Demand forecasting using artificial neuronal networks and time series: application to a French furniture manufacturer case study
Julie Bibaud-Alves, Philippe Thomas, Hind Bril

To cite this version:
Julie Bibaud-Alves, Philippe Thomas, Hind Bril. Demand forecasting using artificial neuronal networks and time series: application to a French furniture manufacturer case study. 11th International Joint Conference on Computational Intelligence, IJCCI’19 (11th International Conference on Neural Computation Theory and Applications, NCTA), Sep 2019, Vienne, Austria. hal-02304581

HAL Id: hal-02304581
https://hal.science/hal-02304581
Submitted on 3 Oct 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Demand forecasting using artificial neuronal networks and time series: application to a French furniture manufacturer case study

Julie Bibaud-Alves\textsuperscript{1,2,3}, Philippe Thomas\textsuperscript{1,2} and Hind Bril El Haouzi\textsuperscript{1,2}

\textsuperscript{1}Université de Lorraine, CRAN, UMR 7039, Campus Sciences, BP 70239, 54506 Vandœuvre-lès-Nancy cedex, France\
{jbibaud-alves, philippe.thomas, hind.el-haouzi}@univ-lorraine.fr

\textsuperscript{2} CNRS, CRAN, UMR7039, France

\textsuperscript{3}Parisot Meubles, 15 avenue Jacques Parisot, 70800 Saint-Loup-sur-Semouse\
{jbibaud-alves}@parisot.com

Keywords: Sales Forecasting, Neural Networks, Multilayer Perceptron, Seasonality, Overfitting, time series

Abstract: One of the most important decision problem of supply chain management is the demand forecasting. At the dawn of Industry 4.0 and with the first encouraging results concerning the application of deep learning methods in the management of the supply chain, we have chosen to study the use of neural networks in the elaboration of sales forecasts for a French furniture manufacturing company. Two main problems have been studied for this article: the seasonality of the data and the small amount of valuable data. After the determination of the best structure for the neuronal network, we compare our results with the results form AZAP legacy system job.

1. INTRODUCTION

One of the most important decision problem of supply chain management is the demand forecasting in order to balance inventory and service levels (Brown 1959), (Chapman et al. 2017). The role and use of artificial intelligence and machine learning methods in supply chain forecasting remains underexplored (Boone et al. 2019). Besides, the important issues to the development of supply chain forecasting are:

- The processes and systems through which the disaggregated forecasts are produced;
- Methods and selection algorithms that are suitable for supply chain data;
- The impact of new data sources from both consumer and supply chain partners;
- The effects of uncertainty and forecasts errors on the supply chain;
- The effects of linking forecasting to supply chain decisions, at both aggregate and disaggregate levels.

Several quantitative approaches exist. For this paper, we chose to focus our work on neural approaches, which are more and more used in supply chain management (Carbonneau et al. 2008). The purpose of this paper is to evaluate the ability of a neural model of time series prediction to predict sales. As this work is being conducted as part of a collaboration project between CRAN laboratory and Parisot company, we apply this approach to Parisot, which is currently using a commercial software to establish its sales forecasts. Two main problems have been studied for this article:

- The seasonality, which will be taken into account through two different ways as input data in the experiments: by assigning a number to the month in question, and by using the average, over the period, of the recorded orders relating to the month in question
- The small amount of data available, constituting a risk of overfitting, which will be taken into account through the algorithm.

The paper begins by presenting a brief state of the art regarding sales forecasts, then the neural network methods used in this work is presented. Then we will present the results before concluding and putting into perspective this work.
2. SALES FORECASTING

In classical forecasting techniques, it is first of all a question of straightening the series in order to eliminate the “accidental” variations whose origin is known. Next, we must determine the typology of the time series. Only after, comes the stage of the selection of forecast techniques adapted to the problems.

Complex time series can be broken down into three components (Chatfield 1996), (Nelson 1973):
- A trend component
- A cyclical component (seasonal) which can itself know a temporal evolution
- A random component (disturbance)

Concerning forecasting techniques, (Giard 2003) proposes the following classification, Figure 1:
- Explanatory models (called exogenous): the forecast is based on values taken by variables other than those we are trying to predict.
- Autoprojective models (called endogenous): the future is simply deduced from the past. We can refer to (Kendall 1976), (Kendall and Stuart 1976).

