Analysis and Evaluation of Business and Sales

Sunil Sangave[1], Anjali Dongre[2], Avdhoot Pawar[3], Aditya Bansul[4], Anubhav Pandey[5], Shriya Deshpande[6]

[1,2,3,4,5,6] *Dep. of AIDS Vishwakarma Institute of Technology Pune*

Abstract — ***Cancellations in bookings show a negative impact in the hospitality industry field while making management decisions. To prevent the negative impact of the cancellations, a lot of policies are implemented along with a few overbooking techniques, this, in turn, can severely damage the income and reputation of that particular hotel. To prevent this situation, machine learning models have been developed. These models use previous data from the hotel and then get trained to predict whether the specific reservation would be canceled. For this, two hotels, Resort hotel, and City hotel have been considered and then the ML models are used to predict how these specific actions taken by the hotel management have a practical effect on the revenue and cancellations of the hotel. This makes the management think again about policies and their decisions. The ML models will help management in predicting the number of cancellations that may occur.***

Keywords — Machine Learning, Artificial Intelligence, Power BI, Logistic Regression, Ridge Classifier, SGD Classifier, Random Forest Classifier, Navies Bayes'.

I. INTRODUCTION

Hotel booking cancellations have emerged as a pervasive challenge within the hospitality industry, leading to substantial revenue losses and operational disruptions. As the market becomes increasingly competitive, understanding and mitigating the factors contributing to these cancellations is paramount for effective hotel management. This research paper delves into the intricate patterns of hotel booking cancellations and aims to identify the key determinants that influence revenue generation.

The landscape of hotel bookings is shaped by a multitude of factors, including booking policies, guest demographics, seasonal trends, and broader economic conditions. The rise of flexible cancellation policies and the influence of online travel agencies (OTAs) further complicate this landscape, necessitating a deeper analysis of consumer behavior and market dynamics. By examining these diverse variables, this study seeks to provide actionable insights that empower hoteliers to refine their revenue management strategies.

Ultimately, this research aims to equip hotel operators with the knowledge to anticipate and address booking cancellations, fostering more resilient revenue streams and improving overall operational efficiency in an ever-evolving industry.

II. LITERATURE SURVEY

The hotel industry has been significantly impacted by the cancellation of bookings, resulting in lost revenue for hotels. In recent years, there has been an increasing trend in hotel booking cancellations, which has led to a need for research to understand the factors contributing to this trend. Several studies have been conducted to identify the reasons behind hotel booking cancellations.

This paper [3] looked into integrating ML technology into the hotel industry to improve service delivery, efficiency, and profitability. Using a comprehensive review of literature between 2010 to 2020, the authors framed their analysis based on three research questions: (RQ1) Which implementation areas of ML apply to hotels, (RQ2) specific ML techniques applied, and (RQ3) countries currently implementing ML in hospitality. Despite the importance of ML applications, substantive literature is scarce, and little usage prevails in the fields of demand and price forecasting, booking cancellations, and operational efficiency. Compared to leaders in ML research, China and the USA, a clear lack of studies is found in the geographic regions of Portugal, Spain, and the UK. Further research on ML design and its impact on service quality, revenue management, as well as essential support systems for the successful implementation of ML, is in order.

The paper [4] presents an optimized artificial intelligence-based model for the prediction of hotel booking cancellations by making full use of PNR data. This requires machine learning techniques in conjunction with artificial neural networks optimized by genetic algorithms. A prediction rate of 98% resulted from just 13 independent variables. The study identifies two gaps in the literature individual-level cancellation predictions and the reasons behind cancellations. Most studies identify cancellation ratios in aggregate terms; thus, the need to demonstrate the effectiveness of PNR data in predicting individual cancellations closes these gaps while promoting simpler models that require less historical data to provide greater updates more frequently.

In this paper [5], an innovative model is proposed that integrates PNR in predicting hotel booking cancellations with the challenges of PNR data availability due to privacy concerns. The authors applied a transparent linear machine learning approach by utilizing Bayesian networks for probabilistic modeling and Lasso regression to reduce ineffective predictors and improve predictive accuracy. The empirical results are impressive and demonstrate striking gains in predictive performance. The two enrichment schemes on PNR-based interactive features emerge as essential cancellation predictors. The research under the current study contributes to the knowledge pool value add for hotel management, offering practical implications to develop effective cancellation policies and pricing strategies.

