
Exploratory Data Analysis & Booking Cancellation Prediction on Hotel Booking Demands Datasets

Pujo Hari Saputro ^{1,*}, Herlino Nanang ²

¹ Department of Informatic, Alma Ata University, Indonesia

² Department of Information System, International Islamic University Malaysia, Malaysia

^{1,*} pujo@almaata.ac.id, ² herlino.nanang@live.iium.edu.my

* corresponding author

(Received December 21, 2020 Revised January 4, 2021 Accepted January 14, 2021, Available online January 15, 2021)

Abstract

Online ordering is the latest breakthrough in the hospitality industry, but when it comes to booking cancellations, it has a negative impact on it. To reduce and anticipate an increase in the number of booking cancellations, we developed a booking cancellations prediction model using machine learning interpretable algorithms for hotels. Both models used *Random Forest* and the *Extra Tree Classifier* share the highest precision ratios, *Random Forest* on the other hand has the highest recall ratio, this model predicted 79% of actual positive observations. These results prove that it is possible to predict booking cancellations with high accuracy. These results can also help hotel owners or hotel managers to predict better predictions, improve cancellation regulations, and create new tactics in business.

Keywords: EDA, Data Analysis, Booking Cancellation, Random Forest, Extra Tree Classifier

1. Introduction

'Application of information systems and pricing techniques to assign the correct resources at the right time to the right customer at the right price' is referred to as revenue management[1]. Originally developed in the aviation industry in 1966, it has increasingly been adopted in other service sectors, such as rental cars, golf courses, and in particular, hotels[2][1]. The former definition was updated by the hospitality industry to: "make available the right space for the right guest and the right price via the right distribution channel at the right time"[3]. Hotels allow reservations online in order to cope with this. The reservation is a contract between the customer and the hotel which gives the customer the right to use the facility at a mutually agreed price in the future. Usually, until the terms of services are included, there is an opportunity or option to revoke the deal or reservation. However the decision to cancel the service prior to its provision puts all the responsibility on the hotel, which must guarantee the customer who made the reservation the availability of the room.

While reservations normally allow guests to revoke a room with and without penalty until both the provision of products and services, hotels should risk the possibility of guaranteeing rooms for customers that meet with their reservations. Hotels could also bear the cost of having vacant rooms at about the same time[4] when guests cancel or do not appear. Hotels introduce overbooking and restrictive cancellation policies to minimize this risk[5], And yet both overbooking and strict cancellation policy can damage the performance of hotels. Overbooking, on the other hand, might compel the hotel to refuse a client a service. This can be a really poor customer experience which can lead to online reviews and a negative effect on social reputation[6]. This relocation may also carry the client to a hotel he/she may like and lead the client to lose future reservations[3].

Several research on booking cancellation, allocation of inventory, cancellation policies and overbooking techniques relevant to subjects[4, 5, 7]. The majority of studies focus primarily on the airline industry, which is quite different from the hospitality market. But, lately, the number of studies related to the hospitality industry has increased. The majority of research used the conventional methods of statistics and just a few benefits of the methodology and techniques of machine learning. Currently four studies are unique to the hospitality industry, considering the presence of many studies on the subject[8, 9, 10, 11]. Most studies consider it as a regression problem in the prediction of booking cancellations. Just a couple of the recently published studies address the subject as a classification problem and focus on forecasting the global cancellation rate rather than the risk of each booking being cancelled. In fact, Morales and Wang[11] stated that it is difficult to believe that one can determine with high precision whether or not a booking would be canceled. António, however, showed that the probability of booking cancellations can be predicted with high precision[8][12]. The percentage of all bookings expected to cancel can be deducted from the demand to calculate the hotel's net demand, i.e. the demand minus bookings that are likely to cancel. Designed with a particular market value, the revenue manager of a hotel will make fair and reasonable choices about demand control and develop overbooking strategies and cancellation policies.

2. Literature Review

Mehrotra[3] described that a key aspect of revenue management is good demand forecasting. Talluri[13] also acknowledged the importance of forecasting in revenue management by stating that revenue management systems require quantity forecasts and that the quality of these forecasts is critically dependent on their performance. The demand forecast was described by authors such as Ivanov[14] or Morales[11] as one of the aspects where forecasting is relevant. Booking cancellations are behind this need to forecast demand, since they do not reflect the actual demand for their services in the hospitality sector, as in other service sectors that deal with advanced bookings, because there are also a large number of cancellations.

Booking cancellations are common in revenue management as it relates, in particular, to the service and hospitality industries. Nevertheless the Internet has increasingly changed the way customers search and buy travel services in recent years. In particular research on control-related problems used to mitigate the effect of cancellations on sales and inventory actual delivery, cancellation policy, and overbooking has increased in this area generally. However there seems to be no amount of studies of booking cancellation predictions for the hospitality industry.

3. Method

3.1. Exploratory Data Analysis (EDA) & Feature Engineering

EDA or exploratory data analysis is a technique for fitting linear and higher functions to relationships, for structuring and transforming variables with arithmetic functions, for splitting relationships into partitions and clusters, for extracting features through statistical results and such. One example of the results of EDA is a simple histogram that describes discrete and continuous variables, schematic plots that provide general & partial relationship characteristics that distinguish 2 habits, simplification of functions from low dimensional relationships and two-way tables such as contingency tables.

On this part, we would like to visualize some features and show statistical relationship with target variable. This analysis will help to get overall view and deep familiarity of the data, detect extreme values and identify obvious errors. First graph fig. 1 is about exploring hotel features which denotes the type of the hotels. According to the below graph, approximately 34% of the data was booked for resort hotels and the rest was booked for City Hotel. Now, let's dive into the target value of data fig. 2. The numbers are similar with hotel

features. While 37% of bookings were canceled, 63% of bookings are not canceled. These numbers also show that there is no balanced problem on the target value.

