International Research Journal of Education and Technology
Peer Reviewed Journal
ISSN 2581-7795

SALES FORECASTING OF NON-STATIONARY TIME SERIES SALES DATA USING DEEP LEARNING APPROACH

Mr. Shanmugam K[1], Anish Akshai S[2], Brahadeeshram R[3] , Harish P[4]

Assistant Professor, Department of Computer Science and Engineering[1]
UG Scholar, Department of Computer Science and Engineering[2]
UG Scholar, Department of Computer Science and Engineering[3]
UG Scholar, Department of Computer Science and Engineering[4]
SRM Valliammai Engineering College, Chengalpattu, Tamil Nadu, India

---***---

Abstract - *Sales forecasting is a critical component of business operations, as accurate predictions can help enterprises to improve their operations and make more informed decisions. With the development of artificial intelligence (AI), there are increasing methods to solve the forecasting problem. One of these methods is the use of the Long short-term memory (LSTM) model, an artificial recurrent neural network (RNN) architecture used in the field of deep learning. The LSTM model is particularly useful for time-series regression problems, such as sales forecasting, as it can process entire sequences of data rather than just single data points. This means that it can remember information from previous data points in the sequence and use it to inform its predictions for future data points. To optimize the performance of the LSTM model for sales forecasting, the proposed system involves a number of steps. First, the data is subjected to Exploratory Data Analysis (EDA) and graphical visualization techniques to identify patterns and trends in the data. This allows the model to better understand the underlying structure of the data and make more accurate predictions. Once the patterns and trends have been identified, the data is pre-processed to optimize the structure of the input data sets. This might involve data cleaning, normalization, or other techniques to ensure that the data is in a suitable format for the LSTM model. The deep learning approach used in the proposed system is specifically designed for forecasting non-stationary time series data. This type of data involves predicting future values based on previously observed values, where the underlying patterns and trends in the data may change over time. To account for this, the neural network model incorporates time trend correction, which helps to adjust for any underlying trends or patterns in the data that may impact future sales. Overall, the proposed system is designed to provide accurate and reliable sales forecasts that can help enterprises to improve their operations and make more informed decisions. By combining the power of the LSTM model with EDA, graphical visualization, and data pre-processing techniques, the system is capable of handling complex time-series data and generating accurate predictions that can help businesses stay ahead of the competition.*

INTRODUCTION

Shopping malls and Big Marts collect and store individual item sales data to forecast future client demand and adjust inventory management accordingly. These data stores hold a significant amount of consumer information and particular item details. By mining the data store from the data warehouse, more anomalies and common patterns can be discovered, which can be used to make more accurate predictions. Most of the buying decisions are not based on well-defined logic. Emotions, trust, communication skills, culture, and intuition play a big role in our buying decisions. Although humans do not follow a well-defined logic, there are some repeated patterns that we follow that vary for different people. Consumers often buy the same things and behave in a similar way. Therefore, it is important to understand these patterns and use them to make more informed decisions. In recent years, deep learning algorithms have become increasingly popular for analyzing sales data. Among these, neural networks are one of the most widely used algorithms. Neural networks are particularly useful because they can create an approximation of any function based on the data that they are trained with. This means that they can learn similar responses for similar inputs, which makes them well-suited for analyzing sales data and making accurate predictions.

Sales forecasts help businesses make better decisions based on future revenue, which can help them to forecast likely profit or loss in a designated period. By using sales forecasts, businesses can adjust their inventory management to ensure that they have enough stock to meet demand without overstocking and incurring unnecessary costs. They can also use the forecasts to optimize pricing strategies and marketing campaigns to drive sales and increase revenue. Overall, analyzing sales data using deep learning algorithms like neural networks can provide valuable insights into consumer behavior and help businesses make more informed decisions. By understanding the patterns and trends in sales data, businesses can optimize their operations and increase their revenue, making them more competitive in the marketplace.

