International Journal on Science and Technology (IJSAT)

E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

A Hierarchical Machine Learning Model for GDS Performance Evaluation and Ranking in Hotel Distribution Systems

Anirudh Reddy Pathe

Data Science
Glassdoor
California, USA
patheanirudh@gmail.com

Abstract

This paper also proposes a hierarchical machine learning (ML) model for assessing and ranking the GDSs' performance within the hotel sector. The traditional approaches that are generally used in the evaluation of the performance include the manual method or by simple statistical models which may not be efficient in the current complex hotel distribution systems. The multi-layered model incorporates both supervised and unsupervised learning using historical data and metrics for evaluating the proposal GDS and reinforcement learning for performance adjustment in real time. The model also handles scalability and enhances the ranking precision by incorporating the multi-source data and using the anomaly detection algorithms. The application of real-time GDS performance over different types of cases indicates better results in terms of ranking accuracy and decision-making speed.

Keywords: Hotel Distribution Systems, Artificial Intelligence, Systems Performance Analysis, Ranking, Outlier Detection, Hierarchical Models

I. Introduction

GDS or group distribution systems, are a centralized infrastructure that has allowed the hotel industry to sell their rooms on multiple online portals as well as agents. The right evaluation of the GDS performance can enhance the maximum profitability, the accuracy of the competitive market analysis, and greater customer satisfaction. But currently, for performance evaluation in GDS, the general model used often proves to be quite simplistic and incapable of accounting for the dynamics of the industry. [1]
With the increasing demands of customers and amount of data available today, new ways of measurement are needed that will be not only historical records but also outlook factors with the real data which will show further development tendencies.
The advancement in the theories in the application of ML techniques to performance evaluation has proven far much better than using ratio analysis since it is faster, dynamic and accurate to the business needs. Therefore, in this paper, we introduce a four-level ML model for assessing and continuing the performance ranking of GDS into account various aspects of effectiveness.

A. Challenges in Hotel Distribution System

1. *Data Complexity:* GDS data contains many factors including sales, rates, market conditions and customer profile that are usually volatile and may contain a good deal of noise.
2. *Scalability:* Hotel markets are therefore very sensitive due to variations such as increases in the prices that are usually influenced by demand, competition and consumer preferences. Consequently, for the purpose of performance appraisal that spans for more than one period, static models prove to be inadequate.
3. *Scalability:* GDS performance can only be measured on the enterprise level: to achieve it, models for RL must be able to handle large amounts of real-time data.

B. Hierarchical Machine Learning Model

The hierarchical ML model, described in this paper involves using multiple layers of ML include of Supervised ML, unsupervised ML, and reinforcement ML to provide a robust, adaptive, and scalable solutions to the evaluation of GDS performance.

***Supervised Learning*:** Employed for assessment of performances in terms of past records, KPIs like booking rates, cancellation rates, and customer feedback.

Unsupervised Learning: Conveys information about possible malfunctions in GDS's performance since it highlighted behaviors that are atypical for it, for example, a significant drop in bookings or a surge in negative feedback.

***Reinforcement Learning*:** Adaptive decision-making implementation within the context of rankings real-time performance data improves the system's capacity to predict the future trends in performance. [2]

C. Objective of Research

This paper aims to:

1. Propose a layered model of ML for comparing GDS performance.
2. Enhance the ranking precision by functioning together with historical data, real time data, and behavioral data.
3. Illustrate the applicability of the model for big data and analyze its implications for the hotel management.

II. Background and Literature Review

This paper aims the conventional approaches to assessing the operational effectiveness of GDSs primarily focus on key performance indicators such as room night sell, ADR and RevPAR. However, such models are not practical because in most cases they are far from ideal since they do not capture real time changes of the market.