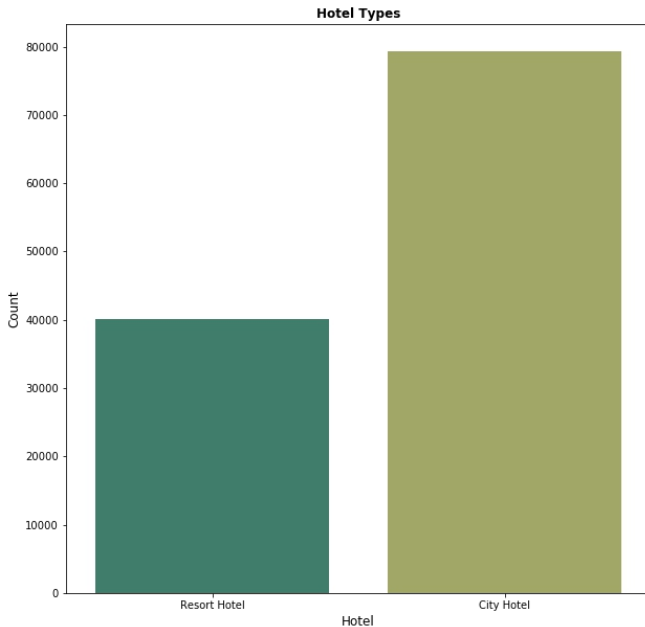


Figure. 1. Hotel types

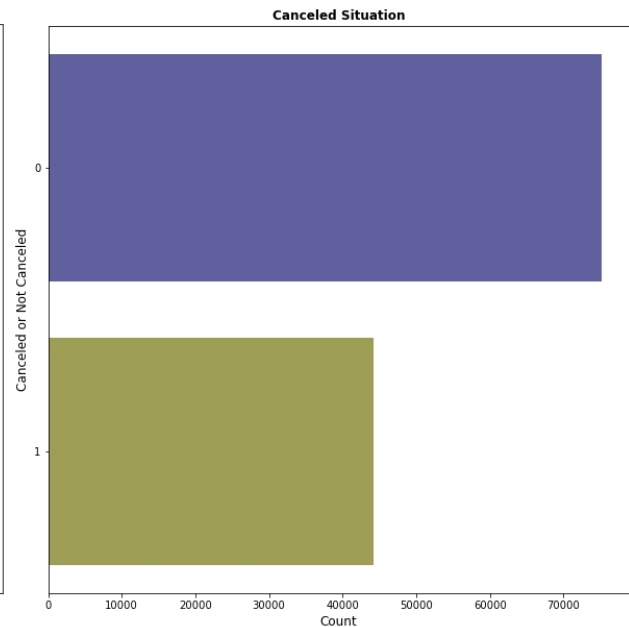


Figure. 2. Canceled situation

Graph below fig. 3 shows the relationship of arrival date (year) to lead time with booking cancellation status. A graph made from the violin plot. The violin plot is a combination of a plot of box and density plot. It shows the distribution of the data. Three violin plots are corresponding to three different years. For canceled booking, means and interquartile ranges are similar in all years. But the shapes of the distributions are quite different from each other. On the other hand, the distribution of not-canceled booking is almost the same. For all years and every booking situation, the small number of large lead time values are pulling the mean up. It shows that the mean is higher than the median.

