

Global Plastic Waste Model based on Linear Programming and Support Vector Regression

Summary

Plastic waste (PW) has been one of intractable environmental dilemma for modern humankind. In order to mitigate this tough problem, we combine production of plastics and management of plastic waste to establish a model, which can evaluate the impacts of plastic waste and find out an achievable minimal level of plastic production that can be reduced with a time line.

Firstly, we develop a plastic waste estimate model based on Product Life Cycle (PLC) theory to calculate the maximal levels of single-use or disposable plastic product waste within the environmental capacity in a certain region or country. Then we independently explore the four possible ends of PW and formulate their impacts on environment by several functions. We use linear programming that includes one particular objective function and four constraints. As a result, we can output the maximal amount of plastic products by analyzing the input digits of any region.

Secondly, to find out to what extent plastic waste can be minimized, we established a model called HSVR which combines happiness-index analysis and SVR method. By analyzing the characteristics of different regions and people's living standards, we find out the relationship between the amount of plastics produced and people's happiness index. Based on this, we can analyze how much plastic production we can reduce without significantly reducing happiness.

Thirdly, we follow the previous happiness index to assess how much human life is changed, and apply the Environmental Assessment of Solid Waste Systems and Technologies (EASE-WASTE) model to evaluating how the the environment is affected by plastic reduction. Moreover, we take computable general equilibrium (CGE) model to quantify the effects of the target plastic production. Then we normalized the outputs and into non-dimensional indices, and calculate the weight for each indicator above by the analytic hierarchy process (AHP). The total impact for achieving the confined level is a linear combination of the three indices.

Furthermore, reduction of plastic production will impact countries unequally because of their different development level. We use dual programming to fine the dual model of the linear programming based model we established first. Then shadow prices of every factor formulated by constraints can be calculated, which indicate the significance of each factors. Then we divided them into environmental and governmental factors. Sensitivity analysis is be used to identify which factors that is most sensitive to a particular country, and the equitable reduction of plastic of the country depends on the categories that its sensitive factors belong to.

Finally, we analyze how to approach the minimum level following the time line by an achievable way. Many unexpected circumstance would definitely appear so we empirically choose most significant possibilities that may delay or accelerate the achievement process. The ultimate result are specified into a memo which will be provided for ICM.

Keywords: Plastic waste; LCA; SVR; EASEWASTE; AHP

Contents

1	Introduction	1
1.1	Background	1
1.2	Problem statement	1
2	Model Analysis	1
2.1	Analysis of task 1	1
2.2	Analysis of task 2	1
2.2.1	Rating-review based reputation model	1
2.2.2	Relations between time and reputation	2
2.2.3	Indication of potentially successful or failing product	2
3	Assumptions and justifications	3
4	Parameter Table	3
5	Model Construction	4
5.1	Rating-Review Based Reputation Model	4
5.2	Reputation-Time model Based on RBF Regression	6
5.3	Model to Indicate Potentially Successful or Failing product	7
6	Strength and weakness	8
6.1	Strengths	8
6.2	Weaknesses	9
7	Conclusion	9

1 Introduction

1.1 Background

For e-commerce platforms such as Amazon, there is no shortage of the customer's ratings and reviews. Ratings are usually expressed in five grades of 1 to 5 stars, while reviews are expressed in text form, giving customers the opportunity to show their feedback. But how to make good use of this huge rating and reviewing data is also a big challenge. Taking good advantage of these data can not only help companies to provide targeted customer service, find potential high sales and low sales goods but also help them adjust operational strategies, or enhance products in a timely manner.[5]

1.2 Problem statement

Now, Sunshine Company plans to launch three new products to sell online: microwave, pacifier and hair dryer. The company has obtained historical ratings and comments from customers of these three products. They hope to dig out effective information from reviews and ratings to help companies better formulate sales strategies and enhance the effectiveness of products. They hope that the following problems can be reasonably solved:

1. What is the relationship between product star rating, review and help votes? Can we quantify the reviews and explore the mathematical relationship among them?
2. Based on task1 to address the following specific requests from the Sunshine Company Marketing Director:
 - (a) How to build a model that can integrate ratings and reviews, and help company get effective information and adjust marketing strategies?
 - (b) How to describe the reputation of the product from the perspective of time, so as to see and predict the change of the product's reputation in the online market?
 - (c) Based on text evaluation and rating, how to find potential successful or failed products?
 - (d) Is there a relationship between a specific star rating and the number of reviews? In other words, what kind of ratings attract customers to comment?
 - (e) Are subject words of a review, such as "happy", "disappointed" related to specific ratings?