Conventional methods of GDS Assessment

1. Key Performance Indicators (KPIs): Metrics like the ADR and RevPAR are recognized to make comparisons of the profitability of hotels and sellers hotels/GDS available in the market.
2. Manual Reviews: The assessment of performance is done through the physical analysis of the reports and can be criticized for being arbitrary and a time-consuming process. [3]
3. Rule-Based Systems: According to such assumptions, it is possible to define predefined parameters for GDS performance, for example, the channel's ranking depending on the number of bookings or

the revenue achieved. However, such systems do not offer flexibility whereby new market trends can be implemented.

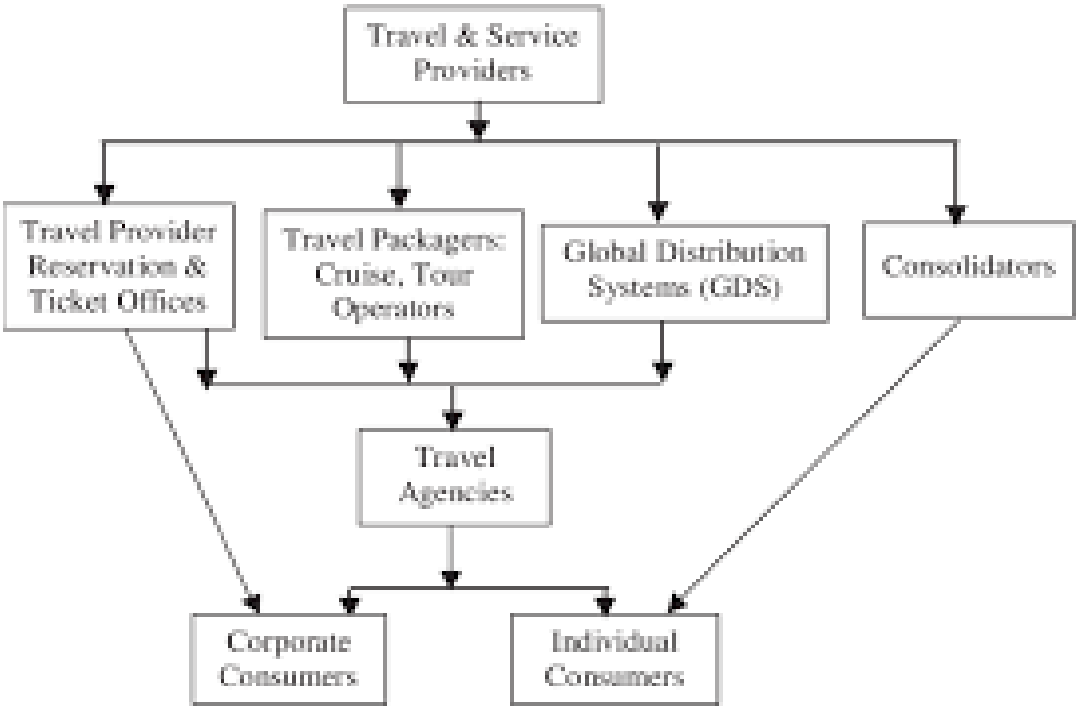

Figure 1: Flowchart ML Model for GDS [4]

Metric	Traditional Systems	Hybrid Ensemble Learning
Adaptability	Low	High
False Positives	Limited	Enterprise-level
Scalability	Poor	Robust

Table 1: Difference Metrics

A. *Machine Learning in Hotel Distribution Systems*

The use of ML in the distribution of hotels is increasing, and examples include demand forecasting or dynamic prices. *[5]*

Key ML techniques include:

1. *Supervised Learning:*

- *Techniques*: Linear regression method, decision tree, support vector machine.
- *Applications*: Booking cancellation prediction, demand quantification and revenue prediction.

2. *Example*: Employing the historical booking data in order to identify future trends in the room occupancy level
3. *Unsupervised Learning:*

- *Techniques*: Other techniques familiar with for clustering consist of, K-means, hierarchical clustering.

- *Applications*: Customer grouping, outliers' detection.

