**Analytic Plan**

Cameron Gallien

DSE6211OM\_FA2023F1\_Machine Learning & Artificial Intelligence

Merrimack University

Prof. Peter Salemi

September 6, 2023

**Introduction**

The ABC Hotel company is looking for a way to identify bookings that carry a high risk of cancellation. This will allow the hotel to prepare for cancellations when bookings that meet the criteria are identified, and target these customers with additional advertisements and offers in an attempt to prevent the rooms from being cancelled. This paper presents a supervised classification process to provide this business need, leveraging a data set containing over 35,000 past bookings for which it is know whether or not the customer’s booking was cancelled.

The target variable that will be utilized for this supervised classification problem is the booking status of the reservation. Each reservation has been marked as “cancelled” or “not cancelled,” which will be used to train the model as to which reservations are more likely to be cancelled. By utilizing this information we will be able to input the data into our algorithm and train it to see which factors contribute the most to future cancellations so that the hotel company can be fully prepared for any situations they might find themselves in.

**Data Processing**

The first step to processing the data will be selecting which variables should be included in our analysis, and which variables should be excluded. At first glance it appears that the only variable worth excluding from our analysis is the booking ID, which is an arbitrarily assigned number to identify each booking. Since this doesn’t have any specific value for the bookings, this will need to be removed.

In addition, there are variables that will need to be converted to factors so that the model can analyze them properly. These variables are type of meal plan, room type reserved, market segment type, and our target variable, booking status.

In addition, the date variable will need to be modified as well. Since there are hundreds of different dates ranging from 7/1/2017 to 12/31/2018, I will create an additional variable for the month that the reservation was made. This will allow the model to pick up on potential seasonal trends for room cancellations, without introducing a high level of dimensionality that could cause overfitting in future use.

**Analytic and Informational Outcomes**

The proposed approach is expected to produce several analytical and informational outcomes. The first potential outcome is the central purpose behind the model: predicting the likelihood that a future reservation will be cancelled. As mentioned in the introduction, this would allow the company to target reservations with a high cancellation probability with targeted ads or exclusive offers to prevent those customers from cancelling. In addition, this will allow the company to prepare themselves for specific scenarios that would cause high cancellation rates by lowering fees, reducing staffing levels, or overbooking the hotel when appropriate.

Another potential outcome of the model would be finding specific variables that lead to increased cancellations. For example, if our model finds that higher room prices correlate to higher cancellation probabilities, then these insights could be used to adjust the future price of hotel rooms.

**Conclusion**

Once implemented, this proposed classification approach will address the stated business needs of ABC Hotels in regards to predicting future cancellations among their hotels. Leveraging a dataset of over 35,000 past bookings, our objective is to train a predictive model to determine which factors contribute most significantly to future cancellations. This predictive model will enable ABC Hotels to assess the likelihood of a reservation being canceled, assigning a probability value between 0 and 1 to each booking. Variables have been carefully selected for use in our model, and additional data processing steps will be taken to ensure that all variables are suited for training as accurate of a model as possible.

In practice, the proposed predictive model will become an invaluable tool for ABC Hotels. It will enable them to make data-driven decisions, optimize their booking management strategies, and enhance customer satisfaction by minimizing cancellations. This will allow ABC Hotels to improve revenue and customer retention while ensuring a seamless and enjoyable booking experience for their guests.