Machine Learning Solution based on Gradient Descent Algorithm for Improved Business Process Outcomes*

Ivana Dimitrovska[1], Toni Malinovski[2], and Dane Krstevski[2]

[1]Faculty of Informatics, FON University, Skopje, Republic of North Macedonia

[2]National Bank of the Republic of North Macedonia, Skopje, Republic of North Macedonia

ivana.dimitrovska@fon.mk, tmalin@nbrm.mk**, krstevskid@nbrm.mk

Abstract: This study aims to provide guidelines that can help organization identify preconditions before they can employ machine learning, as well as provide evidence that machine learning can be used to improve business process outcomes. It considers supervisory learning as a business learning strategy, and employs machine learning solution based on gradient descent algorithm in large enterprise company in North Macedonia. The solution was designed to streamline the business process, automate the activities, and provide resilience to employees' subjectivity, wrong decisions, and human errors. The machine learning solution was used in production for ten months, including period of changes in the business process, and its average accuracy was 95.018% compared to the employees' decisions. Hence, it verifies the appropriateness of the chosen approach, with predictive accuracy that can be meaningful in practice. Although it is a specific case study, it provides valuable information that organizations can use while undertaking similar initiatives.

Keywords: machine learning, linear regression, gradient descent algorithm, business process improvement

*The published version on this paper in International Journal of Business Innovation and Research, Vol.24, No.4, is available on the following link: https://www.inderscience.com/info/inarticle.php?artid=114047 .
The views expressed in this paper are those of the authors and do not represent the views of any institution.

**Corresponding author.

1. Introduction

Companies take different initiatives to improve, redesign or even completely restructure their business processes, aiming to reduce process completion time, decrease cost and increase output quality. Commercial companies strive to improve the way they do business and produce their final service or products more effectively, and optimally increase their profits (Harmon, 2014; Jeston & Nelis, 2014; Yousfi, et al., 2016). Similarly, nonprofit organizations are concerned with efficiency and productivity, while trying to achieve their strategic goals and objectives (Harmon, 2014). Thus, process thinking, rather than focus on singular task or function, creates an approach that can deliver increased business values while trying to achieve process excellence within the company (Davenport, 2013). In addition, the information technology (IT) is identified as key enabler for business process improvement or redesign, which can help eliminate human labor or intermediaries, change process sequence, capture process information and provide analysis, as well as monitor process status and outcomes (Davenport, 2013; Harmon, 2014; Jovanoski, Malinovski, & Arsenovski, 2017, Metzger et al., 2015). Latest IT trends such as cloud and ubiquitous computing (Schulte et al., 2015; Yousfi, et al., 2016), internet of things and big data analytics (Riggins & Wamba, 2015; Zancul et al., 2016), as well as service-oriented computing (Tan & Zhou, 2013) are increasingly adopted by industry to increase visibility, agility and efficiency of business processes (Metzger et al., 2015).

On the other hand, machine learning (ML) has progressed dramatically over the last two decades, from laboratory curiosity to a mature technology that has widespread commercial use (Bottou, 2015; Jordan & Mitchell, 2015). ML enables computers to automatically improve (learn) from large amounts of data without being explicitly programmed to do so, while absorbing new behaviors and functions over time. It has already been successfully applied in different areas, such as self-driving cars (Jain et al, 2015), new medical treatments (Miljkovic et al., 2016), large-scale data processing solutions (Meng et al, 2016), as well as practical software for speech recognition, natural language processing, robot control (Jordan & Mitchell, 2015), etc. In like manner, companies are begging to understand the importance of ML and are trying to employ different ML approaches to analyze or manipulate their growing data assets in order to improve their business processes. ML can facilitate development of complex models and algorithms that can provide predictive analytics and uncover hidden insights, learned from historical relationships and trends in the datasets (Bottou, 2015; Jordan & Mitchell, 2015; Kelleher, Mac Namee, & D'Arcy, 2015), which can be further used for business process improvement. Hence, ML can bring significant value to businesses, from advances in their core functions to transformation of operations, services and products they provide. There are many examples that elaborate business benefits of ML, such as bankruptcy prediction using ML and financial expertise (Yu at al., 2014), ML techniques for customer churn prediction (Vafeiadis et al., 2015), ML approach for stock price prediction (Leung, MacKinnon, & Wang, 2014), prediction of repeat visits and likelihood of purchase for online customers (Shmueli & Koppius, 2011), etc.

Even though ML is an innovative approach, it is not an ultimate solution and not every business problem should be addressed with ML. It is based on algorithms that learn from business datasets, which do not rely on specific programing, and therefore companies must address number of issues in advance. Each company that tries to employ ML will be faced by a challenging task that encompasses: (1) identification of a problem that can be enhanced via ML solution, (2) determination of relevant data that should be collected, (3) choice of an appropriate learning method, (4) practical ML implementation, and (5) proper evaluation of ML benefits to the business (Jordan & Mitchell, 2015; Wagstaff, 2012).

This study researches the benefits of ML usage into business context, possibilities for process improvement for increased outcomes, as well as identification of possible pitfalls during implementation. In our methodological approach, we propose an empirical predictive model that utilizes an algorithm based on linear regression analysis, which is further verified via practical ML implementation into core business solution in a large enterprise company in North Macedonia. Hence, we provide an example for ML method that enables transformation of large datasets into intelligent information with increased business value. Having in mind the importance of proper evaluation after the ML implementation (Jordan & Mitchell, 2015; Yu at al., 2014; Wagstaff, 2012), we give adequate measurements for usability of the solution and acquired benefits, such as precision and accuracy during execution, achieved business benefits, possible performance issues, etc. Although it is a specific case study, we believe that the learned experiences can be generalized to other organizations and contexts, and thus provide value to the business community.

