

Temür, Ayşe Soy; Yıldız, Şule

Article

Comparison of Forecasting Performance of ARIMA LSTM and HYBRID Models for The Sales Volume Budget of a Manufacturing Enterprise

Istanbul Business Research (IBR)

Provided in Cooperation with:

Istanbul University, Business School

Suggested Citation: Temür, Ayşe Soy; Yıldız, Şule (2021) : Comparison of Forecasting Performance of ARIMA LSTM and HYBRID Models for The Sales Volume Budget of a Manufacturing Enterprise, Istanbul Business Research (IBR), ISSN 2630-5488, Istanbul University Press, Istanbul, Vol. 50, Iss. 1, pp. 15-46,
<https://doi.org/10.26650/ibr.2021.51.0117>

This Version is available at:

<https://hdl.handle.net/10419/250251>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by-nc/4.0/>



Comparison of Forecasting Performance of ARIMA LSTM and HYBRID Models for The Sales Volume Budget of a Manufacturing Enterprise*

Ayşe Soy Temür¹ , Şule Yıldız² 

Abstract

This study aims to create a monthly sales quantity budget by making use of the previous income data of an enterprise operating within the construction sector, which is considered the locomotive of the economy. For estimating time-series of sales as a linear model *ARIMA* (*Auto-Regressive Integrated Moving Average*), as nonlinear model *LSTM* (*Long Short-Term Memory*) and a *HYBRID* (*LSTM and ARIMA*) model built to improve system performance compared to a single model was used. As a result of the study, *Mean Square Error (MSE)*, *Root Mean Square Error (RMSE)*, *Mean Absolute Percentage Error (MAPE)* and *Mean Absolute Error (MAE)* values obtained from each of the methods used in the application were compared, and a monthly sales volume budget was created for 2017 with all the methods used. When the *MAPE* and *MSE* values obtained from each of these methods were compared, the best performance was the *Hybrid* model that gave the lowest error, and in addition, the fact that all of the application models got very realistic results by using the historical data showed the success of the predictions..

Keywords

Sales Budget, Time Series Forecast, Hybrid Model, ARIMA, LSTM

Introduction

The construction sector is considered the “locomotive of the economy” because it affects the demand for goods and services produced by many sub-sectors connected to it. Also, due to its contribution to employment and its relationship with other sectors, it performs a fundamental role in achieving and maintaining the economic growth rate desired to be achieved (Kılıç & Demirbaş, 2012).

* This study is derived from Ayşe SOY TEMÜR's PhD thesis titled "Comparison of ARIMA, LSTM and Hybrid Models: Production Business Example in the Formation of Sales Budgets of Businesses," which was completed under the consultancy of Assoc. Prof. Dr. Şule YILDIZ, Sakarya University, Institute of Business Administration.

1 **Corresponding Author:** Ayşe Soy Temür (Instr. Dr.), Duzce University, Kaynaşlı Vocational School, Foreign Trade Department, Duzce, Turkey. E-mail: aysesoy@duzce.edu.tr ORCID: 0000-0003-4455-5035

2 Şule Yıldız (Assoc. Prof. Dr.), Sakarya University, Faculty of Business, Department of Business Administration, Sakarya, Turkey. E-mail: kasapoglu@sakarya.edu.tr ORCID: 0000-0002-4630-0637

To cite this article: Soy-Temur, A., & Yildiz, S. (2021). Comparison of Forecasting Performance of ARIMA LSTM and HYBRID Models for The Sales Volume Budget of a Manufacturing Enterprise. *Istanbul Business Research*. Advance online publication. <http://doi.org/10.26650/ibr.2021.51.0117>



Economic uncertainties are the most critical elements affecting demand in the construction industry. In some periods, suppressing the demand in the sector or disrupting the continuity in manufacturing also has a diminishing effect on fixed capital investments in the sector. This situation causes producers to stay away from fixed capital investments and keep their assets liquid at all times. The contraction of the construction sector may cause the sub-industries producing construction materials to be forced to work at low capacity in this period. This adverse process also impacts the manufacturing industry, the national economy, and as a result, employment (Düzyol, 1997). Due to these disadvantages, businesses want to predict the future to compete in increasingly competitive conditions and to adapt to the continually changing world. Therefore, demand forecasting is one of the critical tools that businesses will use to attain their short- and long-term goals. With a good forecast, future uncertainties are reduced. The achievement of this depends on the effective establishment and operation of the forecasting system.

As a result of the estimations made, the operating budget consisting of management plans related to the activities of the companies for the future periods is prepared. These budgets show the resources that businesses need to achieve their goals, how these resources will be obtained and where they will be used. Although there are exceptions, sales budgets form the basis of the budget system in businesses. Forecasting sales for the next period is significant for a sound budget system and therefore a reasonable sales budget.

The methods used in sales forecasting are generally divided into two groups as traditional and advanced methods. Methods based on expert opinions and personal experiences are expressed as traditional methods, while methods based on statistical and mathematical data are referred to as advanced methods.

The set of values observed at regular time intervals and consecutive time periods are expressed as time-series. The techniques of forming estimates and policies regarding the values that may occur in the future with the use of the data that emerged in the past are time-series analyses. The primary purpose of time-series analysis is to develop a mathematical model based on available data. Due to the difficulty in evaluating the precise nature of a time-series, it is often challenging to produce appropriate predictions. (Khandelwal, Adhikari, & Verma, 2015). There are several methods in time-series analysis, and each method has several advantages and disadvantages. One of the most significant benefits of time-series methods is their ability to solve problems where the data set, which is very complicated for traditional technologies, is not linear, missing or multidimensional. (Yılmazel, Afşar, & Yılmazel, 2018). Because of these advantages, time-series are frequently used in forecasting.

The reliability of demand forecasting methods is the essential factor in the success of that method. Whether the method is successful or not is determined by comparing the estimated results obtained with the actual values. The fact that the difference between the predictions made and the values emerging starts to open means that it gives a critical alarm that the pre-

dicted environment has changed. Considering the fast consumption sector, following the demand trend in this sector and the positioning of the products according to these estimates will enable both the customers to find the product they are looking for and to not make incomplete or excessive stocks. Also, it will provide various benefits such as making the budget right, managing the cash flow correctly and not missing out on opportunities (Ahmetoğlu, 2018).

Many methods that can be called linear, nonlinear and hybrid methods have been developed to increase the accuracy of predictions. Linear methods establish a linear relationship between observed data points and future data points. The ARIMA model from linear methods and artificial intelligence-based models from nonlinear methods have been used frequently in demand forecasting in recent years. ARIMA, which is one of the linear statistical methods, has been used in many applications in relevant literature. This method used in stationary time-series is used to deduce the factors in the time-series.

The Artificial Neural Networks (ANN) method has many important features such as generalization, working with an unlimited number of variables, and learning from data. Providing important benefits due to these features, ANN stands out as an effective demand forecasting technique. ANN is similar to qualitative methods in terms of learning the relationship and degrees between factors, and quantitative methods in terms of using statistical data. But people do not learn new meanings for every word they have just learned. Based on similar words that existed earlier, they add meaning to the word they just learned. The biggest deficiency of traditional ANNs is the absence of this meaningful attribute present in humans. Repetitive ANN (RNN), on the other hand, creates a loop and enables the use of historical information, so that it can classify by making meaningful connections between the frames. In this sense, RNNs make up for this shortcoming in traditional ANNs (Olah, 2015).

RNNs can make sense of events that occur sequentially, as they can form a loop. Due to this advantage, they are widely used in many areas such as language modeling, translation, image captioning, speech recognition, etc. in recent years. However, although RNNs give successful results in some problems thanks to their connection with the past and their interpretation features, some situations such as which activities to remember, and how long to remember, are not known (Bengio, et al., 2015). While some information is important for activities, some information may be unnecessary. Therefore, there is no need to save the entire history. In the classification phase, this information may not be available if the information needed has already been created. Therefore, RNNs need a different architectural structure to predict previous events. In such problems, Long Short-Term Memory (LSTM) networks are used that work better and can learn a particular type of RNN and long-term dependencies. (Olah, 2015). Unlike RNN, “Vanishing gradient” problem that may arise during training can be solved with LSTM. For this reason, LSTMs are generally preferred in many activity classifications. The use of LSTMs in the vast majority of successful studies with RNNs in recent years is an indication of this situation.

Both linear and nonlinear methods are not suitable universal models that can be used in all conditions. For this reason, in the literature, “hybrid” methods are also used, which combine the advantages of more than one individual model to overcome this limitation and increase the performance of the forecast. The idea behind hybrid models is that each unique method in hybrid methods is used to capture different prediction models. For example, in the case where a time-series includes both linear and non-linear models at the same time interval, neither linear nor nonlinear models alone are sufficient. For this reason, hybrid models may have been created by combining homogeneous, i.e. differently configured neural networks or heterogeneous, i.e., linear and non-linear models. Models that combine ARIMA with ANN are an excellent example of a hybrid model. While ANN is concerned with non-linear models in such hybrid models, the ARIMA model is concerned with linear models.

In this study, ARIMA was preferred as it is a linear model for sales prediction frequently used in social sciences, and LSTM, which makes up for the deficiencies of RNN as a non-linear model, and which has widely been used in recent years, was preferred. The application of the same data to more models rather than a single model can increase the accuracy of the estimation. Therefore, a hybrid model was created to overcome the limitations of LSTM and ARIMA models. As the data set, 96 monthly sales data of an enterprise operating in the construction sector in the period between 01.2010 and 12.2017 were used.

The primary purpose of the study was to produce accurate predictions about future periods by using the data of the previous period, regardless of the method used. Therefore, analyses were conducted firstly with the ARIMA model, secondly with the LSTM model and lastly with a hybrid model which was created by combining the advantages of the two models as mentioned above. The results obtained from all of the methods were compared with each other and with real values. With this study, it was aimed both to benefit the company in making future production and sales plans by predicting the sales at a decent and minimum error level, and to contribute to the literature in terms of similar studies to be carried out from now on.

