STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **8**, March 2020, pp 272–280.
Published online in International Academic Press (www.IAPress.org)

Business Analytics using Dynamic Pricing based on Customer Entry-Exit Rates Tradeoff

Hamed Fazlollahtabar [1,*], Minoo Talebi Ashoori [2]

[1] *Department of Industrial Engineering, School of Engineering, Damghan University, Damghan Iran.*

[2] *Purdue University Northwest, IN, USA*

Abstract This paper concerns with an integrated business process to be applied as a decision support for market analysis and decision making. The proposed business intelligence and analytics system makes use of an extract, transform and load mechanism for data collection and purification. As a mathematical decision optimization, dynamic pricing is formulated based on customer entry-exit rates in a history-based pricing model. The optimal prices for products are obtained so that aggregated profit is maximized. A case study is reported to show the effectiveness of the approach. Also, analytical investigations on the impacts of the sensitive parameters of the pricing model are given.

Keywords Business intelligence and analytics, Dynamic pricing, Customer entry-exit rates

AMS 2010 subject classifications 91B26 , 90B60

DOI: 10.19139/soic-2310-5070-551

1. Introduction

Business Intelligence that applies data analytics to generate key information to support business decision making, has been an important area for more than two decades. In the last five years, the trend of Big Data has emerged and become a core element of Business Intelligence research. In this article, we review academic literature associated with Big Data and Business Intelligence to explore the development and research trends [1]. Firms continuously report increased competitive value gains from the use of business intelligence and analytics (BI&A), however, little is known about how insights from BI&A are transformed to added value to date [2]. Business Intelligence Systems provide capabilities such as business information analysis in order to support and improve decision-making management in a wide range of business activities and also provide for managers easy to use, timely and appropriate information at different organizational levels and enable them to make better decisions. Caseiro and Coelho [3] proposed a model to investigate the direct effects of Business Intelligence (BI) on performance, and the indirect effects, through network learning (NL) and innovativeness (INNOV).
Product pricing decisions are main core of any business marketing and have a direct impact on corporate business strategy. Price is the benefit that consumer pay for benefits of having or using the product [4]. Simply pricing means set the prices for goods or services. Pricing is an activity that should be repeated and the process is continuing. The continuing resulting from environmental changes and market conditions instability that creates price adjustment. Pricing is done to maximize profit, increasing market share, and quality leadership or increase market prices [5].

*Correspondence to: Hamed Fazlollahtabar (Email: hfazl@du.ac.ir). Department of Industrial Engineering, School of Engineering, Damghan University, Damghan Iran.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

Reasonable price is the price that satisfies the needs of the seller and the buyer at the same time.
The model in [6] investigated the instantaneous fuzzy economic order quantity model by allocating the percentage of units lost due to deterioration in an on-hand inventory by framing variable ordering cost. The objective was to maximize the fuzzy net profit so as to determine the order quantity, the cycle length and number of units lost due to deterioration in fuzzy decision space.
The problem of optimal estimation of linear functional which depends on the unknown values of a stochastic sequence with stationary increments was considered in [7]. Formulas for calculating the mean-square errors and the spectral characteristics of optimal estimates of the functional were proposed under condition of spectral certainty.
Measuring the accuracy of diagnostic tests is crucial in many application areas including medicine, machine learning and credit scoring. In [8], authors presented Nonparametric Predictive Inference (NPI) for the best linear combination of two biomarkers, where the dependence of the two biomarkers is modelled using parametric copulas.
Having been in touch with technical side of many companies in various sectors, we know industry is facing a period of unprecedented change. A list of common technological challenges in the sector that companies are facing in adapting to the change was compiled in [9]. The purpose was to discuss and communicate areas where technological challenges in the Software Development Sector lie.
The aim here is to propose an integrated business intelligence and analytics for product pricing purpose. The business process is designed and extract, transform and load mechanism for data collection and purification is developed. Then, dynamic pricing model is formulated based on customer entry-exit rates for aggregated profit computation.
The paper is organized as follows. The problem is defined and modeled in Section 2. The proposed process of business intelligence and analytics is presented in Section 3. Implementation of the proposed model is conducted as a case study in Section 4. Finally, we conclude and express limitations and directions for future research in Section 5.