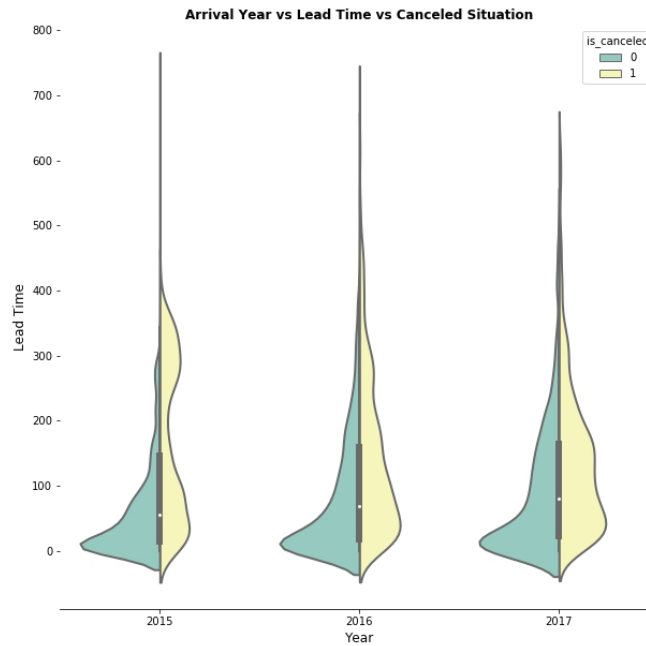


Figure. 3. Arrival year vs Lead time vs Canceled situation

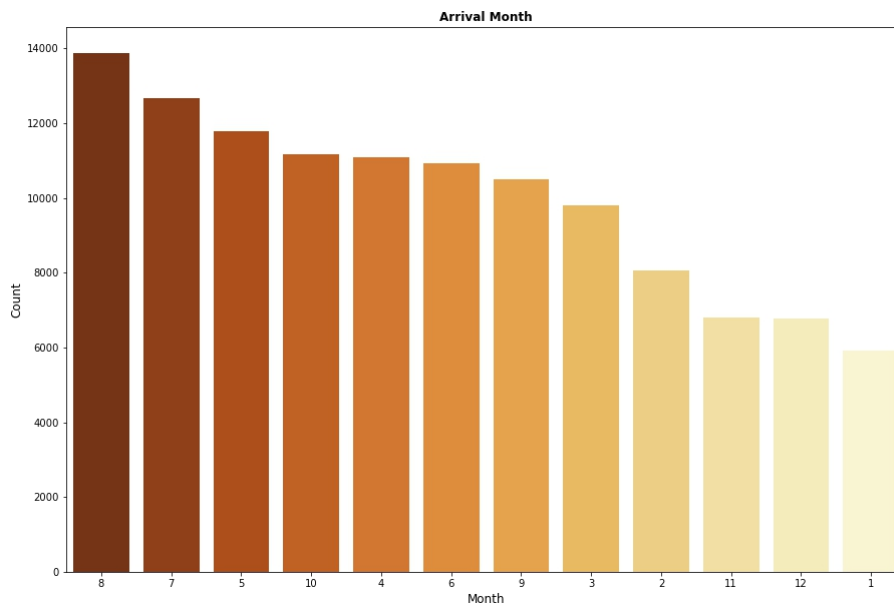


Figure. 4. Arrival month

Fig. 4 is showing the number of bookings for each month. According to that, August is the busiest month and January is the most unoccupied month. It is half as busy as August. Another important features which are related to time are stays_in_week_nights and stays_in_weekend_night features. The below table shows the relationship between these two features. According to that, there is some missing data. 715 values are inputted zero both weekend and weeknights. However, this missing data is small enough to neglect.

Table. 1. Night-time stays feature

stays_in_weekend_nights	0	1	2	3	4	5	6	7	8	9	10	12	13	14	16	18	19	Total
0	715	4569	2361	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7645
1	16451	7326	6533	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30310
2	17956	8979	6749	0	0	0	0	0	0	0	0	0	0	0	0	0	0	33684
3	11564	6152	4542	0	0	0	0	0	0	0	0	0	0	0	0	0	0	22258
4	4482	2412	2669	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9563
5	830	1188	8655	310	94	0	0	0	0	0	0	0	0	0	0	0	0	11077
6	0	0	851	301	347	0	0	0	0	0	0	0	0	0	0	0	0	1499
7	0	0	446	399	184	0	0	0	0	0	0	0	0	0	0	0	0	1029
8	0	0	393	131	132	0	0	0	0	0	0	0	0	0	0	0	0	656
9	0	0	81	63	87	0	0	0	0	0	0	0	0	0	0	0	0	231

The above table brings an idea about creating a new feature. Which is indicated just_stay_weekend, just_stay_weekday and stay_both_weekday_and_weekday. These 715 values which are not assigned any feature, indicated as undefined_data. Next graph (Fig. 5) is about the weekend_or_weekday feature's relationship with arrival_date_month. Below bar graph shows that most bookings were made to stay only for weekdays or both weekdays and weekends. On the other hand, the numbers of staying in just the weekend category are quite low compared to other categories.

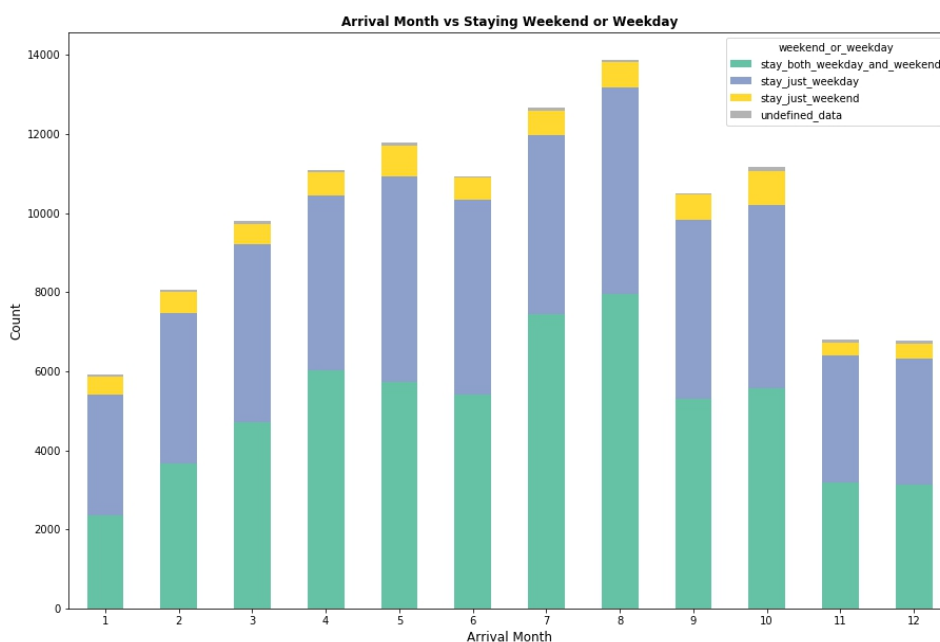


Figure. 5. Arrival month vs Staying weekend or weekday

The below graph (Fig. 6) gives information about the location which bookings made in. According to that, there is an apparent difference in booking location between Portugal and the others. Approx. 40% of all bookings made in the same location: Portugal. The statistics show that online hotel and airline reservations have increased in recent years. Most people complete their reservation via their smartphones. And the other graphs (Fig. 7) summarise these statistics. More than 45% of bookings are made via Online Travel Agents and around 20% of bookings made via Offline Travel Agents. Less than 20% of bookings made directly without any agents.