The study consists of a total of three parts, apart from the introduction and conclusion. The first part of our research is the literature section and it includes specific examples of projects which used ARIMA, LSTM and some Hybrid models. The second section describes the methodology used to create predictions. In the third and last section, the methods adopted in the analysis are explained and the findings obtained from the analysis results are included.

Literature Review

ARIMA models in the literature have been applied to predict future values of various time-series data such as wind speeds, global temperature values, water quality, housing prices, electricity prices, sugar prices, sales data, stock market data. The popularity of these models is due to their statistical properties as well as their dependence on the Box-Jenkins metho-

dology. These models can help understand the dynamics of data in a particular application (Smith, 2018).

On the other hand, LSTM networks are used mostly for more in-depth learning and achieve greater success with large data sets. However, though limited in number, examples of LSTM being trained with few data sets are also available in the literature. For instance, Namin and Namin (2018) succeeded in estimating the economic and financial time-series they performed using an LSTM network model with an error margin of 13% -16% (Namin & Namin, 2018).

In recent years, it has been suggested that hybrid methods created especially by combining the benefits of models give much better results compared to the methods used alone. Some of the studies on hybrid models that combine the advantages of two or more individual models in the literature can be summarized as below.

Wedding and Cios (1996) presented a methodology integrated with radial-based function networks and Box-Jenkins ARIMA models (Wedding & Cios, 1996). Luxhoj et al. (1996) estimated the 24-month total sales of a firm using the hybrid econometric neural network approach (Luxhoj, Riis, & Stensballe, 1996). Tsaih et al. (1998) combined rule-based systems and hybrid artificial intelligence methods to predict the daily price change direction of the S&P 500 Stock Index (Tsaih, Hsu, & Lai, 1998).

Tseng et al. (2002) proposed a hybrid model consisting of a combination of seasonal ARIMA and a neural network back propagation algorithm (BP) known as SARIMABP (Tseng, Yu, & Tzeng, 2002). In his study of time-series forecasting using the ARIMA and ANN hybrid model, Zhang (2003) used Wolf's sunspot, Canadian lynx and British pound / US dollar exchange rate data. Zhang stated that neither ARIMA nor ANN was suitable for all real-time series with the hybrid model developed in this study and that there were linear and nonlinear correlation structures between observations in these series and therefore a hybrid model should be used to estimate both linear and nonlinear components of a time-series (Zhang G. , 2003).

Armano et al. (2005) proposed a hybrid model approach based on ANN combined with a genetic algorithm in stock market prediction (Armano, Marchesi, & Murru, 2005). Yu et al. (2005) presented a new nonlinear hybrid prediction model by combining ANN and generalized linear auto regression (GLAR) to obtain more accurate prediction results and improve prediction performance (Yu, Wang, & Lai, 2005). Kim and Shin (2007) investigated the effectiveness of a genetic algorithm-based hybrid model based on ANN, such as adaptive time-delayed neural networks (ATNN) and time-delayed neural networks (TDNN), to determine transient patterns in stock market prediction (Kim & Shin, 2007).

Zhou and Hu (2008) proposed the hybrid model estimation approach based on Gray and Box-Jenkins ARMA methods (Zhou & Hu, 2008). Khashei et al. (2008), in their work using

ANN and fuzzy regression methods, proposed a hybrid model that provides more accurate results with missing data sets. In the proposed model, the advantages of ANN and fuzzy regression were combined to overcome the limitations in both ANN and fuzzy regression. To show the suitability and effectiveness of the method, it was used in the estimation of the price of gold (Gram/US \$) and exchange rates (US/Iranian Rial). The results showed that the proposed model could be an effective way to improve estimation accuracy (Khashei, Hejazi, & Bijari, 2008). Aladağ et al. proposed a new hybrid approach with the Elman's recurrent neural networks (RNN) and seasonal ARIMA (SARIMA) models. In this proposed hybrid model, Canadian lynx data for the period 1821-1934 were used, which consisted of the annual number of lynx traps in the Mackenzie River area in Northwest Canada. Although the data used were minimal, the hybrid method gave the best estimation accuracy in the application results (Aladağ, Eğrioglu, & Kadilar, 2009). Koutroumanidis et al. (2009), in their studies to examine the role of forests in firewood production in Greece, predicted the future situation of the sale prices of wood produced by the Greek state forest farms. They used ARIMA, ANN and Hybrid models for forecasting and obtained the best estimate results using the ARIMA-ANN hybrid model (Koutroumanidis, Ioannoub, & Arabatzis, 2009).

Koutroumanidis et al. (2011) aimed to establish confidence intervals for predicted values of a time-series in their studies for predicting stock market prices with a hybrid method. The daily closing prices of the shares of Alpha Bank from 28/01/2004 to 30/11/2005 were used as the sample of the study. For the estimation, ANN was applied to the raw data and then the market prices were estimated using the Bootstrap method. Estimation accuracy was measured by using different criteria, and satisfactory results were obtained (Koutroumanidis, Ioannou, & Zafeiriou, 2011). He and Deng (2012) developed a hybrid model using ARIMA and ANN to estimate air pollutant factors in their study. Firstly, ARIMA and ANN were used to estimate the time-series and then a re-estimation was done with the hybrid model that was developed. When the results were compared, it was evident that the hybrid model performed better (He & Deng, 2012).

Babu and Reddy (2014) investigated the nature of volatility by using experimental and simulated data sets such as sunspot, electricity price and stock market data. They first used moving average filter, and later ARIMA and ANN models were applied. A hybrid model was proposed along with ARIMA and ANN models used in the application and some existing HYBRID ARIMA-ANN models. The results from the data sets showed that the hybrid model had a higher estimation accuracy for both single-step and multi-step predictions (Babu & Reddy, 2014). Hocaoglu et al. (2015) used the hybrid model they created using ANN and regression methods for load estimation in the energy sector. When the error results were compared, it was concluded that the hybrid system had the lowest errors (Hocaoglu, Kaysal, & Kaysal, 2015).

In Ateşongun's (2015) study, using ARIMA, ANN and the hybrid model created by these two models, he made predictions on the number of wild cats in Canada, sunspots, airline passenger numbers, Australia's New South Wales region's hourly electricity prices and wheat

efficiency in Turkey. The results reached were compared both with the methods used and with the previous studies (Ateşoğlu, 2015).

Pablo et al. (2016) proposed a hybrid approach to the reconstruction of the time-series with the creation of ANN and Monte Carlo Simulation. With these models, they tried to estimate the daily milk sales of a dairy company. The results showed that the proposed method could reconstruct the past and predict the future from the known time-series segment (Pablo, ve diğerleri, 2016). Sugiartawan et al. (2017) used a hybrid model that they created with wavelet transform and LSTM to predict the number of tourists coming to Indonesia over a monthly period. The prediction results of the proposed hybrid model were compared with other RNN algorithms, namely ELMAN RNN and Jordan RNN and the hybrid of Elman's wavelet and the hybrid of Jordan's wavelet. They stated that the hybrid model created by the wavelet transform and ANN could be used to predict the trend evolution with the original ANN, which gives Elman and Jordan better training time than RNNs and could predict the number of incoming tourists more accurately than other hybrid methods (Lin, Guo, & Aberer, 2017).

Yu et al. (2018) reached the conclusion of Convolutional Neural Networks based on deep learning to estimate second-hand housing prices in Beijing (Sugiartawan, Pulungan, & Sari, 2017). Lin et al. (2017), inspired by the recent successes of artificial neural networks, proposed TreNet, a completely new hybrid neural network, to predict the trend of time-series. They used three different data sets in their study including electricity consumption, chemical sensor records subject to dynamic gas mixtures at variable concentrations, and daily stock transaction information on Yahoo Finance and the New York Stock Exchange. As a result of the study, they used TreNet's time-series (Convolution Neural Network-CNN) and LSTM models and the Autoregressive Moving Average (ARMA) model from the time-series. The results obtained from CNN, LSTM and ARMA models were compared. They applied a logical regression model to compare the three models used. They concluded that the prediction accuracy of the LSTM, which takes the time-series into account, was better than other methods (Yu, Jiao, Xin, Wang, & Wang, 2018).

Temür et al. used monthly house sales data from past periods as a time-series in their studies for the prediction of house sales in Turkey. In their study with ARIMA, LSTM and the Hybrid model obtained from these two models, they obtained the best estimate result from the Hybrid model (Soy Temür, Akgün, & Temür, 2019).

Methodology

In this part of the study, information about ARIMA, LSTM, which are used in the prediction of sales quantities, and the basic principles and modeling processes of the hybrid model obtained by rational combining of these models are given.

ARIMA

ARIMA is a statistical analysis, it uses time series data. The ARIMA predicts future values by examining the differences between values in the time series. This method is a univariate model developed by George Box and Gwilym Jenkins in 1970 and is also known as a method that gives very successful results in short-term forecasts. One of the critical assumptions of the method is that the applied series is a discrete and stationary series consisting of observation values obtained with equal time intervals. Stationariness means that the series is free from periodic fluctuations (Bircan & Karagöz, 2003). Stability; It requires basic statistical properties such as mean, variance, covariance, or autocorrelation to be constant over time periods (Babu & Reddy, 2014).

An ARIMA model consists of 3 components Auto regression (AR), Integrated (I), and Moving average (MA). Each component is a parameter. To represent these parameters, ARIMA models use a standard notation p , d , and q . This standard notation indicates the type of ARIMA model used (Fattah, et al., 2018).

In the determination process of the model, first of all, by applying the stationarity and seasonality analysis of the series, it is determined whether the series is stationary or not. If the series are not static, this process is continued by applying the d th degree differencing process until a static series is reached. D , which shows the degree of differencing; gets the value of 0 in static series, 1 in series that become static with first-degree differencing, and 2 in series that are made static with second-degree differencing. (d) is a positive integer, and the value of d is usually taken as 1 or 2 in practice. However, if the series are static, there is no need to apply any differencing to the series (Smith, 2018). If the differencing operation is performed (d) times, the integration parameter of the ARIMA model is set to (d). Then, identification is performed on the stationary data obtained. In this process, the parameters of the autoregressive (AR) and moving average (MA) operations shown in equation (1) are determined as (p) and (q), respectively (Newbold, 1983).

p : degree of autoregressive model (AR)

d : differencing degree

q : degree of moving average model (MA)

$$y_t = \alpha_1 w_{t-1} + \alpha_2 w_{t-2} + \dots + \alpha_p w_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (1)$$

For the t time here, y_t denotes the linearized real data, w_t weight values, while ε_t denotes the moving average error. As shown in the formula, a linear relationship was established between actual data y_t to be predicted, the observed (p) data ($y_{t-1}, y_{t-2}, \dots, y_{t-p}$), and (q) error data ($\varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-q}$). Differencing removes the changes in the level of a time series, eliminating trend and seasonality and consequently stabilizing the mean of the time series.