2. Theoretical background

ML represents a field of study that focuses on computer algorithms for data transformation into intelligent actions. It originated in environments where available datasets, statistical methods and computing power have simultaneously rapidly evolved over time (Bottou, 2015; Lantz, 2013). In like manner, modern datasets used by business processes are rapidly growing in size and complexity, and there is a pressing need to develop solutions that can harness this wealth of data and create additional business value for the organizations. Business intelligence (BI) and analytics of large and complex datasets can be used to determine future trends based on past observations of business process execution (Chen, Chiang, & Storey, 2012; Metzger et al., 2015). BI can forecast the impact of business process changes to certain extend, but ML enables computers to learn from experience and utilize it so its performance improves up on similar experiences in the future (Lantz, 2013). Thus, ML can create an additional value while transforming business data assets (Chen & Zhang, 2014; Gomez-Uribe & Hunt, 2016; Leung, MacKinnon, & Wang, 2014), which can further change the way companies are doing business and provide competitive advantage.

Typically, learning process is not complete until the learner is able to use its abstracted knowledge for future action. ML methods are generally divided in three subdomains: (1) supervised learning, (2) unsupervised learning and (3) reinforced learning (Garcia-Sillas et al., 2016; Jordan & Mitchell, 2015; Lodhia, Rasool, & Hajela, 2017). Supervised learning uses algorithms that are trained via classified examples where input data and correct output are known. Unsupervised learning does not require labeled data for training and needs inputs without any desired outputs, while the algorithm must discover patterns in the data on its own (Lodhia, Rasool, & Hajela, 2017; Solanki & Dhankar, 2017). Reinforcement learning learns from feedback received through interactions from an external environment, while trying to retro-feed the utilized model in order to improve (Dunjko, Taylor, & Briegel, 2016; Lodhia, Rasool, & Hajela, 2017). Some researches consider an additional domain, semi-supervised learning, which tackles the problem of classification when only a small subset of observations have corresponding class labels (Attoh-Okine, 2017; Kingma et al., 2014; Solanki & Dhankar, 2017). In each ML case, the machine is able to learn and absorb new functions over time from the utilized dataset (structured and unstructured), which may increase in the future, without explicitly being programmed to do so.

The most widely used ML method is the supervised learning method (Garcia-Sillas et al., 2016; Jordan & Mitchell, 2015; Lodhia, Rasool, & Hajela, 2017; Solanki & Dhankar, 2017). Supervised learning systems can enhance spam classifiers of e-mail, face recognizers over images and medical diagnosis systems for patients (Jordan & Mitchell, 2015), but also can be used to predicts company performance based on available reports (Qiu, Srinivasan, & Hu, 2014), datamining for business analytics (Shmueli, Patel, & Bruce, 2016), detection of behavioural deviations while monitoring business processes (Cabanillas et al., 2014), etc. Generally, supervised learning facilitate two learning tasks: classification and regression. Through regression, a learning process can be performed, which enables learning by demonstration from a given data set (Garcia-Sillas et al., 2016; Lantz, 2013). Regression algorithms are part of the massive set of ML algorithms that are developed to build machine-studying models, which further employ an iterative machine gaining knowledge of the desired process. Solanki & Dhankar (2017) list some of ML regression algorithms, such as simple least square regression, linear regression, analytical regression, stair regression, multivariate changing regression and locally estimation based regression.

This study considers supervisory learning as a business learning strategy. During business process improvement projects we may not know exact inner relations of the large dataset that we are processing, but we usually do know the desired output. Therefore, a properly chosen algorithm can learn to predict the correct answers from the available business training set. Since linear regression algorithms are widely used methods for various predictive purposes (Devore, 2015; Fumo & Biswas, 2015; Harrell, 2015; Lantz, 2013), it is a legitimate choice to use one of them as a mechanism for minimizing the error produced as a difference between the real output and obtained value after each algorithm iteration during improvement of the chosen business process. Similarly to Schmidt et al. (2014) that try to discover how data affects the perceived advantages in each business process via linear regression, as well as Hopkins & Ferguson (2014) that asses regression analysis in family business research, we explore ML with linear regression based on gradient descent algorithm, as valuable technique that can be used to improve business process outcomes. As certain studies suggest for similar algorithms (Allen-Zhu & Hazan, 2016; Johnson & Zhang, 2013; Shamir & Zhang, 2013), we use gradient descent in practical ML implementation to minimize the error of the proposed model and the chosen business training dataset and thus provide additional value to the business processes.

3. Gradient descent based ML for business process improvement

3.1. Design

The main purpose of this study is to use ML for improvement of business process outcomes while extracting knowledge from large amounts of business data, and further utilize the extracted knowledge for predictive analytics. Still, a starting key point is a decision whether a specific business objective can be approached via ML implementation. Having in mind the potential ML misuse (Jordan & Mitchell, 2015; Ridgeway, 2013; Wagstaff, 2012), we believe that the following preconditions should be fulfilled before one can employ ML to improve business process outcomes:

- Basic relationship between data input and output cannot be clearly understood,
- The business process cannot be clearly defined with behavioral statements and rules, and therefore it cannot be translated into a programming code,
- Business case is defined with specified input and output parameters,
- Business case uses large and comprehensive dataset,
- The functional relationship between the data input and output is constantly evolving, which requests real-time recalculation and fast adaptation,

- There is a need for parallel architecture in a request for rapidly delivered and compact solutions.

Even though certain problems in the business processes can be tackled with traditional programming tools, the business process improvement strategies that face the stated preconditions can significantly benefit from utilization of ML solutions. Furthermore, the implementation of the gradient descent to update the parameters of a specific ML model, can help minimize the difference between the expected (or ideal) outcomes and the actual business outcomes in the working environment. Hence, we have decided to leverage a ML solution based on gradient descent algorithm, attempting to minimize the error in the ML model's predictions.

