European Research Studies Journal
Volume XXIV, Special Issue 2, 2021

pp. 627-636

Forecasting Sales in the Supply Chain Based on the LSTM Network: The Case of Furniture Industry

Submitted 19/03/21, 1st revision 18/04/21, 2nd revision 10/05/21, accepted 30/06/21

Damian Pliszczuk[1], Piotr Lesiak[2], Krzysztof Zuk[3], Tomasz Cieplak[4]

Abstract:

Purpose: *The aim of the article is to develop an algorithm for forecasting sales in the supply chain based on the LSTM network using historical sales data of a furniture industry company.*
Design/Methodology/Approach: *Machine learning was used to analyze the data. The method of predicting the behavior of sales value in a specific time horizon in terms of a time series was presented. The LSTM network was used to predict sales. The network used is a special case of recursive neural networks with an important difference in the repeating module. Due to the fact that the activities are carried out on time series, the data was analyzed in terms of the stationarity of such series or trends and seasonal effects. The data used in the analysis includes the daily sales values of a group of certain furniture collections over a specified time horizon. The stationarity of the time series can have a significant impact on its properties and behavior prediction, where failure to bring the time series to the correct form of stationarity can lead to false results.*
Findings: *The result of the research was the analysis of sales forecasting in the supply chain based on machine learning. As a result of the data transformations, the algorithm was able to recognize and learn long-term relationships.*
Practical Implications: *The presented method of predicting the behavior of sales value in a specific time horizon allows for a look at the forecasting of demand in terms of the supply chain. The sales data of a company from the furniture industry were used for the analysis.*
Originality/Value: *A novelty is the use of the LSTM network trained on real transaction data of a furniture company that has based its business on the supply chain and cooperates with its suppliers and recipients in Central and Eastern Europe.*

Keywords: *Machine learning, time series, LSTM, supply chain, forecasting.*

JEL codes: *C45, C61, E27.*

Paper type: *Research article.*

---

[1]*Corresponding author, Netrix Group sp. z o.o. , Lublin, Poland,*
e-mail: damian.pliszczuk@netrix.com.pl
[2]*University of Economics and Innovation in Lublin, Lublin, Poland,*
e-mail: piotr.lesiak@wsei.lublin.pl
[3]*same as in 2, e-mail: krzysztof.zuk@wsei.lublin.pl*
[4]*Faculty of Management, Lublin University of Technology, Lublin, Poland,*
e-mail: t.cieplak@pollub.pl

1. Forecasting Demand in the Supply Chain

Forecasting future demand is essential to making supply chain decisions. Undoubtedly, historical information about demand can be used to forecast future demand and such analysis affects the functioning of the entire supply chain. This section will outline the general principles of supply chain forecasting and the tools used to do so. Demand forecasts form the basis of all supply chain planning. Consider the division of processes performed in the push / pull supply chain (Chopra and Meindl, 2007). Pull processes are initiated by an order placed by the customer, while push processes are initiated and executed while awaiting the order. When considering these two options for push processes, you should plan your level of activity, be it production, transportation, or any other planned activity. With pull processes, you need to plan for the level of available bandwidth and inventory, but not the actual amount of products to be made. In both cases, the first step to be carried out is to forecast customer demand.

To be competitive, a company must use a forecast of future demand to determine the quantity of products to be pushed and to determine the production capacity needed in its plants (pull process). Further down the supply chain, another company that produces one of the components also needs forecasts to determine its own production and inventory levels. For the same reason, the suppliers of this company also need forecasts. When each stage of the supply chain makes its own separate forecast, these forecasts are often very different (Chopra and Meindl, 2007; Jaggi and Kadam, 2016). This may result in a mismatch between supply and demand. When all stages of the supply chain work together to produce a shared forecast, it is much more accurate. The resulting forecast accuracy enables supply chains to be both more sensitive and more efficient in serving their customers.

