

# XAI in hotel booking cancellation

---

## XAI in hotel booking cancellation

[Abstract](#)

[Introduction](#)

[Dataset description](#)

[Data process](#)

[Make model confess](#)

[XAI\(eXplicability Artificial Intelligence\)](#)

[Discussion](#)

# Abstract

---

Machine learning models are popular tools in business and research over the past years. Machine learning works well for prediction tasks, but many machine learning models are poorly interpreted. A range of interpretable frameworks such as Lime make machine learning predictions more convincing. This report takes the hotel booking cancellation problem as an instance and uses a random forest model to make predictions on whether a user will cancel a hotel booking. The random forest model has a satisfying accuracy of 89% and an F1 score of 84%. Interpreting the importance of features through the lime framework, we found that whether or not a deposit was paid, and the user's country were the most important factors in deciding whether to cancel a hotel reservation.

# Introduction

---

Machine learning models are popular tools in business and research over the past years. Models assist our customers to make decisions on which hotel to book or set the optimal price for flight tickets based on real-time demand. Recommendation softwares advise us where to eat or which lookout is worth watching. But those decisions are difficult for automated systems to take by which is suit for human. How to make sure the model does not encompass unwanted bias or discriminate clients, and how to explain to CEO that the price should be less than what the concurrence proposes. It is interesting and practical problem for everyone.

In rule-based systems, it is possible to track the decision tree and explain that decisions come from a combination of known facts. ML systems use a different strategy which is not trivial to define how the system computed the outcome. There are some methods to make model interpretable, such as Lime. Those methods could identify the importance of the input variables and make model prediction interpretable.

## Dataset description

---

This data article describes two datasets with hotel demand data. One of the hotels (H1) is a resort hotel and the other is a city hotel (H2). Both datasets share the same structure, with 31 variables describing the 40,060 observations of H1 and 79,330 observations of H2. Each observation represents a hotel booking. Both datasets comprehend bookings due to arrive between the 1st of July of 2015 and the 31st of August 2017, including bookings that effectively arrived and bookings that were canceled. Since this is hotel real data, all data elements pertaining hotel or costumer identification were deleted.

## Data process

---

We only keep the data that entries without missing value that 118,902 entries left. There are 24 columns which having a significant correlation with the target variable "is\_canceled" were kept. The data processing is from [jthomasmock](#)'s Github web page.

The categorical variables: "meal", "market\_segment", "distribution\_channel", "reserved\_room\_type", "assigned\_room\_type", and "customer\_type" were one-hot encoded for the ML model.

The string variables or categorical variables, such as "country" or "hotel", were coded into numbers, and new variables were created, such as "is\_family", which expresses if adults were traveling with children. Most of the feature names are self-explanatory.

Dataset were split into a train (80%) and test (20%). The target variable "is\_canceled" was excluded from the data and saved as a y variable. In total, there were 79,306 canceled and 39,596 non-canceled entries. Class 0 was defined as "not canceled" and Class 1 as "canceled".

## Make model confess

We use cross-validation method to train and test the random forest model. The random forest model has a satisfying accuracy of 89% and an F1 score of 84%. The details of model performance could be seen below [model performance](#) . Moreover, we can find that 904 instances were wrongly classified as class 1, and 1804 instances were wrongly classified as class 0 from [confusion matrix](#).

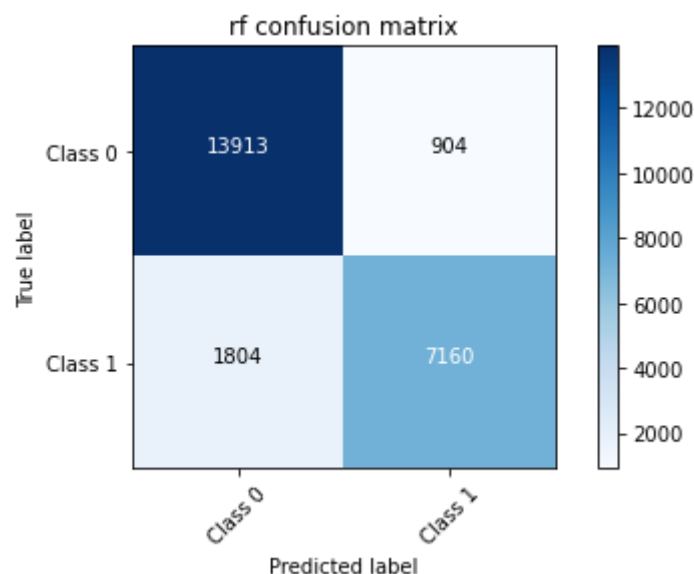
```
Confusion matrix, without normalization
              precision    recall  f1-score   support

   Class 0       0.89       0.94       0.91     14817
   Class 1       0.89       0.80       0.84      8964

 accuracy              0.89     23781
 macro avg           0.89     0.87     0.88     23781
weighted avg           0.89     0.89     0.88     23781

Cohen Kappa Score: 0.7526651191238782
ROC AUC Score: 0.9527114592465081
Default Recall Score: 0.7987505577867023
f1: 0.8409678177120037
```