2. Problem definition and modeling

There are two firms indexed by $i = A, B$. The firms produce homogeneous services (or goods). Consumers gain utility $\bar{u}$ from consumption, which means that $\bar{u}$ is the maximum amount that they are willing to pay for the service [10]. The following rates are defined:
(a) Consumers enter the market at a rate $\beta \geq 0$ (entry rate).
(b) Existing consumers exit the market at a rate $\delta \geq 0$ (exit rate).
New consumers purchase when they enter the market and thereafter become locked in with the firm they initially chose to purchase from. Existing customers purchase at a unit rate. More specifically, in a period of length dt, conditional on not exiting the market, an existing consumer purchases with probability $1dt$. In aggregate, since there are $N(t)$ existing (locked-in) consumers, the expected number of purchases by locked-in customers is $N(t)dt$. This means that, in a period of length dt, new consumers buy $\beta N(t)dt$ units (since there are $\beta N(t)dt$ entering consumers during dt) and existing consumers buy $N(t)dt$ units, whereas consumers who exit the market do not buy during the time interval dt. Thus, the proportion of products purchased by new consumers is [10]:

$$\frac{\beta N(t)dt}{\beta N(t)dt + N(t)dt} = \frac{\beta}{\beta + 1} \in [0, 1) \tag{1}$$

To ensure that the market is fully covered, we assume that the utility rate $\bar{u}$ is sufficiently high such that all consumers subscribe to the offered service (buy the offered goods). A sufficient condition for the market to be covered under the two pricing regimes to be analyzed (history based pricing and uniform pricing) [10]. Note that, we consider the rate of willingness to pay is bounded from below. Formally, $\bar{u} > 2 + \frac{1}{\beta}$.
Firm i has a market share of $\sigma_i^L(t)$ at time t among locked-in consumers and $\sigma_i^L(t)$ among new consumers with the shares for firm j being $\sigma_j^L(t) = 1 - \sigma_i^L(t)$ and $\sigma_j^N(t) = 1 - \sigma_i^N(t)$ respectively. With differentiated pricing firm i sets the price $p_i^N(\sigma_i^L)$ for new consumers and $p_i^L(\sigma_i^L)$ for locked-in consumers. These prices are contingent on the state variable σ_i^L Note that $\sigma_i^{N_i}$ is not a state variable since it can be changed instantaneously through current

period prices. Under uniform pricing we consider $p_i^N(\sigma_i^L) = p_i^L(\sigma_i^L) = p_i^U(\sigma_i^L)$. Firm is instantaneous profit and time discount rate are denoted by $\pi_i(t)$ and ρ^f. Assuming zero production cost, at each period t_0, firm i chooses a continuous time price strategy $(p_i^N(\sigma_i^L) = p_i^L(\sigma_i^L))$ to maximize the value function given by:

$$V_i(t_0) = \int_{t_0}^{\infty} e^{-\rho^{f(t-t_0)}} \pi_i(t)dt, i = A, B \tag{2}$$

where,

$$\pi_i(t) = N(t_0)e^{n(t-t_0)}[p_i^N(t)\beta\sigma_i^N(t)] + p_i^L(t)\sigma_i^L(t) \tag{3}$$

At the time of adoption, the new consumers who are entering the market choose a brand to maximize a discounted stream of utilities with a discount rate ρ^c . Each new consumer entering at t0 chooses a brand i according to

$$\arg\max U_i = \bar{u} - \tau_i - p_i^N(t_0) + \int_{t_0}^{\infty} e^{-(p^c+\delta)(t-t_0)}[\bar{u} - p_i^L(t)]dt. \ i \in A, B \tag{4}$$