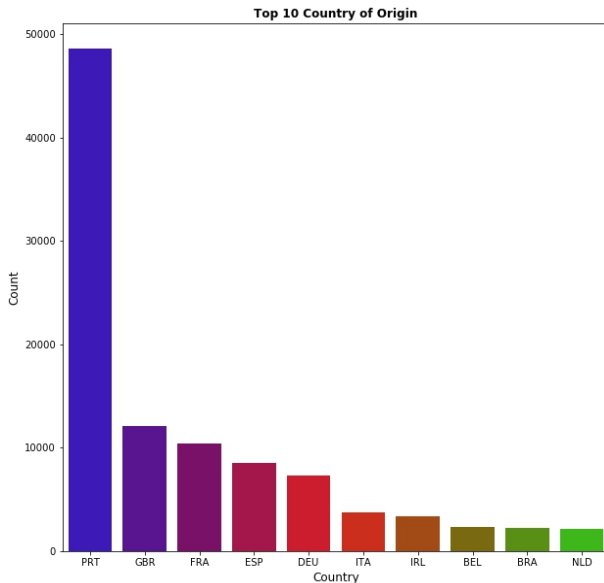


Figure. 6. Country of origin

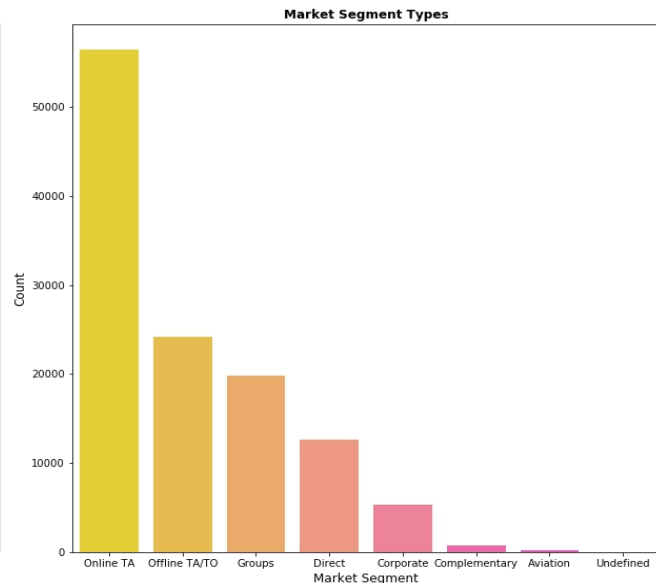


Figure. 7. Market segment types

Table. 2 shows if there is any difference between assigned and reserved room types or not. The results are shown as a percentage. Average 84% of bookings keep their reserved room and the rest of the rooms have been changed. Every row represents the reserved type and distribution over the columns shows what was the assigned room despite the reserved type.

Table. 2. Comparison between room type

assigned_room_type	A	B	C	D	E	F	G	H	I	K	L	P
reserved_room_type												
A	86.0	1.0	2.0	9.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
B	10.0	88.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
C	1.0	0.0	95.0	1.0	0.0	0.0	1.0	1.0	1.0	0.0	0.0	0.0
D	2.0	0.0	0.0	92.0	4.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0
E	0.0	0.0	0.0	0.0	91.0	6.0	2.0	0.0	1.0	0.0	0.0	0.0
F	0.0	1.0	0.0	0.0	1.0	93.0	4.0	0.0	0.0	0.0	0.0	0.0
G	0.0	0.0	0.0	0.0	0.0	1.0	97.0	0.0	1.0	0.0	0.0	0.0
H	0.0	0.0	0.0	0.0	0.0	0.0	2.0	97.0	1.0	0.0	0.0	0.0
L	17.0	17.0	17.0	0.0	0.0	17.0	0.0	17.0	0.0	0.0	17.0	0.0
P	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0

Below (Fig. 8), another exploratory analysis made for diving deep into the relationship between ADR, arrival month, and booking cancellation status. As explained in the previous graph of arrival month (Fig. 4), August is the most intense month of bookings. Besides the highest Arrival Daily Rate has occurred in that month too. Except for rush months like August, July, and September, canceled bookings have higher ADR than not canceled bookings. Maybe this highest rate could be one of the reasons for canceled bookings.

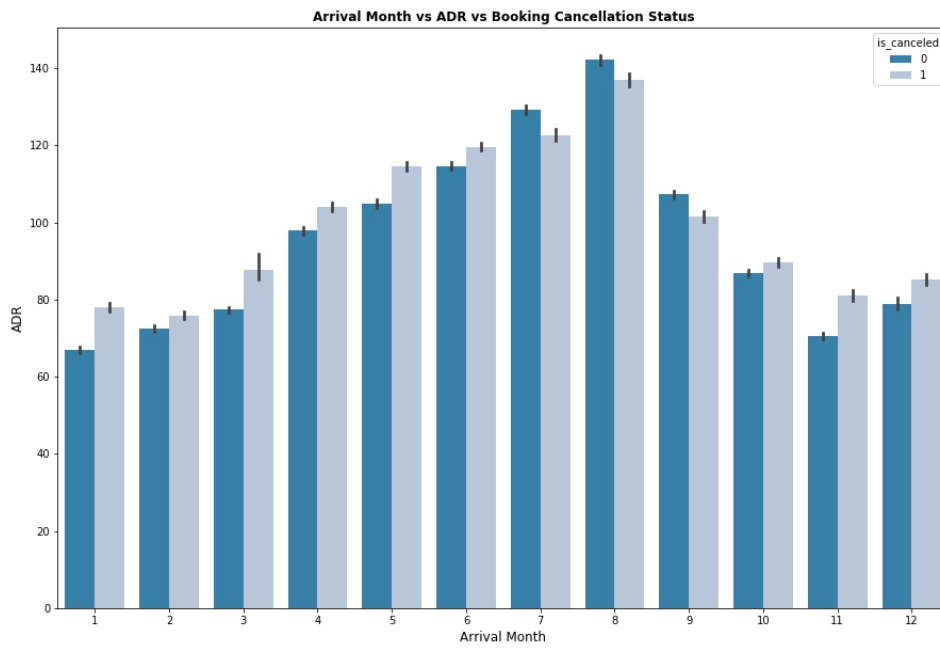


Figure. 8. Arrival month vs ADR vs Booking Cancellation Status

The last graph is about the relationship between special requests and cancellation booking status. Nearly half bookings without any special requests have been canceled and another half of them have not been canceled.