2 Model Analysis

2.1 Analysis of task 1

2.2 Analysis of task 2

2.2.1 Rating-review based reputation model

We can infer the reputation of a product from its star rating and reviews. The better the reputation of a product, the higher the potential sales volume of the product.

Here, we can't simply measure the reputation of a product by star rating or evaluation alone, because there are often situations of "five star with negative feedback" or "one star with praise".

Therefore, we integrate star ratings and reviews to build a reputation model. By comparing the reputation of products, Sunshine Company can enhance the stock and publicity of products with high reputation.

In similar way, the reputation calculation needs to first carry out LDA processing on the reviews of a product, and then convert them into discrete level values.

In order to consider the credibility of the reviews, we also consider several fields such as vine, verified_purchase, helpful_votes, and total_votes, which can measure the accuracy of the review, because reviews with product experience are more credible, and reviews of high helpful votes are more referential.

By modeling and calculating the pacifier samples with more than 25 comments from July to August 2015, we find that in July to August, the reputation of sample 'wubbanub soft toy and pacifier' is higher and the sales prospect is the largest; the reputation of 'Philips aventbpa free another pacifier, 3 + months, green, 2 count' is the lowest and the sales prospect is the worst.

2.2.2 Relations between time and reputation

The reputation of products will fluctuate with the change of time, and the latent regulation may have great inspiration for sunshine company. We can use RBF neural network[1] to fit (reputation time) function. After getting the fitting curve, we can directly observe the change trend of reputation. Take the pacifier for example,

- First of all, we select the sample with the most reviews in the pacifier--'philips avent bpa free soothie pacifier, 0-3 months, 2 pack' with 833 reviews
- Then we use the reputation transformation model to get the reputation of all samples
- Then, taking one day as the unit time, RBF neural network is used to fit the reputation and time
- Finally, according to the fitting curve, we find that the reputation of Philips pacifier has a minimum value on May 30, 2015 and a maximum value on May 10, 2015. In the next question, the causes of the maxima and minima are analyzed.

2.2.3 Indication of potentially successful or failing product

As mentioned above, reputation will fluctuate with time, so the curve we fit will naturally have inflection point. Therefore, we define that if the curve will usher in a maximum value, the product will become a potential successful product; if the curve will usher in a minimum value, the product will become a potential failure product.

In order to facilitate processing, we only consider a set of maxima and minima of the curve when modeling. We can cluster the stars and reviews near the maxima and minima of the curve to find the topic words and stars with the highest word frequency. When these themes and stars appear, it means that the reputation of the product will probably usher in a turning point.

In our experiment, it is found that when 'baby cant' and 2-3 stars appear, the reputation of the product will decline, so the product will become a potential failure product; when 'baby love' and 4-5 stars appear, the reputation of the product will rise, and the product will become a potential success product.

