RMS ALGORITHMS DURING THE PANDEMIC:

CHALLENGE OR OPPORTUNITY ?

Authors: Anna Raykhlina, Ludwik Bielczyński

Editor: Yelyzaveta Horbenko

Graphic editor: Adelina Dorokhova

Gratitude for the help: Piotr Olesiński

INTRODUCTION

With the start of the COVID-19 pandemic, the hospitality business was one of the sectors which were hit the hardest. Unknown problems, complex and changing challenges and an endless struggle for survival present millions of hoteliers with extreme difficulties today. Most of us have never imagined that 2020 would bring one of the worst public-health crises in modern history and turn our life upside down.

We truly believe that awareness of the situation and proper preparation can bring a perfect solution, help to save many hotel properties from closure and breathe life into them again when the vaccine is generally available, restrictions are lifted and travellers are back.

To get deeper into this, **Anna Raykhlina (AR),** our Key account manager at YieldPlanet, interviewed **Ludwik Bielczyński (LB)** - our Data Scientist. The person who oversees our R&D (Research and development), who is responsible for the algorithmic side of YieldPlanet's Price Optimizer, and who brings together statistics and machine learning to our customers.

In this practical e-book Ludwik lets us plunge into his day-to-day operations to understand better **how the algorithms work in Price Optimizer**, what **challenges COVID-19 has brought** and **how we can use this knowledge to react upon the crisis**.

Anna Raykhlina,
Key Account Manager,
YieldPlanet

Ludwik Bielczynski,
Data Scientist, YieldPlanet

AR: Going straight to the point, could you please draw a general picture of Price Optimizer? Which algorithms are used there?

LB: In **Price Optimizer**, the heart of the system consists of the **Price Recommendation algorithm**. It enables the revenue manager to implement **short- and long-term dynamic pricing strategies** for specific season and room-type group combination.
In consequence, the system acts **automatically almost in real time**: the prices and the products availabilities are being adjusted in selected distribution channels (**over 400 channels,** including Online Travel Agencies, wholesalers, and Global Distribution Systems) based on the actual internal situation of the hotel (from the Property Management System of the hotel) and its performance on the market (obtained from our partners OTA Insights and HQ+).

Several factors are used for the adjustment:
- the actual occupancy,
- the forecasted final occupancy, and cancellation probability,
- the divergence of the actual from the expected occupancy pick-ups,
- the competition prices for a similar product.

Besides the actual occupancy, revenue, and the competition prices, all other input data are prepared by different forecasting and predictive algorithms: Occupancy Revenue Forecasting Algorithm - ORFA, Cancellation Probability Forecasting from Actual Reservations - CPFAR, and Occupancy Pick-up Predictive Algorithm - OPPA.

AR: Is there a difference between predicting and forecasting?

LB: Yes, the predictions are estimates of data which did not occur, indiscriminately future, past, and present. So, the prediction is a broader term that incorporates forecasts. Forecasting is based solely on time series (a set of values ordered by time): **based on the past observations we estimate the future outcome.**

On the example of our system, we can forecast how many guests will arrive on a selected future stay-date, or how many of the reservations on the books will be cancelled. In terms of predictions, we can estimate if our actual occupancy pick-ups are deviating from what should be expected on that report-date.

AR: So, talking about the forecasting... Why is it so important?

LB: As said before, forecasting is the process of predicting a future outcome. In the hospitality industry, it is usually directly linked with the demand estimate based on the historical sales. Demand forecasting plays a key role in revenue decisions. It helps to manage pricing strategies, distribution (including overbookings) and restrictions (min/max length of stay, closed to arrival/departure).

However, forecasting should not be only limited to the demand forecasting as **it is a powerful tool that helps with assessing opportunities and risks** during any decision making and planning.

For example, Facebook created a simple open source forecasting tool (FBProphet) to base possibly all their decisions on hard evidence [1] .

AR: I know that there are different models that can be used for forecasting. Could you please tell us more about it?

LB: A tool is meant to carry a specific task. The same way, the model for forecasting depends on the goal that should be reached. The forecasting is performed to understand better the market situation of the hotel, to assess risks of decisions that need to be made, and to estimate final outcomes of those decisions. The line is blurred between which model should be used to fulfil each of those tasks. Most commonly, because of their simplicity and intuitiveness **naive models** (no offence, it is their name) are broadly used, such as **the comparison of the actual occupancy to the last year's occupancy** at a corresponding DTA and stay-date combination.

