

# Improving Sales Forecast Accuracy Using Deep Learning Models: A Comparative Study of LSTM, GRU, and Transformers

Vishnu N.Dabhade<sup>1</sup> Dr.Bharat R.Naiknaware<sup>2</sup> Akshay P.Deshpande<sup>3</sup>

<sup>123</sup>Dr.G.Y.Pathrikar College of CS and IT MGM University Chhatrapati Sambhajinagar.431003

**Abstract**— Accuracy in sales forecasting is an essential element of effective business planning and strategic decision-making. Existing conventional and deep learning models have difficulty in handling the non-linear sales trends, seasonality, and prolonged sales cycles existing in actual sales data. In light of these challenges, this paper seeks to examine the efficacy of deep learning models in enhancing sales forecast accuracy by comparing the efficacy of Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Transformer deep learning models in sales forecasting based on sales data supplemented by additional influence variables like promotions, prices, and seasonality. The different sales data supplemented by additional influence variables, along with a comparative study of various deep learning models, enable the creation of a comprehensive and uniform research framework. In addition to these, extensive data pre-processing, normalization, and sequence creation strategies are employed in this study to ensure a level playing field. The efficiency of the discussed deep learning models is determined based on conventional sales forecasting accuracy metrics, which include Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and Coefficient of determination ( $R^2$ ) of actual and forecast sales. The experimental evaluation reveals that, although LSTM and GRU models.

**Index Terms**—Sales forecasting, extreme gradient boosting, Auto-regressive Integrated Moving Average, support vector machines, machine learning, and long-and short-term memory.

## I. INTRODUCTION

Sales forecasting can be defined as the prediction of actual data for future demands based upon past data. The use of forecasting has great significance in sales prediction. Sales forecasting helps effective decision-makers determine the cost of materials and sales price [1]. In the present scenario, sales prediction is utilized to be more precise and quicker compared to the manual process of predicted

analysis [2]. Management of business acts as a significant part of sales forecasting, which helps in knowing future sales revenues. The sales forecast is updated on a daily, monthly, quarterly, Semi-annual, and yearly basis. The future sales would examine, stock inventory cash flow, and plan for an expansion of this business [30].

Deep learning is a categorization of machine learning. Another categorization of machine learning is artificial intelligence. Sales forecasting can be stated as a process that estimates future sales from the past data of sales. The convolution neural network acts as a structure of machine learning, which determines the sales prediction of this data without requiring humans to extract the sales prediction from the data.

## II. OBJECTIVE

1. To enhance sales performances through advanced analytical models used in improving the accuracy level of sales forecasting.
2. For understanding and analysing the sales loss to identify the forecasting inaccuracies, gaps in demands, and ineffectiveness that affect sales.
3. To study the limitations of the traditional sales forecasting approach in handling the non-linear patterns within the sales data.
4. The development of sales forecasting models based on deep learning methods LSTM, GRU, and Transformer.
5. It should compare and assess the accuracy of prediction for LSTM, GRU, and Transformer models based on typical metrics of performance such as MAE, RMSE, MAPE, and  $R^2$ .
6. To evaluate the resilience of the proposed models with respect to different sales aspects such as seasonality and promotion.

7. In order to determine the best possible architecture that a deep learning model can take in a real-world scenario for sales forecasting[8].

### III. LITERATURE SURVEY

Zhan-Li<sup>1</sup> This paper used a real dataset from a fashion factory. Datasets will be used for retail fashion design. The methodology of those several forecasting methods is based on back propagation neural network technologies. It should be the main contribution of the article to use an extreme machine learning model. Real data is used on Honk-Kong and we compare this data to other sales data. Using for this dataset various types their attribute, colour, size and price of that jacket. Extreme models give best results for this article. In this, the extreme learning model is more accurate and efficient in prediction for this jacket dataset [1].

Sharma Suresh Kumar<sup>1</sup> This article proposes historical data for various models - average and moving average artificial neural networks. Data collection from various sources. Sales forecasting at review is a traditional sales forecasting method that involves average, smoothing and regression analysis. The best article result in this paper is a combination of moving averages and artificial neural networks; the more accurate result is in terms of that accuracy [2].

