Annals of "Dunarea de Jos" University of Galati
Fascicle I. Economics and Applied Informatics
Years XXV – nº1/2019
ISSN-L 1584-0409 ISSN-Online 2344-441X
www.eia.feaa.ugal.ro

DOI https://doi.org/10.35219/eai1584040915

Machine Learning in Tourism Revenue Management

Maria-Cristina ENACHE[*]

ARTICLE INFO

Article history:
Accepted January 2019
Available online April 2019

JEL Classification
K22, M21

Keywords:
Business, IT, Models

ABSTRACT

Machine learning algorithms increase the efficiency of revenue management systems. Real-time data processing, customization, and automation are the key features that make it possible to overcome the performance of old systems in determining the price and time for a satisfactory offer and maximize revenue. Good practice is hiring external scientists to build segmentation and forecasting features. Such solutions require the collection of user information, which is hard to do without custom-built behavior and a market tracking engine.

1. Introduction

Income management in tourism aims to increase a company's revenue through demand management decisions, such as dynamic prices and capacity allocation. The classic revenue management in tourism is where a service provider sells a fixed number (per capita) of perishable service products (air seats, hotel rooms, etc.) through a booking process that ends at a fixed time (booking horizon). An income management system collects and stores booking records and market information and uses them to anticipate demand and learn customer behaviors. Then, during the booking horizon, choose the optimal orders based on these entries to maximize revenue. Controls are in the form of dynamic pricing and capacity allocation, which are the prices and availability of different tariffs.

The statistics in the figure below show the global revenues of international tourism from 2015 to 2017. In 2017, international tourism revenue amounted to US $ 1.34 trillion. According to the World Bank, the revenues from international tourism are expenses incurred by international visitors, including payments to national carriers for international transport. Any other anticipated payments for goods or services received in the destination country are included in these receipts.

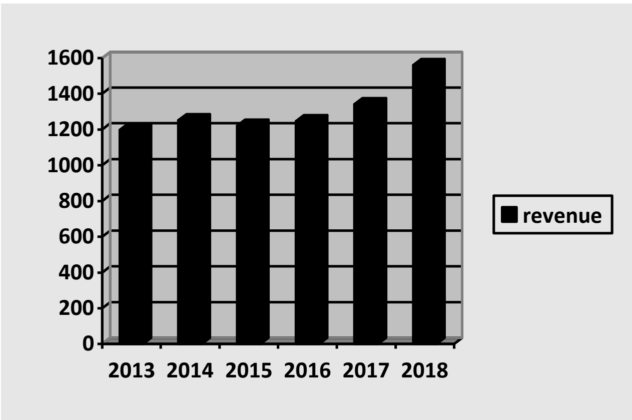

Fig 1. -International tourism revenue from 2013 to 2017 (in billion U.S. dollars)

The increase in tourism revenue is based on the multitude of booking facilities offered by online systems, increasing the involvement of intelligent technologies in tourism activities from booking to accommodation.

[*]Dunarea de Jos University of Galati, Romania. E-mail address: maria.podoleanu@ugal.ro