EXISTING SYSTEM

Multiple Machine Learning solutions in Industry exist where interpretability is required. In retail, this is especially important when dealing with cold-start forecasting of promotional sales. In the planning phase of these promotions, retailers produce sales predictions that are scrutinized by both forecasting experts and managers. The system combines the predictive benefits of Gradient Boosted Decision Trees regressors and the interpretability of contrastive explanations. These are implicitly generated by the manner we shape data. The existing method presents the cold-start forecasts in relation to the observed promotional sales of other products, which we call neighbors. These are selected based on a measure of closeness to the predicted promotion, which is derived from the variable importance calculated during the training of the regressors. The existing system for cold-start forecasting of promotional sales in retail is based on the use of Gradient Boosted Decision Trees regressors and contrastive explanations for interpretability. This approach has some advantages, such as the ability to handle large datasets and provide explanations for the predictions made by the model. Additionally, it can be useful for providing a measure of closeness between the predicted promotion and other products, which can aid in decision-making for the retailer. However, there are also some limitations to this approach. One of the main disadvantages is that it may not be suitable for dealing with non-linear relationships between variables, as decision trees are generally not well-suited for modeling complex interactions. Additionally, this approach may struggle with time-series forecasting, which is a key aspect of sales forecasting in retail. This is because decision trees are not designed to capture temporal dependencies in the data, which are often important for predicting future sales.

In contrast, the Long short-term memory (LSTM) model proposed in our project is specifically designed for time-series forecasting and can capture both short-term and long-term dependencies in the data. This makes it well-suited for the sales forecasting task in retail, where historical sales data is often used to predict future sales.

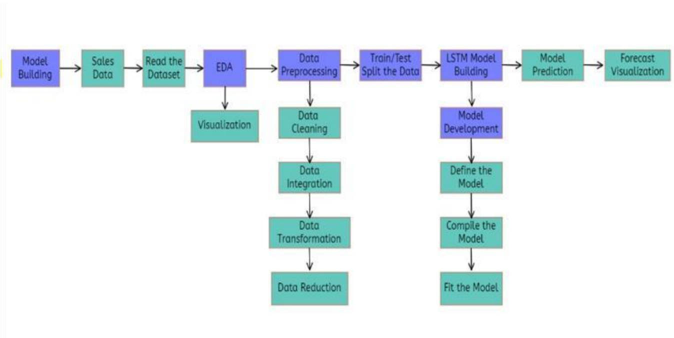
1. PROPOSED SYSTEM

The proposed system presents a unified approach for prediction and modelling of time series collections that naturally tackles prediction of new series and missing data imputation using both high-dimensional metadata and shared seasonality structure. Forecasting entire seasonal profiles for unobserved time series is novel. The key to this approach is harnessing repeated patterns over fixed periods of time. Via matrix factorization, we extract a low-dimensional representation of these repeated patterns seen across periods and time series, enabling us to efficiently form long-range predictions over the next period, as well as impute missing values. The model includes a subnetwork block for the prediction weight for the time trend term which is added to the predicted sales value. The time trend term is considered as a product of the predicted weight value and normalized time value. The predicted sales values and the time trend term are combined in the loss function. As a result, one can receive an optimized trend correction for nonstationary sales for different groups of data with different trends. The proposed sales forecasting system using LSTM can learn the complex relationships between different factors that influence sales patterns, making it more effective in forecasting non-stationary time series. LSTM models are particularly good at capturing trends, seasonality, and irregularities in time-series data. An LSTM model consists of a network of nodes that can store information over time, making it particularly useful for analyzing time-series data. In the context of sales forecasting, an LSTM model can learn patterns in the historical sales data and predict future sales based on those patterns. The model can take into account various factors that influence sales, such as seasonality, promotional activities, and economic indicators. In summary, using LSTM models for sales forecasting can improve the accuracy of predictions and make the forecasting process more efficient. The proposed system can automate the forecasting process and adapt to new data, making it a valuable tool for businesses to make informed decisions. The proposed system can be used in various industries, including retail, e-commerce, and manufacturing, where sales forecasting plays a critical role in business operations. By accurately predicting sales, businesses can optimize their inventory management, production planning, and marketing strategies. The system can also help businesses to identify the factors that drive sales and make data-driven decisions. In addition, the