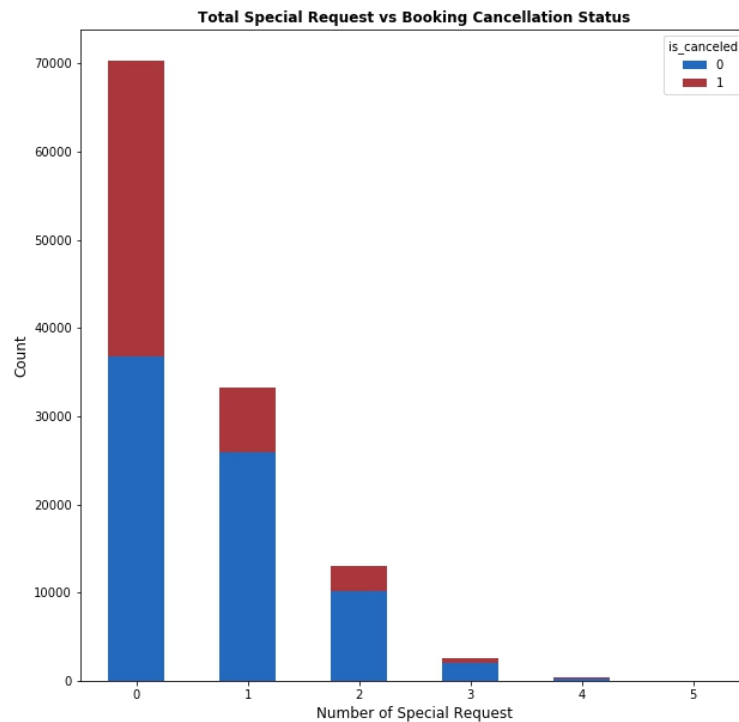


Figure. 9. Total special request vs Booking cancellation status

3.2. Dealing with missing data

On the first part, if there any missing data will be checked. Such that, the company feature's 94% is missing. Because of that, this feature will be eliminated. On the other hand, the children and all_children features have only 4 missing data. This missing data will replace it with zero. Another missing data has occurred in country and agent features. Since missing data of a country is less than 1%, these data will be replaced with the most frequent value. However, the agent's missing features are more than the country. For this feature, missing data will be imputed as 0.

Another part is analyzing categorical features. Categorical labels converted into numerical form. This will help to be more understandable and implementable into machine learning algorithms. Some features are not ordinal such as country. In that case, *One-Hot Encoding* could be chosen. Due to the high number of categories, this method could incur higher computational cost. To help reduce that, a Label Encoding method will be used.

After encoding the categorical data, two data frames will be created. One data frame has only categorical data and another has numerical data. These two different data frames will be used to create a correlation matrix. *Spearman* method will be used for categorical data correlation matrix and *Pearson* method will be used for numerical one.

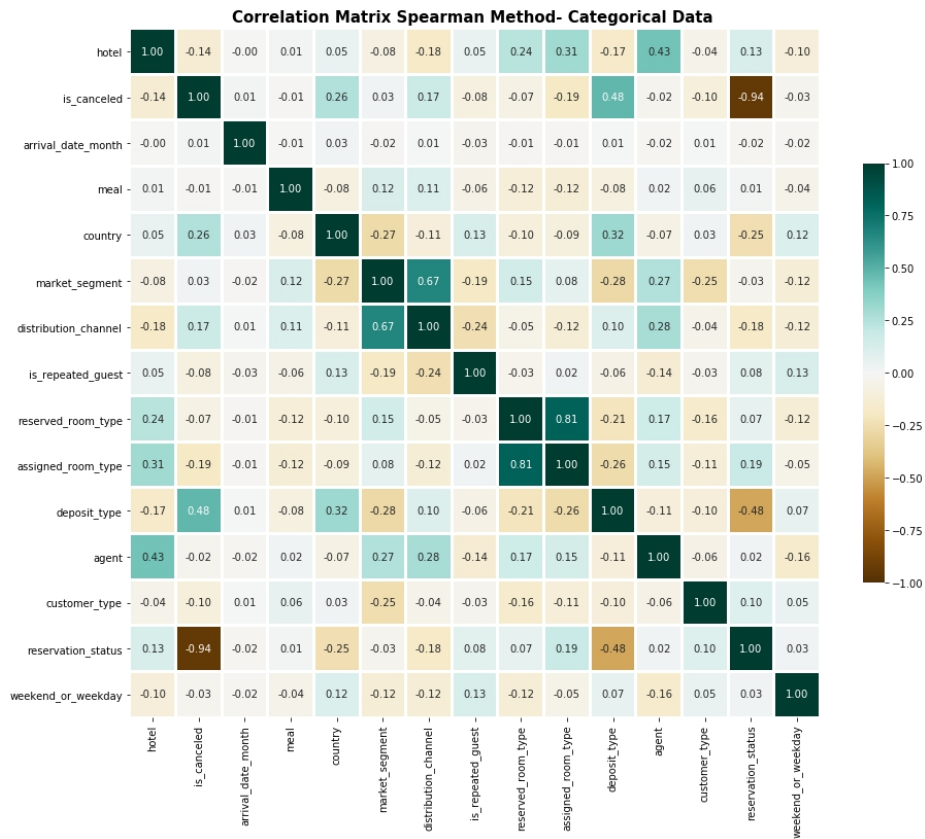


Figure. 10. Correlation matrix spearman method (Categorical data)