This paper [6] presents a prototype decision support system based on an automated machine learning model to address booking cancellations in the hospitality industry. The system succeeded in achieving over 84% accuracy, 82% precision, and 88% AUC using handling changes in patterns of cancellations and adaptation by introducing the measure of Minimum Frequency for precision evaluation. Using hotel-level predictions of net demand and optimizing price strategies, hotels reduced cancellations by 37 percentage points and recovered roughly €39,000 in revenue. This study is very important in highlighting the practical applicability of predictive analytics in service industries by demonstrating how open-source tools and automated systems could be applied to enhance continuous performance.

This research is important in indicating the potential of predictive analytics and machine learning to improve hotel booking cancellation predictions, enhance operational efficiency, and build more effective revenue management. Implementing personal name records (PNRs) along with advanced modeling techniques, this evidence would be beneficial for hoteliers in many ways and is a call for a data-driven approach toward the reduction of cancellations and improvement of customer satisfaction in the hospitality industry.

III. METHODOLOGY/EXPERIMENTAL

3.1. Logistic Regression: Logistic Regression is a technique of statistical classification, the one where binary classification is predicted by fitting data to a logit function. It's particularly well-suited for situations where the dependent variable is categorical example, yes/no, true/false, 0/1. It's not just like ordinary linear regression, where the output may take on any real value. Suppose, instead, you wanted to build a model that predicted a continuous outcome, like asking how likely it was that someone would vote for a particular candidate-and you'd end up constraining the output to be between 0 and 1. This is very useful with probability estimation.

For example, we have two classes Class 0 and Class 1 if the value of the logistic function for an input is greater than 0.5 (threshold value) then it belongs to Class 1 otherwise it belongs to Class 0. It's referred to as regression because it is the extension of Linear regression but is mainly used for classification problems.

The logistic function is therefore defined as:

$$P(Y=1|X) = \frac{1}{1+e^{-(\beta_0+\beta_1 X_1+\beta_2 X_2+\ldots+\beta_n X_n)}}$$

Where:

- P(Y=1|X) is the probability that the dependent variable Y equals 1 given independent variable X.
- B_0 is the intercept of the model.
- $B_1, \beta_2, \ldots, \beta_n$ are the coefficients for each independent variable $X_1, X_2, \ldots, X_n$.
- e is the base of the natural logarithm.

3.2. Ridge Classifier: The Ridge Classifier is a classification algorithm based on the Ridge Regression technique. It can help the training while dealing with multicollinearity; in this case, the independent variables of the data are highly correlated with each other. Ridge Classifier is an analogy of logistic regression when there is a regularization term added to the loss function to prevent overfitting for cases with high dimensionality. The regularization pushes the coefficients of less significant features towards zero so that the model is generalized.
For instance, if you have to predict whether customers would churn or not (1 or 0) based on their

account balance and usage patterns, Ridge Classifier can come in handy when such features are highly correlated. This will regulate the model by not overemphasizing feature values that could be causing noise due to multicollinearity.

Assuming coefficients and the regularization parameter α are known, the Ridge Classifier tries to fit the data with loss minimization incorporating a penalty for large coefficients that stabilizes the model.

The mathematical formula of the Ridge Classifier is defined by the regularized least squares problem**:**

$$\min_{\beta}\left(\sum_{i=1}^{n}(y_i - X_i\beta)^2 + \alpha\sum_{j=1}^{p}\beta_j^2\right)$$

Where:

- y_i is the actual label (binary or categorical).
- X_i is the feature vector for instance iii.
- β represents the model coefficients (weights).
- α is the regularization parameter (also called the "ridge penalty") that controls the amount of shrinkage applied to the coefficients.
- p is the number of features.

The regularization term $\alpha\sum j=1p\beta_j^2$ prevents the model from becoming too complex by penalizing large coefficients.

3.3. Stochastic Gradient Descent Classifier: The SGD Classifier is linear. Because it uses the SGD optimization technique to fit the model, it works very efficiently for large-scale data sets and can easily be adapted to several problems in the scope of machine learning, including classification and regression problems. A user of SGDClassifier can choose from a variety of loss functions: hinge loss (for SVM), log loss (for logistic regression), and others.