Chi-Jie Lu<sup>1</sup> In this article, we employed the data clustering-based forecasting model for computer retail sales. The primary model employs the clustering machine learning algorithms and accuracy to elevate the sales forecast of personal computers, notebook computers, and liquid crystal displays. The model to be used can be applied to the different computer products, such as personal computers, notebook computers, or liquid crystal displays. The proposed clustering-based forecast model incorporates different methodologies in the sales forecast of computer products. The different methodologies employed in the study, including the clustering methodologies, support vector machines, as well as extreme machine learning, can be actual data. Improvement can be made by the use of clustering methodologies and machine learning to optimize the sales forecast of the computer products [3].

Dassanayake<sup>1</sup>'s ARIMA model in this article is used to forecast sales for FMCG. Sales forecasting

provides the better basis for efficient sales, better product returns, and hence efficient product planning also. ARIMA model uses time series method and FGMS are used with average absolute error. Dataset present in daily sales information to forecast future sales differently for fast-moving consumer goods. The ARIMA model has an average absolute error of 0.66 units. ARIMA model is the best result for that sales data prediction [4].

Fan, Zhi-Ping<sup>1</sup> In this paper, the combination technique of the Bass/Norton model with sentiment analysis techniques to predict the sales of products based on the data from the products' sales is highlighted. Then, the Naive Bays algorithm was used to obtain the coefficient of limitations within the Bass/Norton model. The data used is the real-time industry data that gives the forecasting capacity of the technique proposed. Sales forecasting is the conventional technique that gives the predictive information about future requirements. In fact, this study may be the time series analysis of the regression analysis to enhance the accuracy of the sales forecasting technique. The data used within the Bass/Norton sentiment analysis study and the past data are the industry data at its current state. A combination technique that utilizes the Bass/Norton approach with the sentiment analysis technique resulted in an improved accuracy within the sales forecasting technique. Naïve Bays approach gives the best performance output of the other technique [5].

In the proposed method, Rossmann used data from pharmaceutical companies. It's implemented using different machine-learning algorithms. The various models used in this regards are ARIMA, ARNN, XG Boost, SVM, and Hybrid. The ARIMA model is used in a composite model based on previous researches done in front mobility. The training dataset contains the past sales and supplement information about the drugs stored. The result obtained is in graphical and tabular format, while the results of other nonlinear modes include neural networks, XG Boost, and support vector machines. The best result of the predicted sales of drugs supported by other machine learning models [6].

Torres<sup>1</sup> offers the actual data employed by the power sector within the Spanish country for forecasting based on deep learning methods. It has the potential employed within the neural network

technique applicable to the study of time series data at a large scale. New big data methods apply time series forecasting methods. It could be achieved with an overall perspective on the universal methodology without any complexities involved within the forecasting technique itself. The actual time series dataset could be employed within the time series forecasting technique within the context of the big data concept. A practical dataset is employed within the comparison regarding the power usage within the Spanish country at the 10-minute sampling rate from the year 2007 to 2016, based on the performance stability within the proposed system compared to other methods [7].

Cheriyani<sup>1</sup>, These data mining techniques are applied to the use of sales prediction to optimize the sales by increasing the efficiency in sales prediction through the use of a gradient boost model. The data applied is utilized in the commercial application. The data applied is utilized in the sale records of the fashion stores for a period of three years, 2015 to 2017. The data previously belonged to the final dataset, which is smaller compared to the original dataset. The data should be utilized in the application used in the study regarding the compressible prediction. The best result in the article is the accuracy level of each class as well as the confusion matrix given in the article [8].

Krishna<sup>1</sup> presented the sales data in retail store and adopted different machine learning techniques for modeling. In the research work, this paper adopts multiple techniques such as polynomial regression, multiple regression Ridge\_regression, Lasso\_regression, and boosting models for modeling and sales forecasting in retail store sales prediction. It is possible to adopt the proposed approach for implementing normal regressions and boosting models for sales prediction of retail stores. The boosting approach achieves the best possible accuracy and is superior to standard regressions models [9].

