Lecture Notes in Management Science (2015) Vol. 7, 45–49 ISSN 2008-0050 (Print), ISSN 1927-0097 (Online)

Performance optimization in retail business using real-time predictive analytics

Nizar Zaarour [1] and Emanuel Melachrinoudis [1]
[1] Northeastern University, 360 Huntington Avenue, Boston, MA 02115 USA
n.zaarour@neu.edu; emelas@coe.neu.edu

Proc. ICAOR 2015
Vienna, Austria

Keywords:
Pattern recognition
Real-time data
Retail

Abstract

Predictive analytics has become more than a necessity in today's competitive market and technology based world. Traditional data analysis tools such as regression analysis, numerical taxonomy, cluster analysis, and other multivariate statistical methods are not sufficient anymore to handle all the complicated intricacies and big data sets. To compensate for these shortcomings, more advanced tools have emerged, which are based on artificial intelligence, pattern recognition, statistics, machine learning, and many other data mining tools. In this paper, we develop a model that focuses on collecting real-time data for the purpose of recognizing patterns in the retail and point-of-service business, and improving the sales and customer satisfaction. We were able to identify a large number of predictors affecting the behavior of sales; however, this particular paper focuses on one specific independent variable, the number of people involved in each transaction in a restaurant business. Breakdown of the model along with the results is presented, and a summary of planned future research, involving the impact of other predictors, is discussed at the end of the paper. One of the major benefits of this research, is the win-win-win approach to make sure that we optimize the satisfaction of all parties involved: management, staff, and customers.

Introduction

Statistical modeling has always been a necessary tool to analyze and predict the behavior of variables, such as sales, demand, prices, inventory, etc. Understanding how a variable is changing by analyzing the changes of a different variable or variables is not an easy task. In addition to figuring out which factors are best suited to predict the behavior of a variable, we have to deal with the daunting, ever-growing amounts of data available to us today. More data has been recorded in the past five years than in all of previous human history, IBM (2013).

According to Mayer-Schônberger and Cukier (2013), big data is being used to transform medical practices, modernize public policy, and inform business decision making. Inferential statistics such as estimating parameters and testing for causal hypotheses are traditionally based on statistical explanatory models, which tend to test these hypotheses and determine whether we have significant results to challenge these claims being made, rather than generating accurate empirical predictions, thus, the need for better predictive modeling. Real time prediction modeling is becoming a necessity in all fields and domains, due to better statistical tools and machine learning, and is replacing the traditional forecasting that relies mainly on past data to predict future results.

Given the ability of collecting more and more data, businesses today are faced with the pressure of analyzing this data correctly and making better decisions involving their bottom line. Data can now be collected in real-time and stored in the cloud. This is a double edge sword, since the expectations are higher than ever to optimize performance and maximize or minimize the company's deliverables. There is a significant amount of work that is taking place today in understanding what is needed to collect, clean, measure, analyze, and predict outcomes based on the data collected. However, it has not always been a smooth ride in terms of understanding the necessity of data mining processes. Friedman (1997) defined the field as vague since it depends on the experience and views of the definer. Weiss and Indurkhya (1998) defined the field as the search for valuable information in large amounts of data.

With the technology evolving, the need for integrating data mining and predictive analytics became more pertinent. Liu *et al.* (2001) illustrated the benefit of time-series data mining by using fast-food restaurant franchise data. They employed the Box-Jenkins seasonal ARIMA models to analyze and forecast the time series. Shmueli and Koppius (2010) highlighted the need to integrate predictive analytics into information systems research. This need was preceded by many other researchers such as Marcoulides and Saunders (2006) who focused on structural equation modeling, Petter *et al.* (2007) who concentrated on formative constructs, and Li and Hitt (2008) whose motivation was the selection bias.

Latest research includes Waller and Fawcett (2013), whose work focused on the transformation of supply chain design and management by data science and predictive analytics. Similarly, articles by McAfee and Borynjolfson (2012), Davenport and Patil (2012), Barton and Court (2012), and Chen *et al.* (2012) all addressed the growing impact of business intelligence and analytics. The rest of the paper will include the following sections: model methodology and results, and conclusion and future research.