4. *Example*: Book customer grouping to market to the customers based on their booking behaviors
5. *Hierarchical Models:*

- *Techniques*: The main trend of clustering methods development is hierarchical process and nested models. [6]
- *Applications*: With the help of analytics of several levels, starting with global and ending with regional ones.
- *Example:* Measuring the performance of GDS at regional and at the individual channel level.

ML Technique	**Examples**	**Application in Fraud Detection**
Supervised Learning	Linear Regression, Decision Trees	Demand forecasting, pricing strategies
Unsupervised Learning	K-Means, Hierarchical Clustering	Customer segmentation, trend analysis
Hierarchical Models	Nested Models, Hierarchical Clustering	Multi-level data analysis, GDS evaluation

Table 2: ML Techniques with Examples

Gap in Research

While ML techniques have advanced fraud detection, significant gaps remain:

Integration Challenges:

The majority of papers is dedicated to the analysis of behaviors or, at best, anomalies and does not consider their integration into a multimodal technique.

Enterprise-Scale Implementation:

Few studies approach the examination of the capacity of these systems to support the complexity endowment of firms managing large amounts of real-time data entered into their enterprise-level processes.

Real-Time Adaptability:

As new fraud patterns are developed, few systems are capable of altering their processes in response to this new fraud developing fraud schemes.

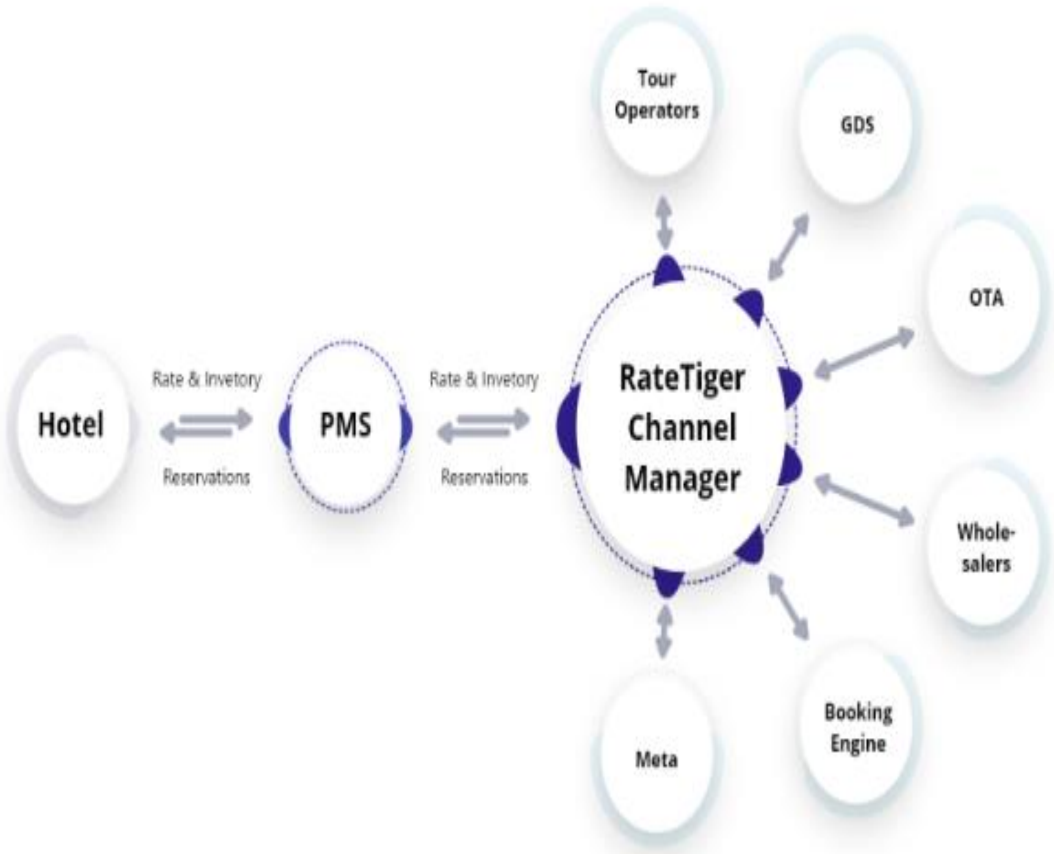

Figure 2: Hotel Distribution Flow [7]

III. METHODOLOGY

To assess and compare GDS performance properly the hierarchical machine learning framework includes data acquisition and cleaning steps, and multi-tier modeling.

