DSSR

Discussion Paper No. 74

Interpretable Perceived Topics in Online Customer Reviews for Product Satisfaction and Expectation

Aijing Xing
Nobuhiko Terui

February , 2018

Data Science and Service Research
Discussion Paper

Center for Data Science and Service Research
Graduate School of Economic and Management
Tohoku University
27-1 Kawauchi, Aobaku
Sendai 980-8576, JAPAN

Interpretable Perceived Topics in Online Customer Reviews for Product Satisfaction and Expectation

Aijing Xing and Nobuhiko Terui[1]

Graduate School of Economics and Management

Tohoku University

February, 2018

Tohoku University, Graduate School of Economics and Management, Kawauchi Aoba-ku, Sendai, 980-8576, Japan; terui@econ.tohoku.ac.jp

[1] Terui acknowledges a grant by JSPS KAKENHI Grant Number (A)17H01001.

Interpretable Perceived Topics in Online Customer Reviews for Product Satisfaction and Expectation

Abstract

Online customer reviews contain useful and important information, in particular, for product management—because customers tend to praise or criticize certain features or attributes of goods in their reviews. We propose a model that extracts the perceived topics from textual reviews by natural language processing under the restrictions of their interpretability and predictability of product satisfaction as current product evaluation and expectation as future possible demand by supervised learning. The empirical analysis on user-generated content of food reviews shows that our proposed model performs better than alternative models, and it suggests product managers the necessity of improving some specific attributes and focus their advertising on these attributes as fulfilling customer needs.

Keywords and Phrases: Food satisfaction, Online customer reviews, Labeled Topic model, Supervised learning, User-generated content, Text data

1 Introduction

With the rise of electronic commerce, online retailers such as Amazon, Walmart, and Taobao have experienced growth in their number of users. Most online retailers have customer feedback systems to receive information, including satisfaction scores (also called product ratings) and textual reviews.

Various methods to utilize online customer review data have been proposed in the literature. Some evidence suggests that consumer reviews are important to get the approval from readers and further facilitate purchase decisions and product sales. As customer-created information, online customer reviews are likely to be more credible than seller-created information (Wilson and Sherrell, 1993). Online customer reviews posted in forums are relevant to consumers because the opinions and reviews are contributed by fellow consumers, who are perceived to have no vested interest in the product and no intentions to manipulate the reader (Bickart and Schindler, 2001). The number of customer reviews affects the profit of book products (Zhao et al., 2013). Chen and Xie (2008) suggest that online reviews can help novice customers identify products that best match their preferences, and they conclude that in the absence of review information, novice consumers may be less likely to buy a product if only seller-created product attribute information is available. Obviously customers prefer to read the customer reviews before purchase decision, because these user-generated information is perceived to be trustworthy and match their preference.

The satisfaction scores were used as indicators of word-of-mouth (WOM) to forecast product sales and revenue. Chevalier and Mayzlin (2006) find that the star rating of online book products has a significant impact on book sales. Liu (2006) shows that movie ratings on the Yahoo! Movies website have a significant effect on box office revenue. Chintagunta et al. (2010) find that the valence of online WOM (mean user rating) has a significant and positive impact on box office earnings.

Performing text analysis of online customer reviews is also useful for improving a seller's marketing strategy and customer management. Company managers can mine textual information about product features from the "voice of the consumer" and hence better meet customer needs when proposing new products (Fay and Xie, 2008, Griffin and Hauser, 1993, Xie and Gestner, 2007, Yohan and Alice, 2011). Lee and Bradlow (2011) and Netzer et al. (2012) exploit text-mined semi-structured product reviews to understand market structures based on the product attributes mentioned in the reviews. Decker and Trusov (2010) estimate consumer preferences for product attributes by text mining pro and con product reviews.

