

Creating product values for a Re-Usable straw

MD.Tanvir Reza Anik

Dept. of Computer Science and Engineering
East West University
Dhaka, Bangladesh
2016-3-60-027@std.ewubd.edu

Mushfique Yasir

Dept. of Computer Science and Engineering
East West University
Dhaka, Bangladesh
2015-3-60-024@std.ewubd.edu

Sadia Ahmed

Dept. of Computer Science and Engineering
East West University
Dhaka, Bangladesh
2017-1-60-019@std.ewubd.edu

Shahariar Bhuiyan

Dept. of Computer Science and Engineering
East West University
Dhaka, Bangladesh
2015-3-60-006@std.ewubd.edu

Abstract—A product value, which is calculated by classified measurements, can come real handy to a business organization, as they would know where, and at what price range, to whom a specific new product launch can be the most beneficial, before even the trial of the product. Keeping that in mind, this research was set out to find the product value of a reusable straw, which can really affect the environment in a good way. A fixed dataset built from a survey using the internet works as the primary source. After a small research on the fields of machine learning prediction, a Bayesian classifier model designed by a c++ program predicts the optimum values of the attributes pre-defined by the researchers.

Index Terms—Classifier, Bayesian, Product Value, optimum, Survey.

I. INTRODUCTION

The term product value is concerned with how much the customers think the worth of a product or a service or a consumable is to their eyes. The developers set the price of their product depending on how the customers are giving value to them, and they always try to produce the higher value product. In that case, customers would take even that high price as a fair price, because they value them. It happens worldwide, and requires to be solved as an important problem of marketing the products. So, our project was to work on finding the values of different categories of products out there! In the end, one would be able to find out how the pricing should be, what are the areas the product will be most sold, what level of people from the society will be consuming the most of the products, and what are more valuable for them than the current state of the product.

II. RELATED WORK

Recommending in context of a user demand, which in other words, personalizing their space is a very good way to keep close relationships and raising loyalty with companies by offering differentiated service based on personalization

function. Being able to relate to this function is the primary motivation of so many papers.

We can look up on some as one of the papers [1] main objective of the paper was to create a web based application system running on a PC to recommend music to users, with a two way communication between them and the system. The researchers looked up on some collaborative companies who shared their product taxonomy table. The procedure uses a multi-attribute decision making to find the utility values of products in same product class of the companies. Based on that value, the system determines which are similar.

The researchers on this paper [2] made some statements about product value, and defined them, before heading into their core methodology. Firstly, it uses cluster analysis and supervised machine learning. Having to cope with high complexity and high dimensionality, conventional methods based on modelling cause effect relations are suspended. Being able to handle large sets of highly complex and high-dimensional data, a combination of Cluster Analysis and SVM is introduced as a possible way to achieve the goal of improved quality monitoring.

[3] suggested a method for extracting consumer purchasing behavior. Utilizing RFID data acquired from individuals in a supermarket, they examined several important methodological issues related to the use of RFID data in support vector machines (SVM's) to predict purchasing behavior. In this paper, they mention that they measure customers' behavior based on age, gender and income, product price, and sale promotion.

another study [4] demonstrates how machine learning can be used to increase our understanding of what customer's value in the engagement-to-loyalty value chain. It also adds to the theory by proposing a framework for measuring and analyzing dynamic customer engagement. This provides methodology for adjusting offers and promotions of products and services to influence customer engagement behaviors in real time, which bridges research gaps on the application of machine learning

in the hospitality and tourism literature. Demonstrated how hospitality venues can co-create value by offering personalized discounts that influence loyalty.

The researchers of this paper [5] used linear regression which is a machine learning algorithm. By using linear regression, they got more accurate predictions of the prices. They build a function to get the results. Then they showed the estimated value which is close to the real values. For coefficient estimates, there are some standard errors. the reported value is p and the three predictors were used to create the y data set. Then the linear regression algorithm was modified.

So, the aim of our research would be to predict the target values of products, and the objectives are to optimize the product values as much as possible with maximum accuracy possible.

III. METHODOLOGY

We made a research with the goal of indicating a proper product value to a product which are selected with a bias, and using real people's opinion about those products. For the small project like of a 3 credit course, this was done in the shortest way possible to include as much simulation as possible within only one product, which is assumed to be a new product that a company tries to put in the market. Participants info was the most important in this case, so we gathered some biometric information such as their age, gender, and a geographic info like their place of residence, and the last factor was the religion. The collected data was quantitative data instead of qualitative. Among the other possible data collection method to this cause could have been observing the space of the internet, and it was doable too, but, for the sake of a small project, we are going for a small area, and the products popular in there. So, we conducted a survey, with a fixed set of questionnaire, and their answers was converted into a mathematical value by calculating in a fixed scale, which was decided once the data were gathered.