The first three terms in (4) constitute the initial utility net of adoption cost and the introductory price. Combined with the last term, (4) implies that consumers choose the brand that yields the highest value net of brand adoption cost, current price, and net of all discounted future prices during the consumers lifetime when the consumers are locked in with the chosen brand. Consumers discount the future by $e^{-(p^c+\delta)(t-t_0)}$. As seen in (4), the discount rate ρ^c and the exit rate δ have similar effects on the consumers brand selection. The parameter δ determines the expected duration of customer relationships and thereby the industry growth rate n. This parameter influences firms return from an acquired customer relationship and thereby the intensity of competition for new customers.
With history-based pricing (HBP), firms differentiate the price targeted for entering consumers from the price offered to locked-in consumers. More precisely, this pricing regime focuses on competition when firms charge two prices: $p^L(\sigma_i^L)$ to their existing (locked-in) customers, and $p^N(\sigma_i^L)$ to new customers. With history-based pricing, the unique model is given by (Shy, 2016):

$$p^{L*} = \bar{u} \, and \, p^{N*} = 1 - \frac{\bar{u}}{\rho^f + \delta}, \, for \, all \, \sigma_i^L \in [0, 1] \tag{5}$$

yielding an average price of,

$$p^{-HBP} = p^{N*}(\frac{\beta}{1+\beta}) + p^{L*}(\frac{1}{1+\beta}) = 1 - \frac{\bar{u}}{\rho^f + \delta}(\frac{\beta}{1+\beta}) + \bar{u}(\frac{1}{1+\beta}) \tag{6}$$

at each point in time.
From (6) we can infer that the price targeted to existing customers, $p^{L*} = \bar{u}$, as well as that targeted to entering consumers without history, $p^{N*} = 1 - \frac{\bar{u}}{\rho^f+\delta}$ are each independent of the rate of consumer entry β. In fact, with history-based pricing the rate of consumer entry only affects the proportions of consumers targeted by p^{L*} and p^{N*} respectively. Thus, the rate of consumer entry only affects the average price associated with history-based pricing. This average price is nevertheless important as it determines the profit rate of firms, firm value, and the rate of consumer surplus. An increased entry rate of consumers (higher β) implies a reduced average price, meaning that an increased entry rate makes the investment effect stronger. On the other hand, a reduced β implies a higher average price as it makes the harvesting affect stronger [10].

3. Business intelligence and analytics process

ETL stands for Extract, Transform and Load. The ETL process typically extracts data from the source / transactional systems, transforms it to fit the model of data warehouse and finally loads it to the data warehouse.
The transformation process involves cleansing, enriching and applying transformations to create the desired output. Data is usually dumped to a staging area after extraction. In some cases, the transformations might be

applied on the fly and loaded to the target system without the intermediate staging area.

Step 1) Extraction
In this step, data is extracted from the source system into the staging area. Transformations if any are done in staging area so that performance of source system in not degraded. Also, if corrupted data is copied directly from the source into Data warehouse database, rollback will be a challenge. Staging area gives an opportunity to validate extracted data before it moves into the Data warehouse.
Data warehouse needs to integrate systems that have different database management system (DBMS), Hardware, Operating Systems and Communication Protocols. Sources could include legacy applications like Mainframes, customized applications, Point of contact devices like ATM, Call switches, text files, spreadsheets, ERP, data from vendors, partners amongst others. Hence one needs a logical data map before data is extracted and loaded physically. This data map describes the relationship between sources and target data.

Three Data Extraction methods:

1. Full Extraction
2. Partial Extraction- without update notification.
3. Partial Extraction- with update notification

Irrespective of the method used, extraction should not affect performance and response time of the source systems. These source systems are live production databases. Any slow down or locking could affect company's bottom line.

Some validations are done during Extraction:

- Reconcile records with the source data
- Make sure that no spam/unwanted data loaded
- Data type check
- Remove all types of duplicate/fragmented data
- Check whether all the keys are in place or not

Step 2) Transformation
Data extracted from source server is raw and not usable in its original form. Therefore, it needs to be cleansed, mapped and transformed. In fact, this is the key step where ETL process adds value and changes data such that insightful BI reports can be generated.
In this step, you apply a set of functions on extracted data. Data that does not require any transformation is called as direct move or pass through data.
In transformation step, you can perform customized operations on data. For instance, if the user wants the revenue for sum-of-sales which is not in the database. If the first name and the last name in a table is in different columns. It is possible to concatenate them before loading.

