OPEN ACCESS
International Journal of Innovations in Science & Technology

IJIST
P ISSN : 2618-1630
E ISSN : 2909-6130
Recognized Journal

Sales Prediction of Cardiac Products by Time Series and Deep Learning

Original Article

Muhammad Waqas Arshad[1] and Syed Fahad Tahir[2]
[1] Dept of Creative Technologies, Air University, Islamabad, Pakistan
[2] Dept of Computer Sciences, Air University, Islamabad, Pakistan
* Corresponding: **Muhammad.waqas.arshad.1@gmail.com,**

Citation |Arshad. W. M, Tahir. F. S, "Sales Prediction of Cardiac Products by Time Series and Deep Learning". International Journal of Innovations in Science and Technology. Special Issue NUML 2022, pp: 1-11

Received | May 31, 2022; **Revised** |June 22, 2022; **Accepted** | June 24, 2022; **Published** | June 26, 2022.

Maintaining inventory level to avoid high inventory costs is an issue for Cardiac Product Distribution Companies (CPDCs) because of the shortage of their products which affect their sale and causes loss of the customer. This research aims to provide a method for predicting the upcoming demand of the Balloon and Stents by using time series analysis (Auto Regression Integrated Moving Average) and Deep learning (Long-Short Term Memory). To conduct this research, data was collected from Pakistan's leading cardiac product distributors to determine the method's performance. The findings were compared using Mean absolute error (MAE) and Root Mean Square Error (RMSE). Resulst conclude that the ARIMA algorithm successfully forecasts cardiac products sale.

Keywords: Cardiac Products, Balloons, Stents, Time Series, Deep Learning, Decision Support.

Acknowledgment.
The authors are thankful to Air University, Islamabad, Pakistan, for their platform to this research . Thanks to Dr. Iqbal Murtza and Dr. Asif Ullah Khan for their suggestions and useful feedback on this research.

Conflict of Interest:
Authors declare no conflict of interest for publishing this manuscript in IJIST.

Project details. This research is a part of the MS Research work carried out by the first author and supervised by the second author.

Author's Contribution. Muhammad Waqas Arshad: carried out data collection, Performed in data analysis and write-up of the manuscript. Syed Fahad Tahir: Supervised the research and did the preliminary review.

Introduction

The precise sales forecast provides a gateway to improve a company's profitability while minimizing expenses. Making the right decision and choosing the best option among the one variable by selecting the best scenario is an essential task. It's no surprise that forecasting has recently sparked a lot of interest. It can also capture the right balance between customer satisfaction and inventory expenses and helps to reduce the risk of mistakes [1], [2]. Due to uncertain demand, sales prediction is challenging for the distributor of cardiovascular products, especially those dealing in balloons and stents. Distributors import these products primarily from countries like USA, Germany, and Australia, which is a time-taking process. The company must maintain its inventory level to manage the market due to its one-year expiry,, especially in an emergency after diagnosis. If the forecasts are affected by the actual outcome, their reliability can be significantly lower. A sale forecast helps shopkeepter to make better decisions that help them to manage their cash flow. It allows sales teams to reach their goals by identifying potential warning signals in their pipeline [3]. The inefficiencies occur while the goods have a splendid call for, and they may be now no longer available. Some challenges as understanding sales data to forecasting sales; however, missing, randomness, and fewer amounts of data created a significant challenge for analysis and trend reporting. Most companies are unaware of present tools and techniques for sale prediction. Manual methods like surveys and comparing records that waste time and cost expensive with the wrong prediction.

LSTM (Long Short-Term Memory model) and ARIMA (Auto-Regressive Integrated Moving Average model) are used to predict the upcoming demands. The quantity of information developing exponentially and rational use of the data has turned out to be the point of interest of firms to serve the destiny and make higher decisions. Accurate sales forecasting is critical for businesses like distributors of health-related products and supermarkets to achieve smart sales, regulate budgets, and develop better plans. This approach provides reliable and accurate estimations of cardiovascular product distribution by utilizing scientific methods to ensure that the estimations are as close as possible to the actual value [4].

The proposed approach consists of ARIMA and LSTM to support a shareholder's decision-making process in the distribution sector for sales prediction of balloons and stents. The results obtained from the four methods are evaluated based on Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). These methods do not pose any restrictions, and it is possible to make estimations using different techniques related to cardiovascular products. The collected dataset has been used for the first time for any research purpose by one of Pakistan's top cardiac vascular products distributors. The company data contains between January 2017 and December 2019. These estimations contribute to many industries related to the health products distribution sector and helped them by utilizing these methods, which perform as a state-of-the-art result for sales prediction. In addition, this study aims to contribute to the literature for the benefit of similar studies to be carried out [5].