In line with our statement that business processes using a large and comprehensive dataset can be improved through ML, the normalization of the dataset is also an important step that must be undertaken during the implementation of the solution. As indication in Jayalakshmi & Santhakumaran (2011), the normalization process for the raw inputs has great effect on preparing the data to be suitable for the training of the ML solution. In addition, Jayalakshmi & Santhakumaran (2011) elaborate different normalization techniques that can enhance the reliability of trained neural networks (specific set of algorithms that have revolutionized ML), such as statistical or Z-score normalization, min-max normalization, median normalization, sigmoid normalization, etc. They also emphasize that performance of the chosen solution to tackle specific problem is dependent on the normalization method. Similarly, proper normalization method can increase the performance of the ML solution, which will ultimately improve the prediction process.

Finally, ML can be beneficial when the input in a specific business process has evolving relationship with the expected output, and thus proactive approach is needed that can quickly adapt with favorable outcomes. Once the ML solution is implemented to improve a business process, the changes in the process itself over time, such as variable inputs or new unpredictable influential factors, must be quickly absorbed by the solution, which should continue to produce similar quality in outputs as before the introduced change.

To support our claims, we present a case study of a large enterprise company in North Macedonia that provides various products and services to large number of customers, where we improved the purchasing model, which is part of the Supply Chain Module in an ERP application, through ML. The solution was deployed in SAP ERP system. Even though ERP solutions are usually associated with business process reengineering initiatives (Ahmad & Cuenca, 2013; Chang, 2016; Hong et al., 2016), we further demonstrate that proper ML implementation in the ERP can provide benefits to the ongoing business processes.

The researched ML solution aims to improve one of the core business processes in the chosen company, which can increase its competency on the market, profit, as well as decrease the process competition time. The business processes involves creation of purchase orders (POs) towards specific supplier of goods or services. Each PO contains 24 items, including PO number, order type, quantity, vendor ID, material number, etc. It is created by authorized employees that have to select a strategy for PO confirmation, processed later by one or more employees from different departments. Having in mind the complexity in this company, proper selection of the PO confirmation strategy is vital for the favorable outcome of this business process. An error will streamline the process in wrong direction and important purchase might be significantly delayed or canceled. Since, the strategy selection is performed by different employees, a ML solution that can automatize this process can improve the businesses outcomes and minimize errors.

3.2. Procedure

Since each PO contains number of input items (variables), a selection of significant ones as inputs into the ML algorithm should be the first step towards the implementation of the ML solution. Through analysis of the previously created POs, we have identified the variables that are significant to the learning process, and omitted ones that don't influence the strategy selection process or jeopardize in any manner. Hence, we have selected 10 variables with ML relevance out of initial 24 variables, such as order type, vendor ID, payment terms, item category, currency, etc. and omitted PO number, document date, item number, material number, etc. Inappropriate selection of input variables can produces inaccurate results, which can diminish the improvement process. Thus, we have omitted variables that have constant values and variables that don't provide relevant information, such as incremental ID numbers, etc.

Based on the variable selection process and the chosen gradient descent for the ML solution that uses an input array of 10 variables, we have constructed the following ML linear equation for the chosen business process:

$$Y(x) = Q_1x_1 + Q_2x_2 + Q_3x_3 + Q_4x_4 + Q_5x_5 + Q_6x_6 + Q_7x_7 + Q_8x_8 + Q_9x_9 + Q_{10}x_{10} \quad (1)$$

This algorithm contains ten coefficients (Q_1, Q_2, … Q_{10}) that should be derived for calculation of the output variable, which represents the PO strategy. Hence, the implemented ML solution should determine the complex relationship between the input array and output, and represent it as these coefficients. Different changes in the input array, should be reflected as appropriate coefficients' change according the learned behavior in the ML solution.

Considering that the chosen business process uses large and comprehensive dataset, which can have different values, we have developed a code book that maps all possible values for each variable into numerical format, so they can be subjected to the referred algorithm. In addition, we have employed normalization techniques based on sigmoid function that is used excessively in neural networks (Jayalakshmi & Santhakumaran, 2011; Tax et al., 2017; Wanto et al., 2017). Hence, we have adapted the sigmoid normalization for the implemented ML solution, based on the following algorithm used to transform the input array for data normalization:

$$Xn = \frac{X - Avg(\bar{X})}{\frac{Dev(\bar{X})}{2}} \quad (2)$$

Since the data distribution is unpredictable in this business process, the normalization process of rescaling the variables in the input array to the range of 0 to 1 according to the sigmoid function is one of the preprocessing steps that leads to improved performance in the chosen ML solution. Similarly to the array of input variables, the output of the ML algorithm, which is the chosen release strategy, was normalized in the same manner using the adapted sigmoid function. Before the ML solution started the calculation process, we have assigned initial values to the coefficients. These coefficients are updated during the calculations and their values are also normalized to improve performance and effectiveness of the solution.

Figure 1 depicts the process for data preparation and normalization. The business data that is stored in the application database is extracted, prepared according to the needs of the ML solution and normalized using the referred algorithm for each component.