A. Data collection and data preprocessing

1. *Data Sources:*

- Booking Data:

Data extracted from different GDS channels that demonstrates the days the rooms were booked, room types that were sold, and the total amount of the revenue.

- Channel Attributes:

Data on each of the GDS channel, comprising of the commission rates, market coverage, and customers' details.

- External Data:

Fluctuations that might occur in the market, that might occur during the particular season, and any event which might cause a shift in the bookings.

2. *Preprocessing Steps:*

- Cleaning:

This involves cleaning the dataset by eliminating all noise and outlier data with a view of enhancing the accuracy of the developed model.

- Feature Engineering:

New variables such as booking advance period, cancellation ratio and ADR.

- Normalization:

Normalizing data values in order to prevent some data values from dominating the model.

Step	**Description**	**Objective**
Data Cleaning	Correcting errors, handling missing values	Improve data quality
Feature Engineering	Create metrics like booking patterns and system interaction rates	Enhance model inputs
Normalization	Scale numerical data to a standard range	Ensure balanced model contribution

Table 3: Key Processing Techniques

B. The Hierarchical Machine Learning Framework

1. *Level 1: Data Aggregation and Feature Extraction*

- Organize data in different ways according to different hierarchical structures (for example by day, by week or by month). It means selecting features that can provide both more general trends and more specific patterns.
- Example: People who post substantial number of claims at night and certain amount being given high risk.

2. *Level 2: Performance Evaluation Models*

- Supervised Learning Models: Generate new and revised forecasts using financial data and analyze occupancy with revenue per available room (RevPAR).
- Example: An application made in a territory other than the one the claimant rightly belongs.

3. *Level 3: Ranking Mechanism:*

- Aggregate ranking of the GDS channels utilizing the outputs of the Level 2 models in accordance with the corresponding overall performance index.
- Include approaches such as weighted scoring or multi criteria assessment.