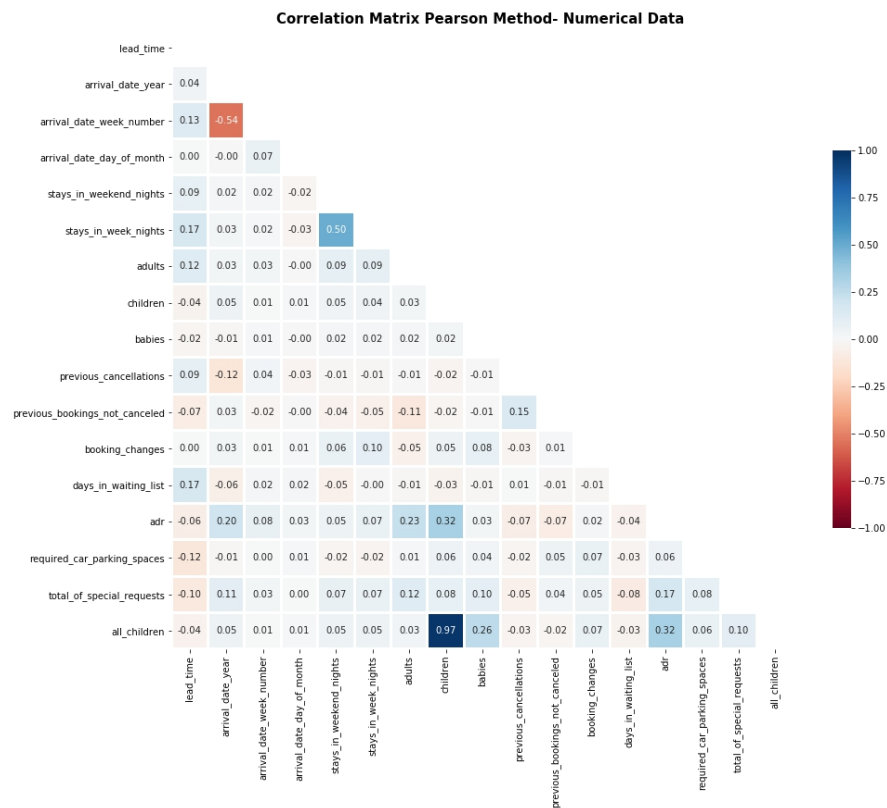


Figure. 11. Correlation matrix pearson method (Numerical data)

The above correlation matrix shows positive or negative relationships between them. In those two heatmaps, the reservation_status feature is drawn more attention because of its negative correlation with the is_canceled feature. The below table shows the relationship with details. This high correlation can cause a wrong prediction or overfitting. Prevent this situation, reservation_status feature will be eliminated. On the other hand, there is another high relationship between the children and the all_children features, since the all_children feature is constituted with the children and the babies features. Therefore the children feature will be eliminated too. Last feature (Table. 3) is reservation_status_date. Since this feature includes date type data and it could not convert another type, this feature will be eliminated.

Table. 3. Reservation status

reservation_status	0	1	2	Total
is_canceled				
0	0	75166	0	75166
1	43017	0	1207	44224
Total	43017	75166	1207	119390

3.3. Hyperparameter Tuning and Feature Importance

In this part, optimum hyperparameters for several tree-based machine learning algorithms will be searched with the help of the Grid Search Algorithm. Hyperparameter tuning will help to make a prediction in the

training part more accurately. Therefore, hyperparameters tuning will be fixed before the training process. Another important work is constituted Permutation Feature Importance graph with the Extreme Gradient Boosting algorithm. This technique calculates feature importance and performance metric to be chosen as the basis of the accuracy score. This graph will help to understand features' contributions to prediction, provide insight into the dataset, and will help to find deemed non-important features if any. The graph below (Fig. 12) shows the feature importance of the features. According to that, 1 out of 29 features are not important to prediction which is babies. It will be eliminated.

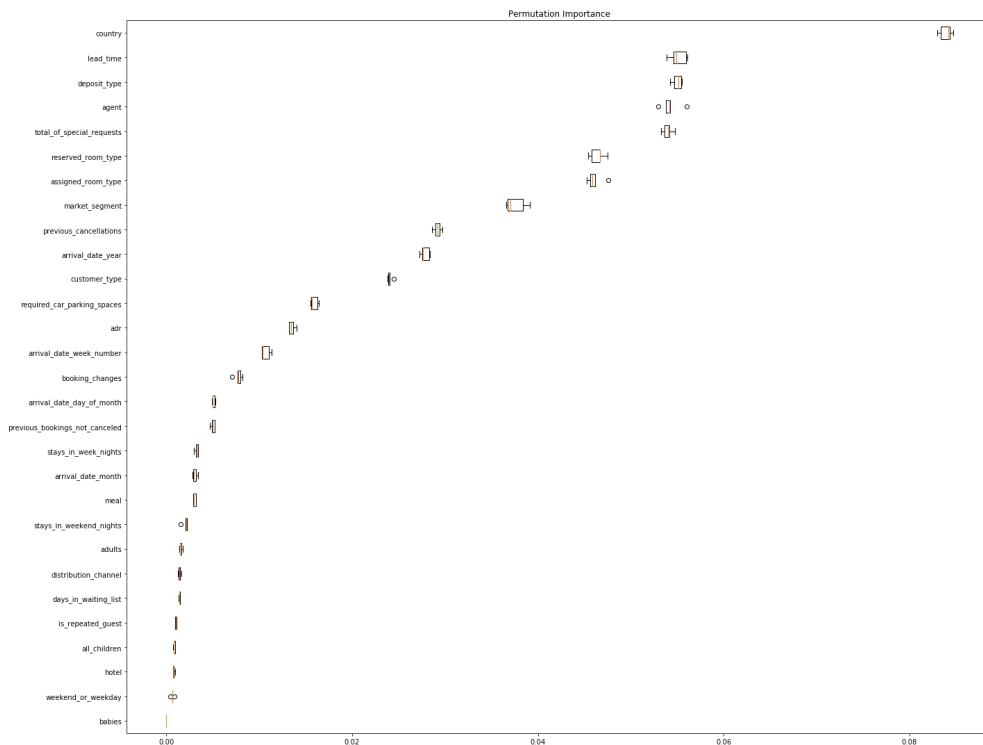


Figure. 12. Permutation importance

3.4. Model Building

In this part, some tree-based algorithms have been used for model building. These are *Decision Tree*, *Random Forest*, *Extra Trees Classifier*, and *Extreme Gradient Boosting*. *Random Forest* and *Extra Tree Classification* algorithms have been chosen as bagging algorithms, *XGBoost* has been chosen as one of the boosting algorithms. *The Decision Tree* algorithm has been chosen as one tree algorithm.