Model Methodology and Results

In this paper, real time data from point-of-service businesses belonging to the same sector was collected. An extensive analysis was done by taking into consideration multiple variables as potential predictors. These variables can be placed in different categories to answer the following questions: how can we optimize the customer's spending while increasing their satisfaction level? How can we maximize the staff performance while keeping them incentivized?

The transaction time of customer experience, including all the steps from arriving to leaving, is very similar in nature to the supply chain management process from order to manufacture to delivery. These steps are recorded first by a point-of-sale order tracking, and are finalized by an invoice. What does good service mean? Does it automatically translate into more revenue (customer spending)? A happy customer can mean many things: more spending at the time of the encounter (tough to measure), a bigger tip (only beneficial to the staff, and not particularly an indication of improved spending), or a returned customer, which can also include better reputation. This could also be materialized by potential new customers through word of mouth or social media. Businesses also ponder the impact of optimal time as a necessary component of good service.

This research focuses on one aspect of these multi-faceted issues, and that is how to optimize spending by looking at the number of customers involved in the transaction. However, since it is very difficult to quantify the impact of time and good service, the model will focus on the number of people being served and will optimize the expected sales. In this research, the authors collaborated with Telative, Inc., a start-up in Boston, MA, which provides real-time data to managers and servers in sit-down restaurants (2015). Telative aggregates data from existing restaurant technologies such as point-of-sale, table management and social sites, as shown in Figure 1, to provide critical data in-the-moment to employees on the restaurant floor. The company wanted to enrich the data provided to employees by providing predictive information to anticipate the needs of restaurant guests and improve the performance of restaurant employees.

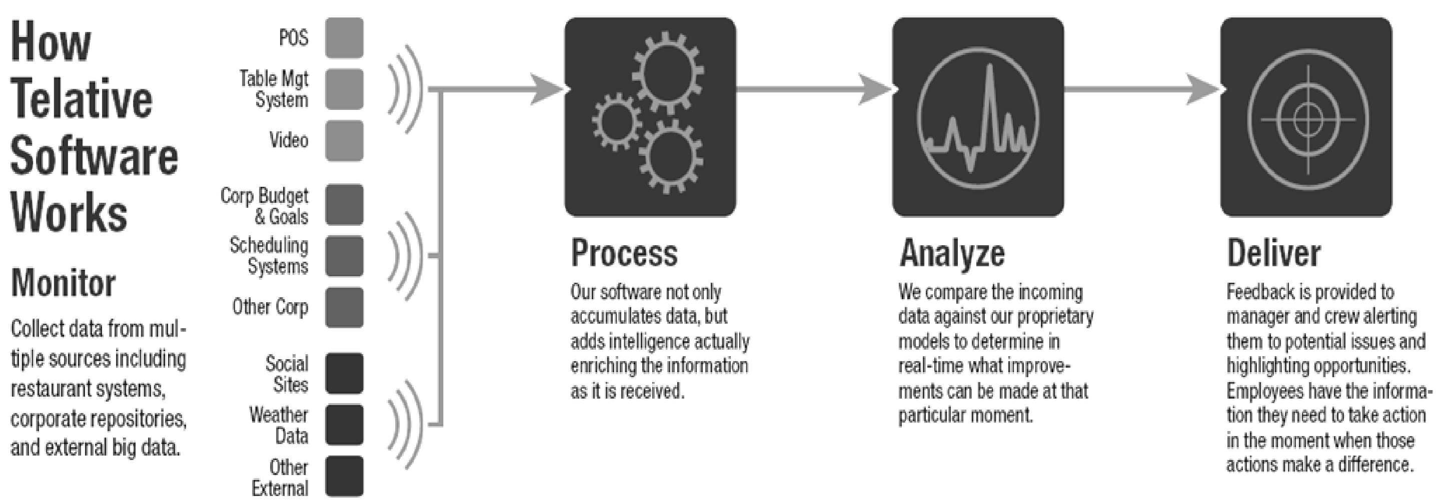

Fig. 1. Telative software work process

Using the Telative software and predictive analytics, we were able to develop a model that allows us to predict the sales based on the number of customers at any given table, regardless of the type of meal, and seasonality. The model also allows us to have a better understanding of where a check is, in regards to where it is supposed to be, and gives direction in terms of how to get to the desired target sales. Even though this model is based on only one predictor, we are in the process of implementing additional independent variables to see the impact on predicting sales. There is no limitation on the number of predictors potentially used, as long as we take into consideration the multicollinearity of these predictors.