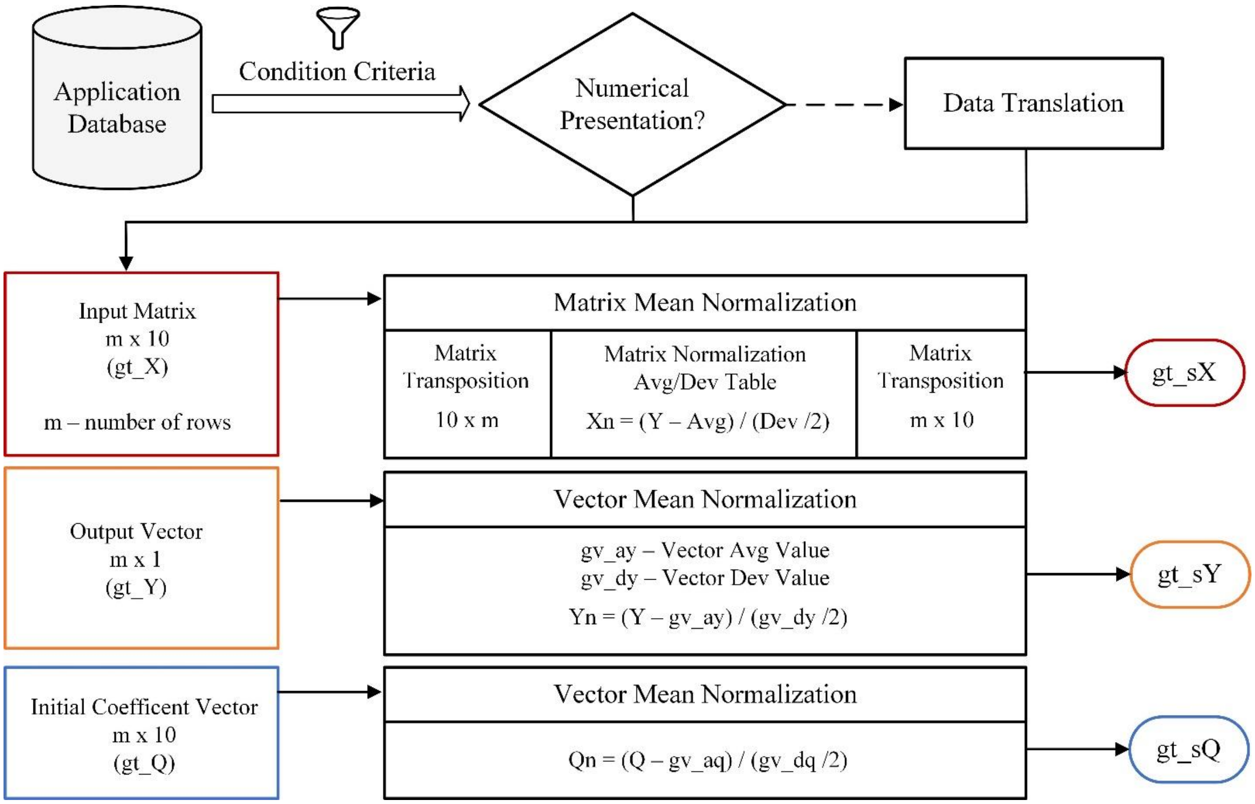


Figure 1. Data preparation and normalization for the ML solution

The input array is represented with a matrix (gt_X) that contains the POs (rows) in the application database and their 10 relevant variables (columns) for the ML solution. To simplify the data normalization of the input matrix, we have transformed the variables into rows (m x 10 => 10 x m), and reverted back after the normalization process of the input. The output and its normalization is also depicted in the Figure 1 (gt_sY), which represents the release strategy for each of the POs present in the application database. This step of the ML process is important when deployed in business environment, since the complex dataset of the business process can be simplified to increase performance during calculations. The coefficients (initial and calculated over time based on the learning process) are also subjected to data normalization, and further fed with the input and output data to the gradient descent based ML algorithm.

Figure 2 illustrates the linear regression used in this ML solution, based on the gradient descent algorithm, which supports supervisory learning as a business learning strategy in this study. This process accumulates historical data of passed activities and updates the value of the coefficients for closer representation of the relationship between the input and output data, towards a desired PO release strategy. As long as the learning process is active in this ML solution, every single iteration updates the values of the coefficients when needed, to improve the output of this business process.