In order to difference the data, the difference between consecutive observations is computed. The linearized real data y_t is replace with its lagged value Δy_t and mathematically, this is shown as:

$$\Delta y_t = y_t - y_{t-d} \quad (2)$$

Δ : It refers to the difference operator series.

While the autocorrelation values of alternative models such as ARIMA decrease geometrically or exponentially, autoregressive fractionally integrated moving average models (ARFIMA OR FARIMA) have hyperbolic decreasing autocorrelation values (Bhardwaj, et al., 2020). In the financial time series within this structure, the existence of long memory is talked about. Long memory models are expressed as FI processes, and d , which has a degree of integration, technically takes a fractional value (Karia, et al., 2013) (Durham, et al., 2019) (Bukhari, et al., 2020).

LSTM

LSTMs are a particular type of RNN designed to learn long-term dependencies. They were first developed by Hochreiter and Schmidhuber (1997). It has a complex structure called the LSTM unit in the hidden layer it contains. A simple representation of this structure is given in Figure 1. Since they work very well on a wide variety of problems, they are widely used today.

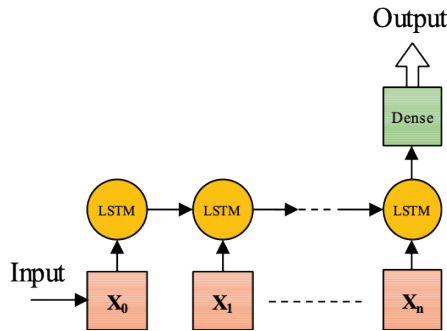


Figure 1. LSTM Architecture (Atienza, Posted on 2017)

To roughly describe, in an LSTM structure, there is also a memory along with the RNN cell. Thanks to this memory, information from the previous time can be retrieved and transmitted to the next one. The model decides which information to take with training. Remembering information for a long time is in practice the default behavior of these networks and not something they try to learn. An LSTM unit is shown in Figure 2.

Updating weights to optimize models of deep learning algorithm, and thus transmitting the entire data set over a single network many times to obtain a better and more accurate prediction model makes sense. However, it is not clear how many epoch numbers will be needed to achieve optimal weights and to train a model with the same data set. Replication (epoch) specifies the total number of recurrences that a particular dataset is used for educational purposes to calculate the network's weight values (w). An epoch signifies that an entire dataset had passed through the network once before and then returned over the network. Different sets of data exhibit different behaviors, so a different number of epochs may be needed to best train networks. Based on this observation, many prediction studies were carried out with different epoch values in the study, and all trial results are presented in the application section.

HYBRID

While the ARIMA models are good at modeling the linear relationship in the time-series, they are insufficient at modeling nonlinear relationships. The LSTM models can model both linear and nonlinear relationships but cannot provide the same results for each data set. Therefore, to reach the best prediction results, hybrid models based on the principle of separate modeling of linear and nonlinear components of time-series are employed. These models, with great success in prediction work for time-series analysis, use multiple learning algorithms to achieve better predictive performance than founder learning algorithms (Opitz & Maclin, 1999). The purpose of these models is to increase the diversity of models and to achieve better results (Adeva, Beresi, & Calvo, 2005) (Oliveira & Torgo, 2014).

When the results obtained with hybrid models and the results obtained by using the models individually were compared, it was observed that they could reduce the general variance or error, even if they are unrelated (Khashei M., 2008). Due to this reason, hybrid models are recognized as the most successful models for forecasting tasks.

This study is based on the hybrid model that Zhang (2003) suggested for a hybrid ARIMA-YSA model for the prediction of time series. This model is based on the assumption that any time series is a mathematical sum of two linear and non-linear model components (Zhang, 2003). Accordingly, the ARIMA-LSTM hybrid model that was created to make predictions by using the historical data of a time series and given a flow diagram in Figure 3, is as follows.

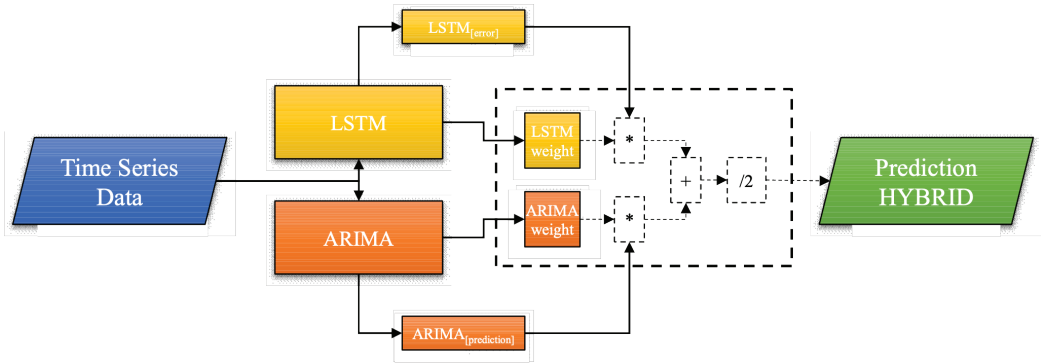


Figure 3. Hybrid Model Diagram

The time-series prediction formula of the generated model can generally be expressed as the sum of linear and nonlinear components, as shown in equation (9).

$$y_t = L_t + N_t \quad (9)$$

L_t shows the linear component of the time-series, whereas N_t shows the non-linear component. In the hybrid model, the linear component of the time-series L_t is first estimated using the ARIMA model, N_t then estimated by the LSTM model. Then, error values of both models are calculated. The formula for this calculation is given in Equations (10) and (11).

$$lstm_{error} = lstm_average[error] \quad (10)$$

$$arima_{error} = arima_average[error] \quad (11)$$

The weights of the models are calculated by using the obtained error values in Equations (12) and (13). Normalization process was carried out in calculating the weight values.

$$lstm_{weight} = \left(1 - \left(\frac{lstm_{error}}{lstm_{error} + arima_{error}} \right) \right) * 2 \quad (12)$$

$$arima_{weight} = 2 - lstm_{weight} \quad (13)$$

The weight values of the models and finally each prediction value of our hybrid model are obtained with the given equation (14).

$$hibrit_{predict}[i] = (lstm_{weight}[i] * lstm_{error}[i] + arima_{weight} * arima_{predict}[i])/2 \quad (14)$$

Success Criteria of Models

No matter which of the prediction methods is used, they do not produce 100% accurate value. In fact, if the future was known 100%, then it would not be a prediction. Therefore, each prediction has a certain error rate. It is one of the most widely accepted criteria in the process of choosing one of the various prediction models for the model to be compatible with the data and high prediction success. The essential criteria used to compare the predictive success of the models is the accuracy of the prediction. This is measured by analyzing the estimated errors (Sarı, 2016).

To measure the model prediction success of the three methods discussed in the study *MSE*, *RMSE*, *MAPE* and *MAE* criteria were used. The formulas used to calculate these criteria are given in Table 1 (Sallehuddin, Shamsuddin, Hashim, & Abraham, 2007)

Table 1
Success Criterion (Correlogram) Formulas

Name	Formulas
MSE	$MSE = \frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2$
RMSE	$\sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2}$
MAPE	$\left(\sum_{t=1}^n \left \frac{y_t - \hat{y}_t}{n} \right \right)$
MAE	$\left(\sum_{t=1}^n \left \frac{y_t - \hat{y}_t}{y_t} \right \right) \frac{100}{n}$

Here y_t represents the real value, \hat{y}_t predicted value, n represents the number of predicted periods. The model with the smallest criteria values obtained with the above formulas should be chosen as the most suitable model.

The Data Set and Methods of the Study

The data used in this study are product sales data of a business operating in the construction industry and are derived from the author's previously published doctoral thesis. Company and product information are reserved for competition elements and at the request of the company authorities. Moreover, the product quantity data have been changed proportionally for privacy. The data were provided by interviews with the sales manager and accounting manager. In 2018, the company suspended production due to the robotization of the production system. For this reason, access to data up to 2017 has been provided. The reason for consider-

ring the year 2010 as the starting year is derived from the fact that previous data were stored in the form of 3-month periods instead of monthly.

Within the scope of the study, the sales volumes of a product used in the construction industry for 2010 (1)–2017 (12) months were provided, and a total of the 96-month data set was used. Since the data is numerically significant, it was processed with models by reducing at the rate of $\times (1000)$. The related data set was represented in a series to be used in the neural network model. The graph of the data represented by years is given in Figure 4.

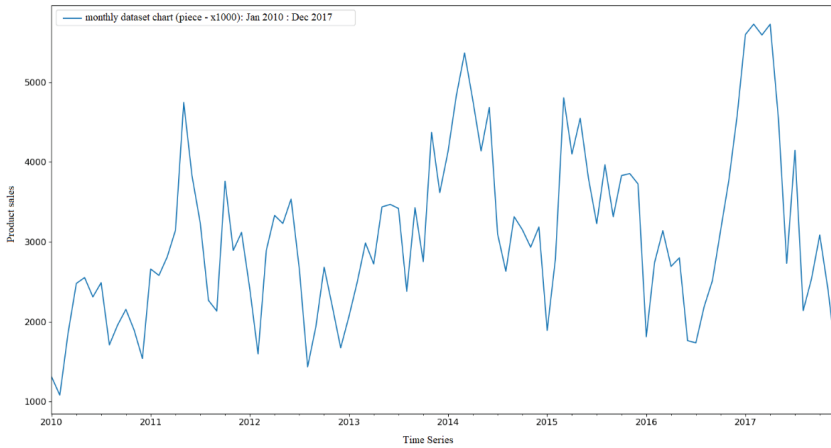


Figure 4. Monthly Product Sales Data Set Chart

A data set in forecasting applications is divided into training and test data. The amount of data that should be for each set is the main problem encountered in forecasting applications, and artificial neural networks should be trained to use as much data as possible. The data not used in the training set are used in the test set. The output obtained by feeding the test data to the network is compared with the actual output values. The main objective is to check whether the power to represent the sample is sufficient.