Figure 1: Forecasting techniques classification

The choice of the technique must minimize the cost of forecasting for a given level of precision, taking into account the type of time series, and the purpose, depending on whether the forecast will be used in the short or long term.

As part of Industry 4.0, researchers are turning to approaches based on artificial intelligence and new technologies increasingly accessible to businesses (Liu et al. 2013). These new approaches seem to offer overall better results than traditional methods (Jurczyk et al. 2016).

3. MULTILAYER PERCEPTRON

3.1 Structure

The traditional multilayer neural network consists of only one hidden layer (using a sigmoidal activation function) and one output layer. It is commonly referred to as a multilayer perceptron (MLP) and has been proven to be a universal approximator (Cybenko 1989), (Funahashi 1989). Its structure is given by:

$$\hat{y} = g_o \left( \sum_{h=1}^{n} w_{oh}^2 g_h \left( \sum_{i=1}^{n} w_{hi} x_i + b_h \right) + b \right)$$

where $x_i$ are the $n_i$ inputs, $w_{hi}$ are connecting weights between input and hidden layers, $b_h$ are the hidden neurons biases, $g_h(.)$ is the activation function of the hidden neurons (hyperbolic tangent), $w_{oh}^2$ are connecting weights between hidden and output layers, $b$ is the bias of the output neuron, $g_o(.)$ is the activation function of the output neuron, and $\hat{y}$ is the network output. The considered problem is a regression one. The $g_o(.)$ is a linear activation function.

Learning of an MLP is performed by using a local minimum search. therefore, the model accuracy depends of the choice of the initial parameter set. The one used here is a modification of the Nguyen and Widrow (NW) algorithm (Nguyen and Widrow 1990). It allows a random initialization of weights and biases, combined with an optimal placement in the input space (Demuth and Beale 1994).

This type of MLP has been extensively used in a broad range of problems including classification, function approximation, regression, times-series forecasting...

The considered problem here is a sales forecasting problem which can be viewed as a special case of times-series forecasting including consideration of seasonality.

Moreover, this problem must be solved by using small datasets which can lead to overfitting problem.

3.2 The seasonality problem

Seasonality is an aspect of the data that makes the prediction task difficult. (Zhang and Qi 2005) studies the effectiveness of seasonal series pre-treatments on the performance of multilayer perceptron. He concludes that even with pre-processing, their performance remains limited.

To take into account the phenomenon of seasonality, we propose to test and compare two approaches:
- We replace the month with a number, Table 1, which is varied (as example, in a first approach,
January is assigned the value 1, February 2..., in a second step, January is assigned the value 12, February 1...).

<table>
<thead>
<tr>
<th>HISTORICAL ORDERS</th>
<th>CDES-1</th>
<th>CDES-2</th>
<th>CDES-3</th>
<th>MONTH NUMBER January1</th>
<th>MONTH NUMBER January2</th>
<th>MONTH NUMBER January3</th>
</tr>
</thead>
<tbody>
<tr>
<td>46 592</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>29 820</td>
<td>46 592</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>37 487</td>
<td>29 820</td>
<td>46 592</td>
<td>0</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>35 861</td>
<td>37 487</td>
<td>29 820</td>
<td>46 592</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>18 772</td>
<td>35 861</td>
<td>37 487</td>
<td>29 820</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>43 712</td>
<td>18 772</td>
<td>35 861</td>
<td>37 487</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>33 106</td>
<td>43 712</td>
<td>18 772</td>
<td>35 861</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>24 074</td>
<td>33 106</td>
<td>43 712</td>
<td>18 772</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>23 813</td>
<td>24 074</td>
<td>33 106</td>
<td>43 712</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>31 578</td>
<td>23 813</td>
<td>24 074</td>
<td>33 106</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>22 093</td>
<td>31 578</td>
<td>23 813</td>
<td>24 074</td>
<td>11</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>41 048</td>
<td>22 093</td>
<td>31 578</td>
<td>23 813</td>
<td>12</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>38 660</td>
<td>41 048</td>
<td>22 093</td>
<td>31 578</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

- We add the average value of sales for the chosen month, Table 3.