The rest of the paper is organized as follows. The prior literature is reviewed in Section 2. Section 3 discusses the proposed methodology by explaining the block diagram. Section 4 describes the dataset and the experiment results, a brief discussion, and an outcome analysis. At last, discussed the conclusion with the feasibility of future work in Section 5.

LITERATURE REVIEW

This section presents previous researches pertinent to our problem statement, which describes a significant part of the literature for time series forecasting and highlights its importance in recent years. An earlier work related to the subject of this research intensely discusses.

Time Series Forecasting.

These observations are recorded continuously over time or at discrete intervals of equal length. Continuous-time series, such as brain activity measurements, are often examined by sampling at equal time intervals to produce a discrete-time series [6]. The correlation could be the sequence in which the data is recorded, the model's linearity, repeated patterns, etc. Techniques for data analysis are provided by time-series analysis [7], [8]. Forecasting predicts future values based solely on previous data [9].

It is possible to project future product sales [6] for which, there are a variety of models that are utilized in different domains and contexts. Examples include ARIMA, ANN, and Support Vector Machines [10], [11]. The majority of models are either linear or non-linear. The assumption of linear behavior binds linear models, but non-linear models can better fit the shape of a more complicated function. Sales, for example, can vary dramatically based on the season, holidays, days after payday, etc. It means that sales are not at the straight line. This raises the question of whether there are many advantages to adopting non-linear models to forecast non-linear functions. It can be advantageous, but it is not always the case. Linear models, such as ARIMA, are well-known for their accuracy and flexibility in dealing with non-linear domains [5] [9].

Forecasting Methods.

An LSTM is a Recurrent Neural Network (RNN) capable of retaining data before future usage [12]. It is mainly utilized in deep learning that performs better on substantial datasets. The outcome of this test is that they can repeat their achievement with a 13 percent –16 percentage error margin [13]. LSTMs are a form of RNN that focuses on learning long-term dependency. It contains a hidden layer and has a complicated structure called the LSTM.

In the RNN cell, an LSTM structure seems to have a memory. Data from the last moment recovered and communicated to the following. The model chooses which data to include in training. In actuality, the habits of these networks is to make sure information for an extended period is not something they strive to learn. A gate is a complete communication layer with a vector for input and a specific value between 0 and 1 as output. When the number is 0, the information gets deleted, but when the number is 1, a cell keeps all of the input data. The data collected there at t time interval and thus the outcome of the preceding unit are represented by X_t. $h_t$is the output of hidden units, whereas h_{t-1} represents their prior output. For such LSTM cells, to compute and utilize the input gate i_t^j (1) the forget gate f_t^j (2) and output gate σ_t^j (3) are calculated.

$$i_t^j = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i)^j \qquad (1)$$

$$f_t^j = \sigma\left(W_{xf}x_t + W_{hf}h_{t-1} + b_f\right)^j \qquad (2)$$

$$\sigma_t^j = \sigma\left(W_{x\sigma}x_t + W_{h\sigma}h_{t-1} + b_\sigma\right)^j \qquad (3)$$

