Predicting Selling Price of First Time Product for Online Seller using Big Data Analytics

Sukhvinder Singh Deora[1] and Mandeep Kaur[2]

[1]DCSA, Maharshi Dayanand University,Rohtak, [2]Panipat Institute of Engineering and Technology, Panipat

Summary

Customers are increasingly attracted towards different e-commerce websites and applications for the purchase of products significantly. This is the reason the sellers are moving to different internet based services to sell their products online. The growth of customers in this sector has resulted in the use of big data analytics to understand customers' behavior in predicting the demand of items. It uses a complex process of examining large amount of data to uncover hidden patterns in the information. It is established on the basis of finding correlation between various parameters that are recorded, understanding purchase patterns and applying statistical measures on collected data. This paper is a document of the bottom-up strategy used to manage the selling price of a first-time product for maximizing profit while selling it online. It summarizes how existing customers' expectations can be used to increase the sale of product and attract the attention of the new customer for buying the new product.

Keywords:

Price prediction, e-Commerce, big data analytics, customer behavior analysis

1. Introduction

World-wide number of online buyers is on an increase as per reports of *Statistica* 2021 and their growth-rate taking 2014 as the base year is shown in Fig. 1. The global online buyers are increasing at an average of 7.17% since 2014. The Indian e-commerce industry is also growing at an exponential rate and expected to reach 200 billion U.S. dollars market by 2027 [1]. They are buying products, services and goods from websites and mobile catalog applications. As the purchasing behavior of the online customers can be analyzed using big data analytics, manufacturers and retailers have started modifying their selling strategies. Use of analytics, machine learning and artificial intelligence can benefit the business organizations in identifying what is driving sales, reduce overall cost incurred on products, improve service line, gold price forecasting [2], predict stock price [3], US stock market [4] and make more informed decisions.

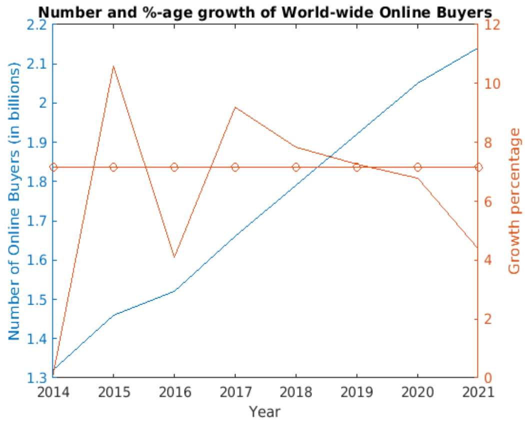

Fig. 1: Number and %-age Growth of Online Buyers in the World

Selling a product from online platforms is a complicated strategic job that is affected by various factors. Earlier, fixing of selling price of item was done on the basis of cost of production, profit to be earned and judgement of seller about its probable cost. Nowadays, setting of selling price is based on users' behavior data along with traditional methods. Analytical methods are used for identifying customers buying choices and influence it for the benefit of the sellers. Due to large number of customers viewing different products, a huge collection of behavioral data records are generated and stored. This data is analyzed using big data analytics, machine learning and conventional statistical measures. Fixing selling price of a new product is trickier due to lack of availability of historical data. This paper is an effort towards developing a novel approach for fixing the selling price of a newly launched item using bottom-up approach. The strategy discussed in the paper was used for selling *Tulsi planters* during the *Navratre* season in 2020.

2. Literature Survey

Price forecasting is a technique of predicting the price of a product/service by evaluating factors [5] like its features, demand, seasonal trends, festival offer, likelihood to purchase [6], demand forecasting, price optimization[7],

Manuscript received February 5, 2021
Manuscript revised February 20, 2021
https://doi.org/10.22937/IJCSNS.2021.21.2.22

competitive product's price and optimal time to buy. Earlier, the initial price fixation was done on the basis of its cost of production, marked price over the cost incurred, competitors' price bracket and merchant's judgement [8,9]. A good prediction model is one that reacts to the change in daily demand and update accordingly [10]. Marginal improvement in fixation of price can be achieved by using regression technique which estimates the relationship between some dependent variable and independent variable(s). A detailed overview of price fixation method for retailers selling products through online portals has been suggested [11]. Use of pricing decision support system that fixes the price of product automatically was also proposed [12] for fashion retailers. Price based methods of managing revenue, promotion and dynamic price fixation [11,13]. Our proposed strategy uses a novel approach to fix the price of first time product and then optimize it using the regression analysis. Various big data analytics tools for predicting prices [14] were proposed and prediction [15]. The problem was solved using a data analytics and relationship of variables with the price prediction.

