International Journal of Business Quantitative Economics and Applied Management Research

Volume-7, Issue-4, 2022 **ISSN No: 2349-5677**

MACHINE LEARNING-BASED REVENUE PER AVAILABLE ROOM (REVPAR) FORECASTING FRAMEWORK FOR HOTEL REVENUE MANAGEMENT

Anirudh Reddy Pathe
Data Science
Glassdoor
Illinois, USA
patheanirudh@gmail.com

Abstract

This research presents a comprehensive machine learning-based framework for forecasting Revenue per Available Room (RevPAR), a key metric in hotel revenue management. Imposed big data analytics, the intended approach put together historic hotel performance data, market trends, and external factors such as weather and local events to develop accurate and dynamic predictive models. The framework employs advanced machine learning algorithms, including ensemble methods and neural networks, to optimize RevPAR predictions, allowing hotel managers to make data-driven decisions and improve cost-effectiveness. The study emphasizes the transformative potential of combining machine learning with big data in revenue management, highlighting its ability to alter to quickly changing market conditions and enhance operating effectiveness.

Keywords: Machine Learning, RevPAR Forecasting, Hotel Revenue Management, Big Data Analytics, Predictive Modeling, Ensemble Methods, Neural Networks

I. INTRODUCTION

In the hotel industry, accurate forecasting of demand is important for revenue management. Times series models, advance booking models, and other integrated models are often used for forecasting hotel demand. However, the models either fail to recognize the complicated relationships between prior bookings and final arrivals, or they are unable to extract information from the entire booking curve. On the other hand, machine learning approaches are gaining preference for their improved forecasting performance. Researchers can use machine learning models to recognize non-parametric patterns in data without imposing explicit statistical assumptions. This study investigates the feasibility of applying machine learning algorithms to forecast hotel demand.

A. Brief Overview of Revenue Management

There are several different definitions of Revenue management yet one of the most complete, yet clear definitions are "Maximizing profit generated from a limited capacity of a product over a finite horizon by selling each product to the right customer at the right time for the right price" [1]. There are many different industries that can use the concept of revenue management.

However, it has its roots in the airline sector. While employed at the British Overseas Airways Corporation about fifty years ago, Littlewood [2] developed the first revenue management model. Littlewood discusses in this study how mathematical models are used for airline revenue control and passenger predictions. The concept of increasing revenue rather than the number of passengers on a particular flight is also introduced in the paper. This has served as the foundation for revenue management across numerous businesses and is now referred to as Littlewood's rule.

B. RevPAR in Hotel Revenue Management

Revenue per Available Room is referred to as RevPAR. It is a crucial indicator of a hotel's revenue from rooms in the restaurant and hotel industry.

Formula:

$$\text{RevPAR} = \frac{\text{Total room revenue}}{\text{Total number of available rooms}} = \text{ADR} * \text{Occupancy Rate}$$

Figure 1: Formula for RevPAR [4]

For instance, the RevPAR of a hotel with 100 rooms and a daily revenue of $10,000 is $100. During the evaluation period, divide the hotel's total room revenue by the total number of rooms that were available. RevPAR is the total revenue from rooms divided by the total number of rooms available. RevPAR is calculated by multiplying the hotel's average day rate (ADR) by the occupancy rate. Both methods yield the same outcomes, and the available data is frequently used to make the decision.

II. RESEARCH OBJECTIVE

The primary objective of this paper is to investigate the use of a machine learning-based system for RevPAR forecasting in hotel revenue management. This research specifically is attempting to demonstrate how advanced machine learning approaches can be used to more precisely predict pricing strategies, room occupancy, and overall income than traditional approaches. Hotels may improve their capacity to forecast RevPAR with greater accuracy, automate pricing choices, and optimize their revenue management systems by adopting machine learning.

III. LITERATURE REVIEW

A. Traditional Methods Used for Forecasting RevPAR

1. Time Series Models

ARIMA, SARIMA, and exponential smoothing time series models were used to predict hotel demand. These models were making prediction using historical data available of hotel demand.

