



Deep Learning Predictive Models for Terminal Call Rate Prediction during the Warranty Period

Aljaž Ferencek

Faculty of Organizational Sciences, University of Maribor, Slovenia

Davorin Kofjač

Faculty of Organizational Sciences, University of Maribor, Slovenia

Andrej Škraba

Faculty of Organizational Sciences, University of Maribor, Slovenia

Blaž Sašek

Faculty of Organizational Sciences, University of Maribor, Slovenia

Mirjana Kljajić Borštnar

Faculty of Organizational Sciences, University of Maribor, Slovenia

Abstract

Background: This paper addresses the problem of products' terminal call rate (TCR) prediction during the warranty period. TCR refers to the information on the amount of funds to be reserved for product repairs during the warranty period. So far, various methods have been used to address this problem, from discrete event simulation and time series, to machine learning predictive models. **Objectives:** In this paper, we address the above named problem by applying deep learning models to predict terminal call rate. **Methods/Approach:** We have developed a series of deep learning models on a data set obtained from a manufacturer of home appliances, and we have analysed their quality and performance. **Results:** Results showed that a deep neural network with 6 layers and a convolutional neural network gave the best results. **Conclusions:** This paper suggests that deep learning is an approach worth exploring further, however, with the disadvantage being that it requires large volumes of quality data.

Keywords: manufacturing, product lifecycle, management product failure, machine learning, prediction

JEL classification: C45, C53

Paper type: Case study

Received: Apr 23, 2020

Accepted: Jul 06, 2020

Citation: Ferencek, A., Kofjač, D., Škraba, A., Sašek, B., Kljajić Borštnar, M. (2020), "Deep Learning Predictive Models for Terminal Call Rate Prediction during the Warranty Period", Business Systems Research, Vol. 11 No. 2, pp. 36-50.

DOI: 10.2478/bsrj-2020-0014

Introduction

Today's business environment is highly competitive thus, businesses need to optimize their costs and improve their profit and/or market share. Warranty claim control, as a part of quality control department, is one of the most important departments since servicing warranties involves additional costs to the manufacturer. If warranty conditions are good that usually indicates higher product quality, thus affects marketing of new products (Murthy & Djameludin, 2002). Higher quality of the product is related with product reliability which is, in a more technical definition, "the probability that the product (system) will perform its intended function for a specified time period when operating under normal (or stated) environmental conditions" (Blischke & Murthy, 2000, p.18). Even though a new product is usually tested before it is released to the public that does not necessarily mean that it will not malfunction under unusual circumstances. To alleviate problems of such scenarios, information about warranties can be used in order to monitor, warn (Lawless, 1998) and even to predict future events.

For warranty claims predictions, empirical models based on past data of products with experience-based correction factors are often used (Kljajić et al., 2000; Škraba et al., 2011; Wu et al., 2011; Dai et al., 2019). Nevertheless, many research in this field focuses on short-term forecasting, since reliable long-term forecasting is usually not possible due to the dataset they operate with. This research first extends previous research by providing new insight into possible machine learning models and second, provides additional understanding and validation on how data quality can affect those models.

In this paper, we set out to develop and validate a prediction model, using deep learning models on a case of a production-oriented company in the field of home appliances. We addressed the problem of improving the product failure forecasting process within the warranty period in the selected company, which has been, up to now, done „manually“, and mainly using spreadsheets in Microsoft Excel. The company was aware that the existing forecasting process was time-consuming, non-transparent and with many potential errors due to many manual data entries. As a result, there were discrepancies in the accuracy of the failure forecasts of individual products. Consequently, this means errors are being made in the reservation of funds to correct defects in products within the warranty period and that resulted as an oversupply of financial resources that cannot be used anywhere else in the company. Our main goal in this paper was to predict call rates 12 months into the future given up to 6 months of initial data for each series of product using machine learning methods.

Related work

When addressing problems that aim to forecast the future, we are talking about many different approaches such as time series, statistical models, simulation models and machine learning, more specifically predictive analytics. When using time series forecasting, we are collecting and analysing data from previous observations with the purpose of developing a model describing the underlying relationship between selected variables (Zhang, 2003). This kind of forecasting is convenient when there is little knowledge available on the underlying data. In the past researches, much effort has been devoted to the improvement of time series forecasting models. Kuremoto et al. (2014) for example proposes an improvement of multi-layer perceptron (MLP) and other artificial neural networks (ANNs) that had been used in time series since 1980s and replaces them using Hinton and Salakhutdinov's deep belief nets (DBN) (Hinton,

2006). Another example of upgrading time series to predict future is with the usage of statistical techniques such as Autoregressive Integrated Moving Average (ARIMA) – (Conejo et al., 2005; De Gooijer & Hyndman, 2006; Hyndman & Khandakar, 2008). To achieve better results with ARIMA, many authors, used or compared machine learning methods to improve predictions. Zhang (2003) described the use of hybrid ARIMA and neural network model that could be used for forecasting the next 12 months, whereas Valipour et al. (2013) compared ARIMA model with the autoregressive artificial neural networks (ANN) with forecasting abilities of 5 years in the future. Simulation models as another method for prediction of future events are focusing on the exploration and optimization of complex systems (Nyhuis et al., 2005). They represent a fairly complex system in a model that can be later applied to reality and often involve different kinds of forecasting, such as probabilistic and mechanistic modelling (Gujral et al., 2019).

Considering large amounts of data, many attributes and incomprehensible parameters, machine learning can be used to extract useful information from the dataset. Machine learning as such has played an important role in processing data, identification of data samples and in making decisions for the last couple of decades (Ge et al., 2017). Predictive analytics more specifically, has been used in many business applications such as customer relationship management, predicting crime-fighting in law enforcement, predicting warranty problems for automobile manufacturers, predicting change in stock price, etc. (Chan, 2007; Maričić et al., 2019). In the past few years, different methods for validating the reliability prediction model of electric products were used such as time series models and the ARIMA model, fuzzy methods, neural networks and expert systems (Li et al., 2010).