Following are Data Integrity Problems:

1. Different spelling of the same person like Jon, John, etc.
2. There are multiple ways to denote company name like Google, Google Inc.
3. Use of different names.
4. There may be a case that different account numbers are generated by various applications for the same customer.
5. In some data required files remains blank
6. Invalid product collected at POS as manual entry can lead to mistakes.

Validations are done during this stage:

• Filtering C Select only certain columns to load
• Using rules and lookup tables for Data standardization
• Character Set Conversion and encoding handling
• Conversion of Units of Measurements like Date Time Conversion, currency conversions, numerical conversions, etc.
• Data threshold validation check. For example, age cannot be more than two digits.
• Data flow validation from the staging area to the intermediate tables.
• Required fields should not be left blank.
• Cleaning (for example, mapping NULL to 0 or Gender Male to ”M” and Female to ”F” etc.)
• Split a column into multiples and merging multiple columns into a single column.
• Transposing rows and columns
• Use lookups to merge data
• Using any complex data validation (e.g., if the first two columns in a row are empty then it automatically rejects the row from processing)

Step 3) Loading
Loading data into the target data-warehouse database is the last step of the ETL process. In a typical Data warehouse, huge volume of data needs to be loaded in a relatively short period (nights). Hence, load process should be optimized for performance.
In case of load failure, recover mechanisms should be configured to restart from the point of failure without data integrity loss. Data Warehouse admins need to monitor, resume, cancel loads as per prevailing server performance.

Types of Loading:

• Initial Load ł populating all the Data Warehouse tables
• Incremental Load ł applying ongoing changes as when needed periodically.
• Full Refresh łerasing the contents of one or more tables and reloading with fresh data.

Load verification:

• Ensure that the key field data is neither missing nor null.
• Test modeling views based on the target tables.
• Check that combined values and calculated measures.
• Data checks in dimension table as well as history table.
• Check the BI reports on the loaded fact and dimension table.

Our proposal model includes three stages that are explained below and illustrated in Figure 1:

Stage 1 - data collection: first the required data from various sources are collected from customers in the marketing environment and the data may not necessarily be accurate (uncertainty).
Stage 2 - extracting data from databases into a single repository: At this stage the data using ETL tools are refined, integrated and stored to analyze business intelligence.
Stage 3 - then the data are used by pricing models under decision support system for analysis and decision-making and provide optimal price.

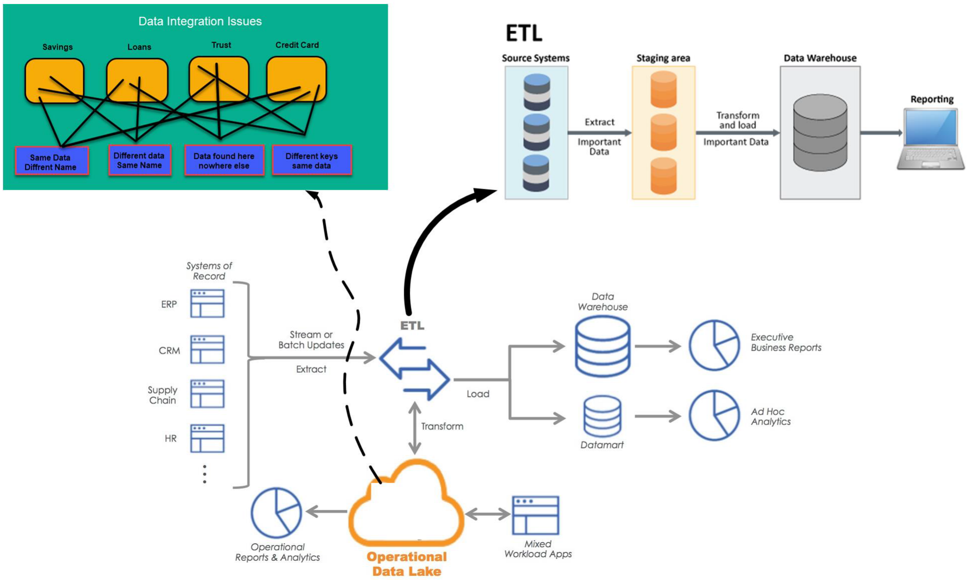

Figure 1. Proposed model

4. Numerical study

In this section, we implement this process for a food distribution company. The main activity of food distribution companies is import, export and distribute a comprehensive variety of products such as tea, oil, rice, sugar, beans, tuna, paper towels, soap and more.