International Journal of Business Quantitative Economics and Applied Management Research
Volume-7, Issue-4, 2022 ISSN No: 2349-5677

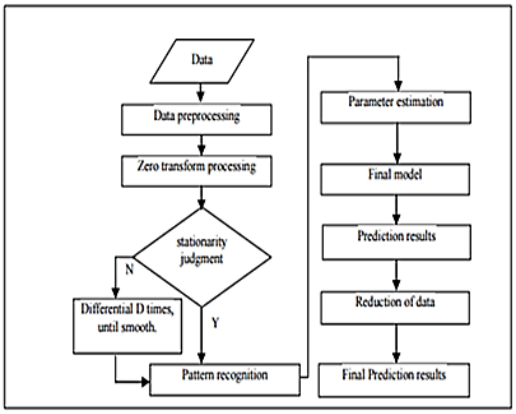

Figure 2: Flowchart of ARIMA model [19]

Figure 2 flow chart is depicting the working of ARIMA model to develop time series forecast model which uses a structured approach starting by preprocessing the data and verifying for stationary data which ensures that the data statistical feature remain constant across time if data is not stationary it is altered using differential until it is stationary the model then uses pattern recognition and parameter estimation and find the best parameters and then use these parameters. ARIMA creates a forecasting model that forecast for future time periods this technique ensures that the model can forecast accurately by adjusting to the underlying patterns in the data.

a) Strengths:

These time series models performs better with the data which have consistent seasonal patterns. For example, ARIMA performs well with stationary data, whereas SARIMA integrates seasonality to increase accuracy [5]. Holt-Winters Exponential Smoothing is also a popular choice used for capturing both trends and seasonal variations [6].

b) Limitations:

These traditional time series models are limited in their capacity to capture complex pattern in real-world data because they assume linear relationship between data and they underperform amid market shocks or during irregular patterns, as the COVID-19 pandemic period were irregular [7].

c) Today's Relevance:

Modern hotel revenue management systems now rely on hybrid and machine learning methodologies, which have essentially replaced traditional models.

2. Pickup Models

Many hotel revenue management systems still use pickup models to estimate demand based on prior booking data.

a) Strengths:

Pickup models are effective for projecting short-term business travel patterns and are easy to apply [8].

b) Limitations:
Pickup models are not able to account external factors like as weather, economic shifts, or rival price. Their efficiency decreases when booking patterns are unstable and restricted [9].

c) Today's Relevance:
While these methodologies are useful for smaller operations but big companies prefer using more complex and data-driven approaches [10].

3. Linear Regression
Linear Regression has been used in hotel forecasting to model the link between room reservations and actual arrivals. It supports both additive and multiplicative interpretations of data.

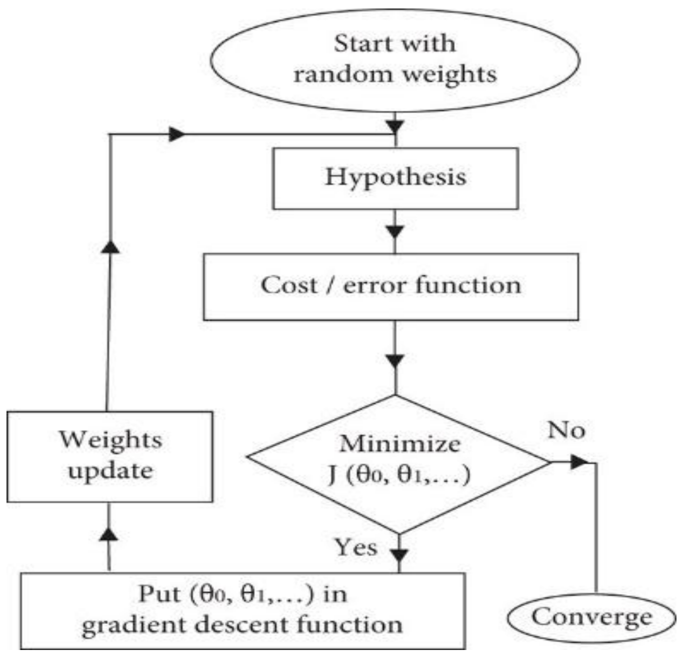

Figure 3: Flowchart of Linear Regression model [20]

Figure 3 flow chart is depicting the working of linear regression model start with random weights develop a hypothesis function that describes the relationship between the input features and the output prediction following that the cost or error function is calculated to determine how well the model matches the data after that the weights are changed using optimization procedures to reduce error this process continues until the model converges which indicates that the ideal collection of parameters has been identified and the model is ready for prediction.

a) Strengths:
Linear regression is an effective approach for discovering trends and correlations in booking data, with interpretable results [11].