This process maximizes the customer experience, while allowing the server to benefit from more tips, and eventually maximizing the sales for the restaurant. Figure 2 illustrates the different areas on the graph of where a check is, where it needs to be, and where it can get to if the server wants to go above and beyond.

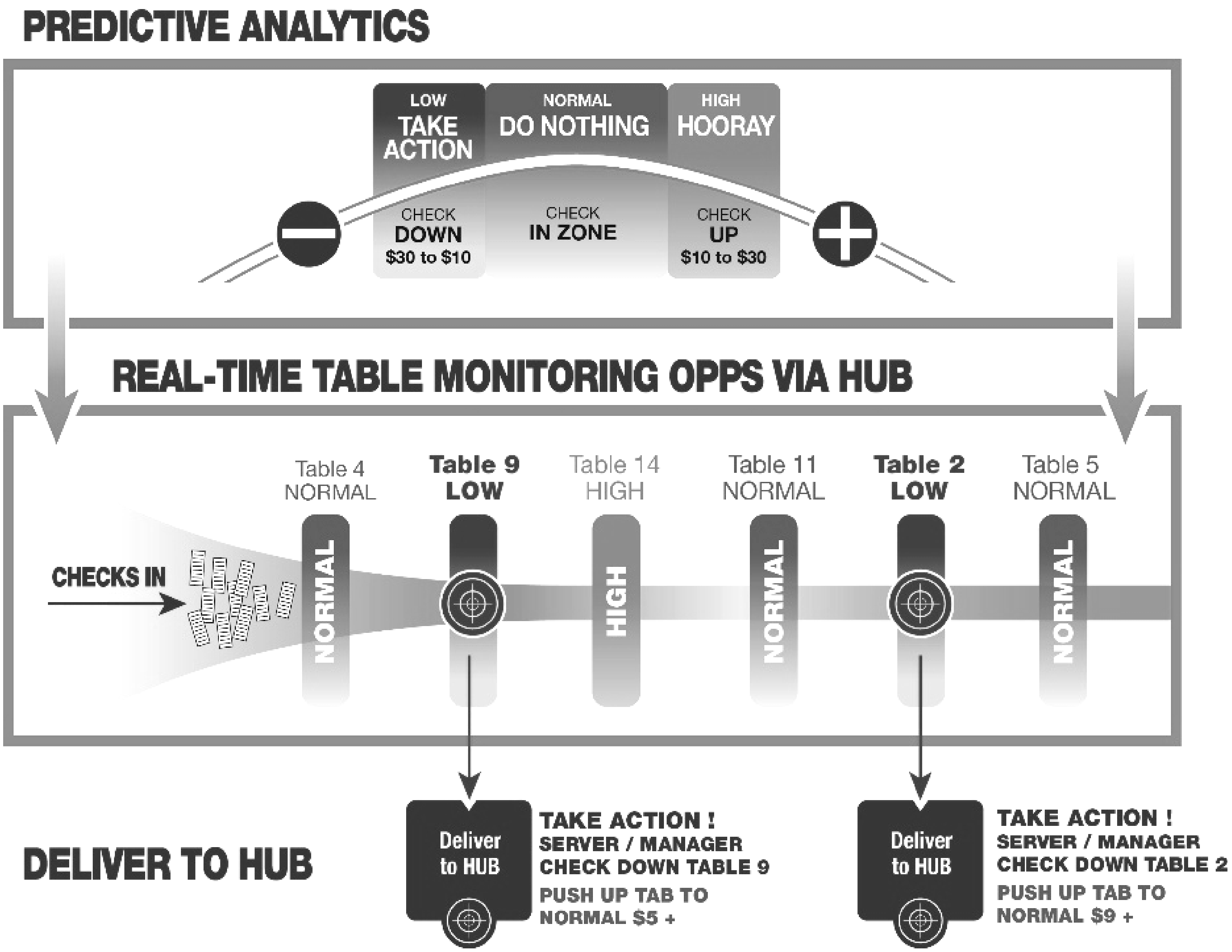

Fig. 2. Real time monitoring

The benefit of such model is that it reacts in real time, and it makes adjustments as the data keeps getting collected. We are not limited to what happened in the past, but instead, we use real-time to analyze, adjust, and make better recommendations. The model is being currently tested and implemented in several locations, and all the results have been positive. Figure 3 shows an example of the implementation of the model and the gain obtained. The "before" case indicates the number of save opportunities, which are calculated by considering the number of cases, for a specific time period, that are underperforming according to the lower tail of our model. The number of saves and misses without using the suggested HUB system, are arbitrary and can go up or down depending on the server. However, most of the time, the number of misses is a lot higher than the number of saves. In other words, looking at these opportunities in real practice, we noticed that in most cases, these save opportunities go unnoticed, and even with the top servers, the ratio of saves to misses is still significantly low. On the other hand, the "after" window shows a similar number of save opportunities, and after only a month of using the system, we noticed a significant increase to the ratio of saves to misses. We ran hypothesis testing of multiple cases of the before and after scenarios, and we were able to confirm our assumption that there is a significant increase in the save to misses ratio in all of the cases.

Fig. 3. Before and after implementing the model

Conclusion and Future Research

This research has great potential by considering many different variables, such as, trends of the servers (years of experience, gender, behavior patterns, etc…), further analysis of seasonality impact (weekday vs weekend, holidays, lunch vs dinner, etc…). In addition, the goal is to continue developing this model to not only optimize the amount spent by customers, but to pinpoint exactly the specific items on the menu to achieve that target by looking at patterns and trends of customers' ordering habits. Furthermore, the model will serve as a basis to optimize sales and customer satisfaction using real-time data for different retail stores and businesses.