Pavlyshenko<sup>1</sup>, In this study, it is suggested that the Rossmann store sales dataset, the machine learning model for this sales prediction analysis, and the stacking methods used to improve the quality of the prediction model for the sales time series algorithm. The application of the logistic regression model in highly imbalanced data. The prediction for this outcome on that validation set applied as input

regression. In this proposal, the present study employs Python libraries such as sklearn, matplotlib, numpy, panda, and seaborn in its examination on this outcome. It must take into consideration the impact of the application in machine learning. The outcome on this study on the application used in sales prediction can be applied into another data set for further investigation since we also considered the approach in application in other case study on different products that can be applied in the field. The stacking methods enhance the prediction approach for sales time series. The generalization in applying a machine learning algorithm on sales prediction, where there's a small amount of data on a particular product that can be applied, which in this proposal can also be new that can be applied in the market [10].

Pavlyshenko<sup>1</sup> in this paper applies the steps of machine learning to improve the accuracy and efficiency of sales forecasting of the data mining technique. The sales forecasting method has been applied to e-fashion stores for three consecutive years. The attributes for sales forecasting include city, types of product and subscriptions, quality, and sales of the quarter. The algorithm is applicable to gradient boost algorithms, and the data set may be applicable to sales data. The data should be applicable to gradient boost algorithms to determine the predictive value and retail of that sales data. This may apply to boost, a decision tree, and a linear model. The gradient boost algorithm has the highest predictive accuracy of sales forecast data, which has an accuracy of 98% in sales forecast data [11].

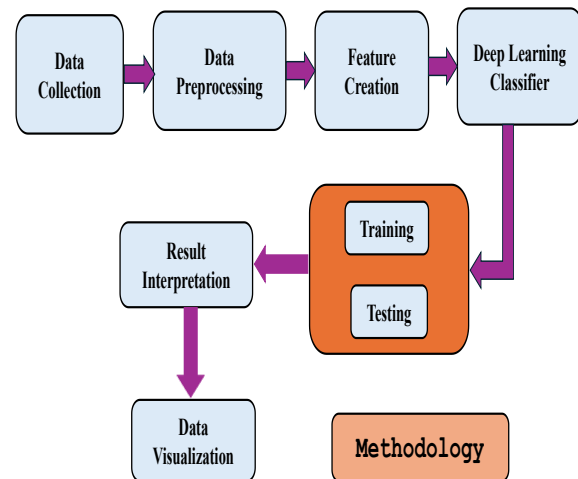
Chandadevi<sup>1</sup>, in proposes a method for the use of an intelligence forecasting system based on the combination of the image feature characteristic of clothes and sales data to forecast and predict future demands within the fashion retailing industry. For this study, we present a dataset for the sales of images used for the processing of old sales image data and the Classify forecast system. The forecast system may enable a forecast of future demands for new fashion-attractive products and thus enable better management of stock levels and cut down on costs to maximize profits. The forecast for the prediction solution provided for the forecast system is developed using machine learning approaches to derive key features of the product image. It may enable a forecast for weekly sales of new fashion based on the best fit within the cluster of product

images. The dataset was employed for two years between 2015 and 2016 for sales data of 290 items of a fashion retail company for sales data. The dataset may also form part of data either for numerical features or images of pixels. It may be altered to a vector for mathematical modelling for the forecast system [12].

In this paper, proposed by Sajawall, there is predicted analyses of sales data by means of a Machine Learning model and usage of these techniques. Sales data from 2013 to 2018 were processed by XG Boost techniques, Time Series, and various regression techniques to derive the best results. It should elaborate and usage of Machine Learning techniques for analyses and improvements in Retail Sales Forecasting and improvement techniques for management and finances and usage and improvement techniques. In this paper, techniques applied include Linear Regression Analysis, Extreme Boost Regression Analysis, andadapelandon Forest Regression Analysis to predict Retail Sales. Analysis by these python techniques, Pandas, Matplotlib, and Sklearn libraries usage includes Mean Average Error and Root Mean Squared Error values to evaluate predictive results into these Models. The result obtained from this article analyses show that XG Boost and Outperform with techniques and usage of Time Series regression and various regression techniques and usage for predictive analyses and Retail Sales Forecasts. XG Boost techniques perform very well with accuracy and show MAE value 0.516 and RMSE value 0.63 with index value [13].