3 Assumptions and justifications

To simplify the problem, we strive for the following assumptions:

- **The mass of annual plastic production equals to that of the annual plastic waste.** On the one hand, every pound of plastic produced will turn into waste. On the other hand, the lifetime of most sorts of plastic is under 20 so that given situation will not change acutely[?].
- **Throughout the life cycle of plastic, the impact of plastic on environment in use phase is negligible.** Main environment issues that are related to plastic merely when it is produced or managed[?].
- **Aiming to make plastic waste safely be mitigated, any waste should not be land-filled.** While it is extremely hard for plastic to biodegrade in solid, landfills are in great shortage.[?].
- **All the plastic waste produced will be managed.** In the long run, the NET plastic waste discarded in natural environment will approach zero by governments and NGOs' efforts.
- **The proper quantity of pollutant emission of plastic waste should depend on the contribution of plastic industry to the gross domestic products.** The pollutant emission of plastic waste is just an ordinary part of the total emission, so the responsibility of environmental protection it takes should also be the proportional part, which is roughly estimated by the percentage of GDP that plastic industry contributes.
- **All the plastic waste is mechanically sorted from plastic waste prior to incineration.** In fact, waste incineration plants are generally very formal and will complete the sorting work. Otherwise, they will be banned soon.
- **The material recycling facilities(MRFs) that sort and upgrade the received waste stream is the only way to recycled sorted PW.** There is no other mature processing method, and this method is used almost everywhere.

4 Parameter Table

Important symbols that in this article are listed in Table 1.

Table 1: Notations

Symbols	Definition
s_i	Annual certain plastic production.
N_i	The number of CO_2 generated by the combustion of 1 molecule specific plastic unit.
EA	Emission to air.
EW	Emission to water.
EC	Emission of CO_2 .
TEA	Total emission to air.
TEW	Total emission to water.
TEC	Total emission of CO_2 .
CEA	Compensatory emission to air of recovery process.
CEW	Compensatory emission to water of recovery process.
SM	Plastic waste into marine.
M	The quantity of annual management of marine plastic.
V_{air}	Total available air of a specific country/region.
V_{water}	Total available water of a specific country/region.
σ	Proportion of incinerated plastic waste.
μ	Compensate.
θ	The percentage of GDP plastic industry contributes, quantified by coefficient.
ι	Tossil fuel consumption percentage. Generally, ι equals 4%.
k	Proportional coefficient of carbon emission per unit of energy consumption

5 Model Construction

5.1 Rating-Review Based Reputation Model

By modeling ratings and reviews, we can get product reputation, and higher reputation means higher potential sales volume, which can help Sunshine Company make better decisions. The modeling process is as follows:

Assume the star rating of the product is R_α , the review is quantified as R_β , the helpful votes of the review is R_h , so that unhelpful votes(R_u) is total votes minus helpful votes, net helpful votes is($R_h - R_u$), We use net helpful votes plus 1 as a factor and multiply the quantified review to judge the credibility of the evaluation.

We use the logarithmic form to make the intermediate calculation smoother, then average the reputation during specific period, and finally refer to the article[4] we magnified the reputation differences between samples in the form of index, and the reputation model is initially established as follows:

$$F_{repu} = e^{\frac{\sum_{i=1}^n (\log_2(R_\alpha + 1) + \beta \times \log_2(R_\beta \times (R_h - R_u + 1) + 1))}{n}} \quad (1)$$

where n is the number of evaluation in a certain period. For the convenience of processing, when the net helpful votes are negative in the model, it will be set to -1, that is, this kind of comment has no contribution to reputation.

Furthermore, on the one hand, the super parameter β should be able to reflect the weight relationship between rating and comment, on the other hand, it should be able to represent the credibility of the review, and the review with the vine or verified_purchas field equals 'y' is a review that has experienced the product, which is more referential. So the expression of β is as follows:

$$\beta = \begin{cases} 1 + W_r & \text{if } vine \text{ or } verified_prurchase = 'Y' \\ W_r, & \text{else} \end{cases} \quad (2)$$

where W_r is the relationship coefficient of rating and review mentioned above. Set 0.61 here.