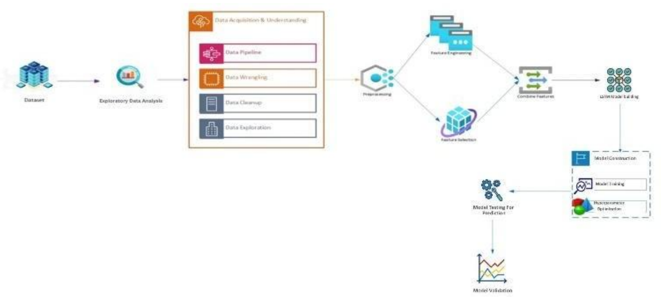
International Research Journal of Education and Technology
Peer Reviewed Journal
ISSN 2581-7795

system can be adapted to new data, which is essential for businesses that operate in a dynamic environment. As new data becomes available, the system can automatically update its models and improve its predictions. This makes the system more efficient and reduces the need for manual intervention. Overall, the proposed LSTM-based sales forecasting system offers a more accurate and efficient way of forecasting sales compared to traditional statistical methods. By leveraging the power of deep learning, businesses can make better-informed decisions and stay ahead of the competition.

FLOW DIAGRAM DESIGN



ARCHITECTURE DIAGRAM DESIGN



LSTM MODEL DEVELOPMENT MODULE

In this module, The LSTM model is defined to train a neural network that can accurately predict sequential data. the development of the LSTM model involves several steps. Firstly, we preprocess the data by converting the time-series data into a supervised learning problem. This involves transforming the data into input/output pairs, where the input is a window of historical sales data, and the output is the sales value at the next time step. Once the data is preprocessed, we define the LSTM model function using the Keras API. The model architecture includes an input layer, a hidden layer with the ReLU activation function, and an output layer with the sigmoid activation function. We also include dropout regularization to prevent overfitting. After defining the model function, we compile the model by specifying the optimizer, loss function, and evaluation metric. We use the Adam optimizer, which is a variant of stochastic gradient descent that adapts the learning rate during training. For the loss function, we use mean squared error, which measures the difference between the predicted and actual sales values. We evaluate the model performance using the mean absolute percentage error metric, which measures the percentage difference between the predicted and actual sales values. Once the model is compiled, we fit it to the preprocessed data using the fit() method. During training, the weights of the neural network are adjusted based on the difference between the predicted and actual sales values, in order to improve the model's accuracy. After training, we can use the LSTM model to make predictions on new data. We can input a window of historical sales data into the model, and it will output a predicted sales value for the next time step. The LSTM model can also be used to impute missing values in the time-series data. Overall, the development of the LSTM model in our project involves data preprocessing, model definition, compilation, training, and prediction. By using the LSTM model, we can accurately forecast sales values and make informed business decisions.

LSTM MODEL PREDICTION MODULE

In this module, The LSTM model that was trained on a historical dataset is now used to make predictions on new data. Making predictions using the LSTM model is important because it allows us to use the insights gained from analyzing historical data to predict future outcomes. the trained LSTM model is used to make predictions on new data, which represents future time periods. The input data is preprocessed in the same way as the training data, which involves scaling and reshaping the data to match the input shape expected by the LSTM model. The model is then used to generate predictions for each time step in the future period. The output of the model is a sequence of predicted values, which can be post-processed and analyzed to gain insights into future trends and patterns. Once the predictions have been generated, they can be visualized and analyzed using various techniques, such as time series plots, histograms, and statistical analysis. This allows us to identify patterns and trends in the predicted data, and make informed decisions about future actions. For example, if the predicted sales trends show an increase in demand for a certain product, a business can adjust its marketing and inventory strategies accordingly. Overall, the LSTM model prediction module plays a critical role in leveraging the power of deep learning to generate accurate predictions for future outcomes. This enables businesses and organizations to make more informed decisions, improve efficiency, and mitigate potential risks. This can be useful in a wide range of fields, from finance to healthcare to weather forecasting. By accurately predicting future outcomes, we can make more informed decisions and take proactive measures to mitigate potential risks. Businesses can gain valuable insights into future sales trends and make more informed decisions about inventory management, marketing campaigns, and resource allocation.