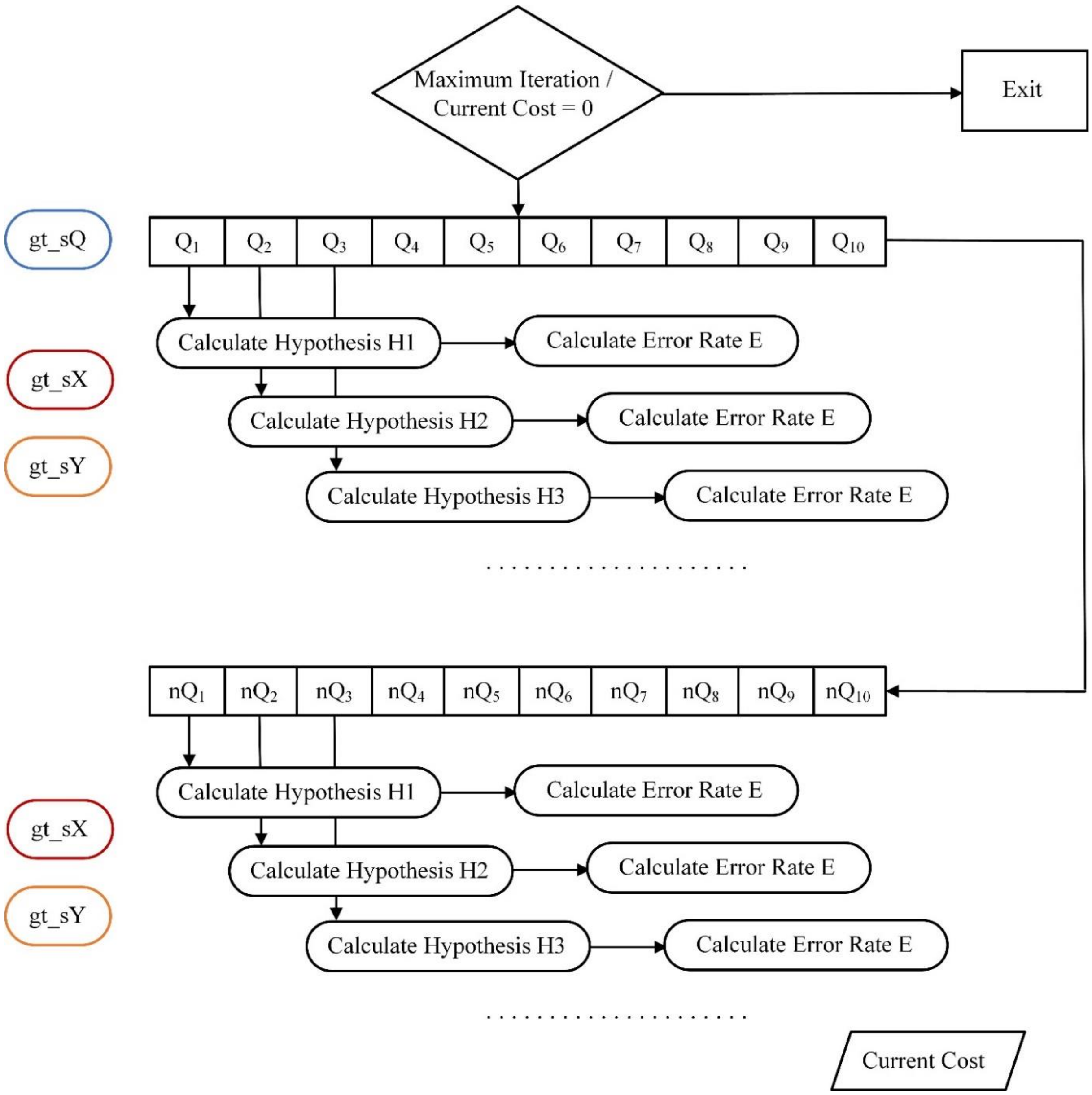


Figure 2. Linear regression based on gradient descent algorithm utilized in the ML solution

As illustrated in the figure, the ML solution uses the normalized dataset to calculate the coefficient values based on the application data set. The input and output for each PO in the database is used for the calculation of the coefficients. The algorithm also calculates the hypothesized output value, which is compared to the actual output in the database in order to estimate the learning rate. The new generated values for the coefficients (nQ_1, $nQ_{2,}$ … nQ_{10}) are also verified using the next hypothesized value and actual output, while trying to minimize the error during calculation. Hence, the learning process is repeated until the error rate is zero or a constant minimal value that is repeated in number of subsequent calculations. Finally, the obtained values for the coefficients represent the compiled version of the ML solution, which can be later used by current POs to predict the output release strategy and thus improve the business process and streamline the decisions. This solution retains its learning capabilities and these values can be improved during the practical utilization according to the changes in the business process.

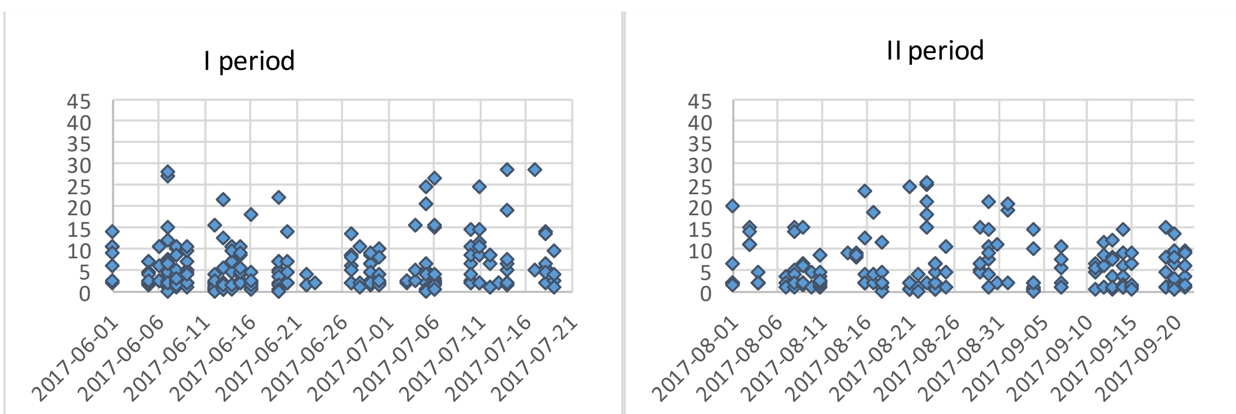
4. Results

The complied version of the ML solution, which was generated by analyzing the existing database, was deployed in everyday activities for the chosen business process. To evaluate the effectiveness of the proposed methods, the ML solution was used to predict an output release strategy, which was recorded and further compared with actual employees' decision, responsible for PO creation. Hence, this ML solution is intended to fully automate the process and streamline the decisions, or to provide a virtual assistant based on a proposal to an employee, so the error during human decision is minimized among different employees.

The ML solution was utilized in practice for 10 months, while processing around 1491 distinct POs as a result of the proposed gradient descent algorithm. We have analyzed the data for the POs, generated coefficients and outputs in four distinct periods, so we can detect the differences between the ML proposed release strategy and actual release strategy chosen by the employees. The four periods had similar distribution of POs and covered business days with increased generation of POs, as following:

- I period: 01 June – 20 July 2017
- II period: 01 August – 21 September 2017
- III period: 01 December 2017 – 04 January 2018
- IV period: 23 January – 17 March 2018

The deviation from the ML predicted release strategy and actual release strategy chosen by the employees in the four periods is illustrated in Figure 3.