The problem addressed in this paper, prediction warranty call rates using some initial data for each production series, was previously addressed in (Kofjač et al., 2014). In Kofjač et al. (2014) authors used a combined model of market absorption and failure process, where forecasting activity for current products was done by using warranty claims for the first few months of service with the purpose of anticipating the final numbers of warranty returns while the warranty was still active. In general, there were two types of data. First type was focusing on products from sales to failure, and the other type was from production to failure, with the consideration of the market absorption time. For claims prediction, the authors proposed the Markov Modulated Fluid Model (Zhang, 1993). The model was verified and validated on the actual data and authors also proposed an application of grid computing due to costs of prediction in means of computational power (Kofjač et al., 2014).

In their later research Kofjač et al. (2016) proposed the use of machine learning methods for forecasting of terminal call rate (TCR), which is an expected maximum value of failed products within the warranty period (Kofjač et al., 2014). They investigated the impact of cumulative density function with ML models on TCR prediction accuracy which were modelled with exponential and logistic models. Their parameters were estimated with regression trees, neural networks and ensembles of regression trees. To evaluate goodness of fit of cumulative density functions, a standard error of the estimate (SEE) measure was used. The best results were achieved by ensembles of regression trees. Because the stance of the study was focused on fundamental research in the prediction of TCR with ML models, authors proposed future studies to improve the prediction accuracy, addressing the impact of other attributes, such as mean time to failure (MTTF) and the optimization of ML models, such as number of instances in leaves for regression trees.

Finally, the last research on the given problem was a project PKP (Po kreativni poti do znanja) funded under Public Scholarship, Development, Disability and Maintenance Fund of the Republic of Slovenia. The project addressed the falling

prices of processing and storage capabilities and the development of new models and techniques, where one could build models as needed using only the most relevant and recent data. The project aimed to develop the prototype application that would provide employees from the quality management department with call rate predictions for the next year to support planning of the expenses. In the core of this prototype was a stacked model that consisted of a linear regression model a simple tree and a random forest regressor. As new data was extracted from the transactional database it was transformed and stored as a matrix of cumulative call rates for each month following a production of an individual series. The model would then be applied to the user's demand and it would locate the most appropriate date range and learn all three models on the provided data, followed by a modulation using the implementation of techniques from the Forecasting terminal call rate with machine learning methods (Kofjač et al., 2016). This process was successfully implemented in the Guided Machine Learning for Business Users (Bourcevet et al., 2019). Models from PKP project would then vote on future predictions for series that are still in the warranty period. Their contribution would be weighted based on how well they performed on test inputs while learning. If prediction were triggered in the following months, when new data was available, the model would incorporate this data and re-learn itself before evaluating predictions for the new time. Given more data with each subsequent execution, the prototype should be less volatile and prone to errors (Arnerić et al., 2018), but this was never empirically tested on unseen data since the project ended before the prototype could be fully implemented and integrated with the transactional database.

This research contributes new findings based on related work, of how successful predictions in the field of warranty claims can be with the use of ML methods. Since both companies and customers can benefit from successful terminal call rate prediction, this research also contributes to practical aspect in terms of reserving more investment funds for company and getting more reliable products or faster and more efficient service procedures for customers.

Methodology

The proposed methodological approach is rooted in Design Science Research (DSR) (Hevner et al., 2004). The core of DSR represents the development cycle, in which an IT artefact is developed, in our case the ML predictive model. The development cycle is nested between the relevancy (real-life business problem), and rigor cycle, assuring scientific rigor by building on theory and prior knowledge. The development cycle follows the Cross-Industry Standard Process for data mining, CRISP-DM. The process or methodology of CRISP-DM is described in the six major steps: Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation and Deployment (Wirth & Hipp, 2000).

In the first two steps, business understanding and data understanding, we have studied previous research results of the warranty claim predictions (Kofjač et al., 2014, 2016). Additionally, we have made several company visits, where we acquired new data and gained a deeper understanding of the business problem. The basic idea of gaining additional insight if another dimension to the data is added (Sašek, 2017).

The third step of the process refers to data preparation, which is usually the most time-consuming. Besides that, in this phase we gain new knowledge of the problem and data, so the first three steps are usually repeated in cycles until the desired result is achieved, that is the final data set.

The data was taken from the transactional database of all the service interventions. We received data for four specific markets - Nordic region, Germany, Russia and

Serbia. There were specifics for each market, which we were made aware of during our visits to the factory (for instance, it has been noticed that in the Nordic countries there is a greater time between the appliance failure and the service procedure than in other markets). Data was of varying quality depending on the market and the time in which it was produced. While newer data was more consistent, there was less of a systematic approach to gathering data in the past, thus limiting the usefulness of the data. Still, almost 50 % of all service inquiries did not have the reason for failure listed and that was the important feature in our proposed models. Data preparation consisted of checking for inconsistent and missing data, which was then either cleaned or removed. The aim was to gain an insight into the relative number of failed appliances (within the first 6 months) and the reason for failure (387 unique reasons for failure). Next, we split the data based on the product hierarchy (product groups) and the country of manufacturing. The product hierarchy refers to the aggregation level of products, and allows us to analyse the TCR on different levels of detail (from groups of products to the specific product). It means that the product hierarchy PH5 has five identification numbers that go five levels in depth. If we take an example of kitchen appliances, we could say that PH1 represents kitchen appliances, PH2 cooling appliances, PH3 refrigerators and so on, until we can classify one product based on its colour. To provide enough training samples for neural networks, which require large quantities of data, we prepared separate matrixes for each market and we fragmented products based on a few different levels of product hierarchy provided by the data owner. This gave us 308 fragments on 4 markets for which we had data available for multiple series produced in the past. In total that gave us 11852 data points. We constructed upper triangular matrix, which is a square matrix where all the entries below the main diagonal are zero that contained TCR calculations. After that, we were able to proceed towards model development.