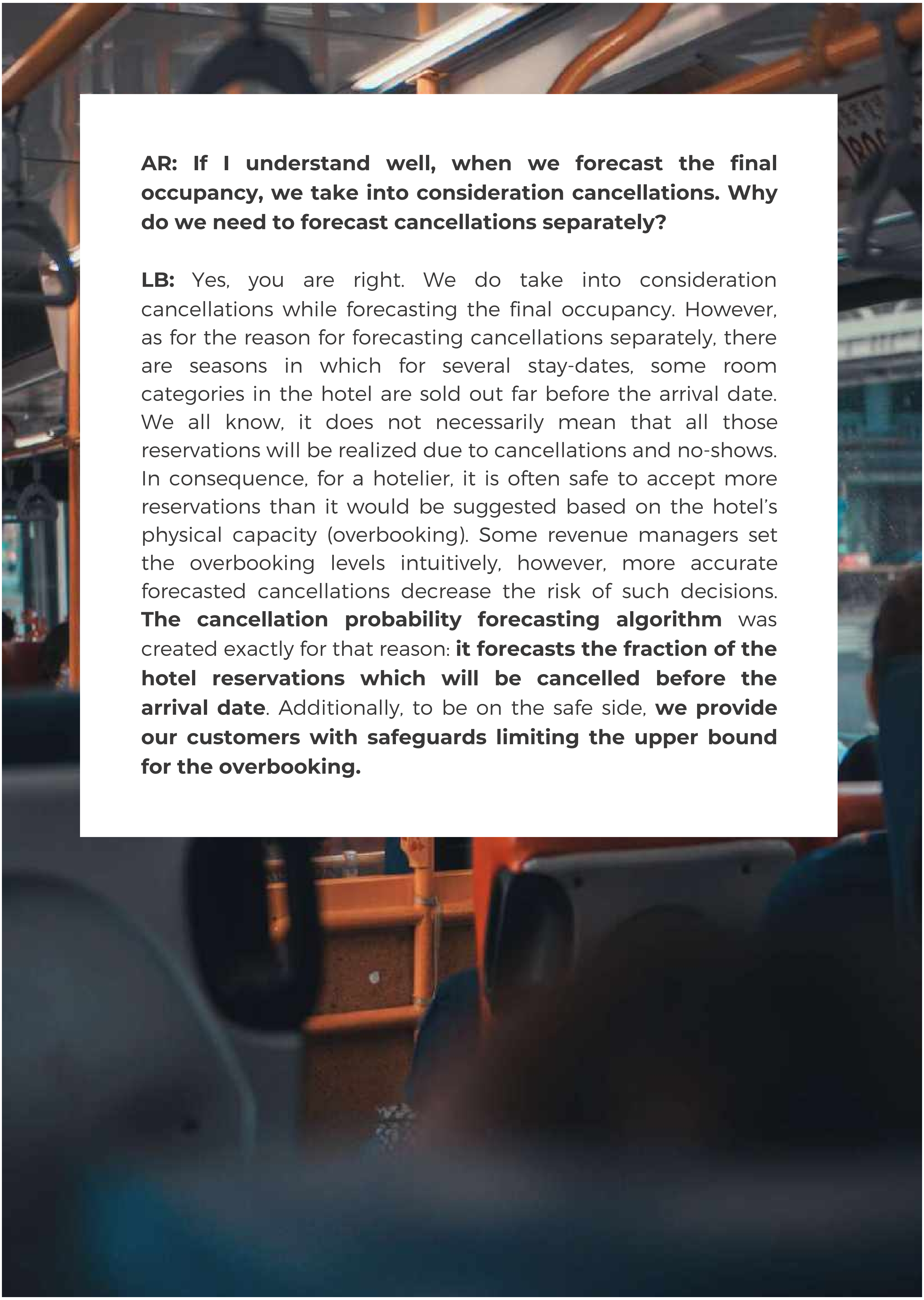
However, obviously such methods have the lowest accuracy, they do not take into consideration any newer scenarios or combination of factors.

In consequence, they are often used as baseline models to benchmark more complex ones.

Another popular group of methods are **heuristics,** which are **strategies derived from previous experiences.** It is well known that they have not been proven to be optimal but due to their computational simplicity and intuitiveness, their results are sufficiently accurate to be used. When high descriptive capacity of relationships between variables, and hypothesis testing are needed, **statistical models** are most commonly employed (e.g. time-series analysis). Those methods are well established and thoroughly tested. However, recently, due to high forecasting accuracy of the **machine learning models**, they became ubiquitous (e.g. gradient boosted trees, and artificial neural networks). One of the few conditions for their use is a low need for intuitiveness and readability as their complexity is often exceeding our understanding.