For every training example, the stochastic gradient descent optimizes the loss function by iteratively updating model parameters as opposed to using all the data together as in batch gradient descent, thus leading to faster computations, but it may introduce more variability in the updates with large sets.

The objective function for SGDClassifier:

$$\min_{\beta}\left(\frac{1}{n}\sum_{i=1}^{n}L(y_i, X_i\beta) + \alpha\|\beta\|^2\right)$$

Where:

- $L(y_i, X_i\beta)$ is the loss function, which could be hinge loss for SVM or log loss for logistic regression.
- α‖β‖2 is the regularization term that prevents overfitting.
- n is the number of samples.
- β represents the model coefficients (weights).
- α is the regularization parameter (L2 penalty).

3.4. Random Forest Classifier: The Random Forest Classifier is an ensemble learning algorithm where several decision trees are constructed during training and the mode of the classes (for classification) or the mean prediction (for regression) output of these individual trees is given. In other words, this method tries to overcome overfitting through the average computation of many decision trees so that improved performance and robustness of a single decision tree can be attained.

Each tree in the random forest is built using a random subset of data, and at each node, it considers a random subset of features for splitting. It is this randomness that prevents correlation among the individual trees, thus maintaining the stability and accuracy of the ensemble of trees.

3.4.1. Working:

Bootstrapping: From the original training dataset, multiple random subsets, known as bootstrapped samples, are generated. These samples are built by sampling with replacement.
Building Decision Trees: We build a decision tree for each of the bootstrapped samples. In this problem, however, when splitting a node, we only consider a randomly selected subset of features rather than all the features if we have more than a thousand. This will make the trees more diverse.
Aggregation (Voting/ Averaging): For classification problems, the random forest output is ascertained based on majority voting among all decision trees. For regression, it calculates the mean of all the tree predictions.

Mathematical Formula for Random Forest:
In a classification problem, the random forest will predict the class label.

$$\hat{y} = \text{mode}\left(\{h_1(x), h_2(x), \ldots, h_n(x)\}\right)$$

Where:

- $h_i(x)$ is the prediction of the iii-th decision tree for input sample xxx.
- y^ is the final predicted class (for classification).

For a regression task, the prediction y^ is the average of all the individual tree predictions:

$$\hat{y} = \frac{1}{n}\sum_{i=1}^{n} h_i(x)$$

Where:

- $h_i(x)$ is the prediction from the i^{th} decision tree.
- n is the number of trees in the forest.

3.5. K-Nearest Neighbor Algorithm: The K-Nearest Neighbor (KNN) algorithm is a straightforward method for classifying data and is classified as a supervised learning technique. It is a parametric-less algorithm, meaning it does not make assumptions about the underlying data distribution. Since it does not learn from the training data, it can be referred to as a lazy or idle learner; however, during classification, the data is stored, and actions are performed on it. This algorithm searches for similarities during classification.

3.5.1. Why KNN?
For instance, if there are two categories, A and B, and a new data point arrives between A and B, KNN can easily classify whether the new data belongs to category A or B.

3.5.2. Working:

1. First, determine the number of neighbors (K) to consider.
2. Calculate the Euclidean distance from every data point to the K estimators.
3. Identify the nearest neighbor based on the calculated distances.
4. Allocate the data points to the group with the highest number of neighbors.
5. The model is now ready.

The Euclidean distance between points A and B is given by the formula:

$$d(P,Q) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

There is no definitive method to determine the optimal K value; experimentation is often required to find the best output, with $K = 5$ being a common choice. A very low K value may lead to outliers and noise in the output.
3.5.2.1. Advantages:

- Implementation of KNN is straightforward.
- No discriminative function is derived from the data.
- With a larger training dataset, KNN becomes more effective.

3.5.2.2. Disadvantages:

- Standardization and normalization must be completed before implementing KNN.
- The KNN algorithm is computationally expensive, as it requires distance calculations for each data point in the dataset.

3.6. Decision Tree: The decision tree algorithm falls under supervised learning and can be used for both regression and classification problems, although it is primarily utilized for classification.

- It maintains a tree structure, consisting of internal nodes, branches, decision rules, and leaf nodes.
- Each node represents a feature from the dataset, while the decision rules are the branches, leading to an output at the leaf nodes.