Büschken and Allenby (2016) recognize that the informative aspects of text data are readily observed and can directly serve as explanatory variables to identify customer behavior (e.g., satisfied versus unsatisfied experiences, helpful and unhelpful levels), and they assume that words within a sentence pertain to the same topic. Their sentence-based text analysis improves the predictive fit of customer satisfaction rating.

Except for the content of customer review, exploring the influences of reviews toward purchase intention is also important. Maheswaran and Levy (1990) find that positively and little detailed framed messages may be persuasive, and negatively and detailed framed messages may also be persuasive. Herr et al (1991) conclude that vivid WOM about product attribute information has a greater impact on customer choices. We assume that readers of these reviews expect to search for valuable information from the experienced customers, such as risk of the product or favorable recommendation, and these information will be important reason for purchase decision of readers.

Following empirical research above, this study aims at information extraction from review data and addresses two questions about customer experience and expectation:

"What product features are mentioned in the online customer reviews posted by satisfied and dissatisfied customers?"

"Whether readers expect to find helpful information about product features in customer reviews."

To answer these questions, we extend topic model for extracting perceived topics containing product features and predict satisfaction rating and helpfulness of product expectation.

Topic models such as Latent Dirichlet Allocation (LDA) (Blei et al., 2003) assumes each customer review consists of several topics with varying probabilities. The main problem of LDA model is the lack of interpretability. Promiscuous words sometimes are clustered in a common topic, which makes the topics hard to interpret. To improve interpretability, several modifications of LDA have been proposed in terms of incorporate supervision. Supervised LDA (Blei and McAuliffe, 2007) accommodates movie reviews (text data) and movie scoring (ordered categorical data), and a regression procedure is used for topic extraction and movie scoring forecasting. In labeled LDA (Ramage et al., 2009), documents are labeled with perceived tags artificially, and the topics can be easily interpreted.

Another focus of this paper is assuming the perceived topics according to prior information-product features mentioned in customer reviews. By reading the subset of customer reviews and observing the descending sort of word frequencies from the whole vocabulary, we employ specific perceived topics about product features of the customer reviews for the labeling.

This paper proposes the supervised–labeled LDA model by combining supervised LDA and labeled LDA to extract interpretable perceived latent topics about product features from online customer reviews and predict customer satisfaction and product expectation. Then the number of word assigned to respective topic taking logarithm is used to forecast customer's satisfaction ratings and helpfulness of product expectation.

In section 2, we first illustrate the layout form of customer reviews, which contains information about textual reviews and numerical variables, and we investigate the features of research object(one brand of potato chips) through reading the subset of customer reviews, observing the descending sort of word frequency list of the whole vocabulary, and referred to the literature about food features, thereby we presuppose the perceived topics. To improve interpretability and make the perceived topics match with product features, we propose the supervised–labeled LDA model in the form of multivariate linear regression and discuss the estimation procedure in section 3. The section 4 reports the empirical analysis. We discuss managerial implications in section 5 and conclude in section 6 finally.

2. Online Customer Review

2.1. Layout Form of Online Customer Reviews

Most online retailers have a similar layout form of online customer reviews, including textual content and product rating. For instance, the Amazon review in Figure 1 includes three parts: (i)Satisfaction Score, (ii)Textual Review, and (iii)Helpfulness Numerator. Specifically, after buying and accepting goods, customers are allowed post their feelings and experiences by writing a textual review. At the same time, these customers may grade products with “stars” ranging from 1 to 5. The reader of this online review evaluate whether this review helpful for his/her product expectation and purchase decision by clicking “Yes” or unhelpful by clicking “No”.

Figure 1. Layout form of Amazon food review

2.2. Textual Reviews and Perceived Topics

The “Amazon Fine Food Reviews” data is available at the website of “Kaggle”. This site is a platform for predictive modeling and analytics competitions, and the datasets are freely uploaded by companies and users.) The data set consists of about 500,000 food reviews by Amazon users up till October, 2012 on products including soda, cocoa, dog food, and so on. Specific brand of potato chips which has largest number of customer reviews is chosen, and it contains 564 reviews. Selecting single product makes it easy to assume perceived topics on the point of product features.