Our final model was a simple convolutional neural network with three convolutional layers and a fully connected layer. We were particularly interested in the performance of convolutional neural networks, given that we have a 2-dimensional dataset that has the same representation as a single-channel (grayscale) image. A Convolutional Neural Network (CNN) is a Deep Learning algorithm where learnable weights and biases are applied on the raw input image (Ji et al., 2013). CNN was recognized as a useful tool when, for example, enhancing reliability of software by predicting potential code defects in its implementation (Li et al., 2017), extracting useful clinical representations from longitudinal electronic health record data by predicting risks with CNN (Che et al., 2017) or for predicting instability mode in power system control (Shi et al., 2020).

We systematically approached hyper-parameter optimisation by running multiple neural networks with different configurations. We ran multiple permutations of the following parameters:

- Depth and type of neural networks: up to 6 hidden layers with varying number of neurons in each layer, convolutional neural network;
- Activation functions: ReLU, ELU, softmax, tanh and sigmoid function;
- Regularization: L2 normalisation and dropout with 50 % dropout probability;
- Learning rate: starting with 0.1 and lowering it by order of magnitude with different configurations.

Using results from previous runs, we excluded some hyper-parameter configurations from the testing at researcher discretion in order to avoid combinatorial explosion.

Results

Our goal was to predict call rates 12 months into the future given up to 6 months of initial data for each series. First, we briefly present the business problem and data, which has already been discussed in (Kljajić et al., 2000; Kofjač et al., 2014, 2016; Škraba et al., 2011).

We received data in 56 different spreadsheets that varied in amount, features and purpose. Spreadsheets were classified in advance by the company experts for 4 specific markets. Data was gathered from service interventions, production and sales data. The complete dataset contained different features that were shattered across spreadsheets. List of features, its explanation and data type can be found in Table 1.

Table 1

Dataset features, explanation of features and their data type

| Feature name | Explanation | Data type |
|----------------------------|---|-------------|
| No_OSS | Service intervention number | Numeric |
| Product_code | Unique identifier for a specific product | Numeric |
| Product_description | Description of a product | Text |
| Serial_number | Serial number of a product | Numeric |
| Date_manuf | Date of manufacturing | Date |
| Date_failure | Date of failure | Date |
| Date_purch | Date of purchase | Date |
| Date_repair | Date of repair | Date |
| Error_location | Descriptive location of the error | Text |
| Error_desc | Description of the error | Text |
| YearsOf_use | Number of years in use before failure | Numeric |
| Tech_nontech | Technical or nontechnical error identifier | Categorical |
| Brand | Brand of a product | Text |
| Work_order | Number of work order | Numeric |
| Error_type | Type of error | Text |
| Product_generation | Product generation (descriptive) | Text |
| Main_error | Description of a main error | Text |
| Graph_symb | Manufactured material for insertion part 1 | Text |
| Graph_symb_2 | Manufactured material for insertion part 2 | Text |
| Inserted_mat_ID | Id of inserted material | Numeric |
| Inserted_mat_desc | Description of inserted material | Text |
| Buyer | Buyer of the product | Numeric |
| Seller | Seller of the product | Text |
| If_subgroup | Subgroup identifier | Text |
| Model_tag | Tag of a model | Text |
| Ph | 8-digit unique designation based on product hierarchy | Numeric |
| Product_name | Name of the product | Text |
| Service | Name of repairer | Text |
| Type_tag | Tag of model type | Text |
| Material_kind | Description of material kind | Text |
| Material_type | Description of material type | Text |
| No_interventions | Number of interventions on a product | Numeric |
| Market | Market shortcode | Text |

Source: Author's work

Since data pre-processing pipelines are used to help automate machine learning workflows (Li, 2019), we had to rewrite them for the purpose of this new analysis. In previous research a dimension in a matrix first consisted of all products produced in a certain month and was now changed to a reason for failure. The result of this data preparation phase were matrixes, as presented in Figure 1, where the first dimension represents the error location and the second dimension represents the time from purchase to failure in months. We made a matrix for every product category in three different product hierarchies (PH3 – PH5) for every market. These matrixes were used as an input for our predictive models. Since the data in matrices on PH5 have duplicated values as ones on PH3 we got more data points we could use for our models. We used this solution to multiply training/testing instances because of the missing or inconsistent data. An example of the matrix with random data is represented in Figure 1.