In a decision tree, two types of nodes exist:
1. Decision Nodes: Used for making decisions, with branches leading to further nodes.
2. Leaf Nodes: Output nodes with no further branches.

A decision tree attempts to provide a solution in all possible ways and is represented graphically as a tree. The CART (Classification and Regression Trees) algorithm is commonly used to build decision trees.

3.6.1. Why Decision Tree?
1. It mimics human decision-making processes.
2. The logic is easily understandable.

3.6.2. Working:
1. The tree starts with a root node containing all variables or features in the dataset.
2. Attribute selection or Principal Component Analysis (PCA) is used to identify the best attributes of the dataset.
3. The root is divided into multiple sub-trees using the best attributes or features.
4. The decision tree with the best feature is selected.
5. This process repeats until no further trees can be constructed from the dataset features.

3.6.3. Rules for Feature Selection:

1. After dividing the dataset based on an attribute, the change in entropy (known as information gain) is measured. Information gain quantifies how much detail a feature provides about a class. The decision tree algorithm aims to maximize the information gain, splitting nodes with the highest information gain value.

Information Gain (IG) can be calculated using the following formula:

$$IG = E(S) - [(WA) \times E(I)]$$

Where:
- E = entropy
- WA = weighted average
- I = each feature
- IG = Information Gain

2. Entropy: The impurity present in a dataset feature, describing randomness. It can be calculated as:

$$s_e - \left(\frac{16}{40} \times s1_e + \frac{24}{40} \times s2_e\right)$$

3. Gini Index: A measure of impurity or cleanliness used in constructing decision trees with the CART algorithm. Attributes with a low Gini Index are preferred over those with a high Gini Index. The Gini Index is calculated as follows:

$$GI = 1 - \sum_j P_j^2$$

Where GI is the Gini Index.

3.6.3.1. Advantages of the Decision Tree:
- Useful for decision-related problems.
- Encourages exploration of all possible solutions.
- Requires minimal data wrangling compared to other models.

3.6.3.2. Disadvantages of the Decision Tree:
- May become complex due to multiple layers.
- Overfitting issues can be resolved using the Random Forest algorithm.
- Requires more time for model training, indicating computational complexity.

3.7. Naive Bayes Classifier

The Naive Bayes classifier belongs to the probabilistic classifier family and is based on Bayes' Theorem. It employs kernel density estimation alongside a simple Bayesian network. Naive Bayes is a highly scalable classifier that efficiently handles multiple variables with linear relationships. It assumes that features are independent of each other when constructing classifiers.

For example, consider a fruit like a mango that is yellow, round, and has a diameter of 14 cm. The Naive Bayes algorithm would consider all these features to contribute equally or independently to determining whether the fruit is a mango.

The Naive Bayes classifier efficiently uses conditional probability, expressed as:

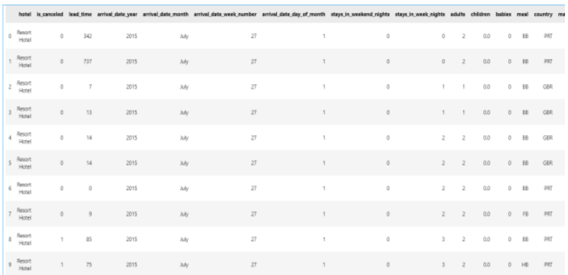
$$k = \frac{c \times i}{l}$$

Where:
- Posterior = K
- Prior = c
- Likelihood = i
- Evidence = l

Interestingly, the accuracy of the model or algorithm is not correlated with the number of independencies.

IV. PROPOSED WORK

The primary aim of the hospitality industry is to provide quality services to its customers which in return leave They will have a positive impact in their minds. This will create some Positive word of mouth and reviews. Customer satisfaction.
This is to become the main aim of the hospitality industry.