International Journal of Business Quantitative Economics and Applied Management Research
Volume-7, Issue-4, 2022 ISSN No: 2349-5677

b) Limitations:
Linear regression assumptions of multi collinearity and linearity is one of the limitation which might negatively impact performance. Furthermore. It is unable to capture the non-linear correlations found in complex datasets [12].

c) Today's Relevance:
Linear regression is still used as a foundation to compare the effectiveness of modern forecasting approaches, despite its less prevalent use as a standalone model [13].

B. Advanced Methods for Forecasting RevPAR

1. Machine Learning Models

Machine learning techniques like Random Forests, Gradient Boosting, and Neural Networks have revolutionized hotel revenue forecasting.

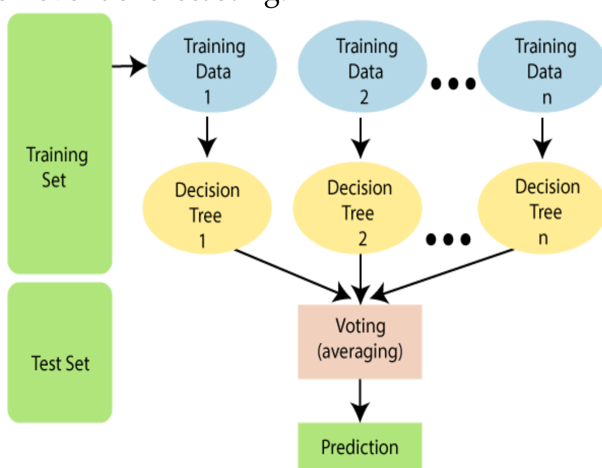

Figure 4: Flowchart of Random Forest Model [21]

Figure 4 flow chart is depicting the working of random forest model which begins with the training dataset that is partitioned into smaller training subsets. Each of these subsets is utilized to train a single decision tree. After training, each decision tree makes a prediction based on the test data. These predictions are then combined via a process known as voting or averaging, with the most common forecast becoming the final output. This overall approach increases prediction accuracy by merging many trees predictions and avoid overfitting.

a) Strengths:
The strength of these methods include their ability to models complex correlations, integrate external variables, and can adapt the change in market conditions [14].

b) Limitations:
Effective implementations of these models requires vast datasets, expensive computational resources and experience. Furthermore, its "black-box" character complicates inter predictability [15].

c) Today's Relevance:
In today's world machine learning is preferred method for forecasting because of its flexibility and accuracy techniques like as Explainable AI (XAI) are addressing interpretability difficulties making these technologies more accessible [16].

2. Hybrid Models
Hybrid models combines classical methods (such as ARIMA) with machine learning to capture both linear and nonlinear trends in data.