AR: Which of those models do we use in Price Optimizer? Maybe let us start with the most commonly used: the forecasting occupancy model?

LB: In Price Optimizer, to provide the highest forecasting accuracy **we use a combination of both - statistical and machine learning models.** At first, the statistical model (additive model) is used to notice general trends, weekly and yearly seasonality. It enables better understanding of the hotel's situation and identifies strategies to be implemented in the pricing algorithm.
Further, we employ an ensemble learning method called **model stacking**, where forecasts from a statistical model are used as one of many features in a far more complex metamodel. As a result of our collaboration with the Warsaw University of Technology, the **Extreme Gradient Boosting model** (one of the machine learning ensemble models) was chosen as the metamodel. This kind of model is extremely useful in finding difficult patterns with complicated non-linear relationships between features. In data science competitions (e.g. Kaggle), it commonly outperforms other models in the category of structured problems. In comparison, neural networks are known for their great results for unstructured problems, like object identification and classification on images (e.g. Chihuahua or Muffin classification[2], or in more ambitious projects: for computer-aided detection of different samples of cancerous tissues[3,4]).



AR: If I understand well, when we forecast the final occupancy, we take into consideration cancellations. Why do we need to forecast cancellations separately?

LB: Yes, you are right. We do take into consideration cancellations while forecasting the final occupancy. However, as for the reason for forecasting cancellations separately, there are seasons in which for several stay-dates, some room categories in the hotel are sold out far before the arrival date. We all know, it does not necessarily mean that all those reservations will be realized due to cancellations and no-shows. In consequence, for a hotelier, it is often safe to accept more reservations than it would be suggested based on the hotel's physical capacity (overbooking). Some revenue managers set the overbooking levels intuitively, however, more accurate forecasted cancellations decrease the risk of such decisions. **The cancellation probability forecasting algorithm** was created exactly for that reason: **it forecasts the fraction of the hotel reservations which will be cancelled before the arrival date**. Additionally, to be on the safe side, **we provide our customers with safeguards limiting the upper bound for the overbooking.**

AR: You said before that for price recommendations Price Optimizer uses pick-up predictions as well? Sounds a bit cryptic. What exactly is being predicted?

LB: We use **pick-up prediction** to identify unusually high or low occupancy pick-ups for a selected stay-date. The prediction is needed as the same pick-up can be low during a high season and high during a low season. We create a predictive model to assess if our actual pick-up for a selected room-type is exceeding upper or lower prediction bounds. **This prediction helps to apply an additional adjustment to the price recommendation.**

AR: Considering the current situation... Which algorithms can suffer the most because of the COVID-19 pandemics?

LB: With the start of the COVID-19 pandemics, the hospitality business was one of the sectors which were hit the hardest [5]. It means that the customers behaviour patterns changed abruptly. There are several recovery scenarios possible, however, it becomes apparent that the optimistic fast containment of the virus and fast recovery is not realistic, and we have to brace ourselves for several infection waves and a slow recovery. In consequence, each business basing their decisions on predictive algorithms is faced with a challenging period [6,7].

All algorithms predicting values which were directly influenced by the pandemic, no matter if statistical, machine learning or AI, if trained on the data from before the start of the pandemic, will have problems with the new situation. There is a misconception that machine learning models respond well to unknown situations. In fact, most of them are trained to predict well in stable and repetitive scenarios.

Let us use the example of the behaviour of the occupancy forecasting algorithm from the start of the COVID-19 pandemic. As said before, the main machine learning model used in this algorithm was designed to capture patterns - general and very complicated patterns which we would easily miss. However, it was trained on historical data in a relatively stable economic period, and its predictive ability is constrained to those scenarios. The lock-down was an unprecedented situation for many countries. In contrast to us humans, that were subjected to information about the infection rate or the changes in the government policies, the algorithm was never subjected to them (in the form of the model independent variables). It means that it could have never learnt that during the lock-down there would be no people arriving at any of the hotels. In consequence, during that phase, the algorithm was overly optimistic and systematically overestimated the number of final occupancies. After a few days, the algorithm learnt that even when there were reservations incoming the hotel can end up completely empty. It resulted in a high instability in the next phase. Sometimes, it was forecasting a normal scenario, in other cases, correctly an empty hotel. Softening the lock-down and the reopening of the hotels has lead to our current situation. **The algorithm knows how the demand in a stable situation, during and after a lock-down looks like, however, without additional features it cannot discriminate between those scenarios. Luckily, there are several options on how to improve such unstable models.**

AR: Improving the forecasting algorithm is surely not easy. Despite the difficult situation we are in now, the fact that there are still several ways to make improvements is extremely encouraging. So, if this is not a secret, could you please share with us what changes can be done to improve the predictive algorithms in PO?