	hotel	is_canceled	lead_time	arrival_date_year	arrival_date_month	arrival_date_week_number	arrival_date_day_of_month	stays_in_weekend_nights	stays_in_week_nights	adults	children	babies	meal	country	ma
0	Resort Hotel	0	342	2015	July	27	1	0	0	2	0.0	0	BB	PRT	
1	Resort Hotel	0	737	2015	July	27	1	0	0	2	0.0	0	BB	PRT	
2	Resort Hotel	0	7	2015	July	27	1	0	1	1	0.0	0	BB	GBR	
3	Resort Hotel	0	13	2015	July	27	1	0	1	1	0.0	0	BB	GBR	
4	Resort Hotel	0	14	2015	July	27	1	0	2	2	0.0	0	BB	GBR	
5	Resort Hotel	0	14	2015	July	27	1	0	2	2	0.0	0	BB	GBR	
6	Resort Hotel	0	0	2015	July	27	1	0	2	2	0.0	0	BB	PRT	
7	Resort Hotel	0	9	2015	July	27	1	0	2	2	0.0	0	FB	PRT	
8	Resort Hotel	1	85	2015	July	27	1	0	3	2	0.0	0	BB	PRT	
9	Resort Hotel	1	75	2015	July	27	1	0	3	2	0.0	0	HB	PRT	

Fig.1. Monitored Data

4.1. Power BI: Power BI is a powerful business analytics service developed by Microsoft. It allows users to connect to their data, share insights across the organization, and model their data in the most appropriate way for business use. It enables making better business decisions by allowing users to connect to various data sources, transform and model data, and create highly interactive reports and dashboards. Power BI is well-suited for its intuitive interface, robust data modeling capabilities, and seamless integration with other Microsoft products.

4.2. Components of Power BI:

- Power BI is a powerful business analytics service developed by Microsoft.
- It allows users to connect to their data, share insights across the organization, and model their data in the most appropriate way for business use.
- It enables making better business decisions by allowing users to connect to various data sources, transform and model data, and create highly interactive reports and dashboards.
- Power BI is well-suited for its intuitive interface, robust data modeling capabilities, and seamless integration with other Microsoft products.

The different components of Power BI are meant to let users create and share business insights in a way that fits with their role shown in Figure 3.

4.3. System Architecture

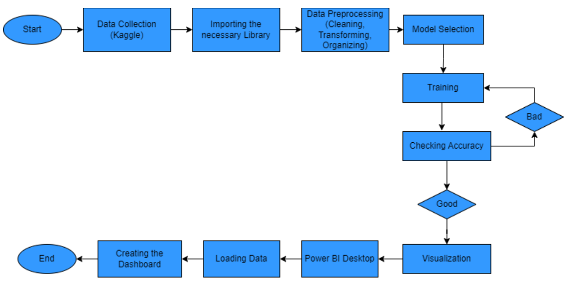


Fig.2. System Architecture

Step-1: First of all, collect the data from the Kaggle and store it in the working directory.
Step 2: Importing the Necessary Python libraries.
Step 3: Next, we perform data preprocessing to prepare the raw data for analysis by executing the following tasks: cleaning, transforming, and organizing the data.
Step 4: Next, we proceed with model selection, evaluating the requirements of our analysis to choose suitable machine learning models.
Step 5: Next, we train the chosen models with the preprocessed dataset, helping them to understand patterns and relationships that will allow for accurate predictions on new, unseen data.
Step 6: After training, we assess the model's performance by examining its accuracy and other important metrics. If the accuracy meets our expectations, we move on to visualization. However, if it falls short, we go back to the model selection phase to consider different models or tweak the parameters.
Step 7; Once we have satisfactory model accuracy, we visualize the results to interpret the findings from our analysis. This includes using various visualization techniques to present the insights gained from the data.
Step 8: Now we are moving to use Power BI Desktop. In which we load the Data and after that, we create the interactive Dashboard (Fig.3 and Fig.4)

Fig.3.Component of Power BI

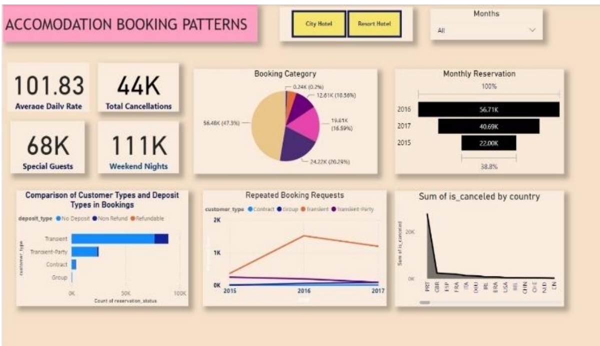


Fig.4.Components of Power BI

V. RESULTS

This evaluation provides insights into the diverse predictions generated by the model, as demonstrated in Figure 5.

```
Average cross validation score: 0.990
Test accuracy: 0.990
F1 score: 0.986
[[15027    23]
 [  221  8607]]
```