3. Proposed Methodology

The proposed price prediction strategy is a pragmatic big data analytics approach to sell *Holy Tulsi Planter (name changed)* using *Quicksell* mobile application during the Navratre festivals in 2020. Key features of the product are that it is a polyester based, white, 10 inch by 10 inch, unbreakable planter as shown in Fig. 2. It was to be launched during days when *Hindus* install *Tulsi* herbal plant at their homes with religious sentiments and pray. The aim of the project was to sell all 1500 pieces manufactured by the company with maximization of profit within the festival week.

3.1 Product Intro

In the bottom-up strategic approach, the target was the conversion of total visits on the product page into sales. For phase 1, the initial visuals of the product were shared using Quicksell, an interactive Google integrated micro website application with old customers of general planters. Initially, the interested customers were identified from the existing database of customers that bought general planters from the seller. It was also used to identify the *peak time (PT)* of the day at which most number of customers were active. The probable customers were told to categorize themselves by recording their choice as *"Interested Customer (IC)"* or "*Not Interested (NI)*" to buy the introduced product by pressing a button. Interested customers were also prompted to record the *expected price (EP)* of the item. However, the behavioral data was recorded for all who visited the catalog irrespective of categorization.

Fig. 2: Image of Tulsi Planter

3.2 Initial Cost

Setting the initial marked price of the *Tulsi Planter* was the most intrusive part of the pricing strategy. The plot diagram of the expected price range predicted by the customers was plotted with the frequency of such customers in the y-axis in Fig. 3. A*verage expected selling price (AVG)* was computed from these expected values, $expSP_i$ as per eq. (1). This value was well above the modest value of seller's judgement price of his product. *Average price of similar products (AP)* offered by other sellers on different platforms was also computed from the *price of similar products,* pSP_j using eq. (2). Initial *price fixation (PF) was done* for selling the product at average of the two averages using eq. (3). It was offered at *peak time* (*pT*) recorded during product intro phase. It was deliberately offered at a price less than the expected price of average customers to the customers so that most of the sales are achieved during flash sale on day 1. After that, the product was available at regular price of the product was shown entire day and sales were recorded.

$$AVG = \frac{\sum_i expSP_i}{n} \quad (1)$$

$$AP = \frac{\sum_j pSP_j}{m} \quad (2)$$

$$ISP = (AVG + SP)/2 \quad (3)$$

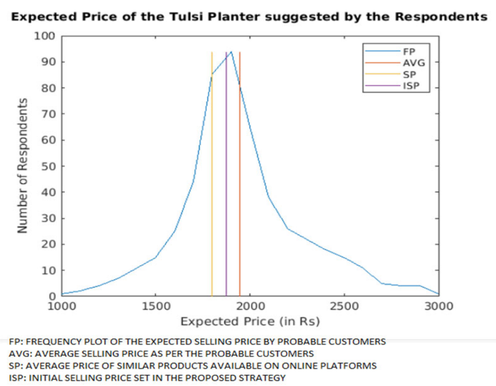

Fig. 3: Expected Price of Product Predicted by Customers

3.3 Flash Sale

In phase 2, flash sale option of 50% of the total inventory of planters was provided for period of *minimum {1 hr, stockForDay >=0}* on the first day with *saleStartTime(SST) = peak_time(pT)* with MRP on product such that selling price, $SP_{day1} = ISP$ after 30% reduction on MRP. After that, the product was provided at the marked price for the rest of the day for the lazy buyers. The *reaction_time* taken by the *Interested Customers* to order the product and the total number of planters sold out was recorded. All relevant behavioral attributes (as mentioned in Table 1) of all the visitors were recorded for applying analytics.

