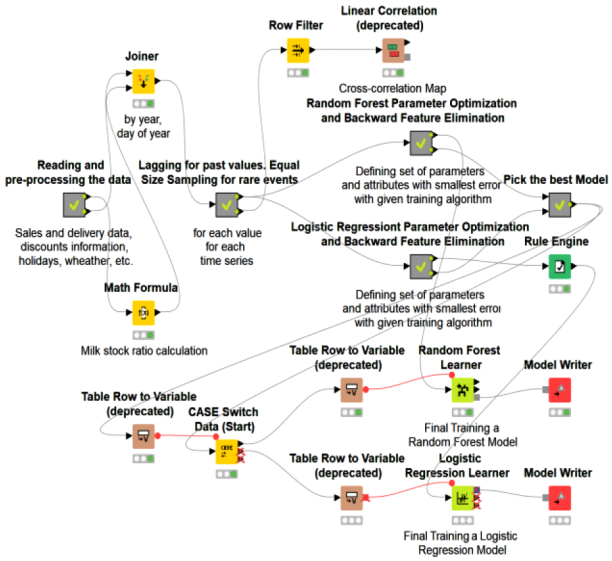

Figure 2. Demand prediction and replenishment models
(top level of KNIME workflow)

After uploading the data are prepared for the further use: new parameters, e.g. Milk ratio (Milk stock ratio) as an additional measure is calculated by dividing milk stocks (in every store in every time frame – day) by optimal order quantity. We did this because retail companies don't use demand prediction itself but as part of decision support to replenish goods. Therefore the ratio can better identify the need for replenishment than just the quantity of items currently available and quantity of items sold.

Actually, we need to forecast whether an action is needed: "No action" or "Replenish". The model evaluated the correlation between the last 10 ratios and provided it in the user-friendly matrix (see Figure 3).

For the purpose to improve the prediction accuracy and to find out which columns are necessary for the model the parameter optimization and backward feature elimination loops was added (see Figure 4). Feature elimination loop work in the way that lets to include in the output table the columns that is best for model construction. We specify an

error threshold and then the level with the fewest features that has a prediction error below the threshold is automatically selected. In any case all columns from the input table that are not present in the selected level are filtered from the input table. In this way only the necessary columns remain in the model what ensure the more accurate results. From the economic side – feature elimination loop helps the company to identify only the real attributes with significant impact to result. It define set of attributes with smallest error possible. Need of this step appear as in the beginning of modelling more than one different attributes are adding.

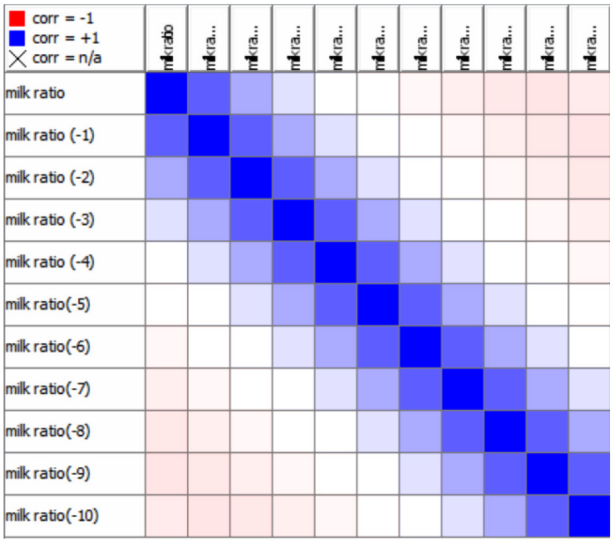

Figure 3. Milk ratios correlation matrix

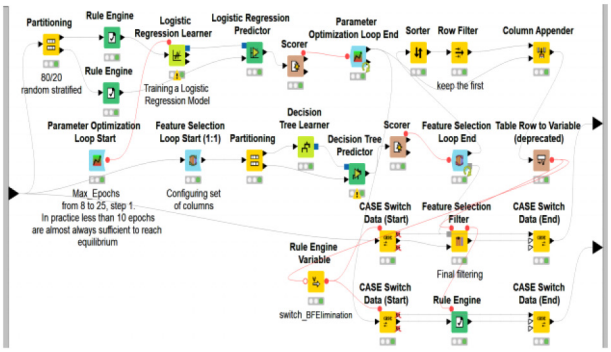

Figure 4. Logistic Regression Parameter Optimization and Backward Feature Elimination (second level of KNIME workflow)

However without separate procedures it would be hard to evaluate possible impact to result. And here feature elimination loop works as combined filter, where the attributes identified as high in level of error are eliminating from the model. In this specific case discounts information, holidays, weather and some milk stock ratios of previous periods (–3,…,–10) were identified as attributes those do not have a significant impact to decision of milk replenishment. All the others attributes were left as significant and meaningful in final calculations. After this step the final calculations are proceed based on eliminations made and results of the models are provided.

The results are showed that in this specific case of Milk product replenishment modelling logistic regression model is by 11% less accurate than the model prepared on behalf of Random Forest method (see Table 2).