Enterprises and supply chain managers should be aware of the following features of forecasts (Khojasteh, 2018; Chopra and Meindl, 2007):

- Forecasts are always wrong and should therefore contain both the expected forecast value and the forecast error measure. Forecast error (or uncertainty in demand) must be a key factor influencing most supply chain decisions.
- Long-term forecasts are usually less accurate than short-term forecasts, i.e., long-term forecasts have a larger standard deviation of error relative to the mean than short-term forecasts. To get closer to this relationship, consider a replenishment process that allows you to respond to an order within a few hours. The short lead time allows the manager to take into account current information, such as the weather, which may have an impact on product sales. This forecast is likely to be more accurate than the week ahead of the demand forecast.
- Aggregated forecasts are usually more accurate than individual forecasts because they tend to have a smaller standard deviation of error from the mean. For example, it is easier to forecast a country's Gross Domestic

Product (GDP) for a given year with a slight error, and it is much more difficult to forecast a firm's annual revenue with a similar error, and it is even more difficult to forecast revenues for a given product with the same accuracy. The key difference between the three projections is the degree of aggregation. GDP is aggregated across many companies, and company revenues are aggregated across several product lines. The larger the aggregation, the more accurate the forecast is.

- In general, the further up the supply chain a business is (or the farther it is from the consumer), the greater the disruption to the information it receives. Joint forecasting based on sales to the end customer helps companies operating upstream in the supply chain reduce forecast error.

In a way, one might be tempted to treat demand forecasting as a kind of art used so as not to leave everything to chance. What the company knows about the previous behavior of its customers is reflected in their future behavior. Demand is influenced by various factors and can be predicted with at least some probability if the firm can identify the relationship between these factors and future demand. To forecast demand, firms must first identify the factors that affect future demand and then establish a relationship between those factors and future demand.

When forecasting demand, enterprises must maintain a balance between objective and subjective factors. In the presented considerations we focus on the methods of quantitative forecasting, however, it should be remembered that companies must take into account human input when preparing the final forecast. The supply chain can experience significant benefits from improving demand forecasting due to qualitative human factors. The company must have knowledge of many factors that are related to the forecasting of demand. Some of these factors are (Zsidisin and Ritchie, 2009; Chopra and Meindl, 2007):

- Previous demand;
- Product realization time;
- Planned advertising or marketing activities;
- State of the economy;
- Planned price discounts;
- Actions taken by competitors.

A business must therefore understand such factors before it can choose an appropriate forecasting methodology. Forecasting methods are usually classified into the four types (Norrman and Jansson, 2004; Chopra and Meindl, 2007) presented in Table 1.

Table 1. *Types of Forecasting Methods*

Type	Description
Qualitative	Qualitative forecasting methods are primarily subjective and rely on human judgment. They are most appropriate when little historical data is available or when experts have market knowledge that could influence the forecast. Such methods may also be necessary to forecast demand for several years in a new industry.
Time series	Time series forecasting methods use historical demand to produce a forecast. They are based on the assumption that a history of past demand is a good indicator of future demand. These methods are most appropriate when the underlying demand pattern does not change significantly from year to year. These are the easiest methods to apply and can serve as a good starting point for forecasting your demand. For an example of analysis based on time series, see the next chapter.
Causal	The causal forecasting methods assume that the demand forecast is strongly correlated with certain environmental factors (e.g. the state of the economy, interest rates). The causal forecasting methods establish a correlation between demand and environmental factors and use estimates of what environmental factors will predict future demand. For example, product prices are strongly correlated with demand. Companies can therefore use cause-and-effect methods to determine the effect of price promotions on demand behavior.
Simulation	Simulation forecasting methods mimic the choices made by consumers that generate demand to obtain the forecast. Using simulations, a company can combine time series and causal methods to answer questions such as what will be the impact of a price promotion or what will be the impact of a competitor opening a new store in a given location.

