Pick-up method + machine learning: a proved efficient approach to forecast hotel demand

1. Introduction

**hotel revenue management is important - demand forecasting is essential**

**why rm is important?**

**why demand is essential?**

* **improving accuracy can increase xx% of revenue**
* **uncertain, can drive pricing xxx**

Hotel revenue consists of price and demand, and demand forecast is essential for successful revenue management since it has high uncertainties and has direct impact on issues such as inventory management, pricing strategies, and marketing plans. Different from industries such as retail where most of the bookings happen instantly when the clients arrive, the hotel industry usually sell the room ahead of the customer arriving, and thus generate valuable advance bookings information.

**introduce advance booking (pick up method)**

Hence, advance booking method, an approach conducting forecast based on current realized bookings and historical booking patterns, is widely used in both academic and the industry. On top of the existed bookings so far (ROH, reservations on hand), the advance booking predicts what will happen from today and the target future by estimating the incremental bookings. In practice, this method takes an average of the incremental bookings in history, or average the incremental ratios in percentage, then add on or multiple to the ROH today. This method is also called “pick-up” method since it estimates the number of incremental bookings “picked up” from today’s reservation. ***Researchers (now add more research about pick up)*** Besides advance booking method, practitioners and researcher have used other methods such as time series, combined methods to forecast hotel demand. ***(now add more research about time series and other)***

**machine learning, on the other hand…**

* **widely used in areas…**
* **why? because…**

In recent years, machine learning method has been picking up attentions among both the industry and academia. Machine learning is a statistical method (**add more definition about machine learning**) and it has the benefit of (flexible, catching patterns, etc. )

**machine learning + hotel**

* **very few research in the hotel industry due to…**

**this research serves as a bridge!**

1. Machine Learning (theoretical)
   1. brief intro of ml
   2. 6 sections to introduce each algorithm
2. Empirical Study
   1. data description

This research uses a one-year long booking records, with the arrival date from December 27, 2017 to December 31, 2018, of a hotel property. For each booking at this hotel, this dataset records the booking date (the day the client makes this reservation) and the arrival date (the day the client checks in). Hence, for each arrival date, the ROH can be calculated accumulatively from the earliest booking date of each day. For instance, if the hotel is predicting the number of final rooms sold on February 2021 while “today” is January 1, 2021, the realized booking (ROH) can be calculated by adding up all of the reservations for February 14, 2021 happened before January 1. In this way, a new variable, the ROHs is calculated as the main independent variable.

Another derivative variable, days before arrival (DBA), is used to describe the time window between the booking day and the arrival day. This study examines 12 different horizons: 1, 2, 3, 4, 5, 6, 7, 14, 21, 30, 60, 90, and beyond. The cutoffs are made in this way since in the hotel industry, the dynamic pricing plan is usually set following those horizons. It is noticeable that during early periods when the booking day is far away from the stay date, the reservations accumulate very slow. Therefore, a wider horizon in earlier periods allows information to accumulate, while when the stay dates are approaching closer, the booking window is broken into smaller horizons for a closer attention.

In short, this study uses ROHs as the single independent variable, on different arrival date and DBA, to make predictions. The dependent variable here is the final accumulated reservations on the arrival day, in other words, the accumulative ROHs when DBA=0.

There are

* 1. methods/models
  2. results

1. Conclusion & Discussion