Figure 1
New matrix as a result of pre-processing pipeline

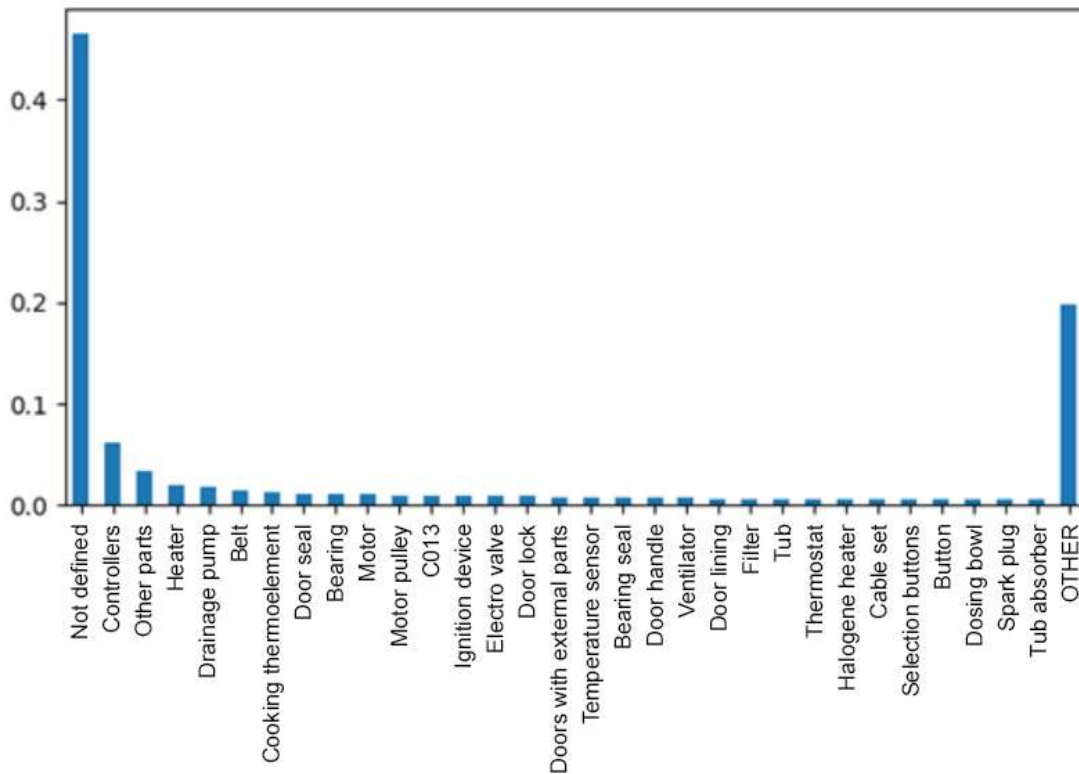
| Error location | Months from sale to failure with according values | | | | | | | | | | | |
|----------------|---|----|----|----|----|----|----|----|----|----|----|----|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| Error 1 | 8 | 8 | 16 | 5 | 4 | 16 | 12 | 4 | 7 | 10 | 9 | 5 |
| Error 2 | 10 | 14 | 5 | 1 | 15 | 6 | 18 | 18 | 6 | 15 | 15 | 19 |
| Error 3 | 18 | 5 | 16 | 6 | 18 | 9 | 3 | 20 | 17 | 16 | 6 | 9 |
| Error 4 | 12 | 18 | 20 | 16 | 3 | 14 | 4 | 18 | 19 | 16 | 1 | 7 |
| Error 5 | 17 | 13 | 11 | 18 | 9 | 17 | 9 | 10 | 4 | 14 | 4 | 14 |
| Error 6 | 4 | 8 | 16 | 18 | 10 | 7 | 7 | 16 | 9 | 2 | 8 | 9 |
| Error 7 | 19 | 10 | 12 | 16 | 2 | 17 | 4 | 14 | 19 | 2 | 17 | 3 |
| Error 8 | 17 | 18 | 20 | 10 | 18 | 9 | 7 | 16 | 2 | 5 | 13 | 4 |
| Error 9 | 9 | 1 | 20 | 1 | 11 | 19 | 8 | 2 | 7 | 15 | 13 | 4 |
| Error 10 | 4 | 7 | 9 | 3 | 20 | 6 | 20 | 2 | 5 | 12 | 14 | 16 |
| Error 11 | 20 | 6 | 7 | 3 | 8 | 1 | 15 | 4 | 18 | 2 | 12 | 14 |
| Error 12 | 10 | 7 | 16 | 12 | 11 | 16 | 8 | 13 | 6 | 8 | 18 | 10 |
| Error 13 | 7 | 4 | 4 | 9 | 16 | 7 | 13 | 13 | 7 | 17 | 19 | 3 |
| Error 14 | 5 | 3 | 15 | 2 | 6 | 11 | 8 | 3 | 10 | 16 | 14 | 2 |
| Error 15 | 8 | 18 | 19 | 16 | 11 | 13 | 2 | 14 | 12 | 14 | 8 | 9 |
| Error 16 | 13 | 14 | 5 | 12 | 15 | 16 | 19 | 20 | 18 | 5 | 5 | 5 |
| Error 17 | 12 | 20 | 11 | 16 | 4 | 12 | 8 | 20 | 18 | 9 | 1 | 19 |
| Error 18 | 16 | 1 | 14 | 8 | 19 | 14 | 8 | 4 | 17 | 12 | 5 | 4 |
| Error 19 | 10 | 13 | 19 | 18 | 9 | 11 | 14 | 12 | 5 | 2 | 5 | 12 |
| Error 20 | 6 | 5 | 20 | 20 | 1 | 19 | 1 | 3 | 12 | 17 | 11 | 7 |
| Error 21 | 1 | 19 | 8 | 2 | 13 | 20 | 20 | 9 | 13 | 11 | 7 | 17 |
| Error 22 | 1 | 17 | 12 | 20 | 17 | 6 | 18 | 2 | 4 | 7 | 17 | 2 |
| Error 23 | 14 | 19 | 11 | 13 | 11 | 4 | 9 | 20 | 6 | 16 | 13 | 20 |

Source: Author's work

Upon inspection of this previously unused dimension of the data, we learned that there are a lot of missing values. More than 46 % of the service entries had no stated reason for the failure. This was varying from market to market, for example, the Serbian market being the worst and German being the best, which suggested that we could possibly extract useful data from certain markets while the model would pay less attention to the markets with a lot of missing data, presented in Figure 2.

Figure 2

Frequency of error location of products (% used on y axis where OTHER is meant as a joined group with less than 0.5% of the data)



Source: Adapted from Sašek (2017)

Before using ANN, Kofjač et al. (2014) applied Markov Modulated Fluid model (MMFM) that achieved a prediction accuracy of 0.79 at best. Kofjač et al. (2014) also proposed the use of other methods such as grid computing due to the computationally costly prediction of MMFM. In their later research, Kofjač et al. (2016) used ML methods (regression tree, feedforward neural network and ensemble of regression trees) to estimate a real-valued variable TCR where input pattern of cumulative failure rates was given. The highest prediction accuracy was given when they used an ensemble of 100 regression trees, with the average mean absolute percentage error (MAPE) of 2.31%. Ensemble learning in that case provided the best results because of the diversity of features and that, along with non-linear characteristics of cumulative density function, which is used to predict TCR, was the reason for using ANN. In addition, since traditional forecasting models have limitations with considering non-linear relationships, ANNs on the other hand, can provide a robust alternative to model and extract unseen features and relationships. In addition, ANN's do not impose any restriction on input and residual distributions.

We generated three different models - a perceptron (which was scrapped immediately since it did not converge), a few variations of deep neural networks and a convolutional neural net. As expected, perceptron did not converge after 100 epoch, since a single layer network is suited for solving simple linear problems, and not, as in our case, more complex non-linear problems.