Source: *Own creation based on Norrman and Jansson, 2004; Chopra and Meindl, 2007.*

Choosing the right method is not an easy matter and largely depends on the specific problem, however, several studies have shown that using multiple forecasting methods to create a combined forecast is more effective than using just one method. In our considerations, we primarily deal with the time series methods that are most appropriate when future demand is related to historical demand, growth patterns, and any seasonal patterns. There is always a random element in any forecasting method that cannot be explained by historical demand patterns. Therefore, any observed demand can be divided into a systematic and a random part (Chopra and Meindl, 2007):

$$D = S + R, \tag{1}$$

where D is the observed demand, S - the systematic component, and R - the random component.

The systematic component measures the expected value of the demand and consists of what we call the level of the current demand, excluding the inflows of holiday and random effects. The random component is the part of the forecast that deviates from the systematic part. The enterprise cannot (and should not) forecast the direction of the random component. All the enterprise can predict is the size and variability of the random component, which is a measure of the forecast error. On average, a good forecasting method has an error that is comparable in magnitude to the random demand component. The purpose of forecasting is to filter out the random component (noise) and estimate the systematic component. Forecast Error measures the difference between forecast and actual demand.

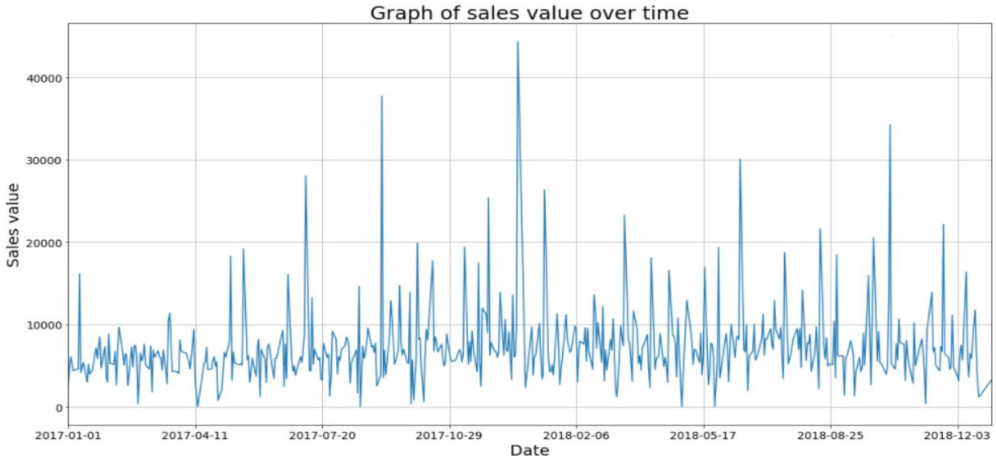
2. Sales Value Forecasting Methods and Analyzes

In this section, we will present how to predict the behaviour of sales value over a specific time horizon - in terms of a time series. The view on forecasting demand is therefore quite simplified and takes place from the top level of the supply chain - without taking into account additional factors such as the supply itself. For this purpose, sales data will be used and an important issue will be an attempt to predict future sales values using only the dependent variable - no other predicates will be introduced (Hastie, Tibshirani, and Friedman, 2009).

2.1 Data Preparation

When considering a business issue, especially in terms of such a complex system as the supply chain, it is worth considering various approaches and models. Due to the fact that the activities are carried out on the previously mentioned time series, it is worth initially analyzing the data in terms of the stationarity of such series or trends and seasonal effects. The data used in the analysis includes daily sales values of a group of certain furniture collections in a specific time horizon - from January 2017 to March 2019 inclusive. As it has already been mentioned, we will not use any additional variables for the analysis, but only the sales value recorded on a given day. In general, the focus will be on sales in 2017-2018 and the first three months of 2019 will be used to test the model created. The distribution of the analyzed variable is presented in Figure 1.