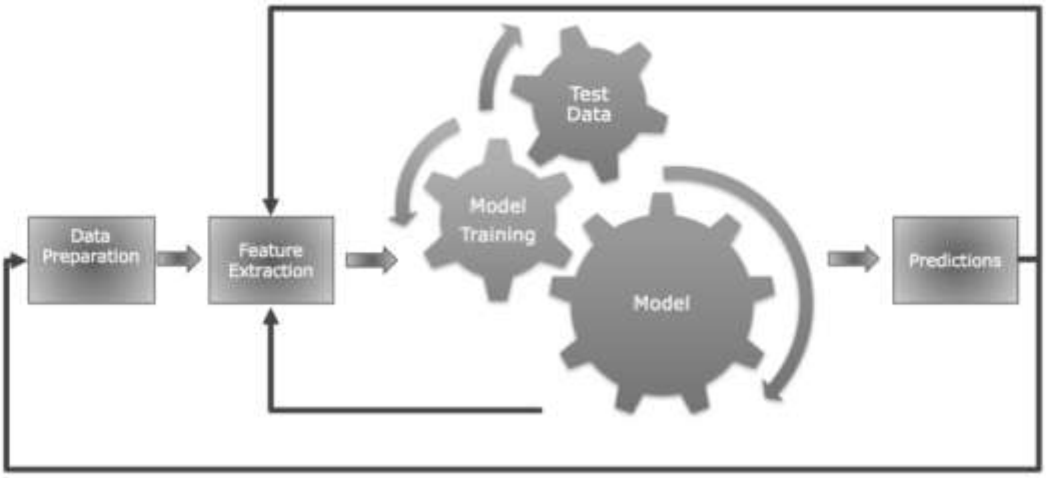

Fig 2. - Generalized Machine Learning Pipeline

An income management system in tourism should take into account the possibility that a reservation can be canceled or that a reserved customer cannot be displayed during the service (no-show), which is a special case of cancellation that occurs in that moment of service. One way to consider cancellations is to work with "net demand" instead of demand. Here, net demand is defined as the number of requests for applications minus the number of cancellations. Alternatively, we may continue to work with the request, but use a "virtual capacity", which is the real capacity plus a buffer representing the number of reservations to be canceled. In both cases, exact cancellation rates are essential for the revenue management system.

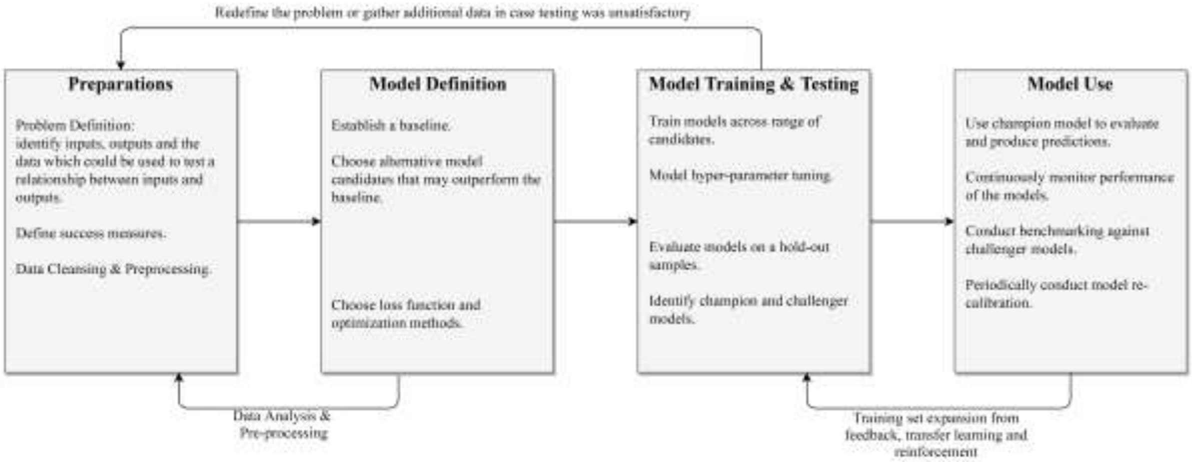

Fig 3. - Machine Learning Pipeline

Airlines and hotels regularly book abroad, which accepts more reservations than actual capacity based on the estimated number of cancellations or in other words accepting bookings to virtual capacity. Obviously, overwriting is only necessary after the number of bookings in progress, that is, reservations that have not yet been canceled among the existing ones, is close to the actual capacity. With standard demand management instruments such as dynamic pricing or capacity allocation control, the revenue management system will increase the price or close down cheaper prices if demand is high, so capacity will not be sold too early. As a result, overbooking is only required in the very late part of the booking horizon. This has led people to believe that the forecast of the cancellation rate is only needed near service provision.
There has been a lot of exaggerated advertising around concepts such as deep learning, ensembles, or bagging lately. These topics are currently some of the most popular in Machine Learning (ML).

In this sense, this paper will present the needs of a business today that can be adjusted with automated learning approaches. This article tends to be an argument or a point of support when we want to answer the following questions:

- Machine learning involves a lot of data?
- Machine Learning involves a high power of the computer?

When it comes to MACHINE LEARNING, we have always considered three main stakeholders, with very different views, who may be involved in a constructive application development process.

- The first interested party is represented by data scientists who are very interested in the power and accuracy of the tools.
- The second category includes the software developer, which is focused on application performance (speed, scalability, etc.).
- The third and most important part involved from a pragmatic perspective is the businessman who actually wants a solution to a certain real problem.

Major software vendors consider MACHINE LEARNING a big money market, and so there are many providers of machine tools and platforms such as IBM for analytics, Oracle or Amazon for cloud solutions, etc. From a business perspective, this tool can help solve problems, but I consider the specific area of knowledge in which it is applied. Know-how in the field is a fundamental ingredient in building solutions for real problems.