FORECAST VISUALIZATION MODULE

The module that visualizes the forecasted data obtained using the LSTM model is responsible for presenting the predictions in a graphical format. This module takes the predicted results from the previous step and creates a plot that shows how the predicted values compare to the actual values over time. Visualizing the predicted results is important because it allows us to see how well our LSTM model performed in forecasting future values. By comparing the predicted values to the actual values, we can assess the accuracy of our model and identify any areas where it may need improvement. The forecast visualization module in our project is designed to provide users with an intuitive and interactive way to explore and analyze the results of our LSTM model's predictions. The module takes the forecasted values generated by the model and displays them in a graphical format, making it easy to visualize how the predicted values compare to the actual values over time. The module uses a popular Python library, Matplotlib, to create visualizations. Matplotlib provides a wide range of tools for creating charts, graphs, and other visualizations, and can be customized to suit the specific needs of our project.

RELATED WORKS

a) Cold-Start Promotional Sales Forecasting Through Gradient Boosted-Based Contrastive, Carlos Aguilar-Palacios, 2020

This Multiple Machine Learning solutions in Industry exist where interpretability is required. In retail, this is especially important when dealing with cold-start forecasting of promotional sales. In the planning phase of these promotions, retailers produce sales predictions that are scrutinised by both forecasting experts and managers. In this paper, we combine the predictive benefits of Gradient Boosted Decision Trees regressors and the interpretability of contrastive explanations. These are implicitly generated by the manner we shape data. Our method presents the cold-start forecasts in relation to the observed promotional sales of other products, which we call neighbours. They are selected based on a measure of closeness to the predicted promotion, which is derived from the variable importance calculated during the training of the regressors. With this information, the expert reviewer adjusts the cold-start prediction by simply varying the contribution of the neighbours. To validate our results, we test our method on a surrogate model, as well as on real-market data. The results on the surrogate model demonstrate that our method is able to accurately identify the features that contribute to sales and then select the closest neighbours to produce a contrastive explanation. The results on real-market data also show that the proposed method performs at a similar level to widespread methods such as conventional CatBoost, NGBoost or AutoGluon, and has the advantage of providing interpretability.

Cold-start forecasting is a difficult but common problem in a variety of disciplines as no historical information exists on the subject of the forecast. This problem is particularly challenging in retail promotions, as the complexity of promotional forecasting is added to the cold-start difficulty. In this paper, we align with authors such as Rudin in the statement that we should move away from black-box models where the decisions made by automated systems require interpretability. Furthermore, in scenarios as such, the lack of interpretability directly affects the usability of a prediction as it inherently produces mistrust. Given the nature of the grocery retail business, forecasting affects many elements of the supply chain, from producers, suppliers, distributors, and customers.

b) Predicting Vehicle Sales by Sentiment Analysis of Twitter Data and Stock Market Values, Ping-Feng Pai, 2018

The Owning to the booming of social media, making comments or expressing opinions about merchandises online becomes easier than before. Data from social media might be one of the essential inputs for forecasting sales of vehicles. Besides, some other effects, such as stock market values, have influences on purchasing power of vehicles. In this paper, both multivariate regression models with social media data and stock market values and time series models are employed to predict monthly total vehicle sales. The least squares support vector regression (LSSVR) models are used to deal with multivariate regression data. Three types of data, namely sentiment scores of tweets, stock market values, and hybrid data, are employed in this paper to forecast monthly total vehicle sales in USA. The hybrid data contain both sentiment scores of tweets and stock market values. In addition, seasonal factors of monthly total vehicle sales are employed to deseasonalizing both monthly total vehicle sales and three types of input data. The time series models include the nave model, the

exponential smoothing model, the autoregressive integrated moving average model, the seasonal autoregressive integrated moving average model, and backpropagation neural networks and LSSVR with time series models. The numerical results indicate that using hybrid data with deseasonalizing procedures by the LSSVR models can obtain more accurate results than other models with different data. Thus, both social media data and stock values are essential to forecast monthly total vehicle sales; and deseasonalizing procedures can improve forecasting accuracy in predicting monthly total vehicle sales. This study presented a framework that consists of both time series forecasting models and
multivariate regression technique to predict monthly total vehicle sales. Deseasonalizing procedures were employed to deal with different types of data. The numerical results indicate that forecasting vehicle sales by hybrid multivariate regression data with desonalizing procedures can obtain more accurate forecasting results than other forecasting models. The superior forecasting performance could be concluded as follows. First, the use of hybrid data containing sentiment analysis of social media and stock market values can improve the forecasting accuracy. Secondly the deceasonalizing procedures both in condition variables and decision variables do help to increase the prediction performance.