Table 2. Logistic regression and Random forest methods accuracy results in prediction models formed

	Predicted data				
	Forecasting methods	Logistic regression		Random forest	
	Actions	No action	Replenishment	No action	Replenishment
Factual data	No action	2491	815	2916	390
	Replenishment	665	3289	307	3647
Accuracy		79.61%		90.40%	
Error		20.39%		9.60%	
Cohen's kappa (k)		0.587478		0.806048	

However the small difference and ideas marked before in the article only confirmed that every case is different and there is no one best method for prediction. Only the precise evaluation of data can shows which method provide the best quality on every specific study performing.

5. Conclusions

In the study Linear regression, Logistic regression, Probabilistic Neural Network, Bayesian Additive Regression Trees, Random Forest and Fuzzy Logic were suggested as the methods those could be used in modelling demand prediction of retail network. Practical application, advantages and disadvantages of each were evaluated and compared. Results of comparison showed all methods analysed have areas to improve. Recurring problems identified: possible deficiencies, high memory requirements, long testing time and possibly large computational cost.

In practical study the model of milk demand prediction and replenishment on behalf of two methods proposed in theoretical part was formed. Results showed that model formed on behalf of logistic regression method was by 11% less accurate than the model created with help of random forest method.

Authors with the help of this practical case also underline that accuracy of the different methods applied in modelling can be evaluated and compared. As every case in the practice is different there is no best method perfectly suitable for all the situations. The authors Sarkar and Mahapatra (2017) confirmed the statement by adding that in real-life situation, it is very difficult to know all the information about the demand. Moreover the information can differ base on the situation, therefore in every case factors affecting the model may vary. Only evaluation of model received can show which method provide the best quality on every specific study performing.

From economic perspective – practical case shows that modeling could help the business to understand better what are the attributes they need to include into considerations before making significant decisions. Discounts information, holidays, weather and some milk stock ratios were not attributes having significant impact to decision of milk replenishment. Their identification helps to adjust the model and by doing this to receive a more precise results.

As practical study evaluated only two methods proposed, future researches could focus on evaluation of more complex models with higher number of methods to include.

References

Agápito, A. D. O., Vianna, M. D. F. D., Moratori, P. B., Vianna, D. S., Meza, E. B. M., & Matias, I. D. O. (2019). Using multicriteria analysis and fuzzy logic for project portfolio management. *Brazilian Journal of Operations & Production Management*, *16*(2), 347–357. https://doi.org/10.14488/bjopm.2019.v16.n2.a14

Aghdaei, N., Kokogiannakis, G., Daly, D., & McCarthy, T. (2017). Linear regression models for prediction of annual heating and cooling demand in representative Australian residential dwellings. *Energy Procedia*, *121*, 79–86. https://doi.org/10.1016/j.egypro.2017.07.482

Ajidarma, P., & Irianto, D. (2019). Application of bayesian additive regression trees to analyze the growth of United States electric automobile industry. *IOP Conference Series: Materials Science and Engineering*, *598*(1). https://doi.org/10.1088/1757-899X/598/1/012035

Anghelache, C. (2015). Analysis of final consumption and gross investment influence on GDP – multiple linear regression model. *Theoretical and Applied Economics*, *22*(3), 137–142.

Berthold, M. R. (2003). Mixed fuzzy rule formation. *International Journal of Approximate Reasoning*, *32*(2–3), 67–84. https://doi.org/10.1016/S0888-613X(02)00077-4

Berthold, M. R., Cebron, N., Dill, F., Gabriel, T. R., Kötter, T., Meinl, T., Ohl, P., Thiel, K., & Wiswedel, B. (2009). KNIME – the Konstanz information miner. *SIGKDD Explorations, 11(*1*)*, 26–31. https://doi.org/10.1145/1656274.1656280

Berthold, M. R., & Diamond, J. (1998). Constructive training of probabilistic neural networks. *Neurocomputing*, *19*(1–3), 167–183. https://doi.org/10.1016/S0925-2312(97)00063-5

Boulesteix, A., Janitza, S., Kruppa, J., & König, I. R. (2012). Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. *WIREs*, *2*(6), 493–507. https://doi.org/10.1002/widm.1072

Burrascano, P. (1991). Learning vector quantization for the probabilistic neural network. *IEEE Transactions on Neural Networks*, *2*(4), 458–461. https://doi.org/10.1109/72.88165

Cankurt, S., & Subasi, A. (2015). Comparison of linear regression and neural network models forecasting tourist arrivals to Turkey. *Eurasian Journal of Science & Engineering, 1*(1), 21–26.