In the literature, it is generally accepted that the data set is divided into 70% or 80% training, 30% or 20% test in determining the training and test data (Prieto, et al., 2019) (Vasques, et al., 2016) (Brownlee, 2020) (Zou & Qu, 2020). In this study, to obtain the most relevant results from the data set, training rates in different combinations were tried. These combinations were 67% training, 33% testing and 73% training, 27% testing. The reason for determining these rates differently from the generally accepted rates was due to the determination of the most appropriate and inappropriate rates by trying many training rates (Brownlee, 2020).

The data were applied to the ARIMA, one of the linear methods commonly used in time-series predictions, and the LSTM network, one of the nonlinear methods recently used for deep learning algorithms. Besides, hybrid methods generally accepted in the literature were examined, and a hybrid method based on ARIMA and LSTM was tested for this study. All methods were coded in Python 3.6 6 (Rossum, 2001) programming language in computer environment, and ready-made libraries were used as open-source code.

Analysis and Findings

In this part of our study, the information about the data used in our applications and the findings obtained from the results of the application are given.

ARIMA

The ARIMA (p, d, q) models were determined by using p [0, 1, 2, 3], d [0, 1, 2], q [0, 1, 2] values for the sales data estimation performed with the ARIMA estimation method, and Correlogram examinations of each of these models were performed separately. The results of the success criteria and model error comparisons of the estimation studies are given in Table 2, and the significance tests were performed both on this table and the output graphics formed by the models. According to the model results in the tables, the model with a lower percentage error rate is accepted as the model with a higher performance in estimation. Accordingly, as a result of the tests, ARIMA (1,0,0) and ARIMA (1,0,1) models with the lowest MAPE values of the study, the training data set of which was 67%, and ARIMA (1,0,0) and ARIMA (1,0,1) models with the lowest MAPE values of the study, the training data set of which was 73%, made the most successful predictions. Among these predictions, 67% of the study-test data set produced lower MAPE values showing that it performed more valuable estimates.

Table 2

ARIMA Models Correlogram Results (67% Training Set)

Comparison of ARIMA Models				
ARIMA	MAE	MSE	RMSE	MAPE
(0,0,1)	775.124	934643.842	966.770	0.255
(0,0,2)	751.472	850727.707	922.349	0.252
(0,1,0)	692.830	766967.805	875.767	0.253
(1,0,0)	672.267	681849.325	825.742	0.234
(1,0,1)	675.996	690921.199	831.217	0.235
(1,0,2)	684.416	707584.581	841.180	0.238
(1,1,0)	688.410	757390.500	870.282	0.251
(2,1,0)	676.638	784725.580	885.847	0.247
(2,2,0)	775.778	935149.511	967.031	0.278

Table 3
 ARIMA Models Correlogram Results (%73 Training Set)

Comparison of ARIMA Models				
ARIMA	MAE	MSE	RMSE	MAPE
(0,0,1)	837.842	1048519.893	1023.973	0.281
(0,0,2)	802.764	944684.711	971.949	0.276
(0,1,0)	699.415	828872.420	910.424	0.266
(1,0,0)	710.461	760316.150	871.961	0.254
(1,0,1)	716.008	773831.273	879.677	0.256
(1,0,2)	723.226	792962.528	890.484	0.259
(1,1,0)	706.616	830713.220	911.435	0.268
(2,1,0)	688.728	858269.317	926.428	0.262
(2,2,0)	781.374	967890.861	983.814	0.293

Graphs for comparing the 31-month forecast and 26-month forecast data obtained with the best estimates with the lowest MAPE values and the 26-month forecast data with the actual test data are given in Figure 5 and Figure 6.

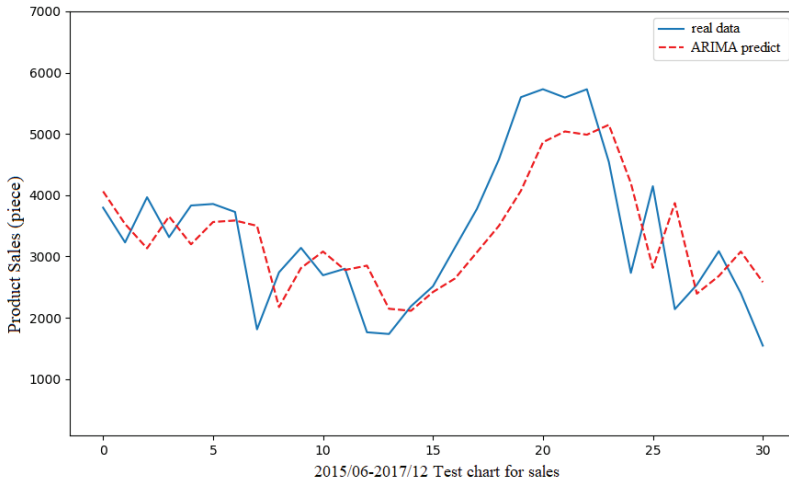


Figure 5. ARIMA (1,0,0) Model and Real Data Set Comparison Chart

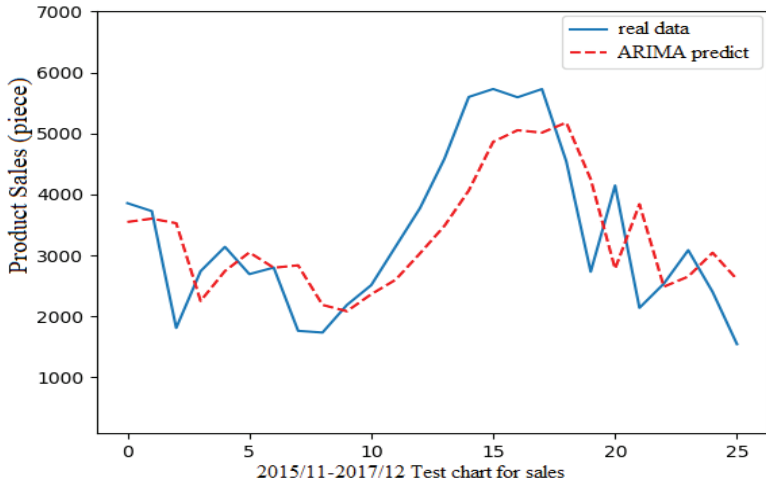


Figure 6. ARIMA (1,0,1) Model and Real Data Set Comparison Chart

When Figure 5 and Figure 6 are analyzed, it is seen that the monthly sales values obtained from the real data and the estimation studies are in overlapping structure and the deviations between them are not excessive. The success seen in the graphics can be observed more clearly than the similarity of the direction breaks as well as the proximity values of the data. The models used here can produce values very close to real data, with an error value of 0.234 and 0.235 MAPE, respectively. This situation indicates the success of the model used in the study.

LSTM Analysis Results

In this part of the study, time-series based Long Short-Term Memory (LSTM) architecture was used. This part coded in Python software was realized with KERAS, a deep learning library used in LSTM model development (Chollet, 2015).

The LSTM network used was run on the monthly sales of the products listed between 2010 and 2017. In this model, as in the ARIMA model, estimates were made by using training combinations of 67% and 73%. The data not used in training were used for testing. Different epoch numbers were tried during the training process and the resulting error values were examined. Error values resulting from epoch numbers according to the training combinations performed are given in Table 4 and Table 5, respectively.

Table 4
LSTM Models Correlogram Results (%67 Training Set)

Comparison of LSTM Models				
Epoch	MAE	MSE	RMSE	MAPE
1000	876.925	1086744.667	1042.470	0.309
1500	1055.303	1543692.874	1242.454	0.364
2000	1223.365	2021929.469	1421.946	0.406
2500	1332.994	2450055.299	1565.265	0.445

Table 5
LSTM Models Correlogram Results (%73 Training Set)

Comparison of LSTM Models				
Epoch	MAE	MSE	RMSE	MAPE
1000	687.389	661910.062	813.579	0.245
1500	795.090	954889.250	977.184	0.275
2000	900.901	1261486.750	1123.159	0.297
2500	1027.180	1895367.375	1376.723	0.318

In applications performed with both 67% training sets and 73% training sets, as shown in Tables 4 and 5, the results produced by models trained with 1000 epoch have the lowest error rate. The graphical representations of the prediction results generated by the network models trained with both training sets are as in Figure 7 and Figure 8.