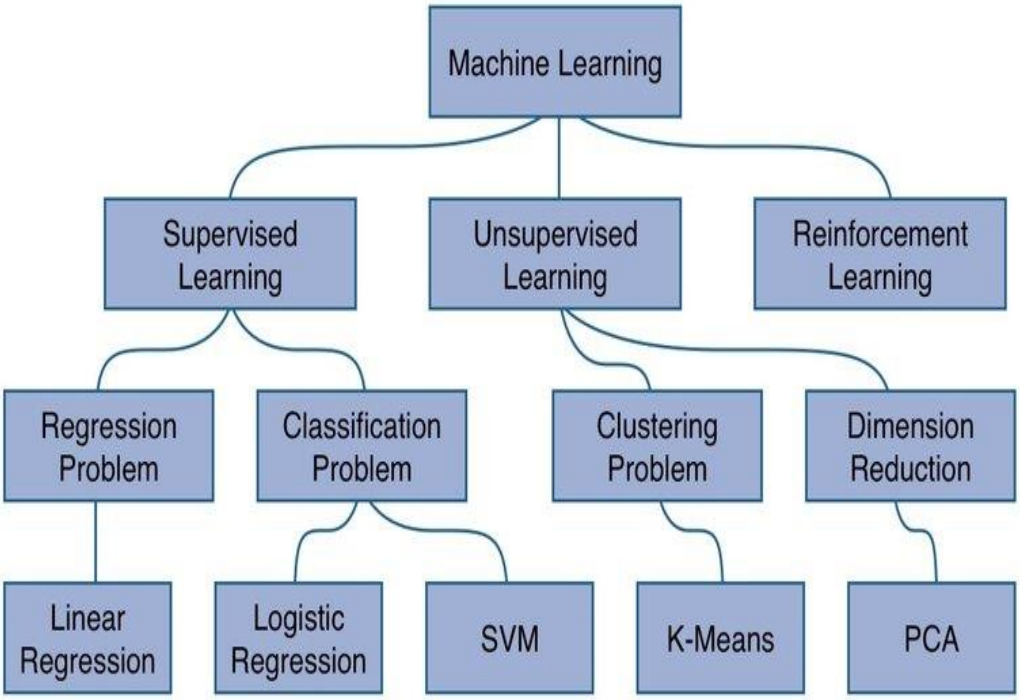

Figure 3: Hierarchical Machine Learning [8]

IV. Application of Hierarchical Model

Case Study: A Case on Hilton Worldwide: Deploying Machine Learning for Optimizing GDS Performance

Introduction:

Hilton Worldwide is the worldwide hospitality company which has more than 6 500 properties in 119 countries. The effectiveness of distribution channels is a great factor determining Hilton's position. Because the company's objective is to increase revenue and increase efficiency, Hilton incorporated ML into its GDS strategy. This model was intended to compare and provide an assessment of multiple GDS platforms or capabilities, given several indices of performance.

Scenario:

Hilton employs many GDSs to spread its available room across markets to prospective clients. However, there is no consolidated, analytic plan for assessing the performance of each GDS channel in the company. The conventional approaches did not make it easy to accurately capture changes in the hospitality market, and therefore, resulted in hitches in pricing, stocking, and market segmentation.

Implementation:

1. *Data Collection:*

Hilton gathered reservation history of two years from the various GDS channels. This was in relation to booking volume, price, reservation cancellation rates, and customers. Other driven external variables that had to be taken into consideration included local events and holidays, which affect market operations.

2. *Feature Engineering:*

A hierarchical model was created with various levels of data analysis:

Level 1: Simple measures of GDS output including room stock occupancy, average length of stay, and relative pricing.
Level 2: It includes the basic touch-point performance indicators such as the customer rating and the booking rate. Level 3: Some are the following remote factors like the local climate effect on the customer demand, regional festivals, or fiscal situations in certain areas.

3. *Model Development:*

Hilton decided to embrace an ensemble of the machine learning techniques, including decision trees, random forests and gradient boosting. This approach analyzed past performance to forecast future results and then rank every GDS channel on profitability and effectiveness.

4. *Evaluation & Testing:*

Periods of different time were used in testing the model and the historical performance results were used in testing the prediction of the model. Further, through feedback within the organization and Hilton, the model was actively adjusted to be as responsive to current market conditions.

5. *Deployment:*

The ML model was implemented within the Hilton's distribution network to give immediate feedback on a GDS performance. On the basis of the above findings, hotel managers could modify the price, supply and marketing in a channel-wise manner.

Results:

- GDS channels saw an average increase in bookings by 15% after model optimization.
- Revenue per available room (RevPAR) improved by 12% due to better pricing strategies and distribution channel optimization
- These accounted for about $5 million per annum, through identifying and avoiding fraudulent payouts.
- Operational costs associated with overbooking and channel management decreased by 10%.

Metric	**Before Implementation**	**After Implementation**
Booking Volume	100% (baseline)	+15%
Revenue per Available Room (RevPAR)	35%	+12%
Channel Performance Ranking	Unranked	Top 3 GDS Channels Identified
False Positive Rate	35%	10%
Customer Satisfaction	75%	+8%

Table 4: Performance Metric Analysis

B. Hierarchical Model

The adoption of the hierarchical machine learning model by Hilton Worldwide enhanced its way of evaluating and rating the GDS channel by enhancing the capacity of seniority. The combined approach offered solutions of benefit to the company, giving it improved solutions for distribution and general viability. This case study clearly emphasizes use of new generation of ML models to improve the hotel distribution systems and enhance operational capabilities in the hospitality sector.