Asking the right questions, gathering accurate assumptions and measuring success is, in most cases, the focus and focus of MACHINE LEARNING stakeholders who have a business mentality. Data scientists and software developers sometimes care about the marketing ad, the references, or the reputation of MACHINE LEARNING tools or approaches. On the other hand, MACHINE LEARNING is commended by business people on the basis of its contributions or proven achievements.

2. Machine learning in revenue management

Next, I will present how R can be used in a MACHINE LEARNING issue in a current business scenario: managing revenue for hotels and optimizing benefits. R is a programming language and an open source environment. He supports statistical computing and automated learning. The Caret package developed by Max Kuhn is my favorite, but there are many more. R is the "weapon" I choose, and the following examples will refer to different libraries. Some similar ones can be found in Python, Java, or other programming languages, but summing up the right questions and finding the right data when implementing a good methodology are more important in putting into use a good MACHINE LEARNING solution.

Machine Learning has at least three applications in benefit management in general and in revenue management strategies for hotels, in particular:

A. **Segmentation of the market**: Different customers are willing to pay different rates for the same room. A hotel manager knows that 50 rooms with 200 euro / day bring the same income as 100 rooms with 100 euros per day, but their main goal is to maximize profit, which involves identifying market segments and finding the right target segments that can be served: ex. 30 rooms with 300 euro / day. The primary goal of the hotel when using segmentation is to describe and anticipate the value created for customers in different segments and target them in an efficient way.

In this sense, MACHINE LEARNING techniques such as classifiers help hotel managers find really profitable customers. These customers are the best and most in line with the hotel's value proposition. Caret from R implements and provides support documentation for the most refined classification tools like Naive Bayes. All CRM data contains insights about most loyal guests and allows hotels to better asses price elasticity within segments. Internal data can be collected via conventional tracking systems like CRMs or internal booking engines. The key variables in this category are:

- occupancy
- room rates
- bookings dates
- geographical information (where guests are arriving from)
- arrival dates
- departure dates
- revenue by day
- room type
- travel purpose
- inventory cost, etc.

This data can often be easily combined with behavioral data from the website. It’s usually gathered and stored in the CRM and PMS. Two main areas of use for internal data are segmentation and yield optimization.

User behavior data - User behavioral analysis requires a profound understanding of behavior patterns The main tasks we need to analyze are:

- web traffic values
- reference sources
- time spent on a website
- the frequency of the visit
- clicks
- device type

Behavior data allows estimates of room demand and price sensitivity. So it is useful for segmenting customers interacting with a website, especially those who have not logged in or visit a website for the first time.

Social media data - User-generated content in social media also provides valuable information about preference customers. This data group includes tweets, posts, ratings, geography, photos and text feelings. Social media is widely used to predict travelers' preferences and to evaluate their interests. Social media data helps and establishes destinations that gain popularity, making tweets, posts and geotags a good variable for forecasting demand.

B. **Forecast -** camera demand is uncertain in the future and the only way to predict with a certain probability is the application of forecasting techniques. Many hotels are starting to consider computerized revenue management systems. Their predictions are not perfect, but they are better than nothing. These systems help to know the behavior of customers, describe the seasonality of demand and predict the number of guests. Hotel managers can also predict costs, identify challenges, and adapt to market, regional or seasonal demand.

The market-related information allows for evaluating demand and supply conditions. This usually includes publicly available data:

- competitor rate information
- hotel occupancy at a destination
- flight demand
- online reviews
- canceled reservation
- payment types
- room-level photos
- guest count
- dates of stay.

The registered application is affected by the business decision and does not reflect the real demand. In the revenue management vocabulary, "free demand" is the amount that could be sold if there were no limitations such as delivery or capacity limitation. Hotels should identify when free demand is over hotel capacity. This is an important part of a hotel's revenue management strategy. Generally, the information comes from the computerized reservation system, and this is called "constrained demand". On the other hand, an ideal situation, when the offer is unlimited, is considered "unlimited / free demand" MACHINE LEARNING tools and statistics can help describe "free demand".