The predicted price for the next day is evaluated using regression trees. We also predicted the selling price that must be offered. Then the probability of customers in the predicted price range is evaluated on the basis of sold planters, probable customers within AVG ± σ, proportion of new customers on catalog page. Flash sale of 50% of the remaining stock was offered to all the customers on discounted price on 2^{nd} day. For the rest of the day, the planter was offered on full marked price again. This process is repeated for all the days of the week to get the maximum profit.

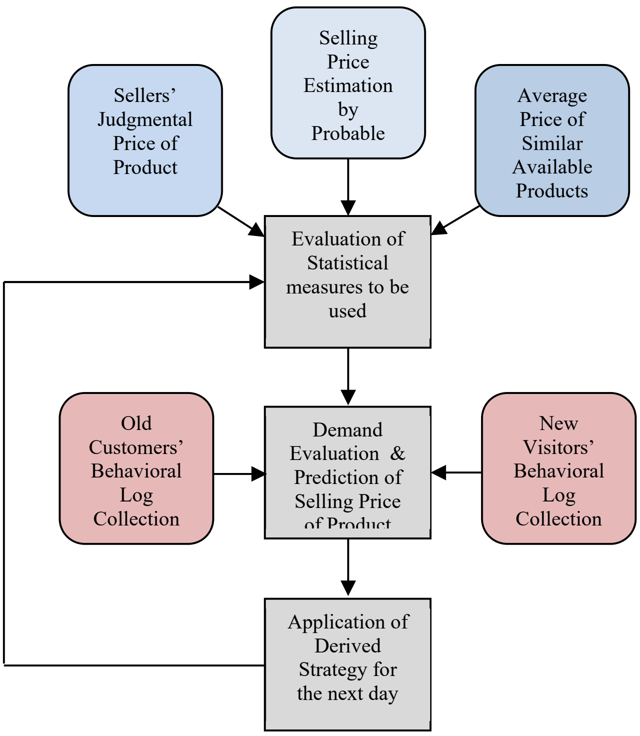

Fig. 4: Iterative Selling Price Prediction Model

3.4 Regression Model used

The proposed iterative strategy (see Fig. 4) evaluates the demand of product and predicts its selling price using non-parametric Classification and Regression Trees (CART) that can handle multiple attributes as shown in Table 1 to predict the selling price. It uses Gini diversity index (see Eq. (4)) for splitting, ensemble technique and multiple tree structures to find the best suitable selling price and probable number of customers for the next day. It identifies selling price that shift marginally close *Not Interested* customers to buyers, customers who predicted low price to buy at offered price and new users to potential customers.

$$Gini_X = 1 - \sum_{\forall xi \in Xi} P(Xi)^2 \qquad (4)$$

Table 1: Parameters considered for Predictive analysis

Variable (X_i)	Description
price offered, MRP	The MRP of the product
discount, disc%	%-age of value deducted from the MRP
discount value, discVal	Monetary value of deduction in price
free delivery condition, delivery	Availability of free delivery option
time spent on the product page, ts_p	Time in seconds spent by the possible customer
reading of customer reviews, $tsRev_p$	Time in seconds spent on reviews reading
%-age of positive reviews, posRev	Proportion of 3-, 4- and 5-star reviews
flash sale time, SST	Time for flash sale offer on the product
concurrent events similar product on platform, #CCE	Whether any other sale options are available
conversion rate of each category of customers, CR_{ec}	%-age of customers of each category who bought the product
proportion of new customers on catalog page, pNC	Records the sharing by the customers to their links and seller side promotion
conversion rate of new customers, CR_{nc}	%-age of new customers who bought the product
time spent by new customers on product page, $tsRev_{nc}$	In seconds
price to order ratio, p2OR	Success rate of the strategy

4. Experimental Results

The initial data analytics was performed using R Studio to predict the selling price on the basis of online behavior of customers. The classifier performance improved with the increase of customers' behavior data added each passing day. Day-wise prediction of price to order ratio is provided in Table 2 and sale achieved in each category is shown in Table 3. The strategy of offering right number of pieces for sale and predicting price as per the customer behavior led to sale of entire stock on the fifth day as per the plan. Fig. 5 shows the hourly percentage of stock sold out from the total offered on the day.