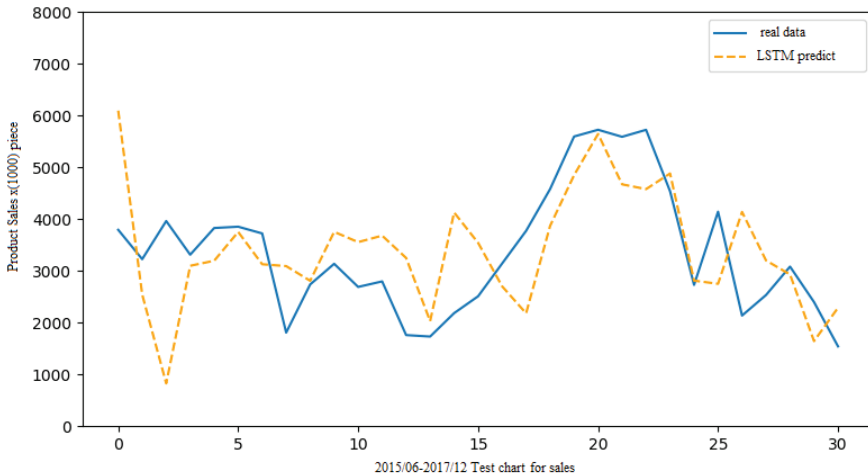


Figure 7. LSTM (1000 epoch) and Real Data Set Comparison Chart (%67 Training Set)

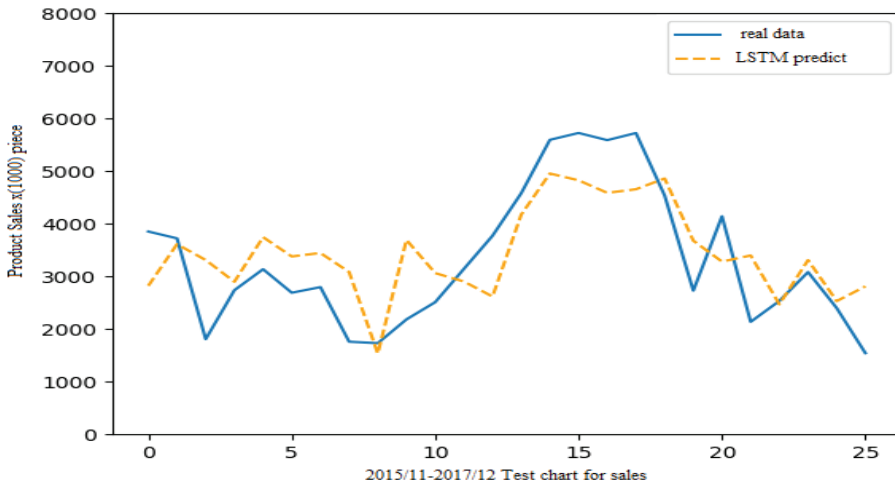


Figure 8. LSTM (1000 epoch) and Real Data Set Comparison Chart (%73 Training Set)

It is possible to obtain more successful results with the use of more data in the LSTM model (Liu, Tan, & Wang, 2019). Nevertheless, as seen in the figures, LSTM models have managed to produce successful results with little data usage. It produced a lower error value compared to the ARIMA model, with a 73% training data set mapping with a 0.245 MAPE value, which is the success criterion assessment. On the other hand, the ARIMA model produced a lower error value in comparison of 67% training data set. As a result of this comparison, it can be said that the ARIMA model for 67% training sets and the LSTM model for 73% training sets were more successful than the other model. The state of having achieved more successful results with the 73% training kit compared to the 67% training kit is proof that higher performance values can be reached due to the increase of the number of data for the LSTM model. However, there are also studies in the literature that prove that success falls after a certain set of data (Liu, Tan, & Wang, 2019) (Liu, Tan, & Wang, 2019).

Hybrid Analysis Results

In this part of the study, estimation experiments were carried out using the hybrid model designed in Figure 3 and different percentile education sets as in other models. All parameters in ARIMA and LSTM models are used for the Hybrid model. A summary of the best estimation results obtained with the hybrid model is given in Table 6.

Table 6
 Hybrid Model Correlogram Results (%67 Training Set)

Comparison of Hybrid Method Results					
LSTM epoch	ARIMA (p, d, q)	MAE	MSE	RMSE	MAPE
1000	(0, 1, 0)	474.901	539710.087	734.650	0.173
	(0, 1, 1)	473.856	505515.043	710.996	0.176
	(1, 1, 1)	547.064	550803.800	742.162	0.203
	(1, 2, 0)	454.536	574853.193	758.191	0.166
	(2, 1, 0)	471.357	523651.351	723.638	0.175
	(2, 2, 0)	455.894	562571.828	750.048	0.167
1500	(0, 1, 0)	441.230	517109.254	719.103	0.158
	(0, 1, 1)	439.884	483443.968	695.301	0.162
	(1, 1, 1)	519.856	520319.460	721.332	0.192
	(1, 2, 0)	425.395	576725.599	759.425	0.151
	(2, 1, 0)	439.576	500267.806	707.296	0.161
	(2, 2, 0)	440.020	582412.655	763.160	0.156
2000	(0, 1, 0)	435.374	484493.392	696.056	0.151
	(0, 1, 1)	438.542	462836.972	680.321	0.157
	(1, 1, 1)	525.968	501871.120	708.429	0.190
	(1, 2, 0)	406.268	556548.101	746.022	0.143
	(2, 1, 0)	436.966	472535.368	687.412	0.156
	(2, 2, 0)	427.079	572413.420	756.580	0.150
2500	(0, 1, 0)	437.287	497100.030	705.053	0.155
	(0, 1, 1)	439.369	474126.965	688.569	0.161
	(1, 1, 1)	538.044	529362.677	727.573	0.198
	(1, 2, 0)	418.051	596092.088	772.070	0.148
	(2, 1, 0)	436.265	481295.563	693.755	0.159
	(2, 2, 0)	441.840	614257.378	783.746	0.157

When the values obtained with the 67% training set in the table are examined, it is seen that much more successful results were achieved compared to the results obtained from both the LSTM network and ARIMA models. The most successful estimation in this training set is the 0.1143MAPE value study produced by the LSTM (2000 epoch) and the ARIMA (1,2,0) model. It can be stated that this error value is much lower than the error values produced by other singular models, and a very successful result was obtained. This situation is an indication that, as mentioned in the examples in the literature, hybrid models give better results than single models and can be used to achieve successful results. The estimation results of the LSTM (2000 epoch) -ARIMA (1,2,0) model were determined as the best estimate and the output graph of the real values are given in Figure 9.

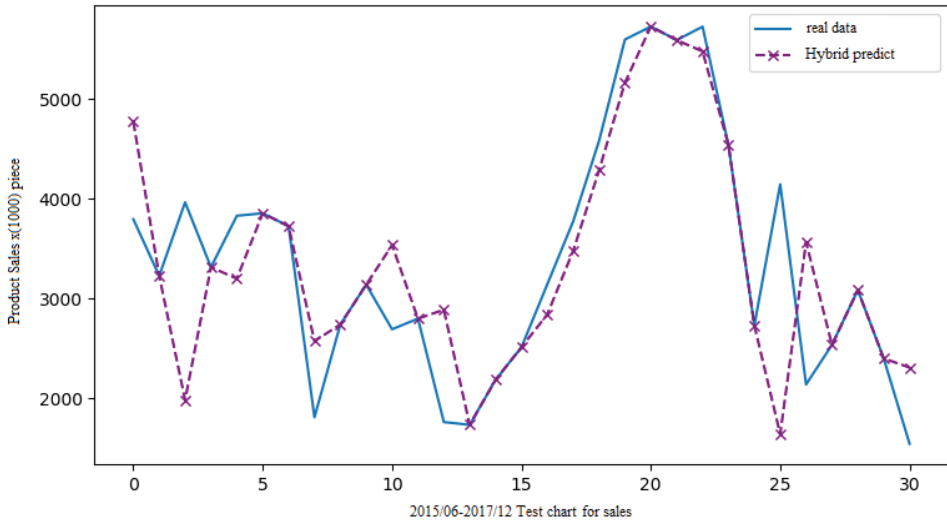


Figure 9. Hybrid Model and Ground Truth Comparison Chart (%67 Training Set)

As can be seen from the figure, the estimate of the model coincides with the real data graph. It is an indication that the prediction is very successful. Especially when looking at trend breakouts, it can be observed that the forecast graph behaves precisely like the real graph, except for minor differences. On the other hand, to improve the model used, the training set was increased to 73% and different attempts were made with LSTM and ARIMA models. A summary of the results obtained in the estimations made is given in Table 7.

Table 7
Hybrid Model Correlogram Results (%73 Training Set)

Comparison of Hybrid Method Results					
LSTM epoch	ARIMA (p, d, q)	MAE	MSE	RMSE	MAPE
1000	(0, 1, 0)	459.247	491694.830	701.210	0.185
	(1, 2, 0)	431.770	444982.797	667.070	0.180
	(2, 1, 0)	473.266	515363.694	717.888	0.188
1500	(0, 1, 0)	461.439	481297.330	693.756	0.181
	(1, 2, 0)	421.336	447043.373	668.613	0.171
	(2, 1, 0)	471.049	503728.652	709.738	0.182
2000	(0, 1, 0)	442.282	459427.418	677.811	0.172
	(1, 2, 0)	409.096	463999.091	681.175	0.165
	(2, 1, 0)	446.925	471641.035	686.761	0.170
2500	(0, 1, 0)	413.791	484388.432	695.980	0.157
	(1, 2, 0)	360.368	474134.222	688.574	0.141
	(2, 1, 0)	432.843	509609.739	713.870	0.161
3000	(0, 1, 0)	408.578	533826.667	730.634	0.155
	(1, 2, 0)	344.138	480333.312	693.061	0.135
	(2, 1, 0)	445.409	578365.642	760.504	0.166

3500	(0, 1, 0)	407.097	537043.602	732.833	0.154
	(1, 2, 0)	338.184	470006.272	685.570	0.133
	(2, 1, 0)	446.167	585519.850	765.193	0.166
4000	(0, 1, 0)	416.356	531732.634	729.200	0.156
	(1, 2, 0)	342.317	452430.129	672.629	0.134
	(2, 1, 0)	447.486	581406.834	762.500	0.163
4500	(0, 1, 0)	408.811	538732.181	733.984	0.157
	(1, 2, 0)	339.349	463405.842	680.739	0.136
	(2, 1, 0)	448.035	588540.125	767.164	0.169

According to the values obtained in the table, it is seen that the error value generated by the hybrid model created by the LSTM 3500 epoch and ARIMA (1,2,0) models is 0.133 MAPE. This shows that a very successful result was achieved compared to the model created with the 67% training set. The graphic representation of this model is as shown in Figure 10.