RM2 is a R package developed by Tudor Bodea (InterContinental Hotels Group), which implements functions used in revenue management (eg, the EM function releases the constraint request using the Wait - Maximize algorithm). On the other hand, "forecast" and "zoo" are just two of the most popular R packages that implement forecasting and time series. The RM2 platform can be explained simply as follows[web3]:

- Inputs are stored as specified by the time-based hierarchical semantic model and takes care of the data in an ensemble of coded weight (s) (0,1). The object is to create the object layer formation in the input layer, and these objects will continue to form networks between other objects available at a particular time stamp and it is further grouped on the object reference objects to create the top layer, to be called the memory layer. If all build a relationship with the object layer, making it easy to associate with a particular visual image or associated memory image. You can say that the memory layer is a complete (hierarchical) information about a particular scenario with time parameters, objects, shapes, colors, labels (names), behavior, derived fields, and result associations. All data is converted into a set of labels.
- During processing, the new assembly of labels is compared to the existing tag assembly of a neural object, if the match reveals the differences and similarities between two label assemblies. Similarities reinforce the relationship and the differences (which could be a combination of input sets) match the input level for similarities. If it is not available, create a new node and auto-tag with an alphanumeric prefix. Above language exposure, objects can create a label network to update the words of natural language and its associations by visual indications. This will allow the machine to understand the conflict models (reasoning) and reach the types of possible models that can cancel the conflict (planning solutions).
- Weights that are essential for state activation work at the individual entry layer and cascades to the highest level to reach cumulative weight to understand the threshold. At each entry, the weight is added to the previous output of the iteration, and when it reaches a certain threshold, the state changes to 0 or 1.

The RM2 design is developed by the machines responsible for making AI accessible and simple for everyone. The platform allows users to manage this unattended machine by monitoring / querying a particular node and understands the likely decisions / actions the machine can take, making AI easy to use and easier for everyone to manage AI without needing AI with very good skill.

C. **Pricing and pricing policy**: The key objective of a pricing strategy is to anticipate the value created for customers and then set specific prices to capture that value. Hotels can adopt different pricing policies to enhance customer value perception.

Tactics involve the creation of price-fixing tools that change dynamically to respond to changes in demand and to make gains on a continuous basis. Price optimization involves the fine tuning of multiple variables such as price elasticity and inventory price to maximize profit and not necessarily revenue. MACHINE LEARNING involves finding patterns in data and using those patterns to predict the future. Learning the pattern of demand and price setting means identifying and recognizing those patterns when we see them again. Optimization plays a very important role in this process. In its center, benefit management involves strategic inventory control to sell it to the right customers at the right time and for the right price. In mathematical language, inventory management can be embedded in an optimization problem. The "mechanic" behind automated revenue management involves some fundamental steps: In high demand weeks, limit your discount rate and group bookings to increase the overall benefit (average rate) and revenue. In low-demand weeks, sell empty chambers at any low price to increase yield and gain factors. Maximizing income requires a balance between benefit and return factors.

MACHINE LEARNING is an iterative process that runs until you get the pattern that makes good predictions. The hotel's revenue model may often imply fine tuning or reconstruction when decreasing its predictive power or accuracy. Income management is not a recent topic. It has been and has been a common business practice over the last fifteen years. MACHINE LEARNING is used, in addition to hotels, in many other industries, such as airlines, rentals, or online inventory.

3. Conclusions

Machine learning does not always require a lot of data to work. The hotel's data volume is not too high. Of course, online reservation systems and referral engines can provide additional data, but today's technologies can produce results on their own.

Machine learning requires a great deal of power on your computer, which is not the case today. At present, an MACHINE LEARNING application can be easily put into use as a micro-service. What we have recently understood is that more innovative methodologies than more revolutionary technologies are needed to solve current problems. A good question and a robust and transparent methodology are the essential ingredients.

Acknowledgement

This work was supported by a grant of the Romanian Ministry of Research and Innovation, CCCDI – UEFISCDI, project number PN-III-P1-1.2-PCCDI-2017-0800 / 86PCCDI/2018 - FutureWeb, within PNCDI III.

References

1. *https://www.altexsoft.com/blog/datascience/machine-learning-redefines-revenue-management-and-dynamic-pricing-in-hotel-industry/*
2. *https://qspace.library.queensu.ca/handle/1974/24906*
3. *https://responsiblemachines.wordpress.com/2017/04/13/how-unsupervised-learning-works-inside-rm2/*
4. *https://medium.com/@samairasandberg05/eight-machine-learning-js-frameworks-to-consider-in-2019-bd816f10dc45*
5. *https://www.altexsoft.com/blog/datascience/data-science-and-ai-in-the-travel-industry-9-real-life-use-cases/*