Further, we optimized the neural network by adding hidden layers, adjusting the number of neurons at each level, applying different activation functions, and regularizations. We used L1 norms+ (absolute difference) to measure accuracy since

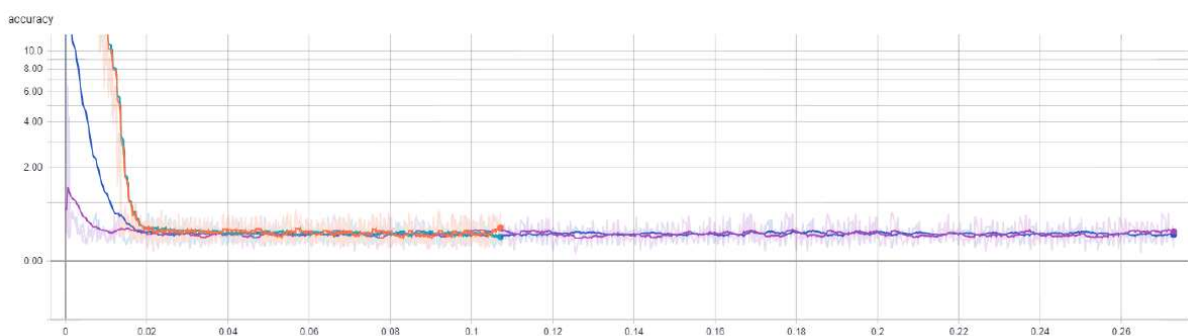
it is most intuitively interpreted by human experts (or users of predictions), in the QA department.

We tested the variations from 1 to 6 hidden layers with variations of number of neurons at each level from 2048, 1024, 512, 256, 128, and 64. For activation function, the sigmoid function was used. The results of learning did not converge after 100 epochs. In fact, the results of various deep neural networks, applying various combinations of layers and neurons, did not defer from the perceptron. Further analysis into the modelled data have shown that there are missing data because of the delay of the market absorption (i.e. no sale is recorded for 6 months, thus the matrix for months 1 to 6 are empty). To address this problem, we have cleaned out the empty data matrices and trained the multilayer NN again. Comparison of convergence of the single layer (orange coloured curve) and the multilayer layer NN (blue and purple coloured curves) is shown in Figure 3. It can be observed that the convergence is faster with the multilayer NN, but needs more time to reach the stable value (the multiple NN stabilizes at 900 seconds of training, while the single layer at 440 seconds).

Our final model was a simple convolutional neural network with three convolutional layers and a fully connected layer. It also incorporated regularization techniques from previous examples. We generated three different models - a single layer perceptron (which was scrapped immediately since it did not converge), a few variations of deep neural networks and a convolutional net. We used L1 norms+ (absolute difference) to measure accuracy since it is most intuitively interpreted by humans that would look at the results (QA department in the selected company). With starting learning speed of 0,101 and prediction accuracy from 0.3 to 0.4, depending on the series, there was no significant difference between a deep neural network and convolutional networks in terms of accuracy but the convolutional network did converge faster, which could mean shorter learning times with less computing resources as can be observed in Figure 3. It can be observed that the convergence is faster with the multilayer NN, but needs more time to reach the stable value (the multiple NN stabilizes at 900 seconds of training, while the single layer at 440 seconds).

Figure 3

Convergence of learning of single and multilayer NN (6 hidden layers)

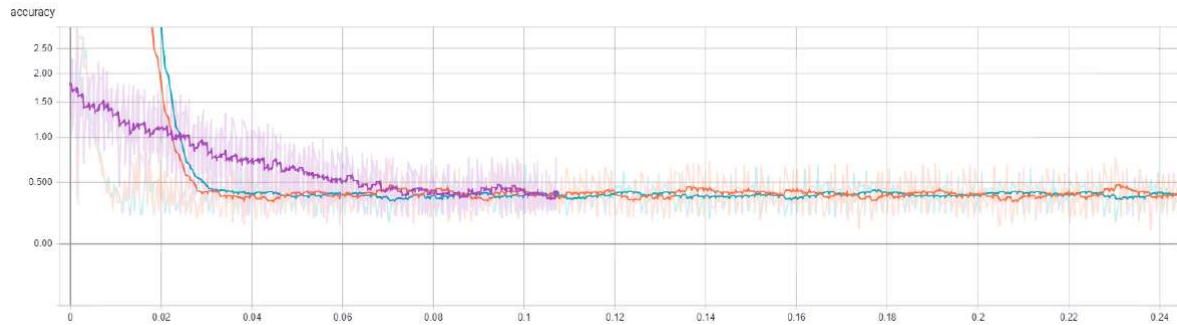


Source: Adapted from Sašek (2017)

Further, we tested how various activation functions contribute to performance (convergence time and accuracy). We tested Sigmoid (shown in Figure 3), ReLu, ELU, and tanh. The prediction accuracy had improved slightly with using ELU function shown in Figure 4.

Figure 4

Convergence of learning with the use of ELU function. Comparing simple (orange) and multilayer NN with 6 hidden layers (purple)



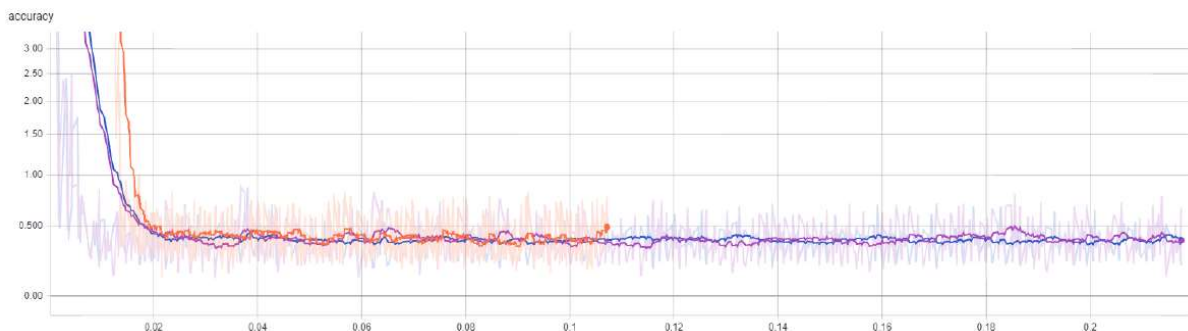
Source: Adapted from Sašek (2017)

Figure 5 compares a simple NN, represented with orange, with a multilayer NN, in purple, where results indicate that the error remains the same as with simpler model when we used Tanh activation function.