Logistic Regression

```
Normalized data:
Average cross validation score: 0.919
Test accuracy: 0.917
F1 score: 0.874
[[15037    13]
 [ 1965  6863]]
Standard scaled data:
Average cross validation score: 0.986
Test accuracy: 0.986
F1 score: 0.980
[[15044     6]
 [  337  8491]]
```

Ridge Classifier

```
Average cross validation score: 0.988
Test accuracy: 0.987
F1 score: 0.982
[[15041     9]
 [  304  8524]]
```

SGD Classifier

```
Before feature selection:
Best parameters: {'max_depth': 5, 'n_estimators': 100}
Best score: 0.9974872300227775
Test score: 0.996
After feature selection
Best parameters: {'max_depth': 5, 'n_estimators': 50}
Best score: 0.9977803749554823
Test score: 0.998
```

Random Forest Classifier

```
Average cross validation score: 0.876
Test accuracy: 0.883
F1 score: 0.838
[[13919  1131]
 [ 1653  7175]]
```

KNN

```
Average cross validation score: 0.958
Test accuracy: 0.956
F1 score: 0.936
[[15016    34]
 [ 1027  7801]]
```

Naïve Baiyes Classifier

Fig.5 Prediction of Models

The analysis of the following figures reveals important insights into customer behavior and booking patterns.

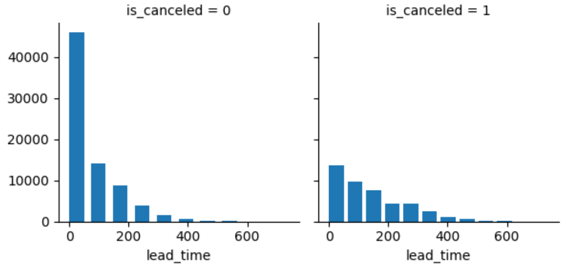


Fig.6. Lead Time vs. Cancellations

Figure 6 is taken from our research depicting the lead time distribution, calculated as the time lag between making a reservation and the guest's date of arrival. Using that graph, we know what kinds of reservations are being made when during the year and how those lead times affect cancellations by enabling us to draw some trends about it. For instance, longer lead times increase the possibility for cancellations. This insight will allow one to understand the behavior of a customer in booking, optimize resource allocation, and strengthen the predictability of cancellations using lead time.

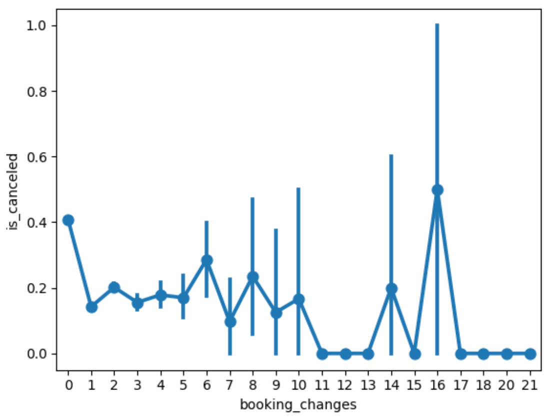


Fig.7. Booking Changes vs Cancellation

Figure 7, shows how many guests change their bookings and the possibility of cancellation. Most booking has no change, and more of them are not cancelled which means guests who do not make changes to their bookings have less possibility of cancellation. But the graph indicates an increased number of alterations in booking probably points to a higher level of possibility of cancellations due to indecision or travel plans changed.