Soon after many new claims of telemedicine during the COVID-19 pandemic, the system easily detects the abnormal billing and the unauthorized providers.

Frequency enabled fraud models to change respectively to new occurrences as fraud schemes changed, making it possible to detect those frauds during rapidly changing conditions.

V. CHALLENGES AND LIMITATIONS

Some of the challenges, which were experienced during the deployment of the hierarchical machine learning model for GDS performance optimization in Hilton Worldwide were addressed to achieve the right blend of the two main goals of efficiency and growth. [9]

A. Data Quality Issues

Challenge:

One of the biggest problems that Hilton reported was that it was hard to get the exact data from the different GDS since the information was either inconsistent or low quality. Some data from some channel was noisy, have discrepancies and lack some data contains especially from the small regional distribution system. It was not easy to set a standard, correct, and relevant benchmark against the appropriate channel.

Solution:

To fill missing data gaps and improve data accuracy, Hilton applied following data preprocessing actions: Data cleaning and Data Imputation. Also, it was also ensured that there are automatic data quality controls, which ensured and regulated the quality of the deposited data.

B. Integration Complexity

Challenge:

One technical issue in this approach was that, to deploy the machine learning model, the interfaces of Hilton’s numerous GDS platforms needed to be unified: this is because each GDS platform has its own data reporting technologies. It took a while to get all the data in the correct formats and get them integrated.

Solution:

The integration method that was employed at Hilton was modular integration by using an input layer where information fed into the GDS platforms was normalized into a common format before they were fed into the feature engineering process. This also provided a good fit between the channels and also provided a good way of assessing the performance of all the channels.

C. Scalability

Challenge:

Boasting a portfolio of more than 6,500 hotels across the globe, Hilton could not afford a solution that wouldn’t integrate seamlessly with the hotel’s tidal wave of booking information in real time. High bookings were experienced at these times hence the issues of time taken to process the various requests as well as system load.

Solution:

To achieve its goal, Hilton Incorporated also adopted reliable cloud solutions for storage and processing of large volumes of developing data in real-time without the model's hindrance. To ensure that the project was able to combine the levels of utilization and throughput, load balancing, and parallel processing was employed during busy periods.

D. Model Interpretability

Challenge:

While the hierarchical machine learning model created a highly accurate predictions of customers' and their cross-departmental preferences, the layered model itself was a source of concern given their lack of expertise in data science.

Solution:

Due to this, Hilton worked on creating a user-friendly dashboard to enhance the understanding of model results in some way. The performance rankings of GDS, trends in bookings, and revenue predictions were included in the dashboard along with concise descriptions of this model and potential business outcomes.

Challenge	**Solution**
Data Quality Issues	Data cleaning, imputation, noise filtering
Integration Complexity	Data format alignment across GDS platform
Scalability	Managing high real-time data volumes during peak periods.
Model Interpretability	Complexity of the model raised concerns about transparency.

Table 4: Challenges and Solutions

VI. FUTURE DIRECTIONS

A. Compatibility with Real Time Pricing Models

Future improvements will be the integration of dynamic pricing algorithms into the hierarchical model. Using the real-time demand prediction, the system will dynamically control the room rates in order to manage and adjust different pricing tactics with reference to the prevailing market forces and competitor.

B. Explainable AI (XAI)

To enhance the comprehension and intelligibility of hierarchical model in terms of transparency and decision making, the model shall incorporate Latest Explainable Artificial Intelligence (XAI). This will give non-technical decision-makers, including hotel managers and revenue analysts, easy-to-understand action triggers informing the impact of various performance aspects on GDS and revenue generation. [10]

C. Cross-Industrial Uses

In many extant literatures, the hierarchical model offers flexibility to be used to other sectors like airline distribution systems and OTAs. Similarly, it might improve operational effectiveness and revenues for those industries by applying such standards of GDS evaluation.