Figure 1. *Graph of sales value*

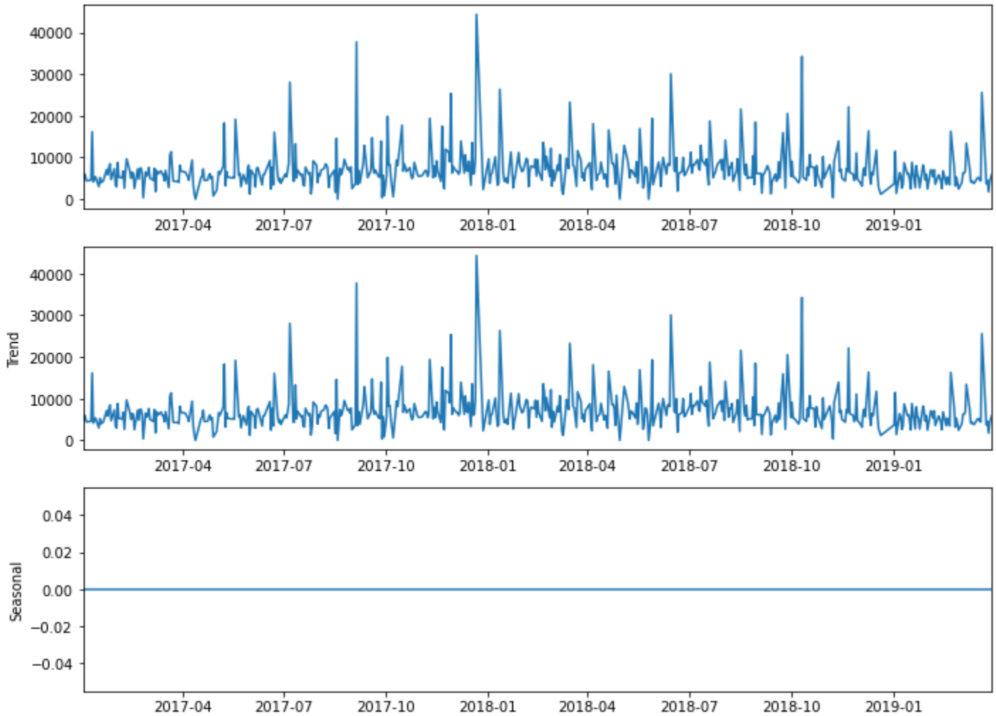


Source: *Own creation.*

The stationarity of the time series can have a significant impact on its properties and behavior prediction, where failure to bring the time series to the correct form of stationarity can lead to false results. A time series is called non-stationary when it is a time series that has no constant mean, variance, or covariance over time (Greunen *et al.*, 2014). The Dickey-Fuller test and the extended version of ADF (Augumented Dickey-Fuller) used in this case (Mushtaq, 2011) are used to determine whether the tested time series is a stationary series or not, in addition to testing the series itself in terms of trend and seasonality. In Python, such a test can be performed using the statsmodels library. The null hypothesis of this test is that the process is not stationary.

When performed on the analyzed dataset, the results are as follows:
ADF Statistic: -18.330827
p-value: 0.000000
Critical Values:
1%: -3.438
5%: -2.865
10%: -2.569

Figure 2. *Trend and seasonality chart*

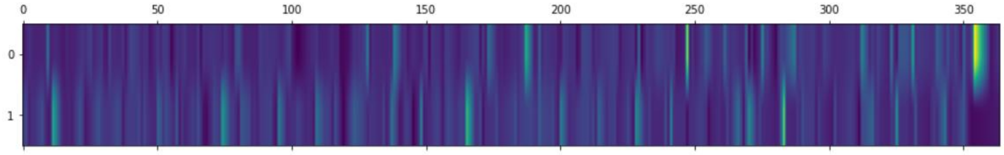


Source: *Own creation.*

The low value of the ADF statistics and the p-value at the significance level of 95% allow us to reject the null hypothesis in favor of the alternative, and therefore the series is stationary. This is also confirmed by the charts presented in Figure 2, which show that the entire series can be treated as a trend and that there is no seasonal effect.

At this point, it is also worth taking a look at the heat map of the sales values present in the series in question. It also indicates that there are no time frames where sales were either significantly high in the two years under consideration or significantly low. The highest value is indicated at the end of the year, which is a direct result of high sales at the end of December 2017 (Figure 3).

