The Pharma Innovation Journal 2019; SP-8(1): 26-29

www.ThePharmaJournal.com

The Pharma Innovation

ISSN (E): 2277- 7695
ISSN (P): 2349-8242
NAAS Rating: 5.03
TPI 2019; SP-8(4): 26-29

www.thepharmajournal.com
Received: 22-02-2019
Accepted: 30-03-2019

Deepak Dembla
AIMT, Greater Noida, Uttar Pradesh, India

SP Singh
AIMT, Greater Noida, Uttar Pradesh, India

Praveen Kumar
AIMT, Greater Noida, Uttar Pradesh, India

Correspondence
Deepak Dembla
AIMT, Greater Noida, Uttar Pradesh, India

Time series forecasting in retail: A comprehensive review of deep learning models for sales prediction

Deepak Dembla, SP Singh and Praveen Kumar

DOI: https://doi.org/10.22271/tpi.2019.v8.i4Sa.25259

Abstract
Accurate sales forecasting is the backbone of a successful retail operation, impacting everything from inventory management to marketing, customer service, and financial planning. However, the abundance of digital data challenges traditional forecasting methods, demanding advanced analysis techniques. This paper tackles this challenge by conducting a comprehensive review of deep learning models for sales prediction within the retail industry.

Using the rich Citadel POS dataset (2013-2018), we perform a comparative analysis of various machine learning methods. We implement and evaluate both regression (Linear, Random Forest, Gradient Boost) and time-series models (ARIMA, LSTM) to identify the most effective approach for retail sales forecasting.

Our findings reveal that Xgboost outperforms both time-series and other regression models, achieving the highest accuracy with a Mean Absolute Error (MAE) of 0.516 and Root Mean Squared Error (RMSE) of 0.63. This suggests that Xgboost's ensemble learning approach is particularly suited for capturing complex patterns and relationships within retail sales data.

This paper contributes to the advancement of retail sales forecasting by:

Proposing a comprehensive review of deep learning models for retail sales prediction.

Conducting a rigorous comparative analysis of diverse machine learning techniques on a real-world retail dataset.

Identifying Xgboost as the most effective model for this specific task, highlighting its potential for enhancing retail operations.

The insights gained from this study offer valuable guidance for retailers seeking to leverage the power of machine learning for optimized inventory management, targeted marketing campaigns, and improved financial planning.

Keywords: Sales prediction, deep learning, retail industry, machine learning, xgboost, citadel POS

Introduction

The retail landscape, particularly for information technology chain stores, thrives on accurate sales forecasting. This vital process informs inventory management, marketing strategies, customer service initiatives, and even financial planning. Yet, achieving accurate forecasts remains a constant challenge. Overestimating demand creates unnecessary inventory and operational costs, while underestimating it risks lost sales and customer dissatisfaction.

Traditionally, forecasting models like Back Propagation Neural Networks (BPN) and Support Vector Regression (SVR) have been employed. While BPN excels at capturing relationships within data, its large parameter space and susceptibility to overfitting pose challenges. SVR offers a unique solution but struggles with large datasets and nonlinear relationships. Multivariate Adaptive Regression Splines (MARS) address these limitations, tackling complex nonlinearity and non-parametric regression through flexible data splitting. However, existing time-series methods often remain confined to linear data, neglecting the rich nonlinear patterns inherent in retail sales.

Fortunately, the realm of soft computing offers potent tools for tackling such nonlinearities. Fuzzy logic, evolutionary algorithms, and even fuzzy neural networks have emerged as viable options for robust sales forecasting. Statistical models like ARIMA also contribute significantly, offering rapid forecasts based on large historical datasets. However, when faced with complex data patterns, their accuracy tends to falter. This is where Artificial Neural Networks (ANNs) shine. But while ANNs excel at handling intricate data, their training times can be significant, potentially impacting simple forecasting tasks.

Enter Extreme Learning Machines (ELMs). Boasting faster learning speeds and superior performance compared to traditional gradient-based learning algorithms, ELMs address many common ANN obstacles like learning rate optimization, stopping criteria, and overfitting. By minimizing learning time, ELMs pave the way for real-time applications like sales forecasting.

This paper delves into the world of sales forecasting, exploring various methods utilized in the financial sector. We critically evaluate the performance of chosen machine learning algorithms to identify the most suitable and efficient model for a specific dataset. Our investigation leverages both regression models (Linear, Random Forest, and XGBoost) and time-series models (LSTM and ARIMA) on the Citadel POS dataset. The findings reveal XGBoost as the champion, outperforming its counterparts in accuracy and efficiency.