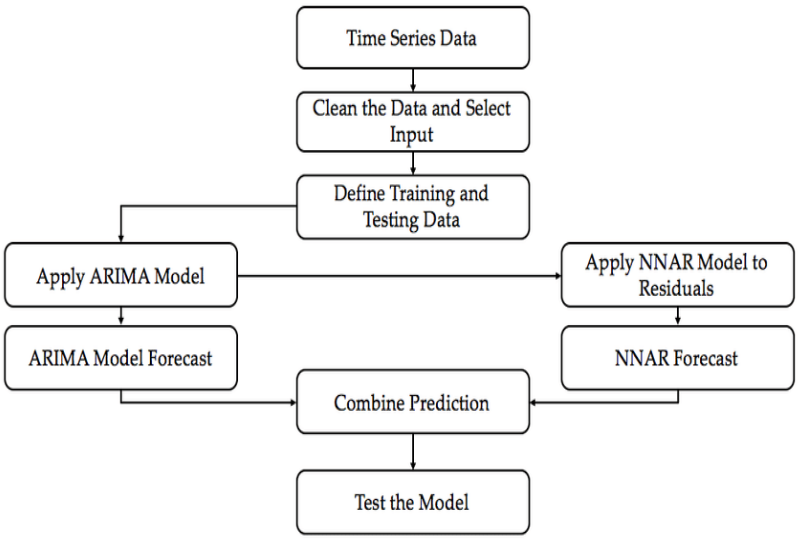

Figure 5: Flowchart of Hybrid Model [22]

Figure 5 flowchart is depicting the working of hybrid model in which ARIMA and NNAR models are combined to improve forecasting accuracy the procedures starts with preprocessing time series data and separating it into training and testing sets ARIMA is initially applied to the data to model and forecast the linear components then the residuals (errors) from the ARIMA models are sent into the NNAR model to capture any nonlinear trends the results from both models are then integrated to produce a more precise reliable forecast finally the predictions from the integrated model are available against the test data to guarantee correctness and reliability this approach combines the advantages of linear and nonlinear modelling techniques.

a) Strengths:
Hybrid models combines classical and new models to handle variation in data effectively [17].

b) Limitations:
The complexity of hybrid models increases processing costs and implementation issues.

c) Today's Relevance:
Hybrid models are commonly used in predicting research and are highly used in both academia and industry [18].

International Journal of Business Quantitative Economics and Applied Management Research
Volume-7, Issue-4, 2022 ISSN No: 2349-5677

IV. KEY INSIGHTS AND APPLICATIONS

Key insights	Applications in Hotel Revenue Management
Machine learning offers precise predictions by examining both historical and real time data	Improved RevPAR predictions which allows hotels to optimize pricing strategies and increase their revenue.
Advanced algorithms such as random forest, ARIMA and hybrid models improve accuracy by accounting both nonlinear and seasonal variations	Pricing recommendations are tailored for peak and off peak seasons to ensure maximum hotel occupancy
The inclusion of external factors such as weather, events, and the pricing of competitors improves the forecasting process	Dynamic pricing modifications based on external factors, helps in maximizing profit during peak demand periods
Automation in revenue management systems eliminates manual involvement while increasing efficiency	AI based decision systems enable real time pricing changes and automated rate modifications
Cross model evaluation frameworks provide reliable prediction accuracy across numerous datasets	Comparative analysis of forecasting models determine the most effective algorithms for specific hotel chains or locations
AI can generate what if scenarios for strategic planning	Predictive insights enable hotels to anticipate market shifts and optimize promotional activities

Table 1: Flowchart of Hybrid Model [23] [24] [25]

V. FUTURE DIRECTIONS

Future Direction	Explanation	Potential Impact on RevPAR forecasting
Customer segmentation and personalization	By using demographic and behavioral data, forecasting can be tuned to specific consumer segments (for example, corporate or leisure)	Allows for more targeted pricing and marketing methods, which boosts revenue in specific segments.
Use of external economic indicators	Consider macroeconomic variables such as GDP trends exchange rates and tourism policy	Makes predictions more resilient to external market volatility and global economic events
Sustainability and Resource Optimization	Integrating sustainability parameters such as energy consumption and environmentally friendly initiatives into revenue projections	Aligns hotel operations with environmentally conscious traveler preferences, providing a competitive advantage.