Figure 3. *Sales value heat map in two years*



Source: *Own creation.*

After examining the seasonality, the data was scaled to the range (0.1) using the MinMax function of the form

$$f(x) = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (2)$$

A further idea behind the proposed solution is to create, on the basis of the analyzed set, a set containing a certain number of time steps backwards, on the basis of which we will predict the value at the current time step. Thus, a timezone has been used which in this case uses the previous two (t-2, t-1) time steps to predict the value of the current step t.

As mentioned earlier, the years 2017-2018 were treated as the training set, of which the last two months of 2018 served as the validation set. The rest of the data, i.e. the first three months of 2019, was taken as test data.

2.2 Forecasting with the Use of Machine Learning

In order to predict sales, the LSTM network was used to combat the problem of short-term memory. This is extremely important because information is lost as a result of the transformations of data that occur as it flows through the network (Géron, 2019). LSTM (Long-Short-Term Memory) is able to recognize and learn long-term relationships, which is used, among others, in forecasting time series. The network used is a special case of recursive neural networks with an important difference in the repeating module. Unlike the standard solution, there are four layers that work together. In the first part (3), the model verifies h_{t-1} and x_t and sets the value to 0 (forget) or 1 (save) for each number in the state c_{t-1}. In the next step, it is determined which information is to be stored in the cell state, and the input layer decides which values will be updated (4).

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{3}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{4}$$

$$g_t = tanh\left(W_g \cdot [h_{t-1}, x_t] + b_g\right) \tag{5}$$

$$C_t = f_t \times C_{t-1} + i_t \times g_t \tag{6}$$

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{7}$$

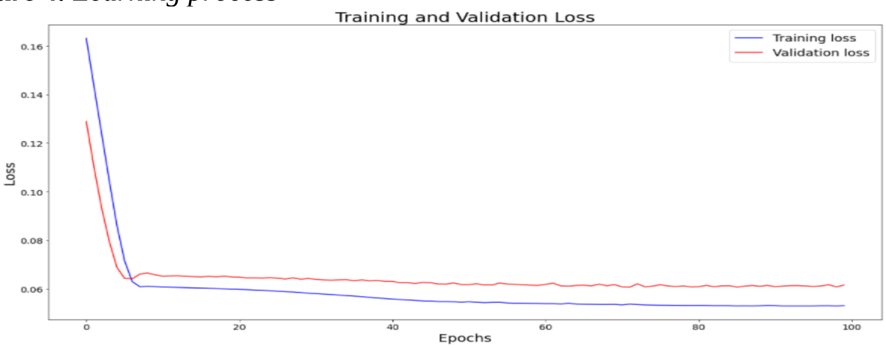
$$h_t = O_t \times \tanh(C_t) \tag{8}$$

Moving on to the next layer, a vector of new candidates is generated at the output, marked as g_t, which will be added to state (5). Now, in order to update the previous cell state to its new state, the two steps mentioned are combined. Equation (6) clearly shows that the value of the previous state c_{t-1} is multiplied by the coefficient f_t which is added to the product of the candidate vectors g_t in order to scale the individual status values to be updated (Zhang *et al.*, 2018). In the last step, the sigmoid layer decides the cell states to exit (7), and the cell states pass through the tanh function and convert the data to values ranging from -1 to 1. The final output (8) is the product of tanh with the sigmoid gate output.

2.3 Network Construction

The library used to build the LSTM network is Keras based on Tensorflow. The model adopted has a sequential structure consisting of two layers using LSTM cells with one dense layer being the output. In order to select the hyperparameters used in the model, the Talos library cooperating with the Keras library used for training was used. The training process of the selected model is presented below.