References

Barton, D., Court, D. (2012). Making advanced analytics work for you. Harvard Business Review, 90, 79-83.

Chen, H., Chiang, R., & Stovey, V. (2012). Business intelligence and analytics: from big data to big impact. MIS Quarterly, 36(4):1165-1188.

Davenport, T., Patil, D. (2012). Data scientist: the sexiest job of the 21st century. Harvard Business Review, 90, 70-76.

Friedman, J.H. (1997). On bias, variance, 0/1 – loss, and the curse of dimensionality. Data Mining and Knowledge Discovery, 1, 55-77.

IBM source (2013). Retrieved from http://www-01.ibm.com/software/data/bigdata, accessed 20 February 2015.

Li, X., Hitt, L.M. (2008). Self-selection and information role of online product reviews. Information Systems Research, 19(4): 456-474.

Liu, L.-M., Bhattacharyya, S., Sclove, S. L., Chen, R., & Lattyak, W.J. (2001). Data mining on time series: an illustration using fast-food restaurant franchise data. Computational Statistics and Data Analysis, 37, 455-476.

Marcoulides, G.A., Saunders, C. (2006). PLS: A silver bullet? MIS Quarterly, 30(2): 3-4.

Mayer-Schônberger, V., Cukier, K. (2013). Big data: a revolution that will transform how we live, work, and think. New York: Houghton Mifflin Harcourt Publishing Company.

McAfee, A., Brynjolfsson, E. (2012). Big data: the management revolution. Harvard Business Review, 90, 60-68.

Petter, S., Straub, D., & Rai, A. (2007). Specifying formative constructs in information systems research. MIS Quarterly, 31(4): 623-656.

Shmueli, G., Koppius, O. (2010). Predictive analytics in information systems research. Robert H. Smith School Research Paper No. RHS 06-138. Available at SSRN: http://ssrn.com/abstract=1606674, accessed 20 February 2015.

Telative. www.Telative.com, accessed 03 March, 2015.

Waller, M.A., Fawcett, S.E. (2013). Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management. Journal of Business Logistics, 34(2): 77-84.

Weiss, S.M., Indurkhya, N. (1998). Predictive data mining. Morgan Kaufmann Publishers, San Francisco.