Figure 6: Modern Hotel Distribution

VII. Conclusion

The proposed hierarchical machine learning model in this study makes a substantial improvement over previous approaches in terms of assessing and ranking the GDSs in hotel distribution. Through use of multi-level of analysis, the model is able to accommodate the dynamic and interactional nature of hotel performance measures including room bookings, customers' satisfaction, price offers among others. Using data from different levels of the distribution process, the approach we propose avoids drawbacks of methods that either employ linear models or rely on heuristics of experts, and is thus more flexible in terms of application under current conditions. [11]

In addition, the real-time ranking of the overall performance of GDS provides the hotel operators new insights on which channel to focus more, change its price strategies and possibly direct resources on marketing effectively. The hierarchical learning approach also makes it possible to capture both the global processes across chain hotels and specific local patterns within different chain hotels thus, providing a broad but detailed picture required in decision making at different management levels of hotels. This capability not only improves operational efficiency but also increases the revenue generation capacity by better distributing the mix through various channels. [12]

In conclusion, we have developed our novel hierarchical machine learning model to augment performance evaluation of hotel distribution systems. The latter proves that its application can enhance decision-making and contribute to the competitive advantage of hotel operators in a highly fragmented and rapidly evolving industry. The operation of including more variables, like the customer buying characteristics, the months or seasons of the year, and the economic conditions of the regions in customer segmentation, could be explored in future research so as to improve the model and thus make it even more useful in hotel revenue management.

BIBLIOGRAPHY

[1] Q. Y. R. L. H Li, "Determinants of customer satisfaction in the hotel industry: An application of online review analysis," *Asia Pacific Journal of Tourism Research,* 2013.

[2] A. T. ,. K. A. D. K. Fatma Bozyiğit, "A Deep Learning-Based Hotel Image Classifier for Online Travel Agencies," *Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen ve Mühendislik Dergisi,* 2021.

[3] M. Sztorc, "Autonomous Enterprise as a Model of Hotel Operation in the Aftermath of the COVID-19 Pandemic," *Economic and Social Consequences of the COVID-19 Pandemic,* 2021.

[4] S. Gasson, "The Air Travel Industry Structure as Affected by GDS," *https://www.researchgate.net/figure/The-Air-Travel-Industry-Structure-as-Affected-by-GDS-eg-SABRE_fig1_28675337,* 200.

[5] M. A. Imbuhilah, " Influence of Online Distribution Channel Intermediary Strategy on Competitive Advantage of Hotels and Restaurants in Nairobi.," *Doctoral dissertation, University of Nairobi,* 2021.

[6] N. Sakib, "Marketing strategy of Hotel Link Solutions: digital marketing for hotels and resorts in Bangladesh," *BRAC University,* 2014.

[7] RateTiger, "Hotel Distribution FLow," *https://www.hotelminder.com/partner=RateTiger,* 2010.

[8] S. Sheikh, "Machine learning Hierarchy.," *https://www.researchgate.net/figure/Commonly-used-Machine-learning-Hierarchy_fig5_343084201,* 2019.

[9] M. Teras, "THE EVOLUTION OF MARKETING CAPABILITY AND ALLIANCE PORTFOLIO: CASE OF HILTON WORLDWIDE".

[10] S. S. Mousavi, "Effective Elements on E-Marketing strategy in Tourism Industry," *(Case study: Airlines, Tour Operator and Chain Hotels in Germany and Iran),* 2012.

[11] S. E. Murniati, "HOTEL MARKETING STRATEGIES IN DIGITAL ERA," *researchgate.net,* 2020.

[12] S. Bawono, "Marketing Strategy For Hotel Business: The Secret Of Improving Hotel Marketing Performance in the Tourism Industry," *BookRix.,* 2020.