### Table 3: Construction of the input file considering the number of the month

<table>
<thead>
<tr>
<th>AZAP FORECASTS</th>
<th>HISTORICAL ORDERS</th>
<th>CDES-1</th>
<th>CDES-2</th>
<th>CDES-3</th>
<th>MONTH NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td>55 805</td>
<td>46 592</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>45 573</td>
<td>29 820</td>
<td>46 592</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>37 487</td>
<td>29 820</td>
<td>46 592</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>35 861</td>
<td>37 487</td>
<td>29 820</td>
<td>46 592</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>18 772</td>
<td>35 861</td>
<td>37 487</td>
<td>29 820</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>43 712</td>
<td>18 772</td>
<td>35 861</td>
<td>37 487</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>33 106</td>
<td>43 712</td>
<td>18 772</td>
<td>35 861</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>24 074</td>
<td>33 106</td>
<td>43 712</td>
<td>18 772</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>23 813</td>
<td>24 074</td>
<td>33 106</td>
<td>43 712</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>31 578</td>
<td>23 813</td>
<td>24 074</td>
<td>33 106</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>22 093</td>
<td>31 578</td>
<td>23 813</td>
<td>24 074</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>41 048</td>
<td>22 093</td>
<td>31 578</td>
<td>23 813</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>38 660</td>
<td>41 048</td>
<td>22 093</td>
<td>31 578</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

### Table 2: Principle of shift of the number of the month

<table>
<thead>
<tr>
<th>HISTORICAL ORDERS</th>
<th>CDES-1</th>
<th>CDES-2</th>
<th>CDES-3</th>
<th>MONTH NUMBER January1</th>
<th>MONTH NUMBER January2</th>
<th>MONTH NUMBER January3</th>
</tr>
</thead>
<tbody>
<tr>
<td>46 592</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>29 820</td>
<td>46 592</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>37 487</td>
<td>29 820</td>
<td>46 592</td>
<td>0</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>35 861</td>
<td>37 487</td>
<td>29 820</td>
<td>46 592</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>18 772</td>
<td>35 861</td>
<td>37 487</td>
<td>29 820</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>43 712</td>
<td>18 772</td>
<td>35 861</td>
<td>37 487</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>33 106</td>
<td>43 712</td>
<td>18 772</td>
<td>35 861</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>24 074</td>
<td>33 106</td>
<td>43 712</td>
<td>18 772</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>23 813</td>
<td>24 074</td>
<td>33 106</td>
<td>43 712</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>31 578</td>
<td>23 813</td>
<td>24 074</td>
<td>33 106</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>22 093</td>
<td>31 578</td>
<td>23 813</td>
<td>24 074</td>
<td>11</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>41 048</td>
<td>22 093</td>
<td>31 578</td>
<td>23 813</td>
<td>12</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>38 660</td>
<td>41 048</td>
<td>22 093</td>
<td>31 578</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

### 3.3 The overfitting problem

One of the main risks encountered in the use of machine learning is the overfitting problem. This problem is related to the fact that the dataset used to learn is generally noisy and generally corrupted by outliers. If the model used includes too many parameters (freedom degrees), the learning step can lead to learning noise at the expense of learning the underlying system. To avoid this, different approaches may be used individually or in combination.

A classical approach is to perform a cross validation. Different approaches may be used. The considered here is the holdout method consisting by dividing the dataset into learning and validation dataset. This approach allows to detect and avoid this phenomenon by conjunction with the second approach: the early stopping. Early stopping consists to stop the learning when the performance of the model begins to deteriorate on the validation dataset.

The early stopping may be automatized by monitoring some parameters of the learning algorithm used. As example, in the Levenberg-Marquardt algorithm, it is possible to monitor the evolution of the gradient value and/or of the evolution of the parameter λ chosen to ensure the inversion of the Hessian matrix (Levenberg 1944), (Marquardt 1963). The learning algorithm used here is derived from the one proposed by (Norgaard 1995) which includes such mechanisms.
Another approach is to use trial and error approach to determine the optimal structure of this model. This is the last approach which is used here because it is the simplest.

The last approach to limit the overfitting problem is to include a regularization effect in the learning algorithm. This effect may be obtained by adding a weight decay term to the criterion to minimize in the learning algorithm (Norgaard 1995). Another approach is to use a robust criterion to minimize in the learning algorithm (Thomas et al. 1999). This is this second approach which is used here.