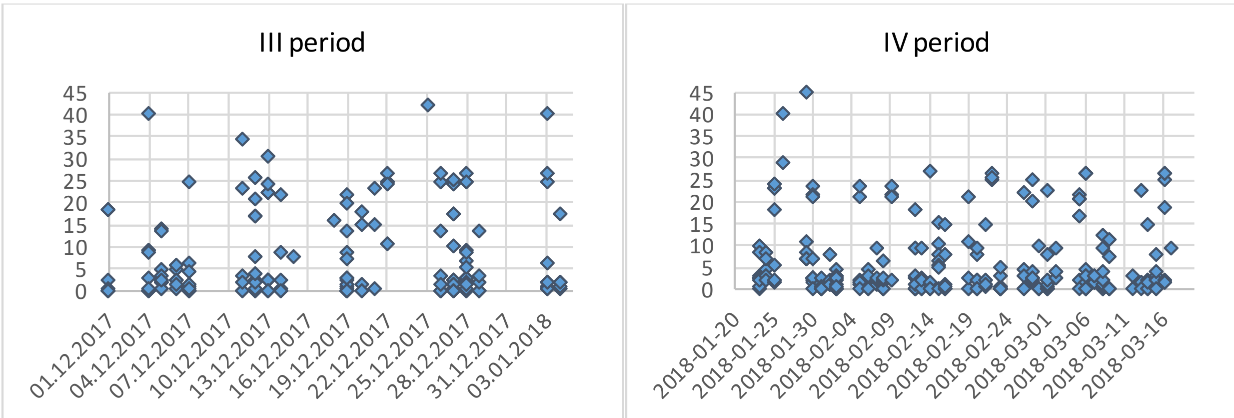


Figure 3. Deviation between ML proposed and actual release strategy in the four distinct periods

During the first two periods the business process was without novelties, while during the third period new release strategies were introduced in the process, which resulted in slight changes in the ML solution. In addition, the comparative statistical information for the predicted and actual release strategy is especially important in this period, since it illustrates the influence of these changes on the effectiveness of the ML solution. The fourth period did not introduce any changes in the process, so the deviations as a result of the previous period were minimized over time.

The results show that most of the predictions for output strategy are within 90-100% accuracy with the actual release strategy chosen by the employees. More precisely, 1262 from 1491 predicted outputs by the ML solution have above 90% accuracy. In average, the ML accuracy reached 95.313% in period I, 95.359% in period II, 93.925% in period III, and 95.132% in period IV of the analysis. Average accuracy for the 10 months was 95.018%. The ML solution demonstrated slight decrease in accuracy during the III period due to changes in the business process, but it regained the baseline performance during the last period. Hence, unpredicted changes in the business environment can slightly influence the effectiveness of the ML solution, but over time it can cope with these novelties while redefining the coefficients in the algorithm to perform effectively.

The deployment of the proposed ML solution in the real-life business environment showed that ML approach to improve business processes provides multiple advantages. In this example, it can be used to streamline PO release strategies based on the historical data, and thus provide competitive value to the business. It also abstracts the influence of the individual employees, including subjectivity, wrong decisions, and human errors. Hence, the organization was enabled to achieve consistent strategy within one of its core business processes and transform business information into knowledge. After a successful deployment and monitoring of its performance, this ML solution can ultimately replace the human labor during decision making process, and fully automate this business process.

5. Discussion and conclusion

ML enables computers to automatically learn from large amounts of data without specific programing, which has led to a widespread of ML applications in different areas. Different studies have already shown that organizations are starting to realize the benefits of ML while transforming their data assets to increase effectiveness in their everyday processes (Gomez-Uribe & Hunt, 2016; Jonsson et al, 2016; Leung, MacKinnon, & Wang, 2014; Wang, Kung, & Byrd, 2018). However, not every business process improvement should be addressed with ML. Each company that considers using ML to improve its processes should identify the business processes that can leverage this approach, as well as evaluate the benefits it will bring to the business (Jordan & Mitchell, 2015; Wagstaff, 2012). As indicated in Wagstaff (2012), the choice of an appropriate learning algorithm should be based on data that originated in the real world, with results communicated back to the origin. In addition, the quantitative improvements in performance should be accompanied by an assessment of whether those gains matter to the everyday activities performed in the company, outside of the machine learning research.

Having in mind that machine learning remains a young field despite its recent practical and commercial successes, this study provides the following contributions:

- Practical guidelines that can help organization identify preconditions that should be fulfilled before they can employ ML to improve business process outcomes,
- Benefits of using a normalization technique that enhances the reliability of trained neural networks in the ML approach, such as sigmoid normalization,
- Empirical evidence that supervisory learning based on gradient descent algorithm can be used as business learning strategy for process improvement,
- Verification of the proposed methods in real business environment with data originated from and fed back to actual core business process in a large enterprise, with elaborated benefits to the company.

As companies are begging to understand the benefits of ML and are trying to employ different ML solutions to improve their business processes, it is vitally important to identify the processes that can leverage such approach. The practical guidelines in this study emphasize that ML should be deployed in business processes with unclear input-output relationship using large and comprehensive dataset, which evolves constantly and cannot be translated into a programming code, etc. The unpredictable distribution of data in the business process imposes the use of a normalization technique, such as one used in neural networks (Tax et al., 2017; Wanto et al., 2017), to improve performance in the chosen ML solution. The results in this study show that supervisory learning based on gradient descent algorithm can be used as a business ML strategy. The ML solutions was used in production for ten months, including period of changes in the business process, and its average accuracy compared to the human decisions was 95.018%, which is extremely high for predictive analysis. It surpasses other techniques, such as statistical analysis in which statistically significant effect does not guarantee high predictive power (Shmueli, & Koppius, 2011), and it also proves that this ML algorithm can be used among the diverse array of machine-learning algorithms that has been developed to cover the wide variety of data and business problems (Jordan & Mitchell, 2015; Lantz, 2013; Solanki, & Dhankar, 2017). Therefore, one of the major contributions of this study is the predictive accuracy of the solution demonstrated in real-life implementation. The chosen algorithm is a statistical tool, but when properly utilized in a ML solution, it provided valuable predictive analytics in a business context.