Table 2: Performance metrics of the strategy

Day	%-age sale conversion of Interested customers	%-age sale conversion of not interested customers	%-age sale conversion of new customers	Total Sale (in Rs '000s)
1	0.34794087	0.03123681	0.0923913	2.488
2	0.22591093	0.02721365	0.12033195	5.072
3	0.10774059	0.02032988	0.10774411	10.3670213
4	0.05509965	0.00815892	0.07022472	20.8315789
5	0.02853598	0.00863447	0.07216495	21.7282609

5. Discussion

The entire process was planned in such a way that the customers who take quick decision will feel that they have got benefit of discounts and the lazy customers will feel that they have been penalized for being slow in decision making during flash sale hour. The price was kept near the average expected value to attract customers' attention during the entire process. Selling price was slowly increased on the basis of dynamic demand prediction for the next day. The strategy was able to convert probable customers (whose expected selling price was marginally less than the offered price) into customers who ordered the product. The psychology of reward and punishment has made customers place order for the product on day-to-day basis. It was helpful in maintaining the daily demand (refer Fig. 5) and meeting it by the end of the day by using AI in providing initial discounts in flash sale, analyzing the customers' behavior and increasing the price slowly to the MRP. It resulted in increase of overall revenue expected by the seller by more than 21.5% and average price by 17.6%. The strategic pricing played a key role in converting new visitors into customers during flash sale on next day due to price drop. Moreover, the proportion of number of orders received from new customers was maintained at 0.10 of the total new visitors of the day. Such customers prefer to visit the next day for flash sale and order the product.

Table 3: Sale achieved in each category of customer with discounted selling price of the day

DN	DSP	TOR	TNIC	IC_OR	TNNIC	NI_OR	TNNC	NC_OR
1	Mask1	750	1894	659	2369	74	184	17
2	Mask2	375	1235	279	2462	67	241	29
3	Mask3	188	956	103	2607	53	297	32
4	Mask4	95	853	47	2819	23	356	25
5	Mask5	92	806	23	3127	27	582	42

Notations Used:

DN: Day Number
DSP: Discounted Selling Price
All the prices have been masked due to privacy of the information.
TOR: Total orders received
TNIC: Total Number of Interested customers
TNNIC: Total Number of Not Interested customers
TNNC: Total new customers of day
IC_OR: orders received from customers categorized as "Interested Customer"
NI_OR: orders received from customers categorized as "Not Interested"
NC_OR: orders received from new customers that visited the product page

6. Conclusion

The proposed bottom-up strategy for predicting the selling price of the first time product and increasing the revenue worked out well. It used customers' behavioral analysis to speculate the daily demand to predict/correct price to make the customers believe that they have got the best deal according to their promptness or lazy action. The overall sale increased by this dynamic pricing approach and playing with the mindset of customers rather than using traditional approach of offering fixed discounts.

In nutshell, the results show that the customers' initial decision on purchasing a product on his expected price can be influenced by using dynamic price fixation, flash sale, rewarding early decision makers with higher discounts and punishing slow decision makers with less discounts using big data analytics on their behavioral logs. Changes in demand in non-festival period can be further taken up as a future scope of investigations of selling similar products.

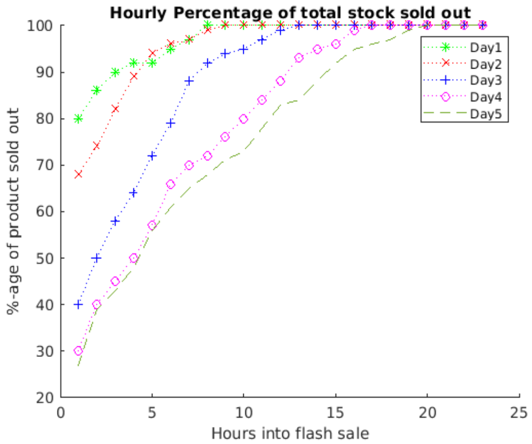

Fig. 5: Hourly Statistics of Percentage of stock sold out. products.

References

[1] https://www.statista.com/statistics/792047/india-e-commerce-market-size/ accessed in Septeember, 2020.