Before model building, data will be split to train and test respectively 70% and 30% ratio. X_{train} and X_{test} data will be standardized with the *Standard Scaler* technique. After that, the *Stratified K-Fold Cross Validation* method will be used for resampling. Cross-validation is an important implementation to avoid overfitting. *Stratified K-Fold Cross Validation* method provides train/test indices to split data into train/test sets. Model parameters have been defined in the previous part.

- Separate target variable for model building

```
y_model = hotel_data_model.iloc[:,1]
```

```
X_model = pd.concat([hotel_data_tunning.iloc[:,0],hotel_data_tunning.iloc[:,2:30]], axis=1)
y_model.describe()
```

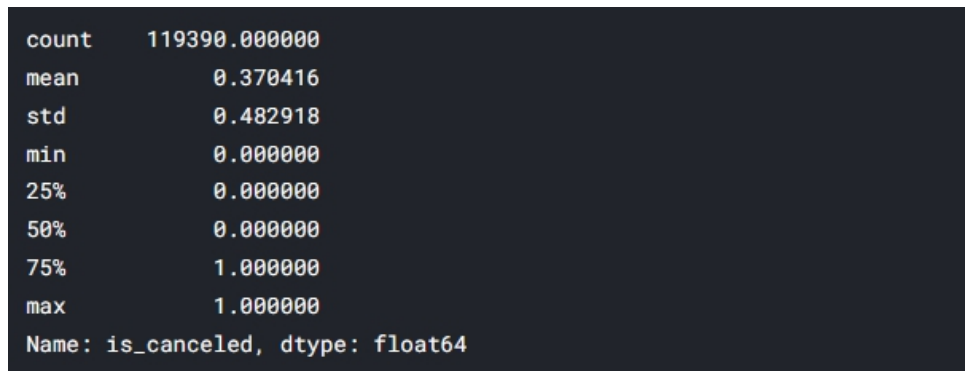


Figure. 13. Separating target variable result

- Split to train and test with 70-30 ratio

```
X_train, X_test, y_train, y_test = train_test_split(X_model, y_model, test_size = 0.3,
random_state=42, stratify = y)
```

- Implement standart scaler method

```
standartScalerX = StandartScaler()
```

```
X_train = standartScalerX.fit_transform(X_train)
```

```
X_test = standartScalerX.fit_transform(X_test)
```

- Stratified K-Fold Cross Validation Method

```
kfold_cv=StratifiedKFold(n_splits=5, random_state=42, shuffle=True)
```

```
for train_index, test_index in kfold_cv.split(X_model,y_model):
```

```
X_train, X_test = X_model.iloc[train_index], X_model.iloc[test_index]
```

```
y_train, y_test = y_model.iloc[train_index], y_model.iloc[test_index]
```

- Decision Tree Model Building

```
dtc_model = DecisionTreeClassifier(criterion= 'gini', min_samples_split=8, min_samples_leaf = 4,
max_features = 'auto')
```

```
#fit the model
```

```
dtc_model.fit(X_train, y_train)
```

```
#Predict Model
```

```
predict_dtc = dtc_model.predict(X_test)
```

- Random Forest Model Building

```
rf_model = RandomForestClassifier(min_samples_leaf = 6, min_samples_split=6, n_estimators = 100)
```

```
# fit the model
```

```
estimator= rf_model.fit(X_train, y_train)
```

```
#Predict Model
```

```
predict_rf = rf_model.predict(X_test)
```

- Extra Trees Classifier Model Building

```
etc_model = ExtraTreesClassifier(min_samples_leaf = 7, min_samples_split=2, n_estimators = 100)
```

```
# fit the model
```

```
etc_model.fit(X_train, y_train)
```

```
#Predict Model
```

```
predict_etc = etc_model.predict(X_test)
```

- Extreme Gradient Boosting Model Building

```
xgb_model = XGBClassifier(criterion = 'giny', learning_rate = 0.01, max_depth = 5, n_estimators = 100, objective ='binary:logistic', subsample = 1.0)
```

```
#fit the model
```

```
xgb_model.fit(X_train, y_train)
```

```
#Predict Model
```

```
predict_xgb = xgb_model.predict(X_test)
```

3.5. Classification Reports and Classification Matrix

The last part is comparison of classification reports of ML models. First comparison the accuracy results.

- Accuracy is a ratio of correct predictions to the total predictions. Its formula is:

$$(TP+TN)/(TP+FP+FN+TN)$$

According to that, *Random Forest* has the highest correct prediction with 88%. Another performance metrics explained below:

- Precision: It is the ratio of correctly predicted observation to the total positive predicted observation. Its formula is

$$TP/(TP+FP)$$

- Recall: It is the ratio of correctly predicted positive observations to the actual positive observations. Its formula is

$$TP/(TP+FN)$$

Table. 4. Random forest classification model

Random Forest	Precision	Recall	f1-score	support
0	0.88	0.94	0.91	15033
1	0.88	0.74	0.83	8845
accuracy			0.88	23878
macro avg	0.88	0.86	0.87	23878
weighted avg	0.88	0.88	0.88	23878

Table. 5. Decision Tree classification model

Decision Tree	Precision	Recall	f1-score	support
0	0.87	0.89	0.88	15033
1	0.81	0.77	0.79	8845
accuracy			0.85	23878
macro avg	0.84	0.83	0.83	23878
weighted avg	0.85	0.85	0.85	23878

Table. 6. Extra tree classification model

Extra Tree	Precision	Recall	f1-score	support
0	0.85	0.95	0.90	15033
1	0.90	0.72	0.80	8845
accuracy			0.87	23878
macro avg	0.88	0.84	0.85	23878
weighted avg	0.87	0.87	0.86	23878

Table. 7. XGBoost classification model

XGBoost	Precision	Recall	f1-score	support
0	0.80	0.93	0.86	15033
1	0.83	0.60	0.70	8845
accuracy			0.81	23878
macro avg	0.82	0.77	0.78	23878
weighted avg	0.81	0.81	0.80	23878

4. Results and Discussion

As we can see from the summary table, *Random Forest* and the *Extra Tree Classifier* share the highest precision ratios. It means that both models predicted around 88% of all the positive labels correctly. On the other hand *Random Forest* has the highest recall ratio. It means that this model predicted 79% of actual positive observations correctly.