The ML solution based on gradient descent algorithm was designed to streamline the business process and provide resilience to employees' inaccurate decisions. Hence, it can minimize employees' subjectivity, wrong decisions, and human errors. After a specific period of successful implementation, it can automate the whole process, which can complement or even substitute human labor where appropriate, to facilitate rapid developments (Brynjolfsson & McAfee, 2014; Brynjolfsson, & Mitchell, 2017; David, 2015).

Entrepreneurs, managers, and workers constantly work to reinvent the relevant processes. When faced with new technologies, they will change the production process, by design or through luck, and find more efficient ways to produce output (Brynjolfsson, & Mitchell, 2017, Manyika et al., 2017). This study helps organizations realize the potentials of ML for enterprise computing, so they can leverage ML for business process improvements, such as better performance, quality and speed, as well as overcome human limits and possibility for errors during operations. Hence, it can help companies that plan use ML to improve their business processes, since it shows that ML can augment human capabilities and even enable development of entirely new products, services, and processes. This study also opens up new avenues for research in addition to recent studies that focus on business prospects of ML.

5. 1. Limitations

The ML solution researched in this study was deployed in large enterprise company using a specific software product, which supports utilization of the decent algorithm and was deployed on a powerful computing architecture. Since ML processing consumption can be excessive, especially the calculation of the coefficient in the algorithm, it may impose certain restriction for similar practical implementation. However, parallel coefficient calculation can be replaced with one or two background executions per day without compromising performance during application usage. It may introduce slight delay in accuracy, but eventually the ML process can compensate this behavior during operations.

This study leverages a standard gradient descent algorithm, but improper selection of the input array in the ML solution, which also leads to improper coefficient calculations, may decrease its accuracy in practice. Thus, the recommendations presented in this study can be valuable to organizations in their initiatives to deploy ML for business process improvement.

References

Ahmad, M. M., & Cuenca, R. P. (2013). Critical success factors for ERP implementation in SMEs. *Robotics and Computer-Integrated Manufacturing, 29*(3), 104-111.

Allen-Zhu, Z., & Hazan, E. (2016, June). Variance reduction for faster non-convex optimization. In *International Conference on Machine Learning* (pp. 699-707).

Attoh-Okine, N. O. (2017) Machine Learning: A Basic Overview. *Big Data and Differential Privacy: Analysis Strategies for Railway Track Engineering*, 59-111.

Bottou, L. (2014). From machine learning to machine reasoning. *Machine learning, 94*(2), 133-149.

Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies.* WW Norton & Company.

Brynjolfsson, E., & Mitchell, T. (2017). What can machine learning do? Workforce implications. *Science, 358*(6370), 1530-1534.

Cabanillas, C., Di Ciccio, C., Mendling, J., & Baumgrass, A. (2014, September). Predictive task monitoring for business processes. In *International Conference on Business Process Management* (pp. 424-432). Springer International Publishing.

Chang, J. F. (2016). *Business process management systems: strategy and implementation*. CRC Press.

Chen, C. P., & Zhang, C. Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *Information Sciences*, *275*, 314-347.

Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS quarterly*, *36*(4), 1165-1188.

Gomez-Uribe, C. A., & Hunt, N. (2016). The netflix recommender system: Algorithms, business value, and innovation. *ACM Transactions on Management Information Systems (TMIS), 6*(4), 13.

Davenport, T.H. (2013). *Process Innovation: Reengineering Work through Information Technology*. Harvard Business Press, Boston, Massachusetts.

David, H. (2015). Why are there still so many jobs? The history and future of workplace automation. *Journal of Economic Perspectives, 29*(3), 3-30.

Devore, J. L. (2015). *Probability and Statistics for Engineering and the Sciences*. Cengage Learning.

Dunjko, V., Taylor, J. M., & Briegel, H. J. (2016). Enhanced learning for agents in quantum-accessible environments. *In ESANN 2016 Proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*. Bruges, Belgium.

Fumo, N., & Biswas, M. R. (2015). Regression analysis for prediction of residential energy consumption. *Renewable and Sustainable Energy Reviews, 47,* 332-343.

Garcia-Sillas, D., Gorrostieta-Hurtado, E., Soto-Vargas, E., Diaz-Delgado, G., & Rodriguez-Rivero, C. (2016, October). Learning from demonstration with Gaussian processes. In *Mechatronics, Adaptive and Intelligent Systems (MAIS), IEEE Conference on* (pp. 1-6). IEEE.

Harmon, P. (2014). *Business Process Change: A Business Process Management Guide for Managers and Process Professionals*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

Harrell, F. (2015). *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*. Springer.

Hong, S. G., Hong, S. G., Siau, K., Siau, K., Kim, J. W., & Kim, J. W. (2016). The impact of ISP, BPR, and customization on ERP performance in manufacturing SMEs of Korea. *Asia Pacific Journal of Innovation and Entrepreneurship, 10*(1), 39-54.

Hopkins, L., & Ferguson, K. E. (2014). Looking forward: The role of multiple regression in family business research. *Journal of Family Business Strategy, 5*(1), 52-62.

Jayalakshmi, T., & Santhakumaran, A. (2011). Statistical normalization and back propagation for classification. *International Journal of Computer Theory and Engineering, 3*(1), 1793-8201.

Jain, A., Koppula, H. S., Raghavan, B., Soh, S., & Saxena, A. (2015). Car that knows before you do: Anticipating maneuvers via learning temporal driving models. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 3182-3190).

Jeston, J., & Nelis, J. (2014). *Business process management*. Routledge.

Johnson, R., & Zhang, T. (2013). Accelerating stochastic gradient descent using predictive variance reduction. In *Advances in neural information processing systems* (pp. 315-323).

Jonsson, L., Borg, M., Broman, D., Sandahl, K., Eldh, S., & Runeson, P. (2016). Automated bug assignment: Ensemble-based machine learning in large scale industrial contexts. *Empirical Software Engineering, 21*(4), 1533-1578.