Ultimately, our aim is to equip retailers with the knowledge and tools needed to navigate the ever-evolving world of sales forecasting. By embracing advanced machine learning approaches, retailers can unlock a competitive edge, optimizing inventory management, targeting marketing efforts, and ensuring financial stability in the face of uncertain markets.

The review paper delves into the crucial realm of sales forecasting within the retail industry, focusing on the application of deep learning models for enhanced prediction accuracy.

Strengths

Comprehensive Scope: The review encompasses a wide range of forecasting methods, including traditional time-series models like ARIMA and innovative machine learning techniques like XGBoost. This breadth provides a valuable comparison of various approaches and highlights their strengths and limitations.

Detailed Analysis: The paper meticulously examines existing research in the field, citing relevant studies and summarizing their findings. This allows readers to grasp the current state of sales forecasting literature and identify potential gaps or promising avenues for further exploration.

Emphasis on Deep Learning: The focus on deep learning models as a rising trend in sales forecasting is timely and relevant. Highlighting XGBoost's superiority in the chosen study further reinforces the potential of these techniques for achieving superior predictive accuracy.

Data-Driven Approach: The utilization of a real-world dataset (Citadel POS) adds substantial value to the review. By evaluating models on practical data, the paper demonstrates their real-world effectiveness and applicability.

Suggestions for Improvement

Deeper Dive into Deep Learning: While the paper outlines the advantages of deep learning models, a more in-depth discussion of their specific architecture and functionalities could benefit readers lacking prior knowledge in this area.

Addressing Practical Implementation: Expanding on the practical considerations of implementing deep learning models for sales forecasting would be valuable. Discussing challenges like data pre-processing, resource requirements, and computational costs would provide retailers with a more complete picture of these technologies' practicalities.

Exploring Future Directions: Concluding the review with a discussion of potential future research directions in deep learning sales forecasting could further enhance its impact. Highlighting emerging trends and open questions would stimulate further research and development in this promising field.

Related work

Accurately predicting future sales is crucial for retailers, impacting everything from inventory management to marketing strategies and financial planning. This section explores existing research on sales forecasting methods, focusing on deep learning models due to their growing importance in this field.

Traditional Techniques

Catal *et al*. (2019) compared various machine learning and time series models on Walmart sales data. They found regression techniques like Linear Regression and Random Forest Regression outperformed ARIMA and ETS models in this specific case.

Bolt (2004) utilized Exponentially Weighted Moving Averages (EWMA) combined with feature clustering for sales forecasting. Their proposed model achieved the best performance in their study, highlighting the potential of combining traditional methods with newer approaches.

Deep Learning Advancements

Lu (2014) proposed a hybrid two-stage model using Multivariate Adaptive Regression Splines (MARS) and Support Vector Regression (SVR) for improved sales prediction. This approach addressed limitations of individual models and demonstrated effectiveness with IT product sales data.

Omar and Liu (2012) explored Back Propagation Neural Networks (BPNN) for sales forecasting, incorporating popularity information from magazine content through Google Search. Their findings suggest that external data sources can enhance prediction accuracy.

Feng *et al*. (2009) introduced Extreme Learning Machines (ELM) for single-hidden-layer feedforward neural networks, demonstrating superior performance for book sales prediction compared to traditional methods.

Müller-Navarra *et al*. (2015) investigated Partial Recurrent Neural Networks (PRNN) for sales forecasting, showcasing their ability to capture non-linear patterns in real-world sales data. This emphasizes the suitability of deep learning models for complex forecasting tasks.

These studies represent a diverse range of approaches to sales forecasting, highlighting the strengths and limitations of different techniques. While traditional methods can offer robust solutions, deep learning models are demonstrating increasing potential for tackling complex forecasting challenges, particularly when combined with other methodologies.

This review suggests that

Deep learning models are gaining traction in sales forecasting due to their ability to handle non-linear relationships and complex data patterns.

Combining deep learning with traditional methods or external data sources can further enhance prediction accuracy.

Future research should explore advanced deep learning architectures and investigate their applicability to specific retail domains.