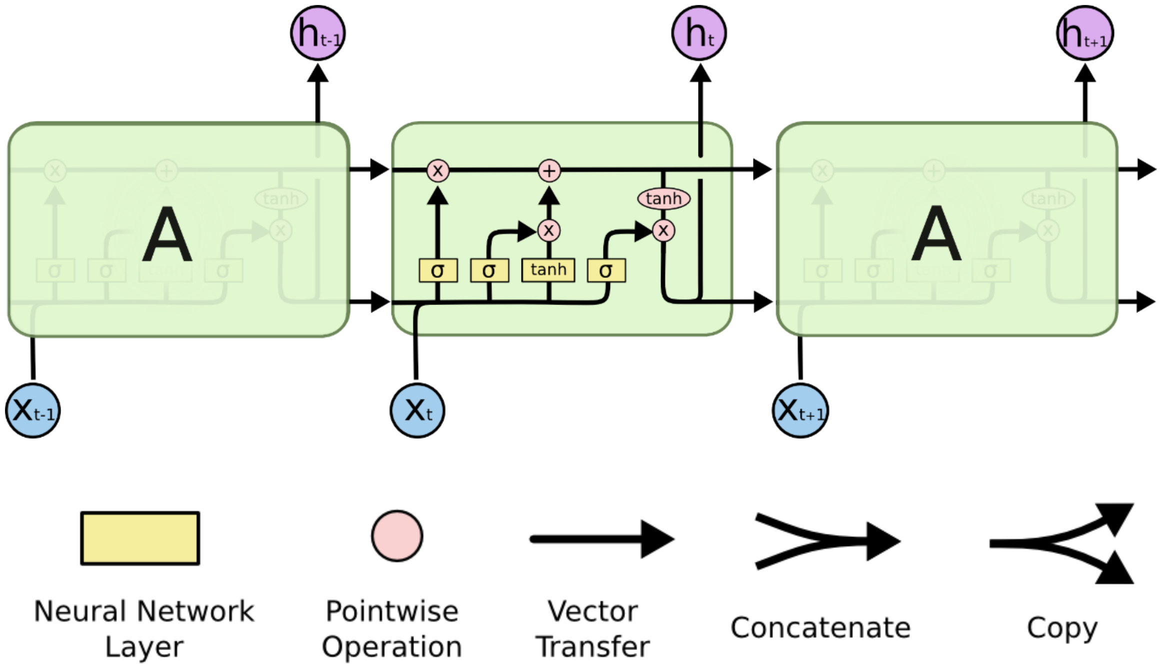


Figure 1. Depicts a simplified illustration of the structures.

The architecture of LSTM describe that each line carries an entire vector, from one node output to others inputs. The pink circles represent pointwise operations, like vector addition, while the yellow boxes are neural network layers. Line merging denotes concatenation, while a line forking denotes copying content, and the copies are going to different locations [14].

The sigmoid function is represented by σ and the w expressions, weight matrices, and the b terms, which are voltage vectors. Apart from classic epoch units, each j. LSTM unit uses (c_t^j) to keep the memory at t time. The equation updates the expression for whom the memory cell is supplied (4).

$$c_t^j = f_t^j c_{t-1}^j + i_t^j c_t^{\sim j} \qquad (4)$$

Equation (5) is used to update the current memory information, and equation (6) is used to generate the outcome for the LSTM unit (6).

$$f_t^j = tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c)^j \qquad (5)$$

$$h_t^j = \sigma_t^j \tanh\left(c_t^j\right) \qquad (6)$$

Propagation on LSTM networks is done by epoch. An epoch assessing the network weighting factor shows the overall number of repetitions of a particular data collection for learning purposes (w). An epoch is when an entire dataset has been sent forward and received back upon that system. It is reasonable to update weights to improve frameworks of deep learning algorithms and transmit the entire dataset through a single platform multiple times to create a faster and more effective prediction system.

The methods are adopted to forecast the value of different series data such as sugar sales, housing sales, and stock quantity. ARIMA model may assist to predit the behavior of a specific application [15] [16]. The ARIMA approach is divided into three steps: examination, identification, and evaluation. Regression monitoring is conducted on the provided time series data over the first stage, term diagnostic inspection. Stationarity is where statistical features like mean and variance are proportional to time passage. When building an ARIMA model, normality of the data is required, which enables the prediction to be highly practical and helpful. Differentiating (d) at a suitable degree is used to render a non-stationary series data stationary and recheck the stability. This procedure is repeated till a stationary sequence is achieved. The differencing level is determined by (d), a positive integer. The ARIMA model's integrated parameters are set to (d) if the differencing procedure is done (d) times. After that, the acquired static data is subjected to authentication. The coefficients of the auto-regressive (AR) & moving average (MA) actions stated in equation (1) are (p) and (q), accordingly, in this process. To use the Autocorrelation Function (ACF) to decide the quantity of linear relation among observational data in a time series detached by a lag p, the partial autocorrelation function (PACF) evaluates how often auto-regressive aspects q are required. The inverse autocorrelation function (IACF) to identify over differencing, we may even recognize the prelim value systems of autoregressive order p, order of differencing d, moving average order q. The variable changes that order of different frequencies between non-stationary toward stationary series [17][18].

p: degree of the autoregressive model (AR)
d: differencing degree
q: degree of moving average model (MA)

$$y_t = \alpha_1 \omega_{t-1} + \alpha_2 \omega_{t-2} + \cdots + \alpha_p \omega_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - .. - - \theta_q \varepsilon_{t-q} \quad (7)$$

Here,y_t, signifies the generalized linear actual data and ε_t represents a moving average error for such t period. A linear connection has been created among actual datay_t, is being estimated, actual (p) data (y_{t-1},y_{t-2},..., y_{t-p}), and (q) error data (,ε_t , ε_{t-1} ,.., ε_{t-q}). , as illustrated in equation 7.