Model Performance



Confusion Matrix

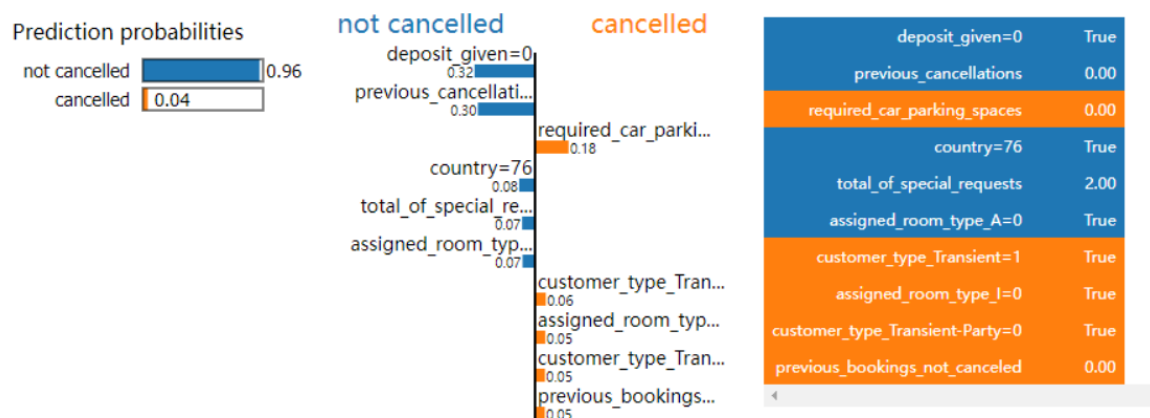
## XAI(eXplicability Artificial Intelligence)

LIME can then decode the categorical coded values and show human-readable labels.

Since we are dealing with tabular data, an explainer is built using "LimeTabularExplainer". The minimal inputs are the original data formatted as a numerical matrix with the class names, the feature column names, the indexes of categorical features, and the names of categorical features. The parameter `kernel_width` defines the size of the local prediction. The bigger the `kernel_width`, the more data points the function will take to generalize. The explanation can be computed along with a visualization when calling the `explain_instance` method. The required inputs are data instances and the prediction function.

The most important feature for the instance 1 is “deposit\_given” set to 0. The country is only the fourth most important feature. And no “previous cancellations” was the second most important feature [Lime explanation1](#). No demand for car parking space contributed to the opposite prediction. Even though the contribution weight has changed.

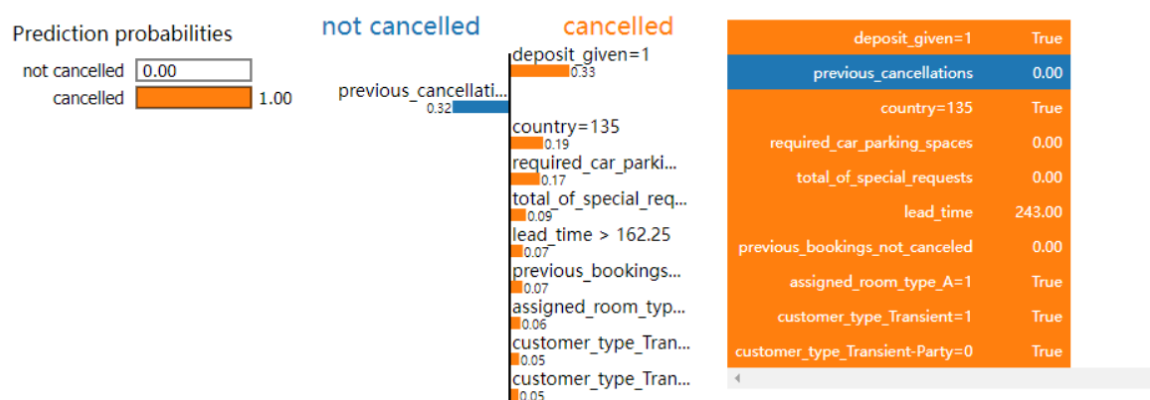
```
Intercept 0.5369475390248153
Prediction_local [0.09247682]
Right: 0.04
```



## Lime explanation 1

For the selected instance from class 1 (canceled), “deposit given” is the most important feature according to the explanation [Lime explanation2](#). Country is placed as the third most crucial feature followed by no required car parking space and “lead time”. Moreover, no “previous cancellations” would push the prediction into “non canceled”.

```
Intercept 0.06212375985322366
Prediction_local [0.82714279]
Right: 1.0
```



## Lime explanation 2

# Discussion

Those powerful XAI tools indicate with consistency which feature values have a substantial impact on the model output. LIME is more general, the explanations are more likely to be biased by the linear surrogate explainer. Explanations indicate the features that are fundamental to the model without placing too much importance on the specific weights and order.