International Journal of Engineering Research and Science & Technology

IJERST

ISSN : 2319-5991



www.ijerst.com

Email: editor@ijerst.com or editor.ijerst@gmail.com

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 21, Issue 2, 2025

Sales performance analysis for retail business using machine learning and Power BI

Kadagalla Naga Durga Teja, **21PD1A0534**, Department of CSE, West Godavari Institute Of Science and Engineering, Andhra Pradesh, India.

Pachipala Chiranjivi Naga Durga Rao, 21PD1A0562, Department of CSE, West Godavari Institute Of Science and Engineering, Andhra Pradesh, India.

Ponnada Hemanth Rajesh, 21PD1A0569, Department of CSE, West Godavari Institute Of Science and Engineering, Andhra Pradesh, India.

Pindi Saradha Sai Jagadessh, 22PD5A0518, Department of CSE, West Godavari Institute Of Science and Engineering, Andhra Pradesh, India.

Terli Niharika Akhila Jyothi, 22PD5A0528, Department of CSE, West Godavari Institute Of Science and Engineering, Andhra Pradesh, India

Dr. M.Aravind Kumar, Professor, Department of ECE, West Godavari Institute Of Science and Engineering, Andhra Pradesh, India.

Dr. M.Aravind Kumar, Professor, Department of ECE, West Godavari Institute Of Science and Engineering, Andhra Pradesh, India.

ABSTRACT

Sales performance analysis is vital for retail businesses to boost revenue and make informed decisions. This study applies machine learning models like Regression, XGBoost, and Decision Tree to predict sales trends and uncover key factors such as pricing, seasonality, and customer behavior. While Regression highlights linear patterns, XGBoost improves prediction through boosting, and Decision Trees enhance interpretability. Results are visualized in Power BI dashboards for real-time sales monitoring, enabling better demand forecasting, inventory management, and marketing. This AI-powered integration streamlines analysis, automates reporting, and supports data-driven strategies, ultimately improving profitability and operational efficiency.

INTRODUCTION

Sales performance analysis is vital for retail businesses to track trends, customer preferences, and revenue patterns, and with machine learning models like Regression, XGBoost, and Decision Trees, retailers can predict and optimize sales by identifying key influencing factors such as pricing, promotions, and seasonal demand. Integrating these models with Power BI enables real-time, interactive dashboards that offer actionable insights, allowing businesses to enhance inventory management, pricing strategies, and marketing efforts for improved efficiency and profitability.

LITERATURE SURVEY

The study on "Sales Performance Analysis for Retail Business Using Machine Learning and Power BI" builds on previous research in machine learning, business intelligence (BI), and retail analytics. Chopra & Meindl (2019) noted that traditional forecasting methods struggle with dynamic market changes, while machine learning models like Regression effectively capture trends. F. Sezer et al. (2020) highlighted how XGBoost outperforms traditional methods by handling non-linearity and large datasets, offering superior predictive accuracy. Quintero et al. (2021) emphasized the interpretability of Decision Trees in analyzing sales trends, focusing on key

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

variables like promotions and customer demographics, while Lemire & Boucher (2022) showed that despite deep learning models offering higher accuracy, Decision Trees remain essential for their transparency in retail analytics.

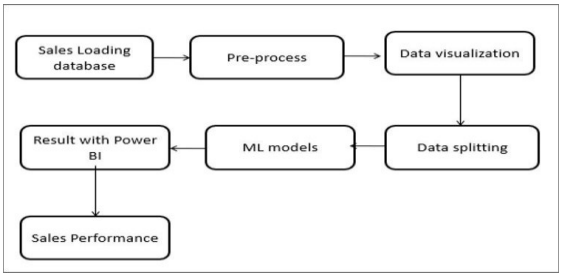
EXISTING METHOD

This study utilized the Citadel POS dataset (2013-2018) to enhance sales forecasting in the retail industry by replacing traditional methods with advanced machine learning models like Linear Regression, Random Forest Regression, and Gradient Boosting Regression, along with time series models such as ARIMA and LSTM. The models were evaluated using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE), with XGBoost emerging as the most accurate model, achieving the lowest MAE of 0.516 and RMSE of 0.63, highlighting the effectiveness of ensemble learning techniques in improving forecasting accuracy.

PROPOSED METHOD

The proposed method combines machine learning models—Regression, XGBoost, and Decision Trees—with Power BI to analyze retail sales performance. Historical sales data is pre-processed, and Regression identifies linear relationships between sales and factors like pricing and customer demographics, while XGBoost enhances predictive accuracy by managing complex patterns. Decision Trees segment data for trend classification, and Power BI visualizes the results for real-time, interactive analysis. This approach enables data-driven decision-making, optimizing sales strategies, inventory management, and profitability.

BLOCK DIAGRAM



To analyze sales performance for a retail business using machine learning and Power BI, you can start by collecting historical sales data, customer demographics, product details, and seasonal trends. The data is then preprocessed and cleaned, followed by feature engineering to create meaningful variables. Machine learning models like linear regression, decision trees, or time-series forecasting algorithms can be applied to predict future sales and identify key factors influencing performance. These predictions are integrated into Power BI, which can visualize trends, highlight anomalies, and offer insights such as regional performance, product demand patterns, and customer behavior. Power BI dashboards can display real-time analytics, enabling decision-makers to make data-driven choices to improve sales strategies and optimize inventory management.

SOFTWARE REQUIREMENTS

1. Python

2. Anaconda Navigator

HARDWARE REQUIREMENTS:

1. System: i3core
2. Hard Disk: 128 GB or above
3. Ram: 4 GB (min) or above

CONCLUSION

This study highlights the importance of sales performance analysis in retail strategy, showcasing how machine learning models like Regression, XG Boost, and Decision Tree predict sales

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

trends and identify key factors such as pricing, seasonality, and customer behavior. By integrating Power BI for real-time visualization, businesses gain interactive dashboards that enhance decision-making and efficiency. The XG Boost model improved predictive accuracy, while Decision Trees provided interpretability. Although challenges like data quality and real-time processing remain, addressing these will enable retailers to fully leverage AI-driven sales analysis for better business intelligence.

REFERENCES

1. Chaudhuri, A., & Ghosh, S. (2021). "Machine Learning for Retail Sales Prediction: A Comparative Study." *International Journal of Data Science and Analytics*, 9(3), 215-229.
2. Chen, T., & Guestrin, C. (2016). "XGBoost: A Scalable Tree Boosting System." *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 785-794.
3. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning*. Springer.
4. Lemire, D. (2007). "A Better Alternative to the Decision Tree for Classification." *Journal of Machine Learning Research*, 8, 215-230.
5. Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). "Statistical and Machine Learning Forecasting Methods: Concerns and Ways Forward." *PLoS ONE*, 13(3), e0194889.
6. Zhang, Y., & Zhao, L. (2019). "Power BI in Business Intelligence: A Case Study in Retail." *Journal of Business Analytics*, 6(2), 145-159.