We looked for people, and that is said to be at least 30, to 50, and they were from a specific area of the country. They were given a google form with our questionnaire attached to it, and they anonymously filled them up. a total of 49 data was gathered.

Once we gathered all the answers filtered from the form, the data was ready for analysis, which, could have been done in many different ways. For analysing quantitative data, we will be using software, instead of figuring and manipulating the values on our human head. Tool used, was Atlas.ti, and R, or both of them, and they are open source software to be found on the internet. The analysis phase will be the shortest, but the most important one, as it will almost visualize the outcome.

The presentation part is the final output. The final data which was found after analysis, were structured in accordance of the research questions, and summaries of each participant answers relative values. Once the data set is presented into a proper formation, the second phase of the research began. We applied proper algorithms based on the background studies previously made from previously completed product value

related projects, and determined the values of our selected products on the selected area, which, the companies can look on to pass in their next production season. our chooses algorithm was the bayesian classifier.

In statistics, Naive Bayes classifiers are a family of simple "probabilistic classifiers" based on applying Bayes' theorem with strong independence assumptions between the features. They are among the simplest Bayesian network models, but coupled with kernel density estimation, they can achieve higher accuracy levels. Bayes' Theorem finds the probability of an event occurring given the probability of another event that has already occurred. Bayes' theorem is stated mathematically as the following equation:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

where A and B are events and $P(B) \neq 0$.

Basically, we are trying to find probability of event A, assuming A is a customers willingness to pay for the product, given the event B is true, where B is assumed as his or her occupation. Event B is also termed as evidence. $P(A)$ is the priori of A (the prior probability, i.e. Probability of event before evidence is seen). The evidence is an attribute value of an unknown instance (here, it is event B). $P(A|B)$ is a posteriori probability of B, i.e. probability of event after evidence is seen.

Now, with regards to our dataset, we can apply Bayes' theorem in following way:

$$P(y|X) = \frac{P(X|y)P(y)}{P(X)}$$

where, y is class variable and X is a dependent feature vector (of size n) where:

$$X = (x_1, x_2, x_3, \dots, x_n)$$

Now, its time to put a naive assumption to the Bayes' theorem, which is, independence among the features. So now, we split evidence into the independent parts.

Now, if any two events A and B are independent, then,

$$P(A,B) = P(A)P(B)$$

Hence, we reach to the result:

$$P(y|x_1, \dots, x_n) = \frac{P(x_1|y)P(x_2|y)\dots P(x_n|y)P(y)}{P(x_1)P(x_2)\dots P(x_n)}$$

which can be expressed as:

$$P(y|x_1, \dots, x_n) = \frac{P(y) \prod_{i=1}^n P(x_i|y)}{P(x_1)P(x_2)\dots P(x_n)}$$

Now, as the denominator remains constant for a given input, we can remove that term:

$$P(y|x_1, \dots, x_n) \propto P(y) \prod_{i=1}^n P(x_i|y)$$

Now, we need to create a classifier model. For this, we find the probability of given set of inputs for all possible values of the class variable y and pick up the output with maximum probability. This can be expressed mathematically as:

$$y = \operatorname{argmax}_y P(y) \prod_{i=1}^n P(x_i|y)$$

So, finally, we are left with the task of calculating $P(y)$ and $P(x_i|y)$ for each individual calculations of the root equation of the bayasian classifier formula.

IV. RESULTS AND DISCUSSION

Our main objective was to find the target attributes of all the selected criterias a customer can be put in, which was previously mentioned as: gender, age, occupation, area, religion and preferred price range. As per the necessity, we ran the Naive Bayes classifier algorithm with an object oriented program, in this case specifically, C++. The algorithm used in the code demonstrate the bayasioan classifier, and makes a prediction.

According to our dataset, the optimal attributes were discovered with an execution time of 0.342 seconds with a gcc compiler. The target attrributes, or the optimal attributes for the product is as below:

Target Gender : Male
Target Age Range: 18-24 years
Target Occupation: Studying
Target Religion: Islam
Target Area: City
Optimal Price Range: 21-30 Taka
Consumer Rate: 77

Bayasian classifier predicted them, and the results describe that, if the new products price range is between 21-30 TAKA, most customers would find the product helpful for them, and 77 per cent of the people would find it reasonable. it also says, the product would be valued properly if it is deployed in a city, where the majority of people are religiously Muslim, the male students aged 18-24 living there, should the target of advertisement of the product.