So, to conclude, to avoid the overfitting risk, the optimal structure is obtained by using trial and error approach. A cross-validation is performed, and an automatic early stopping is performed. Last, a robust criterion is included in the learning algorithm.

4. APPLICATION AND EXPERIMENTAL RESULTS

4.1 industrial context

This paper is a case study, based on the data from Parisot – Saint Loup’s Unit. It’s an octogenarian French furniture industry. She sells furniture kits made of particle board in France and around the world. The catalog consists of furniture for bedroom, living room, kitchen, bathroom, in single package or multi package.

This company has had to face in recent years many changes. Sudden adaptations have not necessarily been optimal and have led to bad long-term practices. These bad practices now prevent the company from remaining competitive.

Today, the forecasts are made using AZAP\(^3\).

AZAP is a publisher of a software suite that assists with optimum Supply Chain management. It covers the management functions of sales forecasting and demand management, industrial planning and inventory level optimization of companies looking to significantly cut costs, optimise stock management and improve customer service promptly.

The need of the company is to improve the sales forecasting function without the need of encapsulated business logic. Thus, the ML could be a good candidate.

4.2 Dataset

We used the summary file of final forecasts and orders recorded by large groups of customers. The data are monthly recorded from 2012 to 2018, constituting 81 values, which is a small dataset. We chose to separate the series into two datasets, one for learning and one for validation. The chosen method is that of 80/20 randomly. So, the learning dataset includes 62 patterns when the validation one includes 16 patterns. The dataset is normalized before to perform the learning.

In order to compare the results between them, we use the root mean square error value (RMSE):

\[
RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^{N} (y(n) - \hat{y}(n))^2}
\]

Where \(N\) is the size of the dataset, \(y(n)\) is the \(n\)th target, and \(\hat{y}(n)\) its estimation.

4.3 Structure determination using a trial / error approach

The first step in this study was to determine the structure of the neural network.

In times-series forecasting, the prediction of the output (here orders) is performed by using past data of orders. The first step is to determine the time window. To do that, we varied the number of inputs considering the data about the orders, as shown by Table 4. This amounts to varying the number of delayed data to be taken into account in relation to the value considered.

Table 4: Construction of delayed data to be taken into account as inputs

<table>
<thead>
<tr>
<th>AZAP FORECASTS</th>
<th>HISTORICAL ORDERS</th>
<th>CDES-1</th>
<th>CDES-2</th>
<th>CDES-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>55 805</td>
<td>46 592</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>45 573</td>
<td>29 820</td>
<td>46 592</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>41 987</td>
<td>37 487</td>
<td>29 820</td>
<td>46 592</td>
<td>0</td>
</tr>
<tr>
<td>46 950</td>
<td>35 861</td>
<td>37 487</td>
<td>29 820</td>
<td>46 592</td>
</tr>
<tr>
<td>39 270</td>
<td>18 772</td>
<td>35 861</td>
<td>37 487</td>
<td>29 820</td>
</tr>
<tr>
<td>51 663</td>
<td>43 712</td>
<td>18 772</td>
<td>35 861</td>
<td>37 487</td>
</tr>
</tbody>
</table>

We varied the frame size from 3 to 6, and, to avoid the problem of local minimum search, this task is performed hundred times and the best result is kept. The number of hidden neurons is tuned to 4. Table 5 presents the best RMSE obtained on the validation data set in function of the number of delayed inputs. This RMSE is calculated with normalized dataset. These results show that using a time window of size

---

\(^3\) http://www.azap.net/en/
In a first step, the seasonality is considered by associating each month to a real value ( ). This approach presents the drawback to introduce an important gap between two consecutive months. As example, by associating 1 to January, 2 to February…, the transition from December to January induces a jump from 12 to 1. To evaluate the impact of this jump on the accuracy of the model, a permutation of these association is performed (in a first experiment, January is associated to 1, in a second it is associated to 2…). As for the determination of the structure, 10 learning has been performed with 10 initialization of parameters for each experiment in order to avoid the local minimum search problem.

Erreur ! Source du renvoi introuvable, presents the RMSE obtained on the validation dataset for these 12 experiments. In order to compare these results with those obtained with AZAP, this RMSE is calculated on non-normalized dataset.