Regularization is normally used to address the problem of overfitting to training data. In our case overfitting was not the problem, but we still wanted to test how it will affect the performance of the model, since the data is asymmetrical. We applied two methods: dropout and L2 regularization. Dropout addresses the problem of slow speed of large networks, which makes dealing with overfitting during training of NN difficult (Srivastava et al., 2014). It randomly drops units from the NN during training, thus preventing overfitting. L2 regularization on the other hand updates the general cost function by adding regularization term where the weight values are pulled towards zero, therefore a smaller weight matrices lead to simpler models which are easier to train (Phaisangittisagul, 2016). None of the regularization techniques contributed to performance quality, the prediction accuracy was even decreased.

Figure 5

Convergence of learning with the use of tanh function. Comparing simple (orange) and multilayer NN with 6 hidden layers (purple)

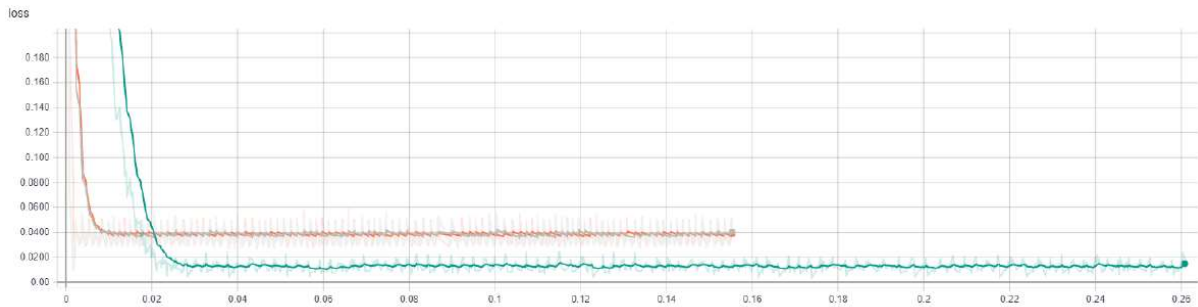


Source: Adapted from Sašek (2017)

Figure 6 shows convergence of learning of the convolutional neural network (in red) and the fully connected neural network (in green) with 8 layers (6 hidden layers, input and output layer) with learning rate 0.0101. Number of neurons in hidden layers are presented in Table 2.

Figure 6

Convergence of learning of the convolutional neural network (in red) and fully connected neural network (in green)



Source: Adapted from Sašek (2017)

Table 2

Number of neurons in hidden layers in fully connected neural network

| Hidden layer | Number of neurons |
|----------------|-------------------|
| Hidden layer 1 | 2048 |
| Hidden layer 2 | 1024 |
| Hidden layer 3 | 512 |
| Hidden layer 4 | 256 |
| Hidden layer 5 | 128 |
| Hidden layer 6 | 64 |

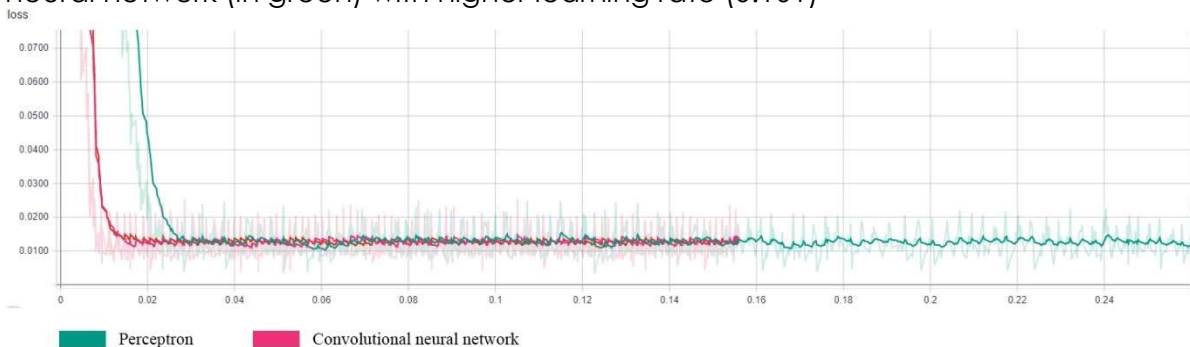
Source: Author's work

Figure 7 represents convergence of learning of convolutional neural network (in red) and the fully connected neural network (in green) with 8 layers (6 hidden layers, input and output layer). Learning rate presented on the Figure 7 was changed to 0,101.

With the starting learning rate of 0.101 and MAPE from 0.3 to 0.4, depending on the series, there was no significant difference between a deep neural network and a convolutional network in terms of accuracy, but the convolutional network did converge faster, which could mean shorter learning times with less computing resources.

Figure 7

Convergence of learning of convolutional neural network (in red) and fully connected neural network (in green) with higher learning rate (0.101)



Source: Adapted from Sašek (2017)

During this stage, we managed to reduce the window of prediction from 36 or 60 to 12 months while maintaining to support the business needs of the company. This new window of prediction enables company to gather less data for the same prediction power which means that they can predict future events three times faster than before.

Conclusion

The problem of call terminal rate prediction for home appliances manufacturing was addressed. In our previous studies, we have tried out different approaches (hybrid models based on discrete and continuous simulation, time series) with little success. In the light of the popularity of machine learning predictive models, new insights and data, we set out to develop the predictive model using deep learning. We employed the CRISP-DM process, learning about business requirements, understanding the data and putting a lot of work into cleaning the data and transforming it into the right format for consumption by neural nets. After that, we developed a few predictive models based on different neural network architectures. The main limitation of our research was the amount and quality of data, which proved to be of issue in training the models.