Table 1: A summary of the literature review, shows the research work related to our paper.

Sr. No	Title	Year	Dataset	Approach	Result
1.	Intelligent Sales Prediction for Pharmaceutical Distribution Companies: A Data Mining Based Approach: Mathematical Problems in Engineering [19]	2014	Dataset provided by one of the leading PDCs in Iran	ARIMA, neural network, and an advanced hybrid neural network approach	Hybrid Approach performs better results.
2.	A comparative study between LSTM and ARIMA for sales forecasting in retail: TRITA-EECS-EX [20]	2018	The sales data is gathered from several groceries stores between 2013 to 2017, publicly available on Kaggle.	ARIM and LSTM	LSTM gives a good result with better accuracy as compared to ARIMA.
3.	Predicting housing sales in Turkey using ARIMA, LSTM, and hybrid models: Journal of Business Economics and Management [17]	2019	The data set include s 124-months ranging from 2008 (1) to 2018 (4) period used for total house sales in Turkey.	ARIMA,LSTM, ARIMA-LSTM (Hybrid)	Hybrid approach more accuracy rate .

4.	On the Application of ARIMA and LSTM to Predict Order Demand Based on Short Lead Time and On-Time Delivery Requirements : Processes [21]	2021	Dataset of a Taiwan's semiconductor IC companies from 1 January 2017 to 31 August 2019	ARIMA and LSTM	LSTM Performs well as compared to ARIMA.

PROPOSED METHODOLOGY

Figure 2 shows the block diagram of the proposed method. The process comprises four main components: Preprocessing, Training, Testing, and Sales Prediction.

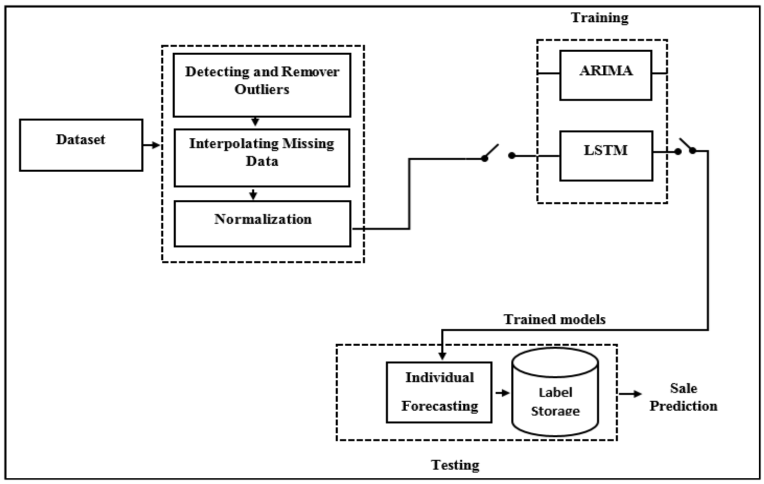


Figure 2: Block diagram of the research methodology used in this paper, taking the dataset, applying through preprocessing and examining through exploratory analysis on the study base, selected model is trained and apply the cross-validation with 5 fold and at last, get the prediction.

Dataset.

Data of sale products was collected from one of the oldest cardiac product distributors, it was not open-source and was used for the first time in any research. They have been doing their business for over 25 years in Pakistan, dealing in all the provinces of Pakistan; Two different kinds of Stents labeled Stent1 and Stent 2, and one type of Balloon was used. The dataset available from 2017 to 2019, as shown in Figure 3, was used representing the attributes as one represents the date in a month, and another shows the sale. The main objective was to predict the sale of an upcoming year as 2020 of this data.