These results show that the position of the jump in the year has an important impact on the results. As example, when it is situated between January and February (line NN FORECASTS Janvier12) the accuracy is 27% worse than that it is situated between May and June (line NN FORECASTS Janvier8).

We can see that NN forecasts with some sequences (Janvier3, Janvier 4 and Janvier 8) are slightly better than the AZAP forecasts.

For all following experiments, the structure of the MLP includes:
- The small amount of data available, which constitutes a risk of over-learning
- 4 inputs: the time window of 3 historical values and an input related to the month: the number or the mean of orders to the month considered.
- 3 hidden neurons.

### 4.5 Comparison

In a second step, the optimal number of hidden neurons is expected, we varied its number from 3 to 6. As for the determination of the number of delays inputs to use, to avoid the local search minimum problem, ten learning has been performed on ten different initial settings of parameters and the best result is kept. Table 6 presents the RMSE obtained on the validation dataset (overfitting). So, in the sequel, the time window is tuned to 3.

<table>
<thead>
<tr>
<th>HISTORICAL ORDERS</th>
<th>CDENS1</th>
<th>CDENS2</th>
<th>CDENS3</th>
<th>MONTH NUMBER</th>
<th>MONTH NUMBER</th>
<th>MONTH NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td>46 592</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>29 820</td>
<td>29 820</td>
<td>46 592</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>43 712</td>
<td>37 487</td>
<td>29 820</td>
<td>46 592</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>18 772</td>
<td>35 861</td>
<td>37 487</td>
<td>29 820</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>33 106</td>
<td>43 712</td>
<td>18 772</td>
<td>35 861</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>25 832</td>
<td>33 106</td>
<td>43 712</td>
<td>18 772</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>31 578</td>
<td>25 832</td>
<td>24 074</td>
<td>33 106</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>22 093</td>
<td>31 578</td>
<td>25 832</td>
<td>24 074</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>41 048</td>
<td>31 578</td>
<td>25 832</td>
<td>24 074</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>38 660</td>
<td>31 821</td>
<td>41 048</td>
<td>22 093</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
</tbody>
</table>
using of the average sales for the considered month improves the accuracy of the model comparatively to the two others. The improvement is of 5.5% comparatively to the preceding approach and up to 8% comparatively to AZAP. This can be explained by the fact that this approach gives a richer information.

Table 8: Results considering the input of the month mean

<table>
<thead>
<tr>
<th>CONSIDERED EXPERIMENT</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>AZAP FORECASTS</td>
<td>5882</td>
</tr>
<tr>
<td>NN FORECASTS Janvier8</td>
<td>5753</td>
</tr>
<tr>
<td>NN FORECASTS mean of the month</td>
<td>5450</td>
</tr>
</tbody>
</table>

5. CONCLUSIONS

This paper presents an initial work of a Ph.D study in collaboration with a French furniture manufacturer. Its goal is to propose a machine learning approach to perform sales forecasting.

A classical neural network model (multilayer perceptron) is used. The main difficulty is related to the small size of the dataset which can lead to the overfitting problem. To avoid it, a combination of different strategies is used (cross-validation, early stopping, robust learning algorithm, optimal structure determination).

The second difficulty is related to the taking into account of the seasonality. Two approaches have been proposed and compared. This study has shown that taking the mean of the month into account as an input is significant to solve the problem of seasonality.

We have to take in account that our result consists of the basic forecasts in the AZAP process, and we compare our results with the final forecasts obtained after the forecaster job. In future works, we must add the information of the commercial and marketing forecasts taking into account the effect of publicity and events. We can also test the impact of agglomerating or disaggregating customer hierarchy data. Last, the cross-validation approach used here is the holdout method which is simple but maybe not the more efficient when the dataset is small. Other cross-validation approaches must be tested and compared in the future such as k-fold or leave one out as example.

REFERENCES


Jurczyk, K., Gdowska, K., Mikulik, J., Wozniak, W., 2016. Demand Forecasting with the Usage of Artificial Neural Networks on the Example of Distribution Enterprise. Int. Conf. on Industrial Logistics ICIL’2016, September 28- october1, Zakopane, Poland.