LB: Sure. As we see, each disaster or crisis leads to a higher uncertainty of predictions. We are the first line of response, as usually algorithms cannot incorporate all the factors.
In consequence, **human oversight over the algorithms should be increased till a new stable situation arises**[6]. It means that all the predictions should be treated with a grain of salt.
As for the data used by the predictive algorithms, one could say, that the censorship of the data by deleting the lock-down period is advisable. That could be a good idea if there would be only one infection wave, and a fast recovery to the initial period. For example, such a situation was observed for mainland China after the SARS outbreak in 2003[8]. However, as we have already seen, there are secondary, and in some countries, even tertiary infection waves.
Another type of censorship would be to decrease the size of the training history only to the data from the beginning of the outbreak. However, such a solution limits the response range of the algorithm only to a very small dataset.
The most complex, however, **the most promising strategy would be to extend the features based on which the model is being trained**, e.g. with the number of active infection cases. If the government decisions are based on those numbers, it will lead our algorithms to learn new patterns and react accordingly to the situation.

AR: This really sounds promising! Could you please raise a bit of a curtain and tell us in a few words what are the recent improvements, and what is planned in terms of research and development for Price Optimizer?

LB: First, we focused on increasing the previously mentioned human oversight. We added a functionality to the Price Optimizer, so **the hotelier can set alerts when an occupancy pickup exceeds a manually selected value.** Simple feature, although, in those uncertain times, necessary to monitor the development of demand and the booking behaviour of the clients.

Additionally, some progressive improvements were brought to our central reservation system. Their aim is to help the reservation department **to work more efficiently and be in sync with the revenue department.**

Moreover, besides increasing human oversight and stabilizing the algorithms, we are constantly extending the possibilities of our system. Recently, we work on providing additional market analysis by estimating unconstrained demand, analysis of the competition prices and sold-outs, and price-demand relationship.

Concluding, **our main goal is always to provide hoteliers with the necessary tools and automate daily routine operations**, so that hoteliers have more time to do what they are best at, assess the risks and explore new business opportunities. Our road map in research and development is always reflecting that goal.

SOURCES:

1. https://facebook.github.io/prophet/
2. https://www.topbots.com/chihuahua-muffin-searching-best-computer-vision-api/
3. https://www.nature.com/articles/s41598-019-55972-4
4. https://www.sciencedirect.com/science/article/pii/S1877050915023613
5. https://www.mckinsey.com/industries/travel-logistics-and-transport-infrastructure/our-insights/hospitality-and-covid-19-how-long-until-no-vacancy-for-us-hotels
6. https://www.technologyreview.com/2020/05/11/1001563/covid-pandemic-broken-ai-machine-learning-amazon-retail-fraud-humans-in-the-loop/
7. https://www.weforum.org/agenda/2020/05/here-s-how-to-check-in-on-your-ai-system-as-covid-19-plays-havoc/
8. https://str.com/data-insights-blog/looking-back-sars-outbreak-and-hotel-performance-recovery

At YieldPlanet we believe that by providing our clients with the right advice we will win this battle together. We hope that plunging into Ludwik's day-to-day operations has helped you tohas helped you to understand better how the YieldPlanet's Price Optimizer algorithms work and how to face all the challenges that COVID-19 has brought. We strongly believe that you can start getting ready for the market bounce back already today and utpace your competitors, having YieldPlanet watching your back.

CONTACT US:

YieldPlanet Poland

Wał Miedzeszyński 630
03-994 Warsaw, Poland
+48 22 769 38 09
salespol@yieldplanet.com

YieldPlanet Spain

Gran Via 1176 BIS 2ndo 9no
08020 Barcelona, Spain
+34 93 566 41 86
spain.sales@yieldplanet.com

YieldPlanet Switzerland

6005 Lucerne,
Switzerland
+41 79 916 11 40
sales@yieldplanet.com

www.yieldplanet.com