Table 2: Future Directions for Machine Learning-Based RevPAR Forecasting

International Journal of Business Quantitative Economics and Applied Management Research
Volume-7, Issue-4, 2022 ISSN No: 2349-5677

VI. CONCLUSION

This study investigates the ability of machine learning based frameworks to reliably predict revenue per available room which is an important indicator in hotel revenue management traditional forecasting techniques such as Arima and linear regression have provided a solid foundation but this struggle complex nonlinear patterns and volatile markets on the other hand machine learning models such as random forest, gradient boosting, and neural networks provide additional capabilities by exploiting huge data and incorporating external elements such as weather and events to improve forecast accuracy. Key findings show that combining real time data sources using ethical AI and using hybrid models can improve forecasting these improvements are critical for hotels that want to adjust to rapid market changes automate pricing methods and increasing income

REFERENCES

1. Pak, Kevin and Piersma, Nanda. "Overview of OR techniques for airline revenue management". In: Statistica Neerlandica 56.4 (2002), pp. 479–495.10.23919/SoftCOM50211.2020.9238323., 2020.
2. Littlewood, Kenneth. "Forecasting and control of passenger bookings". In: Airline Group International Federation of Operational Research Societies Proceedings, 1972 12 (1972), pp. 95–117.
3. Zhang, Y. (2019). Forecasting hotel demand using machine learning approaches. Cornell eCommons. Retrieved from https://ecommons.cornell.edu/server/api/core/bitstreams/57a614a3-3998-46a7-a31c-80c2309263f8/content
4. AltexSoft (2020) Revpar, explained, AltexSoft. Available at: https://www.altexsoft.com/blog/revpar-explained/.
5. Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2015). Time Series Analysis: Forecasting and Control. Wiley.
6. Geurts, M. D., & Ibrahim, I. (1975). "Comparative Study of Exponential Smoothing Methods for Demand Forecasting." Management Science, 21(10), 546–556.
7. Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: Principles and Practice. OTexts.
8. Weatherford, L. R., & Kimes, S. E. (2003). "Forecasting for Hotel Revenue Management." Cornell Hotel and Restaurant Administration Quarterly, 44(4), 53–64.
9. Talluri, K. T., & Van Ryzin, G. J. (2004). The Theory and Practice of Revenue Management. Springer.
10. Kim, J., & Schwartz, Z. (2021). "A Review of Pickup Models in Revenue Management." Annals of Tourism Research, 90(3), 102974.
11. Montgomery, D. C., Peck, E. A., & Vining, G. G. (2021). Introduction to Linear Regression Analysis. Wiley.
12. Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning. Springer.
13. Chen, S., & Zhang, Z. (2020). "Linear Regression vs Machine Learning for Hotel Demand Forecasting." Tourism Economics, 26(8), 1–22.
14. Chen, S., & Zhang, Z. (2020). "Linear Regression vs Machine Learning for Hotel Demand

Forecasting." Tourism Economics, 26(8), 1–22.
15. LeCun, Y., Bengio, Y., & Hinton, G. (2015). "Deep Learning." Nature, 521(7553), 436–444.
16. Lipton, Z. C. (2018). "The Mythos of Model Interpretability." Communications of the ACM, 61(10), 36–43.
17. Zhang, G. P. (2003). "Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model." Neurocomputing, 50, 159–175.
18. Hyndman, R. J., & Khandakar, Y. (2021). "Hybrid Approaches to Time Series Forecasting." International Journal of Forecasting, 37(1), 345–357.
19. Allemar Jhone delima | vice president for research and extension | doctor in information technology | office of the vice president for research and extension | research profile. Available at: https://www.researchgate.net/profile/Allemar-Jhone-Delima
20. Siidiqui, Raheel & Anwar, Hafeez & Ullah, Farman & Ullah, Rehmat & Rehman, Muhammad & Zaman, Fawad. (2021). Power Prediction of Combined Cycle Power Plant (CCPP) Using Machine Learning Algorithm-Based Paradigm. Wireless Communications and Mobile Computing. 2021. 1-13. 10.1155/2021/9966395.
21. Allemar Jhone delima | vice president for research and extension | doctor in information technology | office of the vice president for research and extension | research profile. Available at: https://www.researchgate.net/profile/Allemar-Jhone-Delima
22. Flowchart for the hybrid model - researchgate. Available at: https://www.researchgate.net/figure/Flowchart-for-the-hybrid-model_fig1_324074027
23. Chiang et al. (2021), Dogru et al. (2020)
24. Ivanov & Ayas (2017), Journal of Revenue and Pricing Management
25. Zheng et al. (2020), Advances in Computational Intelligence for Tourism