#### IV. METHODOLOGY

Under this proposed methodology, we are researching different steps, such as data gathering. There are several steps in this methodology. Data pre-processing, feature extraction, deep learning classifiers, data training, data testing, result interpretation, and data visualization: There are several sources of data gathering. These sources are data sources, text sources, image sources, and number sources.



##### Data Collection

Thus, in this data collection method, we are appropriately chosen for this kind of research objective. It is applied in collecting this data and picking and in this process, the required data was collected in a standard manner. There exist numerous ways through which datasets can be classified, such as textual data, numerical data, label data, and organizational data in the form of industry data.

##### Data Pre-Processing

In this data pre-processing technique, the selected data gets pre-processing. Here, in this pre-processing data collection of past sales data from online databases, the data should have the following features, if not, they would affect sales, such as promotion, marketing, and economic variables. With the aim of this pre-processing technique, the data gets cleaned and processed to deal with the missing values by the methods mean, median, and interpolation. It has to eliminate the duplicate files and data from this dataset. Points to be identified and dealt with for this forecasting. The technique of splitting is applied on this dataset. The data preprocessing technique applied here involves scaling and data normalization, and the feature selection.

##### Feature Creation

In this proposed method for creating features, we are actually searching for this new feature, and this particular data we have chosen to predict sales. Generally speaking, the reason for creating features is to develop a new attribute for this data. Creating features is also referred to as featuring engineering. The procedure for creating a new attribute data

feature for this data to improve reprocessing to give better results to this deep learning model includes creating a new attribute data.

#### Deep Learning Classifier.

In this method, deep learning classifiers are used. A deep learning classifier is a type of deep learning method. Deep neural networks for the classification of these problems. A deep learning classifier consists of deep neural networks that are formed by layers of interconnected neurons. Deep learning classifier for sales prediction includes development of a deep learning classifier that uses predictions on historical information and other relevant factors. The deep learning classifier chooses various machine learning algorithms to form the neurons of the deep neural networks. Convolutions neural networks are appropriate for this reprocessing of the numerical information for spatial information.

#### Training Data

In this training, we design a method of selecting data. Training data is the process of using a machine learning model to train a predicted model. It is important to gather this dataset, which includes the standard dataset. It will consist of a feature of the input of this attribute along with a target of this label. It can apply data processing methods to add some noises to the data by considering missing values as clean data. After gathering the training data, the machine learning model will then test the data.

#### Testing Data

The data for the testing is the context for the machine learning analysis of the data. The data for the testing can be applied for the data that the machine learning model has been exposed to during the course of its testing. The application of the data for the purposes of the testing has been distinguished from that of the validation data. The machine learning model is now ready for use in the real world. The data for the testing plays a critical role in the sales prediction. The accuracy and reliability play a critical role for the sales prediction model.

#### Result Interpretation

In this case, the interpretation of the result is the outcome of the sales forecasting model and an important step in comprehending the prediction of the sales forecasting model. The foremost concern

of the sales forecasting model lies in its accuracy in the predictive work. It might measure the average error in the prediction of the units of the sale. The mean squared error calculates the average error in the squared form. The root mean squared root gives a measure in the units of the sale data, and lastly, the squared measure calculates the explanatory power in the variance of the sale.

#### Visualization of Data

One of the most important proposed data visualization methods is for sales prediction. A graphical representation of data helps in understanding, analyzing, and interpreting the trend and insight that could be held within the data. A machine learning model is a powerful tool that conveys complicated information in a visually appealing and understandable manner.