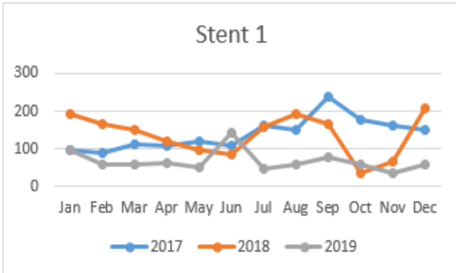
[s0, ...,sd, sd+1, ..., sd+p]

[s1, ..., sd+1, sd+2, ..., sd+p+1]

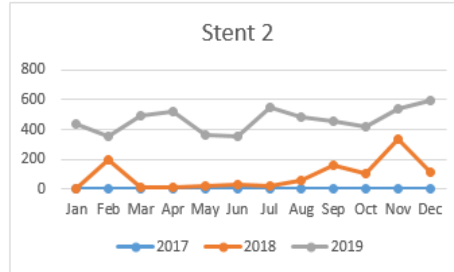
[s2, ..., sd+2, sd+3, ..., sd+p+2].

[st−d−p, ..., st−p−1, st−p, ..., st]

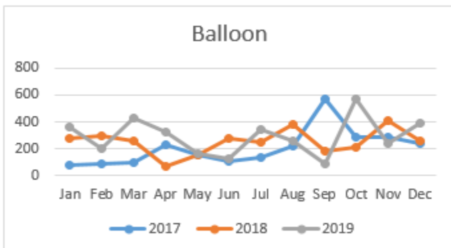
Where d is the number of previous time steps the model and p is the number of future time steps that the model finds. Each row's first d columns were utilized for training.



(a)



(b)



(c)

Figure 3: Representing the trend of the products x-axis represents the months, and the y-axis shows the sales in number, point' a' shows Stent 1, point 'b' Stent 2, and 'c' represents the product for Balloon of the year 2017 to 2019, the blue color shows the year 2017, orange for 2018 and grey for 2019

Preprocessing.

The preprocessing component includes these processes: identifying and eliminating outliers, interpolating missing values, & data normalization.

Detecting and Removing Outlier.

An outlier can impair sales data by producing false expectations about forecast; it would be an observation that differs significantly from those in the remaining observations, raising suspicion that variation is created by a different cause, such as new clients, excess inventory, or force sale. We select a range of data that you consider typical of the entire data; outliers are the numbers outside that range. Data visualization supports outlier analysis by making detected outliers more obvious than creating a scatter plot of our data. Once the plot is constructed, it reveals data points that don't seem to match the rest.

Interpolating Missing Data.

For most data sources, insufficient data is an unavoidable problem. Missing values can lead to complications because subsequent data processing and decision-making procedures typically rely on entire datasets. Therefore, missing data need to be substituted with suitable ones. Such insights will assist us in projecting future values. The time series rationality, changing trends, and missing values must be determined. The incomplete observation gets filled in throughout this research by averaging the last four nearby data points in the same time series after removing the outlier.

Normalization.

The normalization of the data may reduce model time training by beginning the training phase for each variable on the same level. In this study, the z-score normalization approach is used to standardize the values of the parameters for input and output shown in equation 8 [22].

$$C' = \frac{C - \mu}{\sigma} \qquad (8)$$

Where C defines the feature value, μ is the mean of the feature values, and σ represents the standard deviation of the feature values.

Training and Testing.

The data after the preprocessing stage was ready for training, selecting different types of models, including ARIMA and LSTM. The training data use to process a model, then test by providing forecasts against the test set. It is simple to assess whether the model's estimations are correct because the cross-validation data already have actual values for the characteristic we want to forecast.

Implementation.

LSTM model is implemented, which predicts the sales for the year 2020. The model is implemented using the python library Keras; the data is trained, indicates the sale by testing, and uses the k-fold cross-validation with 5-fold. Dropout 0.1, learning rate 0.01, epoch 2000, and activation function sigmoid is used as a hyperparameter. The ARIMA implementation uses the IBM SPSS tool, which predicted the sales for 2020; In terms of the data utilized in the experiment, it was discovered that the actual trend of the stock's time series is not stationary. The time series follows a random walk pattern that changes randomly, with no discernible overall trend or seasonality pattern. The measured values for p, d, and q must be specified to apply the ARIMA model. The data normalize and difference with the value 1, converting the data from non-stationary to stationary. The model is tested with cross-validation using 5-fold as k=5 gives the best result compared to k=4 and 6. There are a total of 17 types of ARIMA models used, from which ARIMA (8,1,0) (0,1,1) and (8,0,0) give the best result.

Prediction.

The outcome is the prediction of sales for each product for 2020. It identifies that ARIMA has a close result to the ground truth values compared to LSTM and the performance measure Mean Absolute Error and Root Mean Square Error gives the less error rate for ARIMA than LSTM. The primary goal is to evaluate each model performance used in this study by these performance measures, such as Mean Absolute Error & Root Mean Square Error, to assess the effectiveness of these algorithms [23].