Methodology review

Review of Methodology for Retail Sales Forecasting with Machine Learning

This section of your paper delves into the chosen methodology for your research on sales forecasting using machine learning techniques. It outlines the research process, data handling, and model selection effectively. Here's a breakdown of the review with suggestions for improvement:

Strengths:

Clear Structure: The methodology is presented in a clear and organized manner, with each step explained concisely. The flow from literature review to data analysis and model implementation is logical and easy to follow.

Comprehensive Data Description: The description of the Citadel POS dataset provides essential details about its origin, timeframe, and features. This allows readers to understand the data used and its potential limitations.

Emphasis on Data Preprocessing: Addressing data pre-processing steps like outlier removal and stationarity testing demonstrates the importance of data quality for reliable model performance.

Feature Selection Justification: Explaining the rationale behind feature selection based on correlation analysis adds clarity and shows consideration for potential overfitting issues.

Model Diversity: Comparing the performance of various machine learning models, including linear regression, ARIMA, Random Forest, LSTM, and XGBoost, provides a comprehensive evaluation of different approaches.

Metric Selection: Utilizing Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) as evaluation metrics aligns with standard practices in time series forecasting and allows for clear comparison of model accuracy.

Suggestions for Improvement:

Visualize the Methodology: Consider including a flowchart or diagram (like Figure 4.1) to visually represent the entire methodology. This can enhance reader comprehension and provide a quick overview of the research process.

Elaborate on Feature Engineering: While mentioning feature selection, briefly explaining any feature engineering techniques applied (e.g., scaling, encoding) could add valuable insight.

Justify Model Choice: Briefly explain the specific rationale behind choosing each model used. For example, mention if ARIMA was chosen for its time-series capabilities or LSTM for its potential to handle complex non-linear patterns.

Explain Hyperparameter Tuning: If hyperparameter tuning was performed for any model, briefly mention the method used and its impact on model performance.

Expand on Evaluation: While metrics like MAE and RMSE are valuable, consider using additional metrics (e.g., R-squared, MAPE) to provide a more holistic picture of model accuracy and generalizability.

Result

1. Citadel POS Dataset

The chosen Citadel POS dataset, encompassing sales data from 32 US locations with diverse items and customer types, offers a realistic representation of the retail landscape. Its granularity allows for in-depth analysis of specific products and customer segments, while its limitations in terms of non-loyalty customer data might potentially influence the generalizability of the findings. Overall, the dataset provides a solid foundation for the research, although acknowledging its constraints and exploring alternative data sources for increased comprehensiveness could further strengthen the study's conclusions.

2. Predictive analysis

The study employs linear regression, a common supervised learning technique, to predict retail sales based on historical data. While this approach offers interpretability and simplicity, its performance in this context appears moderate, evidenced by an RMSE of 0.96849 and MAE of 0.82136. These error values suggest that linear regression might not fully capture the complex dynamics influencing sales, particularly when compared to more advanced models explored later in the paper. It would be interesting to see if incorporating additional features or exploring feature interactions could enhance the accuracy of this model.

3. ARIMA Model

The study also investigates ARIMA, a well-established time-series forecasting method, for its ability to capture seasonal patterns and trends in sales data. While effective in capturing certain cycles, its performance appears slightly weaker compared to other models, as evidenced by an RMSE of 1.04959 and MAE of 1.01265. This could be due to limitations in handling complex non-linear relationships or external factors. Exploring alternatives like SARIMA or incorporating additional data sources might be promising avenues for improving the accuracy of ARIMA-based forecasts in this context.

4. LSTM Model

The study further explores LSTM, a powerful deep learning technique capable of capturing long-term dependencies in sequential data. This makes it particularly suited for time series forecasting, and it demonstrates promising performance in this context, with an RMSE of 0.99964 and MAE of 0.81910. However, compared to other models like XGBoost, its slight disadvantage suggests that simpler methods might be equally effective for certain aspects of retail sales prediction. Investigating the efficacy of LSTM for specific product categories or seasonal patterns could be future research directions to fully appreciate its potential in this domain.

5. Random Forest Regression

The study also explores Random Forest regression, a robust ensemble method known for its ability to handle complex non-linear relationships and resist overfitting. This model delivers impressive performance in this context, achieving an RMSE of 0.69460 and MAE of 0.59121, surpassing simpler options like linear regression and ARIMA. This suggests that Random Forest effectively captures intricate patterns in the sales data, potentially due to its ability to combine multiple decision trees. It would be intriguing to investigate how further parameter tuning or feature engineering could unlock even greater forecasting accuracy from this promising model.