### V. CONCLUSION

In this research work, we are undertaking research on how we can enhance sales productivity. We are undertaking research on different models that can enhance sales prediction in practical scenarios. It can apply different methodologies to determine which model gives the best sales prediction. It can apply on the sales forecasting dataset in computer retailing. The model of clustering offers the best accuracy model on sales prediction other than notebook computers and an LCD item. The best outcome is obtained when a support vector machine can be obtained on other models in machine learning. Ross Man Wall Mart is applying a dataset on Drugs Company for forecasting sales. Applying a model on drug data using a machine learning model, but the best results are obtained on the XG Boosting algorithm other than the machine learning model. The datasets applied on retail sales. Six years dataset on applying retail sales forecasting data. Six-year retail sales data apply in this dataset. XG Boosting algorithm applies on best results other than applying on machine learning models. Several difficulties can occur on the sales prediction process using technology. Different difficulties occur on the current market trend scenario along with new product emergence.

### REFERENCES

- [1]. Sun, Zhan-Li, Tsan-Ming Choi, Kin-Fan Au, and Yong Yu. "Sales forecasting using extreme

- learning machines with applications in fashion retailing." *Decision Support Systems* 46, no. 1 (2008): 411-419.
- [2] Sharma, Suresh Kumar, and Vinod Sharma. "Comparative analysis of machine learning techniques in sale forecasting." *International Journal of Computer Applications* 53, no. 6 (2012): 975–8887.
- [3]. Chen, I-Fei, and Chi-Jie Lu. "Sales forecasting by combining clustering and machine-learning techniques for computer retailing." *Neural Computing and Applications* 28 (2017): 2633-2647.
- [4] Dassanayake, Dilini Madushika Gamage Tennekoon, Ashani Nugaliyadde, and Yasitha Mallwaraachchi. "Increasing sales productivity through sales
- [5]. Fan, Zhi-Ping, Yu-Jie Che, and Zhen-Yu Chen, "Product sales forecasting using online reviews and historical sales data: A method combining the Bass model and sentiment analysis," *Journal of Business Research* 74 (2017): 90–100.
- [6]. Gurnani, Mohit, Yogesh Korke, Prachi Shah, Sandeep Udmale, Vijay Sambhe, and Sunil Bhurud. "Forecasting of sales by using a fusion of machine learning techniques," in 2017 *International Conference on Data Management, Analytics, and Innovation (ICDMAI)*, pp. 93–101. IEEE.
- [7] Torres, José F., Antonio Galicia, Alicia Troncoso, and Francisco Martínez-Álvarez. "A scalable approach based on deep learning for big data time series forecasting." *Integrated Computer-Aided Engineering* 25, no. 4 (2018): 335–348.
- [8]. Cheriyan, Sunitha, Shaniba Ibrahim, Saju Mohanan, and Susan Treasa. "Intelligent sales prediction using machine learning techniques," in the 2018 *International Conference on Computing, Electronics, and Communications Engineering (iCCECE)*. pp. 53-58. IEEE,
- [9] Krishna, Akshay, V. Akhilesh, Animikh Aich, and Chetana Hegde. "Sales-forecasting of retail stores using machine learning techniques," in the 2018 3rd international conference on computational systems and information technology for sustainable solutions (CSITSS). IEEE, 2018.
- [10]. Pavlyshenko, Bohdan M. "Machine-learning models for sales time series forecasting." *Data* 4, no. 1 (2019):
- Electronics, and Communications Engineering (iCCECE), pp. 53-58. IEEE, 201
- [12]. Giri, Chandadevi, and Yan Chen. "Deep Learning for Demand Forecasting in the Fashion and Apparel Retail Industry." *Forecasting* 4, no. 2 (2022): 565-581.
- [13]. Sajawal, Muhammad, Sardar Usman, Hamed Sanad Alshaikh, Asad Hayat, and M. Usman Ashraf. "A Predictive Analysis of Retail Sales Forecasting Using Machine Learning Techniques," *Lahore Garrison University Research Journal of Computer Science and Information Technology* 6, no. 04 (2022): 33–45.
- [14]. Sengar, Ravindra Singh, Dr. F. Ahmed and Bhopal SIRT, "Review on Trends in Machine Learning Applied to Demand and Sales Forecasting", *SMART MOVES JOURNAL IJOSCIENCE* 5 no. 6, 2019.
- [15]. Wang, Haiping. "An Insurance Sales Prediction Model Based on Deep Learning," *Rev. d'Intelligence Artif.* 34, no.3 (2020): 315–321.
- [16]. Dong, Shi, Ping Wang, and Khushnood Abbas, "A survey on deep learning and its applications." *Computer Science Review* 40 (2021): 100379.
- [17]. Ramya, B. Sri Sai, and K. Vedavathi. "An advanced sales forecasting using a machine learning algorithm." *International Journal of Innovative Science and Research Technology* 5, no. 5 (2020): 342–345.
- [18]. Rohaan, D., E. Topan, & Catharina GM Groothuis-Oudshoorn. "Using supervised machine learning for B2B sales forecasting: A case study of spare parts sales forecasting at an after-sales service provider." *Expert Systems with Applications* 188 (2022): 115925.
- [19] Bandara, Kasun, Peibei Shi, Christoph Bergmeir, Hansika Hewamalage, Quoc Tran, and Brian Seaman, "Sales demand forecast in e-commerce using a long short-term memory neural network methodology," in *Neural Information Processing: 26th International Conference, ICONIP 2019, Sydney, NSW, Australia, December 12–15, 2019, Proceedings, Part III* 26, pp. 462-474. Springer International Publishing,
- [20].Cheriyan, Sunitha, Shaniba Ibrahim, Saju Mohanan, and Susan Treasa. "Intelligent sales prediction using machine learning techniques," in the 2018 *International Conference on Computing, Electronics, and Communications Engineering (iCCECE)*, pp. 53–58. IEEE
- [21]. Xue, Ning, Isaac Triguero, Graziela P. Figueredo, and Dario Landa-Silva. "Evolving deep