Mean Absolute Error.

The mean absolute error would be a performance metric for models. The mean fundamental error for a testing dataset is the average of the actual values of the particular prediction errors for all occurrences in the test set. The Mean Absolute Error (MAE) measures sums up and evaluates a Machine Learning model [24].

$$MAE(E,F) = \left(\frac{1}{n_{samples}}\right) \sum_{i=o}^{n_{samples}-1} |E_i - F_i| \qquad (9)$$

Ei provides the predicted values, whereas Fi indicates the actual values in equation 9.

Root Mean Square Error.

The RMSE is calculated by squaring the difference between predicted and observed values and then averaging them across the sample. A mean absolute error is calculated as follows: The RMS is calculated for a collection of n values {q1, q2... qn} in equation 10.

$$RMSE = \sqrt{\sum_{o-1}^{n} \frac{(\widehat{k_o} - k_0)^2}{n}} \qquad (10)$$

Where 'n' is a no of observations.$\widehat{k_o}$are a predicted value andk_oDefines observed values. The Root means square error (RMSE) is a quadratic scoring method that determines the average size of the mistake. Both are used simultaneously, and the RMSE value is always greater than or equal to the MAE value. RMSE & MAE are error measures; however, the model works well if the standard for error is near zero [25].

RESULTS

Each model estimates sales for the three products given as datasets, yielding an RMSE and MAE error value. The most negligible value in the tables for MAE and RMSE shows that the ARIMA model has a minor error. Whereas a higher value of performance evaluator indicates that the LSTM technique has a higher error for all products. In Table 1. Both the MAE and RMSE vary from 0 to ∞. They are negatively-oriented scores: Lower values are better, so we conclude that the model with a minimum RMSE and MAE gives a better prediction.

Table 1: Comparative performance for the three products STENT1, STENT2, and BALLOON using the LSTM and ARIMA, based on Mean absolute error and Root mean square error

Product	Model	MAE	RMSE
Stent1	LSTM 1	0.229	0.3
	ARIMA(8,1,0)	0.204	0.279
Stent2	LSTM 2	0.235	0.335
	ARIMA(0,1,1)	0.119	0.215
Balloon	LSTM 3	0.382	0.477
	ARIMA(8,0,0)	0.333	0.426

Figures 4, 5 and 6 illustrate a sales prediction for the year 2020 of the products STENT1, STENT2, and BALLOON by selected models use in this thesis; a grey line shows a sales prediction by ARIMA, orange line for LSTM and blue line for actual data, so as a result, it indicates that the ARIMA is the best approach for the small dataset of sales prediction for cardiac health products. The relative importance of the data and the consistency of directional breaks explain the graph's performance. The model applies here may yield numbers quite similar to the actual data and demonstrate the model's effectiveness.

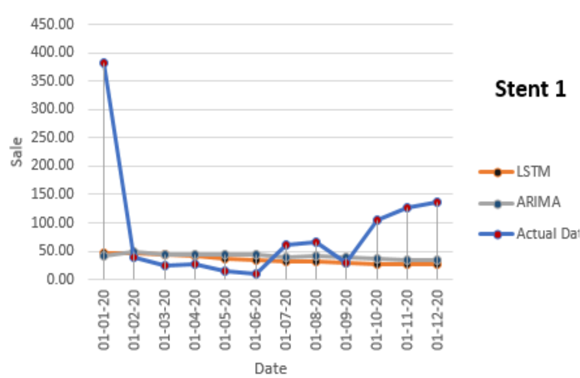


Figure 4: Sales prediction of STENT1 for the year 2020 by selected models

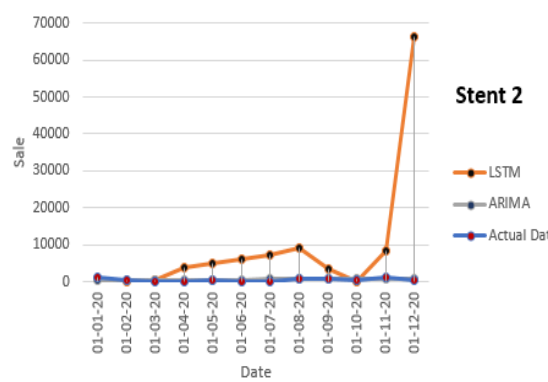


Figure 5: STENT2 sales prediction for the year 2020 for all selected models.