In order to test our model, we first counted more than 25 pacifier samples from July to August 2015, which are:

- 'philips avent bpa free soothie pacifier, 0-3 months, 2 pack, packaging may vary', marked as Sample1
- 'wubbanub soft toy and pacifier', marked as Sample2
- 'wubbanub brown monkey pacifier', marked as Sample3
- 'wubbanub infant pacifier - giraffe', marked as Sample4
- 'philips aventbpa free soothie pacifier, 3+ months, green, 2 count', marked as Sample5
- 'wubbanub brown puppy pacifier', marked as Sample6

Then we calculated the average reputation value of six samples in these two months, and the results are as follows:

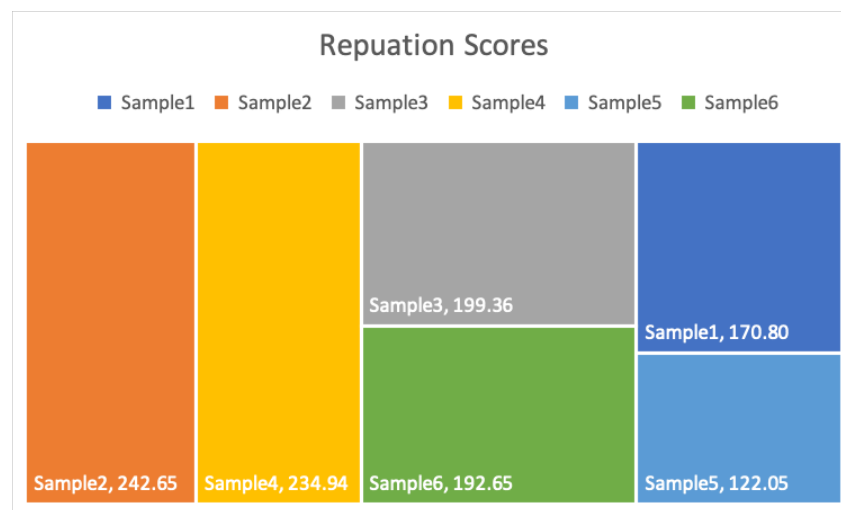


Figure 1: The reputations of six samples calculated from July to August 2015

As shown in the figure:

1. Sample2 and Sample4 belong to the same class, with high product reputation value, which means better sales prospects;
2. Sample3 and Sample6 belong to the same class; Sample1 and Sample5 belong to the same class, with low reputation value, which means that the sales prospects are not clear.
3. In fact, we found that Sample5's rating is indeed full of 1-star and 2-star, and there are many bad comments in the reviews.

5.2 Reputation-Time model Based on RBF Regression

It is helpful for Sunshine Company to adjust its strategy and increase its profit by exploring the changing law of product reputation over time. Inspired by paper[3], we can use neural network to fit and predict reputation.

In this paper, we use RBF neural network as the regression fitting of time based reputation, because RBF has a strong nonlinear fitting ability, can map any complex nonlinear relationship, and the learning rules are simple, which is convenient for computer implementation. The establishment process of our RBF based reputation time model is as follows:

1. We first filter the data and select the most representative samples of each type of product:
 - **Pacifier:** 'philips avent bpa free soothie pacifier, 0-3 months, 2 pack, packaging may vary' with 833 reviews on 557 dates.
 - **Hair Dryer:** 'remington ac2015 t|studio salon collection pearl ceramic hair dryer, deep purple', with 587 reviews on 436 dates.
 - **Microwave:** 'danby 0.7 cu.ft. countertop microwave' with 394 reviews on 224 dates.
2. By substituting the selected samples into the reputation model, we get the discrete sequence of sample reputation changing with time.
3. Building RBF network:
 - (a) Take the first half of the sample as the training set and the second half as the test set, then normalize them.
 - (b) Set the number of neurons in the network to increase gradually, the most is the number of training samples, and set the spread speed to 0.1, start training.
 - (c) The RBF network was tested and the fitting result was obtained.