These studies confirm that bookings with a high risk of being cancelled can be detected. This encourages hotel management to take steps to deter possible cancellations, such as improved facilities, discounts or other incentives, from being provided. However since others are oblivious to this sort of bid, this should not be extended to all consumers. However this prediction model has much to be learned from building and deploying it. More relevant results, such as the amount of room nights expected to also be cancelled in any of the following weeks, can be achieved by running the model against all bookings every day. Hotel operators, estimating their net demand, will deduce this value from their demand. Hotel managers may build stronger overbooking and cancellation strategies if presented with a more specific value of net demand, resulting in lower costs and reduced risk.

As with every other issue in predictive analytics, designing a model for forecasting booking cancellations requires data to follow all of the quality data characteristics: precise, consistent, neutral, relevant, acceptable and timely[15-16]. Any of the data sets had variables with incorrect values, as previously stated. Model output

may be influenced by this lack of consistency. For this cause, hotels who wish to build prediction models need to ensure that a policy on data quality is in effect.

5. Conclusion

This study aims to explore some of the functions of predictive analytics in scientific analysis from a scientific point of view, including defining which features lead to predicting the likelihood of book cancellation. The implementation of data visualization and data analytics techniques, along with the use of the shared knowledge filter, made it possible to recognize the predictive significance of a feature. Different features were found to vary in value depending on the hotel, and certain features for some of the hotels are not needed.

We need to realize that one model may be extended to all hotels by developing a model to identify bookings that are likely to be cancelled and by building a stronger net demand forecast. The creation of the model found that features according to the hotel had different weights and different significance, meaning that every model did not match all hotels and thus each hotel will have its own version. These forecasting models allow business owners to reduce the loss of revenue from booking cancellations and to minimize the risks associated with overloading (relocation costs, cash or service compensations, and, particularly important today, social reputation costs). Booking cancellation models often allow hotel managers, without increasing confusion, to enforce less strict reservation policies.

References

- [1] Kimes, SE and Wirtz, J. 2003. Has revenue management become acceptable? Findings from an International study on the perceived fairness of rate fences. *Journal of Service Research*, 6(2): 125–135. DOI: <https://doi.org/10.1177/1094670503257038>
- [2] Chiang, W-C, Chen, JC and Xu, X. 2007. An overview of research on revenue management: current issues and future research. *International Journal of Revenue Management*, 1(1): 97–128. DOI: <https://doi.org/10.1504/IJRM.2007.011196>
- [3] Mehrotra, R and Ruttley, J. 2006. *Revenue management* (second ed.). Washington, DC, USA: American Hotel & Lodging Association (AHLA).
- [4] Talluri, KT and Van Ryzin, G. 2005. *The theory and practice of revenue management*. New York, NY: Springer. DOI: <https://doi.org/10.1007/b139000>
- [5] Ivanov, S. (2014). *Hotel revenue management: From theory to practice*. Varna, Bulgaria: Zangador.
- [6] Guo, X., Dong, Y., & Ling, L. (2016). Customer perspective on overbooking: The failure of customers to enjoy their reserved services, accidental or intended? *Journal of Air Transport Management*, 53, 65–72. <https://doi.org/10.1016/j.jairtraman.2016.01.001>.
- [7] Hayes, DK and Miller, AA. 2011. *Revenue management for the hospitality industry*. Hoboken, NJ, USA: John Wiley & Sons, Inc.
- [8] Antonio, N, de Almeida, A and Nunes, L. 2017a. Predicting hotel booking cancellation to decrease uncertainty and increase revenue. *Tourism & Management Studies*, 13(2): 25–39. DOI: <https://doi.org/10.18089/tms.2017.13203>
- [9] Huang, H-C, Chang, AY and Ho, C-C. 2013. Using artificial neural networks to establish a customer-cancellation prediction model. *Przeład Elektrotechniczny*, 89(1b): 178–180
- [10] Liu, PH. 2004 Hotel demand/cancellation analysis and estimation of unconstrained demand using statistical methods. In: Yeoman, I and McMahon-Beattie, U (eds.), *Revenue Management and Pricing: Case Studies and Applications*. Cengage Learning EMEA. pp. 91–108.
- [11] Morales, DR and Wang, J. 2010. Forecasting cancellation rates for services booking revenue management using data mining. *European Journal of Operational Research*, 202(2): 554–562. DOI: <https://doi.org/10.1016/j.ejor.2009.06.006>

- [12] Antonio, N, de Almeida, A and Nunes, L. 2017b. Using data science to predict hotel booking cancellations. In: Vasant, P and M, K (eds.), *Handbook of Research on Holistic Optimization Techniques in the Hospitality, Tourism, and Travel Industry*. Hershey, PA, USA: Business Science Reference. pp. 141–167. DOI: <https://doi.org/10.4018/978-1-5225-1054-3.ch006>
- [13] Talluri, K. T., & Van Ryzin, G. (2004). *The theory and practice of revenue management*. Boston, MA, USA: Kluwer Academic Publishers.
- [14] Ivanov, S., & Zhechev, V. (2012). Hotel revenue management—A critical literature review. *Turizam: Znanstveno-Strucnicasopis*, 60(2), 175–197
- [15] Rabianski, J. S. (2003). Primary and secondary data: Concepts, concerns, errors, and issues. *Appraisal Journal*, 71(1), 43 (13).
- [16] M. F. Kokasih and A. S. Paramita, “Property Rental Price Prediction Using the Extreme Gradient Boosting Algorithm,” *IJIS Int. J. Informatics Inf. Syst.*, vol. 3, no. 2, pp. 54–59, 2020, doi: 10.47738/ijiis.v3i2.65.