Figure 4. *Learning process*



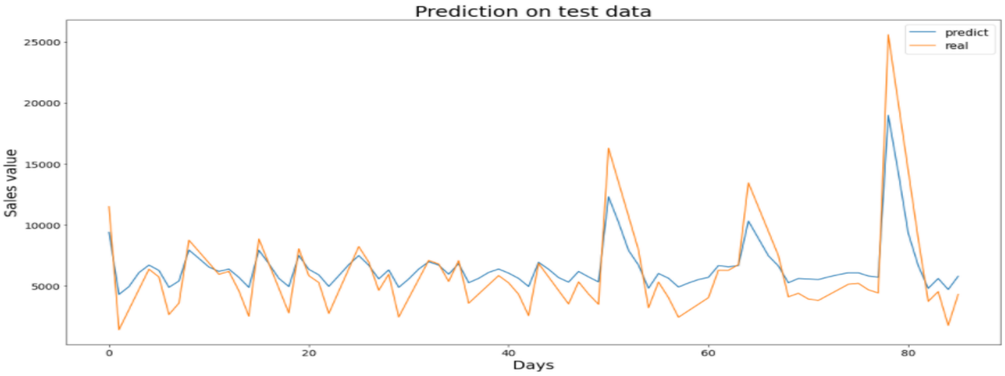
Source: *Own creation.*

The obtained model performance is similar in the training and validation set. Learning and validation errors stabilize after approximately 60 epochs. The plotted losses are normalized values (Figure 4). It follows that the networks have not been overtrained and ensure a good fit with the data.

2.4 Forecasts

Then, on the basis of the model, a prediction was performed based on the previously determined test data, the results of which are presented in the figure below.

Figure 5. *Plot of real and forecasted*



Source: *Own creation.*

Thus, you can see a clear good prediction of the model in terms of the trend of jumps or falls in sales. Of course, the reverse transformation to the MinMax function ensured the return to the initial data. The mean square error for the obtained prediction was 3197.75.

3. Conclusions

The article presents forecasting of sales in the supply chain based on machine learning. Demand forecasts form the basis of all supply chain planning. The use of machine learning models allows you to take into account historical data and the impact of various factors on revenues, and thus make a much more precise price adjustment.

The presented method of predicting the behaviour of sales value in a specific time horizon allows for a look at the forecasting of demand in terms of the supply chain. The analysis used sales data of a company from the furniture industry, and an important issue was the prediction of future sales values using the dependent variable. The LSTM network was used to predict sales. As a result of the data transformations, the algorithm was able to recognize and learn long-term relationships.

References:

Chopra, S., Meindl, P. 2007. Supply chain management : strategy, planning, and operation. Pearson Prentice Hall.
Géron, A. 2019. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems, 2nd Edition. O’Reilly Media.
Greunen, J., Heymans, A., Heerden, C., Vuuren, G. 2014. The Prominence of Stationarity in Time Series Forecasting. Studies in Economics and Econometrics, 38(1), 1-16. https://doi.org/10.1080/10800379.2014.12097260.
Hastie, T., Tibshirani, R., Friedman, J. 2009. The Elements of Statistical Learning. New York, NY: Springer.
Jaggi, H.S., Kadam, S.S. 2016. Integration of Spark framework in Supply Chain Management. Procedia Computer Science, 79, 1013-1020. https://doi.org/10.1016/j.procs.2016.03.128.
Khojasteh, Y. 2018. Supply Chain Risk Management. Singapore: Springer Singapore.
Mushtaq, R. 2011. Augmented Dickey Fuller Test. SSRN Electron.
Norrman, A., Jansson, U. 2004. Ericsson’s proactive supply chain risk management approach after a serious sub-supplier accident. International Journal of Physical Distribution & Logistics Management, 34(5), 434-456. https://doi.org/10.1108/09600030410545463.
Zhang, W. et al. 2018. LSTM-Based Analysis of Industrial IoT Equipment. IEEE Access, 6, 23551-23560. https://doi.org/10.1109/ACCESS.2018.2825538.
Zsidisin, G.A., Ritchie, B. 2009. Supply Chain Risk, 124. Boston, MA: Springer US.