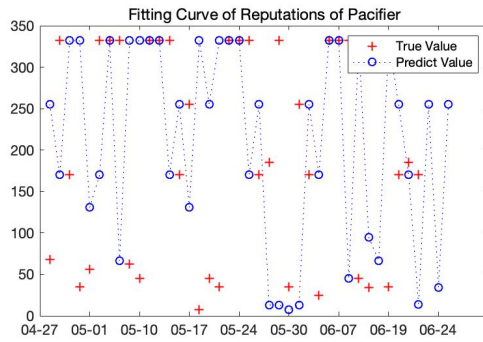


Figure 2: Fitting Curve of Reputations of Pacifier

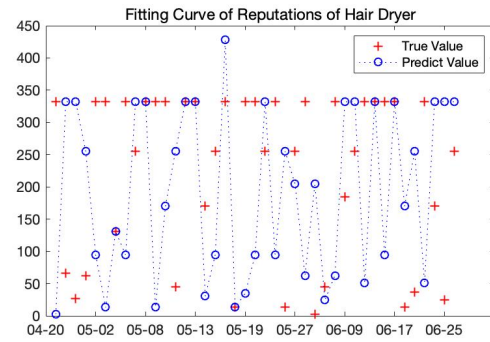


Figure 3: Fitting Curve of Reputations of Hair Dryer

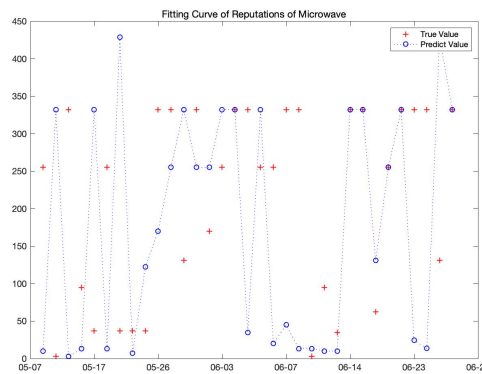


Figure 4: Fitting Curve of Reputations of Microwave

From the figure, we can observe the following phenomena:

- For pacifier, the minimum reputation appears around May 30, 2015, for hair dryer the minimum appears around May 19, 2015 and microwave's minimal reputation was around June 7. The causes for the minimum will be explained in the next question.
- For pacifier, the maximum reputation value appears around May 10, 2015, for hair dryer the maximum appears around June 9, 2015 and microwave's maximal reputation was around May 26. Again, we will discuss the causes for the maximum in the next question.

5.3 Model to Indicate Potentially Successful or Failing product

In our modeling, the theme words and ratings near the maximum value of fitting curve are the signs of potential successful products, while the theme words and ratings near the minimum value of fitting curve are the signs of potential failed products.

Take pacifier for example, according to the Fig.2, The maximum value of fitting curve appears on May 10, 2015 (marked as t_1), The minimum value of fitting curve occurs on May 30, 2015 (marked as t_2), we extract stars and reviews at t_1 and t_2 's 7 nearby time points respectively, We conduct LDA text analysis and clustering, and the results are as follows:



Figure 5: World Cloud for Reviews around May 10, 2015



Figure 6: World Cloud for Reviews around May 30, 2015

From the clustering results, we conclude that:

- If the reviews are full of words like 'baby love', 'easy to hold', 'onto love' with 4-5 star ratings, it means that the reputation of the product will usher in a maximum value and it is a potential successful product. So it's recommended for Sunshine Company to grow the portability of pacifiers.
- If the reviews are full of words like 'baby cant', 'baby not', 'without soothing' with 2-3 star ratings, it means that the reputation of the product will usher in a minimum value and it is a potential failing product. And we recommend Sunshine Company to fix the soothing problem with the specific pacifier.

6 Strength and weakness

The strengths and Weaknesses of our model are summarized as follows:

6.1 Strengths

Model complexity. Our model incorporates relevant research results and considers many details, making the model complicated enough to include enough factors. Therefore, our model can analyze the situation in specific regions and draw targeted conclusions, which is conducive to practical use.

Agreement between experimental data and real data. In the process of estimating maximal plastic waste, our model based on PLC theory carefully consider the environmental impact of plastic in different phases of its life cycle rather than merely in the phase of management. What’s more, despite the lack of data, our machine learning model still fits the data well. The results of tests on real data sets also validate the performance of our model, where experimental data and real data agrees well.