Stage 1 - Data collection

Data are collected from customers of a food distribution company during a year. This company was studied during a year in which its products are sold to 76 customers. The company sells products such as tea, oil, rice, tuna, paper towels, tomato paste, ketchup, mayonnaise, macaroni, and dishwashing liquids. Some information is shown in Table 1.

Table 1. Products and companies

Product name	Company name	Discount rate
Oil	Company 1	5%
Tea	Company 2	3%
dishwashing liquids	Company 3	4%
ketchup	Company 4	2%
tuna	Company 5	4%
Rice	Company 6	6%
macaroni	Company 7	4%
paper towels	Company 8	3%
mayonnaise	Company 9	2%
tomato paste	Company 10	3%

Table 2 shows entry rates for new and locked in customers as well as exit rate for products of a food distribution company.

Table 2. Entry and exit rates of customers

Product Name	Entry rate of lock in customers	Entry rate of new customers	Exist rate of customers
Oil	47	12	5
Tea	24	8	3
dishwashing liquids	53	15	3
Ketchup	59	14	4
Tuna	49	9	3
Rice	48	6	4
Macaroni	32	5	5
paper towels	27	4	3
Mayonnaise	34	5	6
tomato paste	30	6	4

Stage 2 - Extracting data from databases into a single repository Now, the data is collected may include repetitive and redundant data or excessive data that need to be refined, integrated and saved using ETL tools to be analyzed in business intelligence database. According to Figure 1, first specific tables from the data source are chosen to load required data for measuring. Then, the data are transferred to the operational level and ETL process that involves three stages of extraction, transformation and loading being done for extraction, cleaning, customization and loading data. The extracted and refined data are transferred into the data warehouse and data mart for analysis by end users.
Table 3 is the output of the ETL process, which includes required data to calculate prices using pricing model based on history.

Table 3. The output of the ETL process

Product Name	Discount rate	Exist rate of customers	Entry rate of new customers	Entry rate of lock in customers
Oil	0.05	5	12	47
Tea	0.03	3	8	24
dishwashing liquids	0.04	3	15	53
Ketchup	0.06	4	14	59
Tuna	0.02	3	9	49
Rice	0.03	4	6	48
Macaroni	0.04	5	5	32
paper towels	0.02	3	4	27
Mayonnaise	0.03	6	5	34
tomato paste	0.04	4	6	30

Stage 3- Applying history-based pricing model
At this stage data are used by pricing model for analysis and decision-making leading to determining products prices. The results of the calculations are presented in Table 4. The unit measure is 1000$. For example, the average price for oil is obtained 0.61 that indicates the average price of 610$.

Table 4. Prices considered for new and lock in customers and the average price of products

Product Name	p^{-HBP}	p^{N*}	p^{L*}	Utility rate New customers	Utility rate Lock in customers	ρ^f	δ	Entry rate of New customers	Entry rate of lock in customers
Oil	0.61	0.59	2.05	2.11	2.05	0.05	5	12	47
Tea	0.37	0.3	2.07	2.15	2.07	0.03	3	8	24
dishwashing liquids	0.33	0.32	2.04	2.09	2.04	0.04	3	15	53
ketchup	0.50	0.49	2.04	2.1	2.04	0.06	4	14	59
Tuna	0.33	0.3	2.05	2.14	2.05	0.02	3	9	49
Rice	0.52	0.5	2.05	2.19	2.05	0.03	4	6	48
macaroni	0.32	0.27	2.06	2.23	2.06	0.04	5	5	32
paper towels	0.36	0.32	2.06	2.28	2.06	0.02	3	4	27
mayonnaise	0.66	0.64	2.05	2.23	2.05	0.03	6	5	34
tomato paste	0.53	0.49	2.06	2.1	2.06	0.04	4	6	30

