



Dr. Jian Wang | September 19, 2013

Revenue Management: An Earliest Use of Data Science

Big data is hot. Data science is even hotter. As an example, Google Trends shows that the search for “data science” was almost five times higher than “big data” in July 2013. However, for those who have been practicing revenue management long, the use of data science is not new at all. Revenue management as a discipline is, at its core, a data science problem.

Revenue management is a methodology to maximize an enterprise’s total revenue by selling the right product to the right customer at the right price at the right time through the right channel. It has been successfully applied to traditional service industries such as airlines and hotels for decades.

In essence, revenue management is a data-driven application consisting of three interdependent elements: customer, product, and price. The customer demand for a product is determined by price, which, in turn, is also influenced by the demand. The supply of a product is often assumed to be limited and perishable. The objective of revenue management is to seek such pricing that future revenue is maximized by balancing demand and supply optimally. One of critical steps is to predict how demand changes over price, and we accomplish this using data science.

So, what is data science? As described by Steve Cover in his blog “What To Do About Big Data,” data science is a multi-faceted discipline combining the traditional fields of study like math, statistics, computer science, and so on. Simply put, it is to extract or “mine” insightful knowledge from data for wise decision making. Under the framework of revenue management, data science utilizes complex algorithms to predict demand pattern from historical and real-time data. For example, from historical data of purchase behaviors, customers are often segmented using such approaches as Chi-Square Automatic Interaction Detection (a.k.a. CHAID). Then, for each customer segment, demand can be forecast using time-series methods like Exponential Smoothing.

Since American Airlines invented the world’s first revenue management system in 1985, the data science problem of revenue management has evolved significantly in two aspects: data and algorithm.

With the advent of IT technology, “big data” is generated and collected at amazing speeds. Big data becomes more and more sophisticated in terms of variety, volume, and velocity. Revenue management not only continues using traditional structured data of pre-sale, during-sale, and post-sale, but it also attempts to utilize the unstructured data of social media in the hope of further improving demand forecasts.

Prior to the age of big data, the amount of data that revenue management could use was relatively small and expensive to obtain, which often resulted in the problem of “data sparsity.” That is, for some customer segmentation, the amount of data available might be limited in order to estimate statistics accurately. To alleviate this problem, complex resampling algorithms, such as bootstrapping and cross-validation, were widely used. The nature of these algorithms is to re-use the same sparse data in an attempt to increase the accuracy of statistics. In reality, no matter how the small data is re-used, the resulting accuracy cannot be as good as using big data, simply because small data always contains less information than big data. So like the kids say in the AT&T commercial, “more is better,” or to answer the age-old question, “bigger is better.”

With the availability of big data, statistics can now be estimated using simple algorithms. In contrast to complex algorithms, simple algorithms are more effective and efficient in handling big data without compromising the accuracy. The data science of revenue management is moving from “small data, complex algorithm” to “big data, simple algorithm.” But, simple is harder than complex especially when it comes to designing algorithms and technology. Albert Einstein explains this all in his famous equation $E=MC^2$.

It seems big data is becoming a commodity, but mining for insights requires special skills and expertise that are very scarce. With the easy access and transparency of big data, revenue management companies have to fight for an important competitive advantage: crunch big data quickly. And like in the old Wild West, only the ones who draw the pistols fast and shoot stand last.