The amount of data. In the process of modeling, we extensively collected relevant industry data, making the model sufficiently universal and extensible. To our knowledge, few studies on the relationship between plastics and the environment have been able to model using so much data.

6.2 Weaknesses

Data accuracy of the model. In order to consider as many factors as possible, our model introduces more parameters. This makes it difficult to obtain our data, especially when related research is very scarce. Further we will try to simplify the model for different scenarios and improve the usability of the model.

Macro analysis. Our model is too detail-oriented and has insufficient performance in macro analysis. In global research, our model encountered some difficulties and showed some uncertainty. In the future, we will try to explore the relationship between the overall situation and the details, and strive to let our model to show the superiority in global problems, too.

7 Conclusion

To mitigate the plastic waste problem, we establish a linear programming base model to estimate the environmental impacts of plastic waste, and get the maximal plastic waste amount without further damage to environment. It is proved that the most extent plastic can be reduced to could be confirmed by HSVR model. A minimal target amount of plastic usage is achieved by expending HSVR model. The equity issue can also be solved by dual programming and sensitivity analysis of factors. Finally, we reckon a timeline to approach this realistic minimal target of plastic using amount.

LETTER TO SUNSHINE COMPANY

To: Marketing Director of Sunshine Company

From: Team 2013573

Subject: A Wealth of Data

Date: March 9, 2020

Dear Director:

Thank you very much for your trust and hire our team to solve this important problem for you. As we all know, there is a lot of valuable information in the product ratings and reviews. We hope that the information our team digs from the data of pacifier, microwave and hair dryer can help your company to make reasonable sales strategy, and enhance effectiveness of products timely. We hope our suggestions can be helpful to your company.

As for how to integrate product ratings and reviews to provide effective information for your company, our team first quantified the text reviews effectively, and then made full use of helpful votes, vine to trace the credibility of the reviews. At the same time, with the thoughts of average, logarithmic smoothing and exponential amplification of differences, a robust model is established, because our model can always give the correct relative reputation value between products. We believe that a good reputation can potentially bring a better sales volume to a product, so we suggest that your company can evaluate the reputation of pacifier, microwave, hair dryer and other products, and then increase the stock of products with high reputation or amplify their publicity. As mentioned in our paper, wubbanub infant pacifier - giraffe is the first choice for online sales of pacifier products from July to August 2015.

Our fitting model uses a powerful RBF neural network to fit the relationship between reputation and time successfully, which can help us better summarize and predict the change of reputation of products. Furthermore we can excavate the potential factors of product success or failure from the maximum and minimum values. As we mentioned in the article, when words like 'baby cant' or 'without soothing' with 2-3 stars appear, we should be extra careful. This is the product the reputation of which will begin to decline. We recommend your company to improve your products in time according to such signs. When word like 'baby love' or 'easy to hold' with 4-5 stars appears, your company can safely increase the inventory of the product, because the product is about to sell well.

We would be very pleased if our suggestion could be of benefit to your company. Wish the given advice above would contribute to your work!

Your sincerely
Team 2013573

References

- [1] M. M. Li and B. Verma, “Nonlinear curve fitting to stopping power data using rbf neural networks,” *Expert Systems with Applications*, vol. 45, pp. 161–171, 2016.
- [2] M. R. Stoline, “The status of multiple comparisons: simultaneous estimation of all pairwise comparisons in one-way anova designs,” *The American Statistician*, vol. 35, no. 3, pp. 134–141, 1981.
- [3] D. Tang, B. Qin, T. Liu, and Y. Yang, “User modeling with neural network for review rating prediction,” in *Twenty-Fourth International Joint Conference on Artificial Intelligence*, 2015.
- [4] T. Subkhankulova and F. J. Livesey, “Comparative evaluation of linear and exponential amplification techniques for expression profiling at the single-cell level,” *Genome biology*, vol. 7, no. 3, p. R18, 2006.
- [5] X. Zhang and C. Dellarocas, “The lord of the ratings: is a movie’s fate is influenced by reviews?” *ICIS 2006 proceedings*, p. 117, 2006.