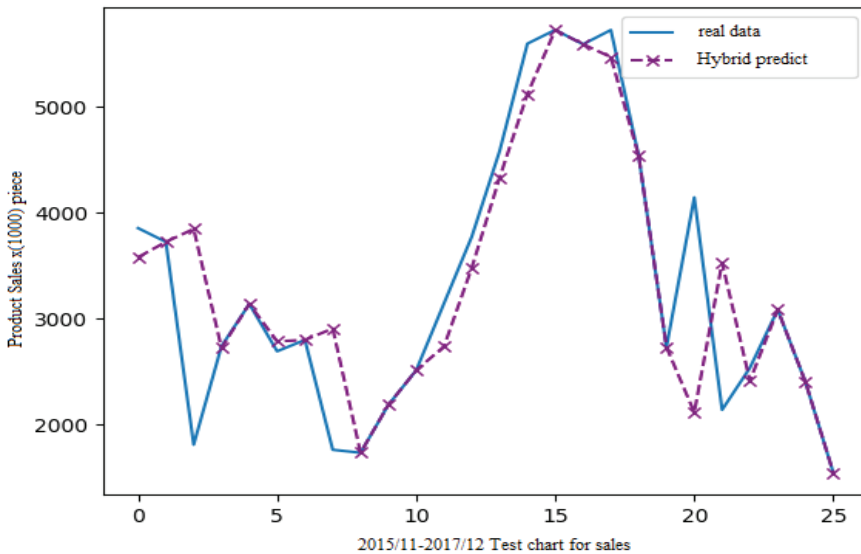


Figure 10. Hybrid Model and Ground Truth Comparison Chart (73% Training Set)

Conclusion

Nowadays, businesses, institutions and nations which want to continue and develop their lives according to competitive conditions and rapidly changing technological developments have to make a future-oriented plan. As stated before, sales quantity budgets formed by sales forecasts should be prepared correctly and reliably, as they are the basis for the preparation of other budgets.

Monthly sales data relating to the years 2010-2017 were used for demand estimation purposes. In sales quantity estimates, ARIMA, LSTM, and a hybrid method made out of these

two models were made use of. As a result of the models, the sales quantity budget was created for 2017, and the comparison was made. To estimate the models and compare these estimates, 67% of the data set were reserved for training of the models, the remaining 33% were used as test data. On the other hand, as a second case, experiments were carried out with a training set of 73% and test data of 27%. After the data were arranged, the test data were estimated by training with ARIMA, LSTM and Hybrid model. Based on the estimation results, all of the models were compared.

Table 8
Comparison of ARIMA, LSTM and Hybrid Models

Comparison of All Models					
Training Set (%)	MODEL	MAE	MSE	RMSE	MAPE
67	ARIMA(1,0,0)	672.267	681849.325	825.742	0.234
	LSTM (1000 epoch)	876.925	1086744.667	1042.470	0.309
	HYBRID (LSTM 2000 epoch-ARIMA(1,2,0))	406.268	556548.101	746.022	0.143
73	ARIMA(1,0,0)	710.461	760316.150	871.961	0.254
	LSTM (1000 epoch)	687.389	661910.062	813.579	0.245
	HYBRID (LSTM 3500 epoch-ARIMA(1,2,0))	338.184	470006.272	685.570	0.133

As seen in Table 8, the RMSE values for the 67% training set of ARIMA, LSTM and hybrid models were calculated as 825.742, 1042.470 and 746.022 respectively, while it was calculated as 871.961, 813.579, 685.570 in the 73% training set. Considering the RMSE values, it is seen that the hybrid model gave better results in both training sets. In terms of MAPE values, the value of 0.133 for hybrid model, 0.254 for ARIMA and 0.245 for LSTM model was calculated in the 73% training set. This proves that, in terms of MAPE values, the hybrid model provided better results for both training sets compared to the other two models.

According to the RMSE values, the most successful model was the hybrid method, and in comparison, to the ARIMA estimation results, it gave better performances of about 16.9% and 15.7% compared to the LSTM estimation results. With the hybrid model which gave more successful results in terms of MAPE success criterion, approximately 43.2% better results were obtained from the ARIMA and 45.7% better than the LSTM. Both this study and other studies conducted in the literature clearly show that making predictions by combining multiple methods that can model different functional relationships in the data set rather than estimating time-series with a single method gives more effective results. The findings obtained as a result of the study also confirm this. Moreover, when the obtained results were compared with the real sales data of the company, it was concluded that the values closest to the actual data were obtained from the hybrid method.

When the MSE values were taken into consideration, the hybrid model achieved a performance increase of approximately 37.4% according to the estimations made with the ARIMA

model and 62.9% according to the estimates made with the LSTM model. Likewise, in terms of MAE success criterion, the Hybrid model was approximately 40.5% better than ARIMA model and 52% better than LSTM model. Both this study and other studies conducted in the literature clearly show that making predictions by combining multiple methods that can model different functional relationships in the data set rather than estimating time-series with a single method, gives more effective results. The findings obtained as a result of the study also confirm this.

In the study, the monthly sales quantity data required to create the sales quantity budget of an enterprise was estimated by using the methods mentioned above. The data used in the estimation phase was not limited to making a single estimate by entering the program, but it was repeatedly tried dozens of times until fair values were obtained. The purpose of not limiting to a single model, using different models and hybrid models created with these different models was to determine the configuration that gives the lowest RMSE and MAE values. For this purpose, ARIMA, LSTM and a hybrid model formed from these two models were used as a model because they are widely used in the prediction studies on social sciences today. The hybrid model provided the best performance among these three models. As can be seen above, the error values found for each method are quite low.

Within the scope of the study, the sales quantity budget data of the enterprise for 2017 are given in Table 9. The graphical representation of these data is as given in Figure 11. While creating the table, sales quantity budget data are presented as the values obtained as a result of a particular mathematical operations due to data privacy. Since the same formula was applied to all data, except for the appreciable width in the y axis, there will be no trend direction difference between the graphical representation of the actual data and the graphical representation of the modified data.

Table 9
2017 (Monthly) Sales Volume Budget According to the Methods Used *

2017	ARIMA (1,0,0)		ARIMA (1,0,1)		ARIMA (1,0,0)		ARIMA (1,0,1)		LSTM 1000 epoch %67 Training Set		LSTM 1000 epoch %73 Training Set		Hybrid (LSTM 2000 epoch-ARIMA (1,2,0) %67 Training Set		Hybrid (LSTM 3500 epoch-ARIMA (1,2,0) %73 Training Set	
	%67 Training Set	%73 Training Set	%67 Training Set	%73 Training Set	%67 Training Set	%73 Training Set	%67 Training Set	%73 Training Set								
January	1506.6	1504.0	1506.6	1504.0	1506.6	1504.0	1506.6	1504.0	1433.0	1545.6	1931.6	1895.7	1433.0	1545.6	1931.6	1895.7
February	1799.7	1799.3	1799.7	1799.3	1799.7	1799.3	1799.7	1799.3	1795.6	1834.8	2175.4	2119.7	1795.6	1834.8	2175.4	2119.7
March	1867.8	1869.6	1867.8	1869.6	1867.8	1869.6	1867.8	1869.6	2089.7	1787.0	2069.4	2069.4	2089.7	1787.0	2069.4	2069.4
April	1845.2	1854.8	1845.2	1854.8	1845.2	1854.8	1845.2	1854.8	1729.9	1698.9	2026.6	2024.8	1729.9	1698.9	2026.6	2024.8
May	1904.8	1916.2	1904.8	1916.2	1904.8	1916.2	1904.8	1916.2	1695.6	1724.1	1680.2	1680.2	1695.6	1724.1	1680.2	1680.2
June	1553.2	1572.3	1553.2	1572.3	1553.2	1572.3	1553.2	1572.3	1807.4	1800.0	1011.2	1011.2	1807.4	1800.0	1011.2	1011.2
July	1041.4	1029.3	1041.4	1029.3	1041.4	1029.3	1041.4	1029.3	1041.7	1362.9	811.7	783.3	1041.7	1362.9	811.7	783.3
August	1432.2	1420.9	1432.2	1420.9	1432.2	1420.9	1432.2	1420.9	1019.2	1215.5	1480.3	1306.5	1019.2	1215.5	1480.3	1306.5
September	885.8	920.4	885.8	920.4	885.8	920.4	885.8	920.4	1532.7	1258.8	818.6	893.7	1532.7	1258.8	818.6	893.7
October	992.2	981.9	992.2	981.9	992.2	981.9	992.2	981.9	1185.9	915.4	1142.2	1142.2	1185.9	915.4	1142.2	1142.2
November	1139.6	1126.6	1139.6	1126.6	1139.6	1126.6	1139.6	1126.6	1085.8	1226.1	888.7	888.7	1085.8	1226.1	888.7	888.7
December	955.6	964.8	955.6	964.8	955.6	964.8	955.6	964.8	609.8	938.1	800.8	572.0	609.8	938.1	800.8	572.0