The results showed that the best two models, deep neural network with 6 layers and a convolutional neural network differed in 1% when predicting TCR at 12 months. The lowest error of prediction was in the range between 30 and 40%, thus being too high and not better than the methods company is currently using. Based on these results we could not conclude that deep neural networks perform better or worse than conventional ML methods for prediction of call rates in home appliances. What we could conclude is that the main factor for performance of our models was inconsistency and quality of our data.

The main contribution of the paper is in the application of deep learning models on a real-life industry case, displaying how standardized, complete and comprehensive data is of great importance in leveraging any kind of predictive analytics. Previous research provided beneficial results for the company based on literature review, nevertheless, there wasn't much research done on using deep learning for the same business problem. Since we had the opportunity to gather more data than previous researchers (Kofjač et al., 2014, 2016), therefore we could extend previous research by providing new models and findings on addressing terminal call rate prediction.

Our study as well as several others (Wu et al., 2011; Ge et al., 2017; 5. Bourcevet et al., 2019; Dai et al., 2019) corroborate the importance of data quality and the data preparation process. Data quality and quantity was the biggest limitation of our work, since deep learning method need large datasets to train the models, and the models are only as good as the input data is. In the future, more effort should be put to quality data gathering, for researchers as well as businesses. Only with quality data we will be able to examine the effect of neural networks and other machine learning methods (Deep belief networks, LSTM networks, recursive networks, etc.). Furthermore, as suggested by (Bohanec et al., 2017a, 2017b), explainability and transparency of predictive models can contribute to better user involvement, which is important for quality data gathering. Not only the predictive power of the models, but also understanding the predictions, especially which attributes or combination of attributes can contribute to the product failure, is important.

References

1. Arnerić, J., Poklepović, T., Wen Teai, J. (2018), "Neural Network Approach in Forecasting Realized Variance Using High-Frequency Data", *Business Systems Research*, Vol. 9, No. 2, pp. 18-34.
2. Blischke, R., Murthy, P. (2000), *Reliability: Modeling, Prediction, and Optimization*, Wiley, New York.
3. Bohanec, M., Kljajić Borštnar, M., Robnik-Šikonja, M. (2017a), "Explaining machine learning predictions", *Expert Systems with Applications*, Vol. 71, pp. 416-428.
4. Bohanec, M., Robnik-Šikonja, M., Kljajić Borštnar, M. (2017b), "Decision-making framework with double-loop learning through interpretable black-box machine learning models", *Industrial Management & Data Systems*, Vol. 117, No. 7, pp. 1389-1406.
5. Bourcevet, A., Piller, G., Scholz, M., Wiesemann, J. (2019), "Guided machine learning for business users" in Pucihar, A. (Ed.), *32nd Bled eConference*, Vol. 1, pp. 257-270.
6. Chan, J. (2007), "A predictive analytic model for value chain management", *Journal of International Technology and Information Management*, Vol. 16, No. 1, pp. 31-42.
7. Che, Z., Cheng, Y., Sun, Z., Liu, Y. (2017), "Exploiting convolutional neural network for risk prediction with medical feature embedding", in *NIPS 2016 Workshop on Machine Learning for Health*, available at: <https://arxiv.org/abs/1701.07474> (April 23, 2020)
8. Conejo, A. J., Plazas, M. A., Espinola, R., Molina, A. B. (2005), "Day-ahead electricity price forecasting using the wavelet transform and ARIMA Models", *IEEE Transactions on Power Systems*, Vol. 20, No. 2, pp. 1035-1042.
9. Dai, A., Zhang, Z., Hou, P., Yue, J., He, S. & He, Z. (2019). "Warranty Claims Forecasting for New Products Sold with a Two-Dimensional Warranty", *Journal of Systems Science and Systems Engineering*, Vol. 28, No. 2, pp. 715-730.
10. De Gooijer, J. G., Hyndman, R. J. (2006), "25 years of time series forecasting", *International Journal of Forecasting*, Vol. 22, No. 3, pp. 443-473.
11. Ge, Z., Song, Z., Ding, S. X., Huang, B. (2017), "Data mining and analytics in the process industry: the role of machine learning", *IEEE Access*, Vol. 5, pp. 20590-20616.
12. Gujral, B., Kothaus, M. C., Lakshmiathy, U., McNicoll, A., Cardona-Torres, R., Sandhu, R., Goedecke, M., Giselbrecht, K. (2019), "The role of computer simulation modeling in product/process optimization", *Monatshefte für Chemie - Chemical Monthly*, Vol. 150, pp. 933-940.
13. Hevner, A. R., March, S. T., Park, J., Ram, S. (2004), "Design science in information systems research", *MIS Quarterly*, Vol. 28, No. 1, pp. 75-105.
14. Hinton, G. E. (2006), "Reducing the dimensionality of data with neural networks", *Science*, Vol. 313, No. 5786, pp. 504-507.
15. Hyndman, R. J., Khandakar, Y. (2008), "Automatic time series forecasting: the forecast package for R", *Journal of Statistical Software*, Vol. 27, No. 3, pp. 1-22.
16. Ji, S., Xu, W., Yang, M., Yu, K. (2013), "3D convolutional neural networks for human action recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 35, No. 1, pp. 221-231.
17. Kljajić M., Bernik I., Škraba A. (2000), "Simulation approach to decision assessment in enterprises", *Simulation*, Vol. 75, No. 4, pp. 199-210.
18. Kofjač, D., Škraba, A., Brglez, A. (2014), "Warranty claims prediction with a combined model of market absorption and failure process", *International Journal of Computing Anticipatory Systems*, Vol. 30, pp. 81-91.
19. Kofjač, D., Škraba, A., Mujanović, A. (2016), "Forecasting terminal call rate with machine learning methods", in Pucihar, A. (Ed.), *29th Bled eConference*, Vol. 1, pp. 503-514.
20. Kuremoto, T., Kimura, S., Kobayashi, K., Obayashi, M. (2014), "Time series forecasting using a deep belief network with restricted Boltzmann machines", *Neurocomputing*, Vol. 137, pp. 47-56.
21. Lawless, F. (1998), "Statistical analysis of product warranty data", *International Statistical Reviews*, Vol. 66, pp. 41-60.
22. Li, G.-D., Masuda, S., Yamaguchi, D., Nagai, M. (2010), "A new reliability prediction model in manufacturing systems", *IEEE Transactions on Reliability*, Vol. 59, No. 1, pp. 170-177.