c) Forecasting Promotional Sales Within the Neighbourhood, Carlos Aguilar-Palacios, 2019

Promotions are a widely used strategy to engage consumers and as such, retailers dedicate immense effort and resources to their planning and forecasting. This paper introduces a novel interpretable machine learning method specifically tailored to the automatic prediction of promotional sales in real-market applications. Particularly, we present fully automated weighted nearest neighbours where the distances are calculated based on a feature selection process that focuses on the similarity of promotional sales. The method learns online, thereby avoiding the model being retrained and redeployed. It is robust and able to infer the mechanisms leading to sales as demonstrated on detailed surrogate models. Also, to validate this method, real market data provided by a worldwide retailer have been used, covering numerous categories from three different countries and several types of stores. The algorithm is benchmarked against an ensemble of regression trees and the forecast provided by the retailer and it outperforms both on a merit figure composed not only by the mean absolute error but also by the error deviations used in the retail business. The proposed method significantly improves the accuracy of the forecast in many diverse categories and geographical locations, yielding significant and operative benefits for supply chains. Additionally, we briefly discuss in the Appendix how to deploy our method as a RESTful service in a production environment. The accurate prediction of promotional sales is of paramount importance for customers, retailers, producers and suppliers. Amongst others, one of the benefits that the authors would like to highlight is waste reduction. This paper presented a new forecasting method inspired in the knowhow of the business analysts, or forecasters, that tackle the problem of sales forecasting on a daily basis. That motivated us to work on a kNN method that produces explainable and easy to modify predictions. Also, we wanted to couple feature selection and prediction, which we accomplished through a cost function that is minimized with a NNLS solver. Our method is founded on online learning, so the predictions are calculated with the latest data available, thus avoiding to retrain the model. Also, the forecast is generated at SKU level meaning that the data always in memory.

d) Online Sales Prediction: An Analysis with Dependency SCOR-Topic Sentiment Model, Lijuan Huang, 2019

This study aims to find a robust method to improve the accuracy of online sales prediction. Based on the groundings of existing literature, the authors proposed a Dependency SCOR-topic Sentiment (DSTS) model to analyze the online textual reviews and predict sales performance. The authors took the online sales data of tea as empirical evidence to test the proposed model by integrating the auto-regressive review information model into the DSTS model. The findings include: 1) the effect of the distribution of SCOR-topic from reviews on sales prediction; 2) the effect of review text sentiment on sales prediction increases as the specific topic probability dominates; and 3) the effect of review text sentiment on sales prediction increases as the rest topic probability evenly distributes. These findings demonstrate that the DSTS model is more precise than alternative methods in online sales prediction. This study not only contributes to the literature by pointing out how the distribution of sentiment topic impacts on sales prediction but also has practical implications for the e-commerce practitioners to manage the inventory better and advertise by this prediction method. While existing studies have been conducted to examine the importance of sentiment topics of reviews, the number of studies is still small on the relationship between the distribution of the topics of reviews and sales prediction. This study some gap in the literature to out how the distribution of sentiment topic is related to sales prediction. In this paper, we conduct a case study in the tea industry and focus on the problem of the

International Research Journal of Education and Technology
Peer Reviewed Journal
ISSN 2581-7795

topics distribution of reviews for predicting product sales performance. The paper firstly developed a DSTS Model to extract the sentiment SCOR- topic information from online textual reviews; Also, we prove that the prediction accuracy is improved through integrating the ARRIM into DSTS model. These endings could explain why the sentiment information embedded in the reviews can help the business to make more accurate forecasting.