4.1. Analysis

With history-based pricing, the average price (p^{-HBP}) is:
(a) decreasing as a function of the customer entry rate β.
(b) increasing as a function of the customer exit rate δ.
(c) increasing as a function of the firms time discount rate ρ^f.
For more weighted index that p^{N*} has in equation. The effect of p^{N*} to calculate the average price is much more than p^{L*} and the average price is closer to p^{N*}.
Assuming a constant discount rate and exit rate, with increasing utility rate the price for new customer's decreases and vice versa.
Assuming a constant utility rate for new customers, whatever the total rate of exit rate and discount rate is more, target price for new customer's increases, and vice versa.
According to above results, average prices fall with an influx of new customers into the market (higher β) since a larger proportion of the consumers are offered introductory discounts. This is the investment effect. With an increased effective discount rate, due to either a higher exit rate of existing customers (higher δ) or firms operating with a higher discount rate, this investment ρ^f effect is weakened.
This model can be used as a tool to predict profits in future periods and company managers can forecast profits for the next year. Some of the advantages of the proposed model are earning more profit, increasing corporate value and avoiding price volatility.

5. Conclusion

An enterprise should be able to pricing its products in a way to achieve more earnings proportional to the value provided to the customer and so would maintain its position to the customers, complementary products, competitors and potential newcomers. Pricing is a significant decision for producers (sellers) and has become a difficult issue on the market today that is changing rapidly over time. In this study, we proposed a model under decision support system to help business corporations. In products pricing and history-based pricing method which was used to determine the price of products, a new paradigm for in a business intelligence mechanism was implemented to maximize revenue and fulfill customer demands and their satisfaction. Also the pricing model was used to predict profits in future periods which could identify the importance of each factor on products profit. We proceed to interpret the main results of history-based pricing for Business Corporation and express limitations of our work and directions for future research. In future research web mining exploration techniques such as correlation rules to collect data can be used. Also the effect of specific features on corporate income and profits can be explored.

Identifying effective risk factors on profit can be considered in order to provide more accurate model for predicting profit.

Acknowledgement

The first author would like to thank Iranian Operations Research Society for supporting this work

Reference

1. Ting-Peng Liang, Yu-Hsi Liu, (2018). *Research Landscape of Business Intelligence and Big Data analytics: A bibliometrics study*, Expert Systems with Applications, Volume 111, 2-10.
2. Katerina Bozic, Vlado Dimovski, (2019). *Business intelligence and analytics for value creation: The role of absorptive capacity*, International Journal of Information Management, Volume 46, 93-103.
3. Nuno Caseiro, Arnaldo Coelho, (2019). *The influence of Business Intelligence capacity, network learning and innovativeness on startups performance*, Journal of Innovation & Knowledge, Volume 4, Issue 3, 139-145.
4. Kotler, P.T. and Armstrong, G. (2006), *Principles of Marketing ,Published by Pearson Prentice Hall, 11th Edition*, p.p 768.
5. Holden, R.K., Burton, M. R. (1965), *Pricing with confidence: 10 ways to stop leaving money on the table*, Volume 1360.
6. Pattnaik, M. (2014). *Fuzzy Economic Order Quantity (FEOQ) Model with Units Lost Due to Deterioration*, Statistics, Optimization and Information Computing Vol 2, No 3, 257-273.
7. Luz, M. and Moklyachuk, M. (2014). *Minimax-robust filtering problem for stochastic sequences with stationary increments and cointegrated sequences. Statistics*, Optimization and Information Computing Vol 2, No 3, 176-199.
8. Muhammad, N., Coolen-Maturi, T. and Coolen, F.P.A. (2018). *Nonparametric predictive inference with parametric copulas for combining bivariate diagnostic tests*, Statistics, Optimization and Information Computing, Vol 6, No 3, 398-408.
9. Arnold, R. and Shadnam, R. (2014). *Innovation Goals in Software Development for Business Applications*. Statistics, Optimization and Information Computing, Vol 2, No 4, 368-383.
10. Shy, O., Stenbacka, R., Zhang, D.H. (2016), *History-based versus Uniform Pricing in Growing and Declining Markets*, International Journal of Industrial Organization , Volume 48, .p 88-117.