23. Li, J., He, P., Zhu, J., Lyu, M. R. (2017), "Software defect prediction via convolutional neural network", in 2017 IEEE International Conference on Software Quality, Reliability and Security (QRS), 25-29 July, IEEE, Prague, pp. 318-328.
24. Maričić, M., Bulajić, M., Radojičić, Z., Jeremić, V. (2019), "Shedding Light on the Doing Business Index: a Machine Learning Approach", *Business Systems Research*, Vol. 10, No. 2, pp. 73-84.
25. Murthy, P., Djameludin, I. (2002), "New product warranty: a literature review", *International Journal of Production Economics*, Vol. 79, pp. 231-260.
26. Nyhuis, P., von Cieminski, G., Fischer, A., Feldmann, K. (2005), "Applying simulation and analytical models for logistic performance prediction", *CIRP Annals*, Vol. 54, No. 1, pp. 417-422.
27. Phaisangittisagul, E. (2016), "An analysis of the regularization between L2 and dropout in single hidden layer neural network", in Al-Dabass, D., Achalakul, T., Prom-On, S., Sarochawikasit, R. (Eds.), 7th International Conference on Intelligent Systems, Modelling and Simulation (ISMS 2016), 25-27 January, IEEE, Bangkok, pp. 174-179.
28. Sašek, B. (2017), "Product failure prediction with deep learning methods", available at: <https://dk.um.si/lzpisGradiva.php?lang=slv&id=67936> (access date)
29. Shi, Z., Yao, W., Zeng, L., Wen, J., Fang, J., Ai, X., Wen, J. (2020), "Convolutional neural network-based power system transient stability assessment and instability mode prediction", *Applied Energy*, Vol. 263, pp. 1-13.
30. Škraba, A., Kljajić, M., Papler, P., Kofjač, D., Obed, M. (2011), "Determination of recruitment and transition strategies", *Kybernetes*, Vol. 40, No. 9/10, pp. 1503-1522.
31. Srivastava, N., Hinton, G.E., Krizhevsky, A., Sutskever, I., Salakhutdinov, R. (2014), "Dropout: a simple way to prevent neural networks from overfitting", *Journal of Machine Learning Research*, Vol. 15, pp. 1929-1958.
32. Valipour, M., Banihabib, M. E., Behbahani, S. M. R. (2013), "Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir", *Journal of Hydrology*, Vol. 476, pp. 433-441.
33. Wirth, J., Hipp, J. (2000), "CRISP-DM: Towards a standard process model for data mining", in 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining, 11-13 April, Practical Application Company, Manchester, pp. 29-40.
34. Wu, S., & Akbarov, A. (2011). "Support vector regression for warranty claim forecasting", *European Journal of Operational Research*, Vol. 213, No. 1, pp. 196-204.
35. Zhang, G. P. (2003), "Time series forecasting using a hybrid ARIMA and neural network model", *Neurocomputing*, Vol. 50, pp. 159-175.
36. Zhang, J. (1993), "Performance study of Markov modulated fluid flow models with priority traffic", in IEEE INFOCOM '93 The Conference on Computer Communications, 28 March-1 April, IEEE, San Francisco, Vol. 1, pp. 10-17.

About the authors

Aljaž Ferencek is a doctoral student at the Faculty of Organizational Sciences at the University of Maribor. He received his master's degree at the same faculty from information systems studies. His research interests include data science, prediction analytics and open government data. He has already published some scientific articles on this subject. The author can be contacted at aljaz.ferencek@student.um.si

Davorin Kofjač obtained his PhD from the University of Maribor in the field of information systems management. He is a researcher and an associate professor at the University of Maribor, Faculty of Organizational Sciences, in the Cybernetics and Decision Support Systems Laboratory. His main research interests are modeling and simulation, decision-support systems, operational research, and artificial intelligence. He has been involved in many EU, NATO, bilateral, and national projects and is the author of more than 120 publications in international journals, monographs, and conferences. He is a member of ACM, INFORMS, and SLOSIM. The author can be contacted at davorin.kofjac@um.si

Andrej Škraba obtained his PhD in the field of organizational sciences–informatics from the University of Maribor. He works as a full professor and a researcher in the Cybernetics & Decision Support Systems Laboratory at the Faculty of Organizational Sciences, University of Maribor. His research interests cover systems theory, modeling and simulation, cyber-physical systems, and decision processes. His work has been published in the following peer-reviewed journals: Simulation, System Dynamics Review, Journal of Mechanical Engineering, Computers and Electronics in Agriculture, Kybernetes, Interfaces, and Group Decision and Negotiation. He is a member of System Dynamics Society and SLOSIM. The author can be contacted at andrej.skraba@um.si

Blaž Sašek is a master's student at University of Maribor, Faculty of Organizational Sciences, where he studies information systems. He is most interested in data science, focusing mainly on machine and deep learning. The author can be contacted at blaz.sasek@student.um.si

Mirjana Kljajić Borštnar received her Ph.D. in Management Information Systems from the University of Maribor. She works as an Associate Professor at the Faculty of Organizational Sciences, University of Maribor and is a member of Laboratory for Decision Processes and Knowledge-Based Systems. She has been involved in many EU and national projects. Her research work covers data-driven decision support systems, data mining, multi-criteria decision-making, and organizational learning. She co-authored several scientific articles published in recognized international journals, including Expert Systems with Application, PLOS ONE, and Industrial Management & Data Systems. She is a editorial board member of Uporabna informatika journal, program committee member of Bled eConference, DataScience conference, International Symposium on Operations Research in Slovenia, and research representative in national Strategic Research & Innovation Partnership Smart Cities and Communities for Artificial Intelligence. The author can be contacted at mirjana.kljajic@um.si