Feng, Y., & Wang, S. (2017). A forecast for bicycle rental demand based on random forests and multiple linear regression. *Proceedings – 16th IEEE/ACIS International Conference on Computer and Information Science, ICIS 2017*, 101–105. https://doi.org/10.1109/ICIS.2017.7959977

Ghatasheh, N. (2014). Business analytics using random forest trees for credit risk prediction: A comparison study. *International Journal of Advanced Science and Technology*, *72*, 19–30. https://doi.org/10.14257/ijast.2014.72.02

Grömping, U. (2016). Practical guide to logistic regression. *Journal of Statistical Software*, *71*. https://doi.org/10.18637/jss.v071.b03

Gupta, B., Rawat, A., Jain, A., Arora, A., & Dhami, N. (2017). Analysis of various decision tree algorithms for classification in data mining. *International Journal of Computer Applications*, *163*(8), 15–19. https://doi.org/10.5120/ijca2017913660

Ihlayyel, H. A. K., Sharef, N. M., Nazri, M. Z. A., & Bakar, A. A. (2018). An enhanced feature representation based on linear regression model for stock market prediction. *Intelligent Data Analysis*, *22*(1), 45–76. https://doi.org/10.3233/IDA-163316

Joubert, M., Verster, T., & Raubenheimer, H. (2019). Making use of survival analysis to indirectly model

loss given default. *ORiON*, *34*(2), 107–132. https://doi.org/10.5784/34-2-588

Kusy, M. (2015). Dimensionality reduction for probabilistic neural network in medical data classification problems. *International Journal of Electronics and Telecommunications*, *61*(3), 289–300. https://doi.org/10.1515/eletel-2015-0038

Linero, A. R., & Yang, Y. (2018). Bayesian regression tree ensembles that adapt to smoothness and sparsity. *Journal of the Royal Statistical Society*. *Series B: Statistical Methodology*, *80*(5), 1087–1110. https://doi.org/10.1111/rssb.12293

Logan, B. R., Sparapani, R., McCulloch, R. E., & Laud, P. W. (2019). Decision making and uncertainty quantification for individualized treatments using Bayesian Additive Regression Trees. *Statistical Methods in Medical Research*, *28*(4), 1079–1093. https://doi.org/10.1177/0962280217746191

Parry, M. E., Cao, Q., & Song, M. (2011). Forecasting new product adoption with probabilistic neural networks. *Journal of Product Innovation Management*, *28*(Suppl 1), 78–88. https://doi.org/10.1111/j.1540-5885.2011.00862.x

Penpece, D., & Elma, O. E. (2014). Predicting sales revenue by using artificial neural network in grocery retailing industry: A case study in Turkey. *International Journal of Trade, Economics and Finance*, *5*(5), 435–440. https://doi.org/10.7763/ijtef.2014.v5.411

Pratola, M. T., Chipman, H. A., Gattiker, J. R., Higdon, D. M., McCulloch, R., & Rust, W. N. (2014). Parallel bayesian additive regression trees. *Journal of Computational and Graphical Statistics*, *23*(3), 830–852. https://doi.org/10.1080/10618600.2013.841584

Ranji, R., Thanavanich, C., Sukumaran, S. D., Kittiwachana, S., Zain, S., Sun, L. C., & Lee, V. S. (2019). An automated workflow by using KNIME analytical platform: A case study for modelling and predicting HIV-1 protease inhibitors. *Progress in Drug Discovery & Biomedical Science*, *2*(1), 4–8. https://doi.org/10.36877/pddbs.a0000035

Ridwan, M. (2018). Prediction of lamp price using adaptive neuro fuzzy inference system. *ICCSET 2018* (pp. 742–751), 25–26 October 2018. Kudus, Indonesia. https://doi.org/10.4108/eai.24-10-2018.2280522

Sarkar, B., & Mahapatra, A. S. (2017). Periodic review fuzzy inventory model with variable lead time and fuzzy demand. *International Transactions in Operational Research*, *24*(5), 1197–1227. https://doi.org/10.1111/itor.12177

Sun, Q., Wu, C., & Li, Y. L. (2017). A new probabilistic neural network model based on backpropagation algorithm. *Journal of Intelligent and Fuzzy Systems*, *32*(1), 215–227. https://doi.org/10.3233/JIFS-151415

Syahputra, R. (2016). Application of neuro-fuzzy method for prediction of vehicle fuel consumption. *Journal of Theoretical and Applied Information Technology*, *86*(1), 138–150.

Weng, B., Lu, L., Weng, B., Lu, L., Wang, X., Megahed, F. M., & Martinez, W. (2018). Predicting short-term stock prices using ensemble methods and online data sources predicting short-term stock prices using ensemble methods and online data sources. *Expert Systems with Applications*, *112*, 258–273. https://doi.org/10.1016/j.eswa.2018.06.016

Yadav, H. B., Kumar, S., Kumar, Y., & Yadav, D. K. (2018). A fuzzy logic based approach for decision making. *Journal of Intelligent and Fuzzy Systems*, *35*(2), 1531–1539. https://doi.org/10.3233/JIFS-169693

Yin, Y., Lee, C., & Wong, Y. (2012). Demand prediction of bicycle sharing systems. (2), 1–5. http://cs229.stanford.edu/proj2014/Yu-chun%20Yin,%20Chi-Shuen%20Lee,%20Yu-Po%20Wong,%20Demand%20Prediction%20of%20Bicycle%20Sharing%20Systems.pdf