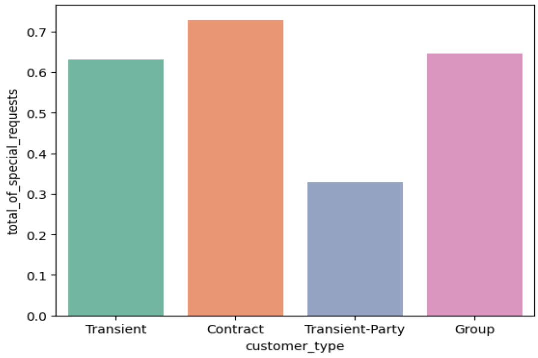


Fig.8. Customer Types vs Special Requests

Fig. 8 identifies several service requests from various segments of customers. With this, it is noted that transient customers often have the greatest number of special requests; this indicates a desire for tailored services. It is valuable business information, hence, especially for businesses that might want to improve customer satisfaction and respond to needs for specific services by recognizing the requirement through understanding the preferences of various groups of customers in their service delivery.

VI. CONCLUSION

In conclusion, our research on hotel reservation cancellations reveals all of the added value in further developing hotel management and revenue strategies. There is a heavy contrast between city and resort cancellation percentages; this calls for tailor-made marketing practices. Our machine learning model might predict potential cancellations based on historical data, thus helping the hotels reduce their losses. Interactive dashboards through Power BI can be included to enable hotel management to make informed real-time decisions. It is then that such data-driven insights and predictive analytics will help improve the guest experience and sustain competitiveness in this emergent hospitality industry.

VII. ACKNOWLEDGEMENT

We are immensely grateful to Dr.Sunil Sangave, Professor at Vishwakarma Institute of Technology, Pune, whose consistent guidance and support were a milestone for this project.

REFERENCES

[1] Antonio, N., De Almeida, A., & Nunes, L. (2017). Predicting hotel booking cancellations to decrease uncertainty and increase revenue.

Tourism & Management Studies, 13(2), 25-39.

[2] Sánchez-Medina, A. J., & Eleazar, C.(2020). Using machine learning and big data for efficient forecasting of hotel booking cancellations. International Journal of Hospitality Management, 89,102546.

[3] Dr. Eid Alotaibi, "Application of Machine Learning in the Hotel Industry: A Critical Review", Journal of Association of Arab Universities for Tourism and Hospitality (JAAUTH), Vol. 18 No. 3, (2020), pp. 78-96.

[4] Agustín J. Sánchez-Medina, Eleazar C-Sánchez, "Using machine learning and big data for efficient forecasting of hotel booking cancellations", International Journal of Hospitality Management, Volume 89, August 2020, 102546.

[5] Shuixia Chen, Eric W.T. Ngai , Yaoyao Ku , Zeshui Xu, Xunjie Gou, Chenxi Zhang, "Prediction of hotel booking cancellations: Integration of machine learning and probability model based on interpretable feature interaction", Decision Support Systems, Volume 170, July 2023, 113959

[6] Nuno Antonio, Ana de Almeida, Luis Nunes, "An Automated Machine Learning Based Decision Support System to Predict Hotel Booking Cancellations", Data Science Journal, Vol. 18 (2019), 08 July 2019.

[7] Adil, M., Ansari, M. F., Alahmadi, A., Wu, J. Z., & Chakrabortty, R. K. (2021).Solving the problem of class imbalance in the prediction of hotel cancelations: A hybridized machine learning approach. Processes, 9(10), 1713.

[8] Antonio, N., de Almeida, A., & Nunes, L. (2017, December). Predicting hotel bookings cancellation with a machine learning classification model. In 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA) (pp. 1049-1054).

[9] Lee, H., Yang, S. B., & Chung, N. (2021). Out of sight, out of cancellation: The impact of psychological distance on the cancellation behavior of tourists. Journal of Air Transport Management, 90,101942.

[10] Song, K. S. (2021). Simultaneous statistical modelling of excess zeros, over/underdispersion, and multimodality with applications in hotel industry. Journal of Applied Statistics, 48(9), 1603-1627.

[11] H. Zhang, Y. Li, and H. Sun, "The impact of online travel reviews on hotel booking intentions and the moderating role of source credibility," Journal of Hospitality Marketing & Management, vol. 28, no. 3, pp. 273-291, 20.

[12] Choi, Y., & Mattila, A. S. (2011). Hotel revenue management and its impact on customers' perceived fairness and satisfaction. Journal of Hospitality & Tourism Research, 35(3), 310-328.

[13] Gao, J., Petrick, J. F., & Wetzels, M. (2016). The influence of tourist behavioral intentions on search engine use: The role of destination familiarity and trust. Journal of Travel Research, 55(7), 935-949