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science, 349*(6245), 255-260.

Jovanoski, D., Malinovski, T. & Arsenovski, S. (2017). Links between strategic goals, information technology and customer satisfaction during business process reengineering. *International Journal of Business Process Integration and Management*, 8(3), 200-213.

Kelleher, J. D., Mac Namee, B., & D'Arcy, A. (2015). *Fundamentals of machine learning for predictive data analytics: algorithms, worked examples, and case studies*. MIT Press.

Kingma, D. P., Mohamed, S., Rezende, D. J., & Welling, M. (2014). Semi-supervised learning with deep generative models. In *Advances in Neural Information Processing Systems* (pp. 3581-3589).

Lantz, B. (2013). *Machine learning with R*. Packt Publishing Ltd.

Leung, C. K. S., MacKinnon, R. K., & Wang, Y. (2014, July). A machine learning approach for stock price prediction. In *Proceedings of the 18th International Database Engineering & Applications Symposium (pp. 274-277)*. ACM.

Lodhia, Z., Rasool, A., & Hajela, G. (2017). A survey on machine learning and outlier detection techniques. *IJCSNS, 17*(5), 271.

Manyika, J., Chui, M., Miremadi, M., Bughin, J., George, K., Willmott, P., & Dewhurst, M. (2017). *A Future that Works: Automation, Employment, and Productivity*. McKinsey Global Institute.

Metzger, A., Leitner, P., Ivanović, D., Schmieders, E., Franklin, R., Carro, M., Dustdar, S., & Pohl, K. (2015). Comparing and combining predictive business process monitoring techniques. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 45*(2), 276-290.

Miljkovic, D., Aleksovski, D., Podpecan, V., Lavrac, N., Malle, B., & Holzinger, A. (2016). Machine Learning and Data Mining Methods for Managing Parkinson's Disease. In *Machine Learning for Health Informatics (pp. 209-220)*. Springer International Publishing.

Meng, X., Bradley, J., Yavuz, B., Sparks, E., Venkataraman, S., Liu, D., Freeman, J., Tsai, D.B., Amde, M., Owen, S., & Xin, D. (2016). Mllib: Machine learning in apache spark. *Journal of Machine Learning Research, 17*(34), 1-7.

Qiu, X. Y., Srinivasan, P., & Hu, Y. (2014). Supervised learning models to predict firm performance with annual reports: An empirical study. *Journal of the Association for Information Science and Technology, 65*(2), 400-413.

Ridgeway, G. (2013). The pitfalls of prediction. *NIJ Journal, 271*, 34-40.

Riggins, F. J., & Wamba, S. F. (2015, January). Research directions on the adoption, usage, and impact of the internet of things through the use of big data analytics. In *System Sciences (HICSS), 2015 48th Hawaii International Conference on* (pp. 1531-1540). IEEE.

Schulte, S., Janiesch, C., Venugopal, S., Weber, I., & Hoenisch, P. (2015). Elastic business process management: state of the art and open challenges for BPM in the cloud. *Future Generation Computer Systems, 46*, 36-50.

Schmidt, R., Möhring, M., Maier, S., Pietsch, J., & Härting, R. C. (2014, May). Big data as strategic enabler-insights from central european enterprises. In *International Conference on Business Information Systems* (pp. 50-60). Springer, Cham.

Shamir, O., & Zhang, T. (2013, February). Stochastic gradient descent for non-smooth optimization: Convergence results and optimal averaging schemes. In *International Conference on Machine Learning* (pp. 71-79).

Shmueli, G., & Koppius, O. R. (2011). Predictive analytics in information systems research. *Mis Quarterly*, 553-572.

Shmueli, G., Patel, N. R., & Bruce, P. C. (2016). *Data Mining for Business Analytics: Concepts, Techniques, and Applications with XLMiner*. John Wiley & Sons.

Solanki, K., & Dhankar, A. (2017). A review on Machine Learning Techniques. International Journal of Advanced *Research in Computer Science, 8*(3).

Tan, W., & Zhou, M. (2013). *Business and Scientific Workflows: A Web Service-Oriented Approach (Vol. 5)*. John Wiley & Sons.

Tax, N., Verenich, I., La Rosa, M., & Dumas, M. (2017). Predictive business process monitoring with LSTM neural networks. In *International Conference on Advanced Information Systems Engineering* (pp. 477-492). Springer, Cham.

Vafeiadis, T., Diamantaras, K. I., Sarigiannidis, G., & Chatzisavvas, K. C. (2015). A comparison of machine learning techniques for customer churn prediction. *Simulation Modelling Practice and Theory, 55*, 1-9.

Wagstaff, K. (2012). Machine learning that matters. In *Proceedings of the 29th International Conference on Machine Learning*, Edinburgh, Scotland, UK.

Wang, Y., Kung, L., & Byrd, T. A. (2018). Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Technological Forecasting and Social Change, 126*, 3-13.

Wanto, A., Windarto, A. P., Hartama, D., & Parlina, I. (2017). Use of Binary Sigmoid Function And Linear Identity In Artificial Neural Networks For Forecasting Population Density. *International Journal Of Information System & Technology, 1*(1), 43-54.

Yousfi, A., de Freitas, A., Dey, A. K., & Saidi, R. (2016). The use of ubiquitous computing for business process improvement. *IEEE Transactions on Services Computing, 9*(4), 621-632.

Yu, Q., Miche, Y., Séverin, E., & Lendasse, A. (2014). Bankruptcy prediction using extreme learning machine and financial expertise. *Neurocomputing, 128*, 296-302.

Zancul, E. D. S., Takey, S. M., Barquet, A. P. B., Kuwabara, L. H., Cauchick Miguel, P. A., & Rozenfeld, H. (2016). Business process support for IoT based product-service systems (PSS). *Business Process Management Journal, 22*(2), 305-323.