V. CONCLUSION

Product values are of upmost importance given the current era of competition in business sectors all over the world. The aim of this research was to predict the target values of a new product in the market, which was happened to be a reusable straw. We made a survey using google form among our networks of people, and asked them a few questions about how would they like to consume it, if at all. Their responses were recorded, which was a number of 49 people from different ages. The data taken from them are used as our primary dataset to run a bayasian classifier algorithm on a c++ program to determine the optimum target values of the product in relation to the survey. the output shows a specific class of people from a specific type of residential area would

be more beneficial to launch at a specific price range, which would attract most number of consumers.

VI. APPENDIX

A. questionnaire

1. What is your age? *
 1. Under 12 years old
 2. 12-17 years
 3. 18-24 years
 4. 25-34 years
 5. 35 or older
 2. What is your gender? *
 1. Male
 2. Female
 3. Non-binary/third gender
 4. I prefer not to answer
 5. Other
 3. What's your work status?
 1. Employed
 2. Self-employed/Freelance
 3. Unemployed
 4. Studying
 5. Retired
 6. Other
 4. What is your religion? *
 1. Islam
 2. Hinduism
 3. Christianity
 4. Buddhism
 5. Other
 5. Where Do you live? *
 1. Village
 2. Town
 3. City
- Reusable Straws:
- The plastic straw, an item that you probably didn't think twice about using a few years ago, has become a source of controversy as concern over pollution from single-use plastics has grown. The plastic straw is unhealthy for human body so, we should minimize the number of using the plastic straw and replace them with ecofriendly Reusable Straws. These reusable straws are lightweight and have a smooth mouthfeel, plus it's easier to see if they're dirty. This Glass Straws, are handmade using durable borosilicate glass, which is more shatter-resistant than regular glass.
6. Do you want this product? *
 1. Yes
 2. No
 7. What do you think the price would be for this product? *

1. 10-20 Taka
2. 21-30 Taka
3. 1-40 Taka
4. 41-50 Taka

B. Informed Consent Form

Informed Consent form:

Why you participate in our research study? The purpose of our research study is to creating product value. Which will help any company to learn their targeted customer, their production rate and optimal pricing for a product.

Researchers:

1. Mushfique Yeasir (2015-3-60-024@std.ewubd.edu)
2. Shahariar Bhuiyan (2015-3-60-006@std.ewubd.edu)
3. MD. Tanvir Reza Anik (2016-3-60-027@std.ewubd.edu)
4. Sadia Ahmed (2017-1-60-019@std.ewubd.edu)

Department of Computer Science and Engineering, East West University

We are asking your independent participation in our research which is Creating product value. Please read the information and if you voluntarily agree to participate then just simply press Yes then we will ask you some question regarding this research. And if you don't want to participate in this survey then press No.

Purpose of this study: This study will help us creating product value algorithm which will help any company to to learn about their product and targeted customer.

Participating in the study you will be asked: If the participant voluntarily agree to participate then the participant will be asked some questions about this topic. The participant has to answer those question through their own experience.

The required time needed to fill up the survey: It takes 5-10 minutes to fill up.

Benefits: The outcome of our study will help any company who want to launch new product in to market. they can create their product value.

Risk: There is no risk for participants to participate in this study.

Voluntary participation: This is voluntary participation. The participant wants to participate in this survey or not its totally up to them. At any point of this survey the participant can back off from this survey. The provided data won't store until the participant clicks the submit button.

- [3] Y. Zuo, K. Yada, and A. S. Ali, "Prediction of consumer purchasing in a grocery store using machine learning techniques," in *2016 3rd Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE)*, pp. 18–25, IEEE, 2016.
- [4] A. Aluri, B. S. Price, and N. H. McIntyre, "Using machine learning to cocreate value through dynamic customer engagement in a brand loyalty program," *Journal of Hospitality & Tourism Research*, vol. 43, no. 1, pp. 78–100, 2019.
- [5] J. An, A. Mikhaylov, and N. Moiseev, "Oil price predictors: machine learning approach," *International Journal of Energy Economics and Policy*, vol. 9, no. 5, p. 1, 2019.

REFERENCES

- [1] S. H. Choi, S. Kang, and Y. J. Jeon, "Personalized recommendation system based on product specification values," *Expert Systems with Applications*, vol. 31, no. 3, pp. 607–616, 2006.
- [2] T. Wuest, C. Irgens, and K.-D. Thoben, "An approach to monitoring quality in manufacturing using supervised machine learning on product state data," *Journal of Intelligent Manufacturing*, vol. 25, no. 5, pp. 1167–1180, 2014.