[2] N. D. G. Vadivu, "Big Data Analytics for Gold Price Forecasting Based on Decision Tree Algorithm and Support Vector Regression (SVR)," Int. J. Sci. Res., vol. 4, no. 3, pp. 2026–2030, 2015, [Online]. Available: https://www.ijsr.net/archive/v4i3/SUB152560.pdf.

[3] S. Jeon, B. Hong, J. Kim, and H. J. Lee, "Stock price prediction based on stock big data and pattern graph analysis," IoTBD 2016 - Proc. Int. Conf. Internet Things Big Data, no. IoTBD, pp. 223–231, 2016, doi: 10.5220/0005876102230231.

[4] Z. Peng, "Stocks Analysis and Prediction Using Big Data Analytics," Proc. - 2019 Int. Conf. Intell. Transp. Big Data Smart City, ICITBS 2019, pp. 309–312, 2019, doi: 10.1109/ICITBS.2019.00081.

[5] M. Price and S. Challenge, "Price Prediction using Machine Learning Regression — a case study," 2021.

[6] P. Y. Zhao and Y. M. Shi, "Predicting the likelihood of purchase by big data," AIP Conf. Proc., vol. 1834, no. April, 2017, doi: 10.1063/1.4981598.

[7] K. J. Ferreira, B. H. A. Lee, and D. Simchi-Levi, "Analytics for an online retailer: Demand forecasting and price optimization," Manuf. Serv. Oper. Manag., vol. 18, no. 1, pp. 69–88, 2016, doi: 10.1287/msom.2015.0561.

[8] Subrahmanyan, Saroja, "Using quantitative models for setting retail prices," Journal of Product & Brand Management 9 (2000): 304-320.

[9] Michael Levy, Dhruv Grewal, Praveen K. Kopalle, James D. Hess, Emerging trends in retail pricing practice: implications for research, Journal of Retailing, Volume 80, Issue 3, 2004, Pages xiii-xxi, ISSN 0022-4359

[10] Christopher, M., "Logistics and Supply Chain Management," Pearson UK, 2016.

[11] Ozer, O., R. Philips, The Oxford Handbook of Pricing Management, Oxford University Press, 2012.

[12] Caro, F., J Gallien, "Clearance Pricing Optimization for a Fast-Fashion Retailer," Operations Research 60(6), 1404-1422.

[13] Talluri, KT, GJ Van Ryzin, "The Theory and Practice of Revenue Management," Springer, 2005.

[14] M. Seyedan and F. Mafakheri, "Predictive big data analytics for supply chain demand forecasting: methods, applications, and research opportunities," J. Big Data, vol. 7, no. 1, 2020, doi: 10.1186/s40537-020-00329-2.

[15] V. Sandhiya, T. Revathi, A. Jayashree, A. Ramya, and S. Sivasankari, "Stock Market Prediction on Bigdata Using Machine Learning Algorithm," vol. 7, no. 4, pp. 10057–10059, 2017.

Sukhvinder Singh Deora received the MSc(Mathematics),& M.C.A. from Kurukshetra University in 2000 and 2002. He did his M.Phil. in Computer Science and completed PhD in 2015. He is currently working as Assistant Professor in Department of Computer Sciences, Maharshi Dayanand University, Rohtak, India. He is a reviewer to many prestigious International Journals and member of editorial board of some Journals. To his credit are many prominent papers in the area of data security, big data analytics and issues related to it. He has also edited Proceedings of National Level Seminars/Conferences. With an exposure of 1.5 years in IT industry his thrust areas include Testing, Java technologies and Database design issues. His interest areas include Big Data Analytics, Network Security, Theoretical Computer Sciences, Data Structures, S/W Testing and Database Designing. He is also a member of ACM, Computer Society of India (CSI) and Indian Society of Information Theory and Applications (ISITA).

Mandeep Kaur is BCom, MCA, MTech (CSE) and MPhil in Computer Science. She is working as Assistant Professor in Computer Applications of Panipat Institute of Engineering and Technology, Samalkha, Panipat, India. She is keenly involved in teaching Cloud Computing, e-Commerce, Analytics for Business. She has written research articles involving latest technologies like IoT, Big Data Analytics, Data Protection, Security and Cloud Computing. She is also a member of ACM-W, Computer Society of India (CSI) and Indian Society of Information Theory and Applications (ISITA).