- CNN-LSTMs for inventory time series prediction," in 2019 IEEE Congress on Evolutionary Computation (CEC), pp. 1517–1524. IEEE, 2019.
22. Loureiro, Ana LD, Vera L. Miguéis, and Lucas FM Da Silva. "Exploring the use of deep neural networks for sales forecasting in fashion retail." *Decision Support Systems* 114 (2018)
- [22]. Park, Byeonghwa, and Jae Kwon Bae. "Using Machine Learning Algorithms for Housing Price Prediction: The Case of Fairfax County, Virginia Housing Data." *Expert Systems with Applications* 42.6 (2015): 2928-2934.
- [23] Rohaan, D., E. Topan, and Catharina GM Groothuis-Oudshoorn. "Using supervised machine learning for B2B sales forecasting: A case study of spare parts sales forecasting at an after sales service provider." *Expert Systems with Applications* 188 (2022): 115925.
- [24]. Pondel, Maciej, Maciej Wuczyński, Wiesława Grynciewicz, Łukasz Łysik, Marcin Hernes, Artur Rot, et al. "Deep Learning for Customer Churn Prediction in E-commerce Decision Support." In *Business Information Systems*,
- [25]. Bandara, Kasun, Peibei Shi, Christoph Bergmeir, Hansika Hewamalage, Quoc Tran, and Brian Seaman, "Sales demand forecast in e-commerce using a long short-term memory neural network methodology," in *Neural Information Processing: 26th International Conference, ICONIP 2019, Sydney, NSW, Australia, December 12–15, 2019, Proceedings, Part III* 26, pp. 462-474. Springer International Publishing.
- [27] Zeynep Hilal Kilimci, A. Okay Akyuz, Mitat Uysal, Selim Akyokus, M. Ozan Uysal, Berna Atak Bulbul, and Mehmet Ali Ekmis, "An improved demand forecasting model using a deep learning approach and a proposed decision integration strategy for the supply chain." *Complexity* 2019, 2019.
- [28]. Venishetty, Sai Vineeth. 'Machine learning approach for forecasting the sales of truck components' (2019).
- [29]. Hinton, Geoffrey E., and Ruslan R. Salakhutdinov. "Reducing the dimensionality of data with neural networks." *Science* 313, no. 5786 (2006)
- [30] A. Amir, W. K. H. W. Khairul, A. B. M. Soom, A. Mat Jasin, J. Ismail, and A. Asmat, "Sales Forecasting Using Convolution Neural Network," *Journal of Advanced Research in Applied Sciences and Engineering Technology* 30, no. 3 (2023): 290–301.