FUTURE GOALS

Regularity slant and arbitrariness and future estimates will offer assistance to analyze deal drops which the companies can maintain a strategic distance from by employing a more focused and efficient strategies to play down the deal drop and maximize the benefit and stay in competition. The exactness of the expectation can be improved in future so that the recognizable proof of deals can be found on the leading. There are several potential areas of improvement for this project. One goal could be to expand the scope of the analysis by using more data sources and incorporating additional variables such as news sentiment or market trends. Another goal could be to optimize the hyperparameters of the LSTM model to improve its accuracy and performance. Additionally, the project could be extended to other industries or asset classes beyond just stocks. Overall, there is ample room for further development and exploration in this area of financial forecasting using machine learning techniques. Expanding the scope of the analysis involves exploring different industries and asset classes beyond stocks. For example, the techniques and models developed for sales forecasting in the retail industry can be applied to other industries such as healthcare, energy, or transportation. Additionally, the models can be applied to other types of time series data such as exchange rates, commodity prices, or weather data. Furthermore, we can also focus on creating a user-friendly interface for the proposed system so that businesses can easily input their data and obtain accurate sales forecasts. Additionally, we can develop a recommendation engine that suggests strategies to maximize profit and minimize sales drops based on the predicted sales trends. Overall, the future goals of this project involve improving the accuracy of sales forecasts and expanding the scope of the analysis while also providing a user-friendly interface for businesses to obtain insights and make informed decisions.

CONCLUSION

Sale forecasting is an inseparable part of every industry especially businesses that work with seasonal items. It is essential to know which model can produce better results because accurate estimation of the future can help businesses to boost their work. So, the approach with the trend correction block in the deep learning model for sales forecasting has been considered. The results show that the forecasting accuracy can be essentially improved for non-stationary sales with time trends using the trend correction block in the deep learning model. the LSTM model with a special loss function and hyper-parameter search is proposed for overcoming challenges in the real-scene sales forecast. By comparing with traditional methods, we find that the proposed model is superior to others, which provide reliable predictions.

REFERENCES

[1] W. Karush, Minima of functions of several variables with inequalities as side constraints, M.S. thesis, Dept.Math., Univ. Chicago, Chicago, IL, USA, 1939.

[2] S. F. Alamdar, M. Rabbani, and J. Heydari, Pricing, collection, and effort decisions with coordination contracts in a fuzzy, three-level Aug. 2018.

[3] F. Martnez, M. P. Fras, M. D. Prez-Godoy, A. J. Rivera, Deal- in time series ing with seasonality by narrowing the training set Aug. 2018.

[4] W. Lomax, K. Hammond, R. East, and M. Clemente, The measurement 1997.

[5] S. M. Lundberg, G. G. Erion, and S.-I. Lee, Consistent individualized feature attribution for tree ensembles,2018, arXiv:1802.03888. [Online] 21) G. Dileep, A survey on smart grid technologies and applications, renene.2019.08.092.

[6] D. Ramchurn, P. Vytelingum, A. Rogers, and E. N. R. Jennings, Putting the smarts into the smart grid: A grand challenge for arti- 10.1145/2133806.2133825.

[7] D. Neves, A. Pina, and C. A. Silva, Comparison of different demand response optimization goals on an isolated microgrid, Sustain. 1016/j.seta.2018.10.006.

[8] Z. Zhao, W. C. Lee, Y. Shin, and K.-B. Song, An optimal power schedul- ing method for demand response in home energy management system, 10.1109/TSG.2013.2251018.

[9] A. K. Barnes, J. C. Balda, and A. Escobar-Meja, "A semi-markov model for control of energy storage in utility grids and microgrids with PV Apr. 2015.

[10] R. L. Silva, M. Rudek, A. L. Szejka, and O.Canciglieri, Machine vision systems for industrial quality control inspections, in Proc. IFIP Int. Conf. 1007/978-3-030-01614-2_58.

[11] L. Wang, Q. Li, R. Ding, M. Sun, and G. Wang, Integrated scheduling of energy supply and demand in microgrids under uncertainty: A robust multi- doi: 10.1016/j.energy.2017.04.115.

[12] S. Veda et al., "Evaluating the impact of the 2017 solar eclipse on U.S. western interconnection operations,"Apr. 2018, doi: 10.2172/1435500