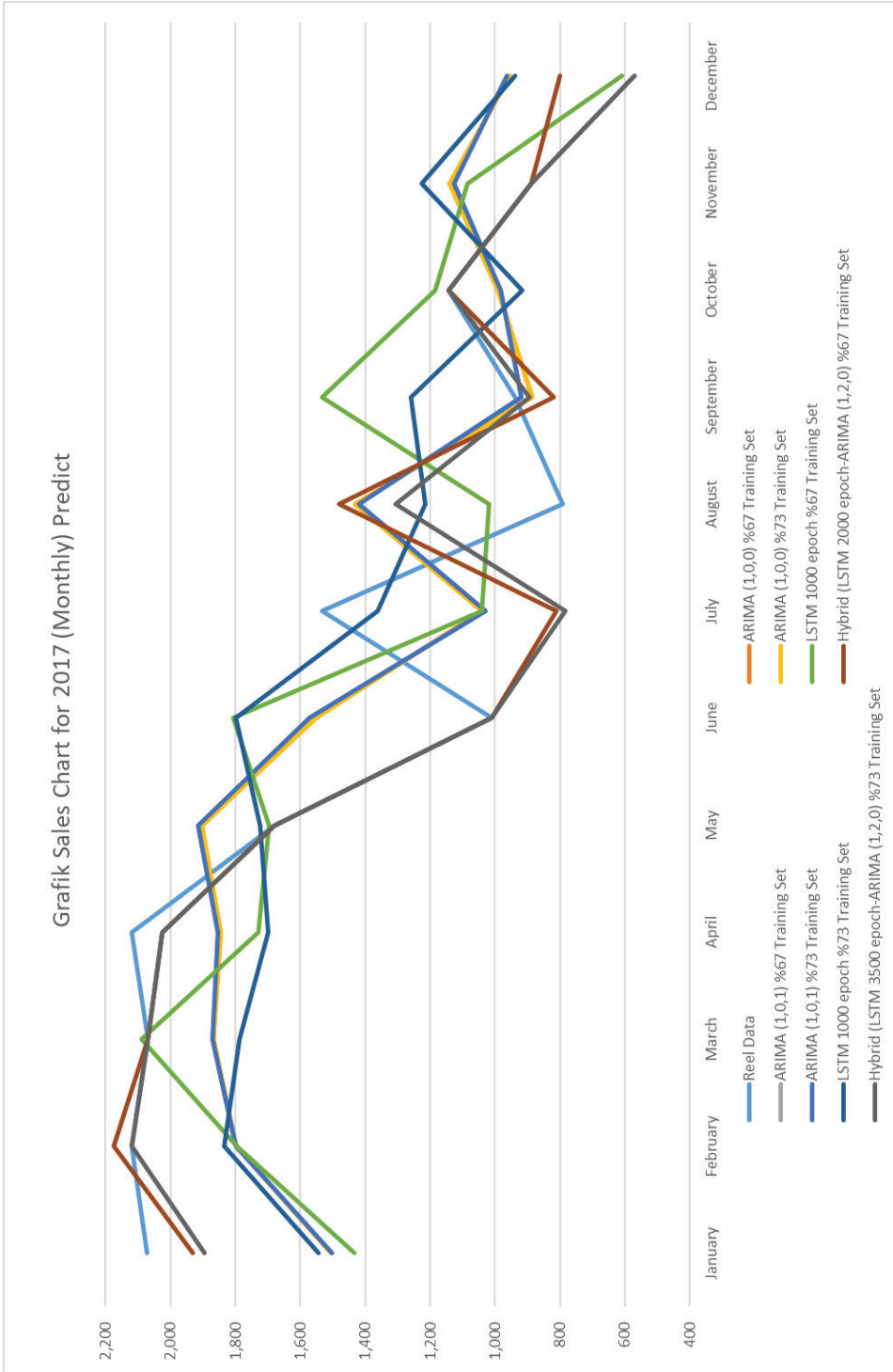


Figure 11. Comparison of Sales Volume Estimates by Methods

The graphic above gives a comparison of the sales estimates obtained from the methods used in the application. Sales data realized in graphical representation are also included in the drawing. As can be seen from the figure, there was not an extreme deviation in all the methods used. The real data graph and the forecast graphs were formed at close levels. In some months, an increase/decrease (such as July-August) occurred in the real data graph, while an opposite result such as a decrease/increase in the forecast graphs was obtained. This can be explained by changes in economic conditions, differences in demand for the industry, etc. As the results obtained are close to each other, all of the methods can be used to create the sales quantity budgets of the business as mentioned earlier. However, it is possible to state that this is the best method since the best result is taken from the hybrid model used in the application and the aim is the most accurate estimate. This situation is an indication that the study has achieved its main purpose.

The results obtained from the methods used in the study are very efficient in terms of performance criteria. In the case of having access to data in later studies, re-trials with further data usage are planned and both creating other budgets in enterprises (production, investment, cash budget, etc.) with the same methods and also the comparison of the results in terms of performance with predictions using different methods are planned.

Peer-review: Externally peer-reviewed.

Conflict of Interest: The authors have no conflict of interest to declare.

Grant Support: The authors declared that this study has received no financial support.

Authors' Contributions : Ayşe SOY TEMÜR has carried out the stages of determining the subject of research, examining the relevant models, collecting their data, recommending the method to be used in the research, carrying out the analysis and test studies, and the interpretation of the data by entering them into the analysis program. Şule YILDIZ contributes to the design of the research, the evaluation and guidance of the study at all of its stages, and the interpretation of the article. All authors contribute to the interpretation of results, edit the text, review it, and approve the article.

References

- Açııcı, K., Aşuroğlu, T., & Oğul, H. (2017). Information retrieval in metal music sub-genres. *2017 25th Signal Processing and Communications Applications Conference (SIU)*. Antalya, TURKEY: IEEE.
- Adeva, J. J., Beresi, U. C., & Calvo, R. A. (2005). Accuracy and Diversity in Ensembles of Text Categorisers. *CLEI Electronic Journal*, 8(2), 1-12. doi:DOI: 10.19153/eleiej.12.1.1
- Ahmetoğlu, M. E. (2018). *Veri Merkezi Sektöründe Satış Tahminleme Modeli ve Özel Kurum Üzerinde Uygulaması*. İstanbul: Bahçeşehir Üniversitesi.
- Aladağ, Ç. H., Eğrioğlu, E., & Kadılar, C. (2009). Forecasting Nonlinear Time Series with a Hybrid Methodology. *Applied Mathematics Letters*(22), 1467-1470. doi:https://doi.org/10.1016/j.aml.2009.02.006
- Albayrak, A. S. (2010). ARIMA Forecasting of Primary Energy Production and Consumption in Turkey: 1923–2006. *Enerji, Piyasa ve Düzenleme*, 24-50.
- Armano, G., Marchesi, M., & Murru, A. (2005). A Hybrid Genetic-Neural Architecture for Stock Indexes Forecasting. *Information Sciences*(170), 3-33.

- Ateşoğlu, A. A. (2015). *ARIMA ve Yapay Sinir Ağları (YSA) Kullanılarak Hibrit Tahmin Modeli Geliştirilmesi*. Ankara: Başkent Üniversitesi Fen Bilimleri Enstitüsü.
- Atienza, R. (Posted on 2017, 4 26). *LSTM by Example using Tensorflow (Text Generate)*. towardsdatascience.com: <https://towardsdatascience.com/lstm-by-example-using-tensorflow-feb0c1968537>
- Babu, C. N., & Reddy, B. E. (2014). A Moving-Average Filter Based Hybrid ARIMA–ANN Model for Forecasting Time Series Data. *Applied Soft Computing*(23), 27-38. doi:<https://doi.org/10.1016/j.asoc.2014.05.028>
- Bengio, Y., Patrice, S., & Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE transaction on neural networks*, 157-166.
- Bengio, Y., Simard, P., & Frasconi, P. (2015). Learning Long-Term Dependencies with Gradient Descent is Difficult. *IEEE Trans. Neural Networks*, 5(2), 157-166.
- Bhardwaj, S., M.Gadre, V., & Chandrasekhar, E. (2020). Statistical Analysis of DWT Coefficients of fGn Processes Using ARFIMA (p,d,q) Models. *Physica A: Statistical Mechanics and its Applications*, 547, 124404. doi:10.1016/j.physa.2020.124404
- Bircan, H., & Karagöz, Y. (2003). Box-Jenkins Modelleri ile Aylık Döviz Kuru Tahmini Üzerine Bir Uygulama. *Kocaeli Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 6(2), 49-62. Şubat 10, 2017 tarihinde <http://dergipark.gov.tr/download/article-file/252067>
- Box, P., & Jenkins, G. (1976). *Time Series Analysis: Forecasting and Control*. Holden-day Inc.
- Brownlee, J. (2020). *Machine Learning Mastery*. <https://machinelearningmastery.com/train-test-split-for-evaluating-machine-learning-algorithms/>, July, 2020.
- Bukhari, A. H., Raja, M. A., Sulaiman, M., Islam, S., Shoaib, M., & Kumam, P. (2020). Fractional Neuro-Sequential ARFIMA-LSTM for Financial Market Forecasting. *IEEE ACCESS*, 8, 71326-71338. doi:10.1109/ACCESS.2020.2985763
- Büyükhahin, Ü., & Ertekin, Ş. (2018). Ampirik Kip Ayırışımı ve Hibrit Metot Kullanarak Zaman Serisi Tahmini. *Sinyal İşleme ve İletişim Uygulamaları (SİU) Kurultayı*. IEEE.
- Chollet, F. a. (2015). Keras. *chollet2015keras*. <https://keras.io>, 2020.04.12
- Contreras, J., Espinola, R., Nogales, F., & Conejo, A. (2003). ARIMA models to predict next-day electricity prices. *IEEE Tans. Power Syst.*, 1014-1020.
- Düzyol, M. (1997). *Türkiye’de Bina İnşaatı Sektörü ve 1990-2010 Dönemi Bölgesel İhtiyaç Tahmini*. Ankara: DPT: 2484.
- Ediger, V. Ş., & Akar, S. (2007). ARIMA Forecasting of Primary Energy Demand by Fuel in Turkey. *Energy Policy*, 1701-1708. doi:<https://doi.org/10.1016/j.enpol.2006.05.009>
- Erdoğan, E. (2007). Electricity Demand Analysis Using Cointegration and ARIMA Modelling: A Case Study of Turkey. *Energy Policy*, 1129-1146. doi:<https://doi.org/10.1016/j.enpol.2006.02.013>
- Fattah, J., Ezzine, L., Aman, Z., Moussami, H. E., & Lachhab, A. (2018). Forecasting of Demand Using ARIMA Model. *International Journal of Engineering Business Management*, 10, 1-9. doi:10.1177/1847979018808673
- Greenwood, J., & Hercowitz, Z. (1991). The Allocation of Capital and The Time Over The Business Cycle. *Journal of Political Economy*(99), 1188-1214. doi:<http://dx.doi.org/10.1086/261797>
- He, G., & Deng, Q. (2012). A Hybrid ARIMA and Neural Network Model to Forecast Particulate. *Matter Concentration in Changsha*. <https://www.isiaq.org/docs/PDF%20Docs%20for%20Proceedings/1F.3.pdf>