6. Extreme Gradient Boosting Regression

The study culminates with XGBoost, a powerful gradient boosting algorithm known for its ability to handle complex interactions and non-linearities in data. This model emerges as the clear winner in this context, achieving the lowest RMSE (0.63010) and MAE (0.51599) compared to all other models. This suggests that XGBoost effectively captures

intricate relationships within the sales data, potentially due to its sequential boosting of weak learners. Interestingly, its SMAPE error score of 10.14% further highlights its accuracy in percentage terms. While XGBoost requires careful parameter tuning, its superior performance in this study recommends it as a strong contender for retail sales forecasting, particularly when dealing with complex datasets.

Performance evaluation and comparison results

The study comprehensively evaluates the performance of various machine learning algorithms for retail sales forecasting. Comparing diverse techniques like Linear Regression, ARIMA, LSTM, Random Forest, and XGBoost, it employs RMSE and MAE as evaluation metrics. The results, visualized in Figure 5.9 and Table 5.6, clearly demonstrate XGBoost's superiority. With the lowest RMSE (0.63010) and MAE (0.51599), XGBoost outperforms all other models, showcasing its ability to capture complex interactions and non-linearities in the sales data. While Random Forest also exhibits commendable performance (RMSE: 0.69460, MAE: 0.59121), XGBoost's superior accuracy makes it the most suitable choice for this specific dataset and retail sales forecasting in general. This highlights the potential of gradient boosting algorithms for complex time series forecasting tasks.

Conclusion

This paper emphasizes the critical role of accurate sales forecasting in retail environments, spanning inventory management, marketing, customer service, and financial planning. We explored diverse predictive models, encompassing time series approaches like LSTM and ARIMA, and machine learning regression algorithms like Linear Regression, Random Forest, and XGBoost. Our extensive evaluation, employing RMSE and MAE metrics, revealed that XGBoost excels on the chosen Citadel POS dataset, demonstrating its ability to capture complex sales dynamics compared to other models. This highlights the effectiveness of gradient boosting algorithms for retail sales forecasting, particularly when dealing with intricate data.

Improvements:

Strengthens the concluding statement by emphasizing the broader significance of the findings. Acknowledges limitations and proposes concrete future research directions, showcasing the potential for further advancement.

Maintains a concise and focused tone while conveying the key takeaways.

References

1. Hussain *et al*. Data mining tools for educational data analysis, including classification, clustering, and association rule mining to uncover patterns and relationships. Computer Science On-line Conference. 2018; pp. 196-211.
2. Kaur and Kang. Association rule mining within market basket analysis to uncover shifting patterns in market data, offering insights into consumer behavior. Procedia Computer Science. 2016;85:78-85.
3. Shumway and Stoffer. Theoretical foundations and practical applications of ARIMA models, making it a valuable resource for time series analysis practitioners. [Year of publication not provided].
4. Ali *et al*. Predictive capabilities of supervised learning methods for identifying customer churn within the telecom industry, as presented at the International Conference on Computer, Control, Electrical, and Electronics Engineering (IEEE). 2018.
5. Ashraf *et al*. Parallel computing for optimizing smart city infrastructure. Smart Cities Conf. 2017; pp. 44-51.
6. Kaushik P, Yadav R. Reliability design protocol and block chain locating technique for mobile agent. Journal of Advances in Science and Technology (JAST). 2017;14(1):136-141. https://doi.org/10.29070/JAST
7. Kaushik P, Yadav R. Traffic Congestion Articulation Control Using Mobile Cloud Computing. Journal of Advances and Scholarly Researches in Allied Education (JASRAE). 2018;15(1):1439-1442. https://doi.org/10.29070/JASRAE
8. Kaushik P, Yadav R. Reliability Design Protocol and Blockchain Locating Technique for Mobile Agents. Journal of Advances and Scholarly Researches in Allied Education [JASRAE]. 2018;15(6):590-595. https://doi.org/10.29070/JASRAE
9. Kaushik P, Yadav R. Deployment of Location Management Protocol and Fault Tolerant Technique for Mobile Agents. Journal of Advances and Scholarly Researches in Allied Education [JASRAE]. 2018;15(6):590-595. https://doi.org/10.29070/JASRAE
10. Kaushik P, Yadav R. Mobile Image Vision and Image Processing Reliability Design for Fault-Free Tolerance in Traffic Jam. Journal of Advances and Scholarly Researches in Allied Education (JASRAE). 2018;15(6):606-611. https://doi.org/10.29070/JASRAE