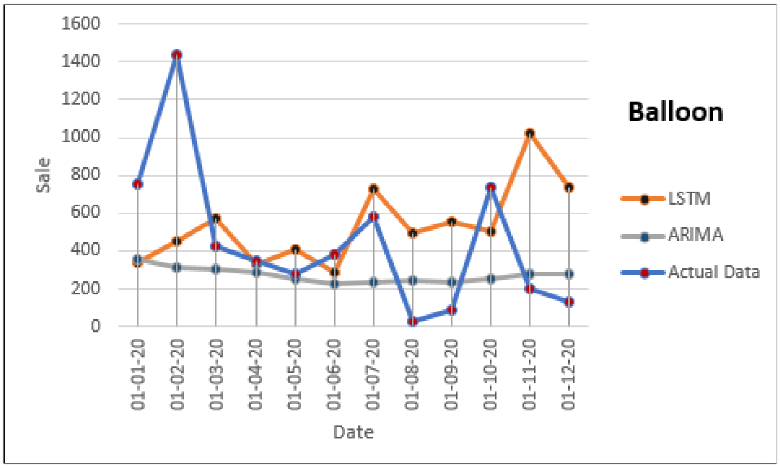


Figure 6: Sales prediction of BALLOON for 2020 by models used in this research.

The experimental findings of our models show that ARIMA has the lowest error rate compared to other models. Resulst conclude that the ARIMA algorithm successfully forecasts cardiac products sale.

Conclusion

Sales revenue analysis can assist in obtaining the information required to predict revenue and profits with sales forecasts the period of a collected dataset from 2017 to 2019, including two stents and one Balloon. This research aims to predict the sale of the products for the year 2020. These proposed approaches contain ARIMA and LSTM, and test upon Cardiac Items (Stent & Balloon) distribution sales data to answer the key variables that affect a solution for predicting sales. Following the application of measures including MAE & RMSE, the ARIMA is determined to be the most suitable algorithm for the data gathered, thus achieving the goal of this study.

These techniques may be applied to an extensive data set and help the organizations in decision making and inventory management by predicting the health products sale base on hospital and location. These performance measures can compare various Machine Learning or Deep Learning.

References

[1] A. Gupta, C. D. Maranas, and C. M. McDonald, "Mid-term supply chain planning under demand uncertainty: customer demand satisfaction and inventory management," *Comput. Chem. Eng.*, vol. 24, no. 12, pp. 2613–2621, Dec. 2000, doi: 10.1016/S0098-1354(00)00617-7.

[2] R. Kaipia, J. Holmström, J. Småros, and R. Rajala, "Information sharing for sales and operations planning: Contextualized solutions and mechanisms," *J. Oper. Manag.*, vol. 52, pp. 15–29, May 2017, doi: 10.1016/J.JOM.2017.04.001.

[3] "Sales Forecasting Methodology: A Beginner's Guide | Anaplan." https://www.anaplan.com/blog/sales-forecasting-guide/ (accessed Jun. 24, 2022).

[4] Y. Niu, "Walmart Sales Forecasting using XGBoost algorithm and Feature engineering," *Proc. - 2020 Int. Conf. Big Data Artif. Intell. Softw. Eng. ICBASE 2020*, pp. 458–461, Oct. 2020, doi: 10.1109/ICBASE51474.2020.00103.

[5] R. A. Hussain. I, Javed. R.R, Qayyum. H., Farman. H, "Detection of Coronary Artery Using Novel Optimized Grid Search-based MLP," *Int. J. Innov. Sci. Technol.*, vol. 4, no. 1, pp. 276–287, 2022.

[6] P. Doganis, A. Alexandridis, P. Patrinos, and H. Sarimveis, "Time series sales