- Hocaoğlu, F. O., Kaysal, K., & Kaysal, A. (2015). Yük Tahmini İçin Hibrit (YSA ve Regresyon) Model. *Akademik Platform*, 33-39. doi:DOI: 10.5505/apjes.2015.94695
- Hochreiter S. (1998). The vanishing gradient problem during learning recurrent neural nets and problem solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based System*, 107-116.
- Hochreiter, S., & Schmidhuber, J. (1997). Long Sort Term Memory. *Neural Computation*, 1735-1780.
- Ioannou, K., Birbilis, D., & Lefakis, P. (2011). A Method for Predicting the Possibility of Ring Shake Appearance on Standing Chestnut Trees. *Journal of Environmental Protection and Ecology*, 295-304. <https://docs.google.com/a/jepe-journal.info/viewer?a=v&pid=sites&srcid=amVwZS1qb3VybmFsLmluZm98amVwZS1qb3VybmFsGd4OjJkYzQwODIyZjE4ZmJmMzQ>
- Kang, E. (2018). Github: https://github.com/llSourceCell/LSTM_Networks/blob/master/LSTM%20Demo.ipynb
- Khandelwal, I., Adhikari, R., & Verma, G. (2015). Time Series Forecasting using Hybrid ARIMA and ANN Models based on DWT Decomposition. *Procedia Computer Science*(48), 173-179.
- Khashei M., H. S. (2008). A New Hybrid Artificial Neural Networks and Fuzzy Regression Model for Time Series Forecasting. *Fuzzy Sets and Systems*(159), 769-786. doi:<https://doi.org/10.1016/j.fss.2007.10.011>
- Khashei, M., & Bijari, M. (2011). A Novel Hybridization of Artificial Neural Networks and ARIMA Models for Time Series Forecasting. *Applied Soft Computing*(11), 2664-2675.
- Khashei, M., Hejazi, S. R., & Bijari, M. (2008). A New Hybrid Artificial Neural Networks and Fuzzy Regression Model for Time Series Forecasting. *Fuzzy Sets and Systems*(159), 769-786. doi:<https://doi.org/10.1016/j.fss.2007.10.011>
- Kılıç, R., & Demirbaş, E. (2012). Türkiye’de Kamu İnşaat Harcamalarının Belirleyicileri ile Ekonomik Büyüme Arasındaki İlişki. *Akademik Yaklaşımlar Dergisi*, 3(2), 84-97.
- Kim, H., & Shin, K. (2007). A Hybrid Approach Based on Neural Networks and Genetic Algorithms for Detecting Temporal Patterns in Stock Markets. *Applied Soft Computing*(7), 569-576.
- Koutroumanidis, T., Ioannou, K., & Zafeiriou, E. (2011). Forecasting Bank Stock Market Prices with A Hybrid Method: The Case of Alpha Bank. *Journal of Business Economics and Management*, 12(1), 144-163. doi:<https://doi.org/10.3846/16111699.2011.555388>
- Koutroumanidis, T., Ioannou, K., & Arabatzis, G. (2009). Predicting Fuelwood Prices in Greece With the Use of ARIMA Models, Artificial Neural Networks and a Hybrid ARIMA-ANN Model. *Energy Policy*(37), 3627-3634. doi:<https://doi.org/10.1016/j.enpol.2009.04.024>
- Lin, T., Guo, T., & Aberer, K. (2017). Hybrid Neural Networks for Learning the Trend in Time Series. *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*, (s. 2273-2279). Melbourne, Australia. doi:<https://doi.org/10.24963/ijcai.2017/316>
- Luxhoj, J., Riis, J., & Stensballe, B. (1996). A Hybrid Econometric-Neural Network Modeling Approach for Sales Forecasting. *International Journal of Production Economics*(45), 175-192.
- Namin, S. S., & Namin, A. S. (2018). *Forecasting Economic and Financial Time Series: ARIMA vs. LSTM*. Lubbock, TX, USA: Texas Tech University. doi:<https://arxiv.org/ftp/arxiv/papers/1803/1803.06386.pdf>
- Newbold, P. (1983). ARIMA Model Building and The Time Series Analysis Approach to Forecasting. *Journal of Forecasting*, 2(1), 23-35. doi: <https://doi.org/10.1002/for.3980020104>
- Olah, C. (2015, August 27). *Understanding LSTM Networks*. Nisan 10, 2019 tarihinde colah.github.io: colah.github.io/posts/2015-08-Understanding-LSTMs/.

- Oliveira, M., & Torgo, L. (2014). Ensembles for Time Series Forecasting. *In Proceedings of the Sixth Asian Conference on Machine Learning*, (s. 360-370). <http://proceedings.mlr.press/v39/oliveira14.pdf> adresinden alındı
- Opitz, D., & Maclin, R. (1999). Popular Ensemble Methods: An Empirical Study. *Journal of Artificial Intelligence Research*, 11, 169–198. doi:<https://doi.org/10.1613/jair.614>
- Pablo, B. J., Hilda, C., Xavier, A., Diego, J. J., Felipe, S., & Henry, B. (2016). Artificial Neural Network and Monte Carlo Forecasting with Generation of L-Scenarios. *2016 Intl IEEE Conferences on Ubiquitous Intelligence & Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People and Smart World Congress*, (s. 665-670). doi:10.1109/UIC-ATC-ScalCom-CBDCCom-IoP-SmartWorld.2016.01110
- Pagourtzi, E., Assimakolopoulos, V., Hatzichristos, T., & French, N. (2003). Real Estate Appraisal: A Review of Valuation Methods. *Journal of Property Investment & Finance*, 21(4), 383-401.
- Papagera, A., Ioannou, K., Zaimis, G., Iakovoglou, V., & Simeonidou, M. (2014). Simulation and Prediction of Water Allocation Using Artificial Neural Networks and a Spatially Distributed Hydrological Model. *Agris on-line Papers in Economics and Informatics*, 6(4), 101-111. <https://ageconsearch.umn.edu/record/196580/>.
- Pascanu, R., Mikolov, T., & Bengio, Y. (2013). On the difficult of training recurrent neural networks. *ICML*, 1310-1318.
- Prieto, M. P., Uribbarri, P. M., & Tardioli, G. (2019). Applying Modeling and Optimization Tools to Existing City Quarters. *Urban Energy Systems for Low-Carbon Cities*, 333-414. doi:10.1016/B978-0-12-811553-4.00010-X
- Rossum, G. V. (2001). *Python*. (T. M. Python) <https://docs.python.org/3/2019.05.17>
- Sallehuddin, R., Shamsuddin, S. M., Hashim, S. Z., & Abraham, A. (2007). Forecasting Time Series Data Using Hybrid Grey Relational Artificial Neural Network and Auto Regressive Integrated Moving Average Model. *ICS AS CR 2007*, (s. 573-605). https://www.researchgate.net/publication/287522177_Forecasting_time_series_data_using_hybrid_grey_relational_artificial_neural_network_and_auto_regressive_integrated_moving_average_model.
- Sarı, M. (2016, Ocak). Yapay Sinir Ağları ve Bir Otomotiv Firmasında Satış Talep Tahmini Uygulaması. *Yüksek Lisans Tezi*. Sakarya, Türkiye: Sakarya University.
- Smith, M. J. (2018). *Statistical Analysis* (2 b.). Edinburgh: The Winchelsea Press, Drumlin Security Ltd. <https://www.statsref.com/StatsRefSample.pdf>
- Soy Temür, A., Akgün, M., & Temür, G. (2019). Predicting Housing Sales In Turkey Using Arima, Lstm And Hybrid Models. *Journal of Business Economics and Management*, 20(5), 920-938. doi:<https://doi.org/10.3846/jbem.2019.10190>
- Sugiartawan, P., Pulungan, R., & Sari, A. K. (2017). Prediction by a Hybrid of Wavelet Transform and Long-Short-Term-Memory Neural Network. *International Journal of Advanced Computer Science and Applications*, 8(2), 326-332. https://www.researchgate.net/profile/Reza_Pulungan/publication/314166897_Prediction_by_a_Hybrid_of_Wavelet_Transform_and_Long-Short-Term-Memory_Neural_Network/links/5950a68f45851543383a3d3b/Prediction-by-a-Hybrid-of-Wavelet-Transform-and-Long-Short-Term-M.
- Suresh, K., & Priya, K. (2011). Forecasting sugarcane yield of Tamilnadu using ARIMA models. *Sugar Tech.*, 23-26.

- Tsaih, R., Hsu, Y., & Lai, C. (1998). Forecasting S&P 500 Stock Index Futures with A Hybrid AI System. *Decision Support Systems*(23), 161-174.
- Tseng, F., Yu, H., & Tzeng, G. (2002). Combining Neural Network Model with Seasonal Time Series ARIMA Model. *Tecnological Forecasting&Social Change*(69), 71-87.
- Vasques, X., Vanel, L., Villette, G., & Cif, L. (2016). Morphological Neuron Classification Using Machine Learning. *Frontiers in Neuroanatomy, 10*, 102-114. doi:10.3389/fnana.2016.00102
- Wedding, D., & Cios, K. (1996). Time Series Forecasting by Combining RBF Networks, Certainty Factors and The Box-Jenkins Model. *Neurocomputing, 10*, 149-168.
- Wu, L., & Brynjolfsson, E. (2015). The Future of Prediction: How Google Searches Foreshadow Housing Prices and Sales. S. M. Avi Goldfarb içinde, *Economic Analysis of the Digital Economy* (s. 89-118). Chicago, USA: University of Chicago Press. doi:http://dx.doi.org/10.2139/ssrn.2022293
- Yılmazel, Ö., Afşar, A., & Yılmazel, S. (2018). Konut Fiyat Tahmininde Yapay Sinir Ağları Yönteminin Kullanılması. *UİİİD-IJEAS*, 285-300.
- Yu, L., Jiao, C., Xin, H., Wang, Y., & Wang, K. (2018). Prediction on Housing Price Based on Deep Learning. *World Academy of Science, Engineering and Technology, 12*(2). doi:doi.org/10.5281/zenodo.1315879
- Yu, L., Wang, S., & Lai, K. (2005). A Novel Nonlinear Ensemble Forecasting Model Incorporating GLAR and ANN for Foreign Exchange Rates. *Computers&Operations Research*(32), 2523-2541.
- Zhang, G. (2003). Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model. *Neurocomputing*(50), 159-175. doi:https://doi.org/10.1016/S0925-2312(01)00702-0
- Zhou, Z., & Hu, C. (2008). An Effective Hybrid Approach Based on Grey and Arma for Forecasting Gyro Drift. *Chaos, Solitons and Fractals*(35), 525-529.