forecasting for short shelf-life food products based on artificial neural networks and evolutionary computing," *J. Food Eng.*, vol. 75, no. 2, pp. 196–204, Jul. 2006, doi: 10.1016/J.JFOODENG.2005.03.056.
[7] D. C. S. Bisht and M. Ram, *Recent advances in time series forecasting*. .
[8] S. Pilemalm, P. O. Lindell, N. Hallberg, and H. Eriksson, "Integrating the Rational Unified Process and participatory design for development of socio-technical systems: a user participative approach," *Des. Stud.*, vol. 28, no. 3, pp. 263–288, 2007, doi: 10.1016/j.destud.2007.02.009.
[9] "time series - Confused on ARIMA's linearity assumption - Cross Validated." https://stats.stackexchange.com/questions/468477/confused-on-arimas-linearity-assumption (accessed Jun. 24, 2022).
[10] I. Khandelwal, R. Adhikari, and G. Verma, "Time Series Forecasting Using Hybrid ARIMA and ANN Models Based on DWT Decomposition," *Procedia Comput. Sci.*, vol. 48, no. C, pp. 173–179, Jan. 2015, doi: 10.1016/J.PROCS.2015.04.167.
[11] I. A. Gheyas and L. S. Smith, "A novel neural network ensemble architecture for time series forecasting," *Neurocomputing*, vol. 74, no. 18, pp. 3855–3864, Nov. 2011, doi: 10.1016/J.NEUCOM.2011.08.005.
[12] "Deep Learning | Guide books." https://dl.acm.org/doi/abs/10.5555/3217337 (accessed Jun. 24, 2022).
[13] A. Soy Temür and Ş. Yıldız, *Comparison of Forecasting Performance of ARIMA LSTM and HYBRID Models for The Sales Volume Budget of a Manufacturing Enterprise*, vol. 50, no. 1. 2021.
[14] P. Newbold, "ARIMA model building and the time series analysis approach to forecasting," *J. Forecast.*, vol. 2, no. 1, pp. 23–35, Jan. 1983, doi: 10.1002/FOR.3980020104.
[15] "Understanding LSTM Networks -- colah's blog." https://colah.github.io/posts/2015-08-Understanding-LSTMs/ (accessed Jun. 24, 2022).
[16] A. Deng *et al.*, "Developing Computational Model to Predict Protein-Protein Interaction Sites Based on the XGBoost Algorithm," *Int. J. Mol. Sci.*, vol. 21, no. 7, Apr. 2020, doi: 10.3390/IJMS21072274.
[17] A. Soy Temür, M. Akgün, and G. Temür, "Predicting housing sales in turkey using arima, lstm and hybrid models," *J. Bus. Econ. Manag.*, vol. 20, no. 5, pp. 920–938, 2019, doi: 10.3846/jbem.2019.10190.
[18] S. Siami-Namini, N. Tavakoli, and A. Siami Namin, "A Comparison of ARIMA and LSTM in Forecasting Time Series," *Proc. - 17th IEEE Int. Conf. Mach. Learn. Appl. ICMLA 2018*, pp. 1394–1401, Jan. 2019, doi: 10.1109/ICMLA.2018.00227.
[19] N. Khalil Zadeh, M. M. Sepehri, and H. Farvaresh, "Intelligent sales prediction for pharmaceutical distribution companies: A data mining based approach," *Math. Probl. Eng.*, vol. 2014, 2014, doi: 10.1155/2014/420310.
[20] A. Elmasdotter and C. Nyströmer, "A comparative study between LSTM and ARIMA for sales forecasting in retail," *Degree Proj. Technol.*, 2018.
[21] C. C. Wang, C. H. Chien, and A. J. C. Trappey, "On the Application of ARIMA and LSTM to Predict Order Demand Based on Short Lead Time and On-Time Delivery Requirements," *Process. 2021, Vol. 9, Page 1157*, vol. 9, no. 7, p. 1157, Jul. 2021, doi: 10.3390/PR9071157.
[22] S. Gopal, K. Patro, and K. Kumar Sahu, "Normalization: A Preprocessing Stage,"

IARJSET, pp. 20–22, Mar. 2015, doi: 10.48550/arxiv.1503.06462.
[23] S. Karasu, A. Altan, Z. Sarac, and R. Hacioglu, "Prediction of Bitcoin prices with machine learning methods using time series data," *26th IEEE Signal Process. Commun. Appl. Conf. SIU 2018*, pp. 1–4, Jul. 2018, doi: 10.1109/SIU.2018.8404760.
[24] W. Thongpeth, A. Lim, A. Wongpairin, T. Thongpeth, and S. Chaimontree, "Comparison of linear, penalized linear and machine learning models predicting hospital visit costs from chronic disease in Thailand," *Informatics Med. Unlocked*, vol. 26, p. 100769, Jan. 2021, doi: 10.1016/J.IMU.2021.100769.
[25] "sklearn.metrics.mean_squared_error — scikit-learn 1.1.1 documentation." https://scikit-learn.org/stable/modules/generated/sklearn.metrics.mean_squared_error.html (accessed Jun. 24, 2022).