First, text mining approach (Netzer et al. 2012) is used to extract useful and meaningful information, such as terms for products and product features in unstructured text. The text mining approach helps us reducing the whole vocabulary and avoids unnecessary computation through two ways, one is that it can automatically disambiguate the words included in the library which contains frequently used but unimportant words such as *I, they,* *for, just, cannot, the, these, that, am, is* and so on. The other one is its transformation from uppercase to lowercase. Below are two typical examples from the online review data, with the red words kept as meaningful words after text mining approach. :

- Customer A

  •Helpfulness Denominator: 13. Helpfulness Numerator: 7 (13 − 6).

  •Satisfaction Score: 1

  •Summary: Gross!

  •Textual Review: I just cannot understand the high praise these chips have received. I ordered the variety pack and am very disappointed. All flavors have a really weird taste to them, except for maybe the plain chips. The salt and vinegar especially tasted gross! I took one bite and had to throw the bag away. Now I'm stuck with a bunch of chips that I'll never eat.

- Customer B

  •Helpfulness Denominator: 35. Helpfulness Numerator: 33 (35 − 2).

  •Satisfaction Score: 5

  •Summary: Fantastic chips!!!

  •Textual Review: I want to start out by saying that i thought at first that a bag with only 120 calories and 4 grams of fat (no saturated or trans) for every 20 chips was going to taste like crap. I must say that not only was i wrong, that this is my favorite BBQ chip on the market today. They are light and you cannot taste any fat or grease after eating them. That's because they aren't baked or fried, just popped as their name suggests. These chips are very easy to dip as well. FANTASTIC PRODUCT!

Customer A gives just a 1-star grade for this product, and 13 viewers regard customer A's textual review as helpful, while 6 viewers evaluate it as unhelpful. Customer A questions the high praise and rating of the product and criticizes the taste and flavor of the chips. On the other hand, customer B praises the chips with a 5-star rating, and she/he discusses about ingredients, packing, cooking, taste, and dipping. Thirty-five viewers regard customer B's textual review as helpful, and only 2 viewers regard it as unhelpful. Based on the two typical examples, we suppose that product features are related with customer satisfaction and product expectation as shown in Figure 2, specifically:

Figure 2. Reviews toward Product Features and Helpfulness Recognition

- Experienced customers praise or complain about this brand of chip by talking about topics about product features that include flavor, taste, ingredients, packing, cooking and consumption experience, private sentiments, and so on.
- The readers of reviews mostly have no consumption experience of the chip, and these novice customers expect to search for important product information from the reviews written by experienced customers. If the review contains informative topics of product features, the readers will recognize the review as helpful.

Our aim is to extract these objective topics from the textual review and explore the connections between topics and satisfaction or the connections between topics and product expectation so that online retailers can improve specific product features. We suppose a certain number of perceived topics about food features. Several food features were mentioned by scholars, such as culinary quality (Chi et al. 2013), taste and flavor (Andersen and Hyldig 2015), service and price (Mason and Nassivera 2013), health/nutrition, and low-fat ingredients (Küster and Vila 2017). At the same time, in order to find the possible food features in the reviews, we read the subset of the customer reviews and observe the descending sort of word frequencies of the whole vocabulary. The top 60 frequently used words are listed in Table 1, these feature-based words are about flavor, taste, packing, weight, food ingredient, cooking method, and purchase. According to these previous related literatures and the product features of reviews, we combine synonymic topics and propose five perceived topics: (1) Flavor and Taste, (2) Packing, (3) Healthy, (4) Money, and (5) Ingredient.

Table 1. Word Frequency of Dataset

3. Model

3.1 Supervised Learning Model with Labeled Topics

For the interpretability of the extracted topics in the text analysis, we first employ a labeled topic model. "Labeled" means that some labeled words are fixed to their respective topics in order to distinguish meanings among latent topics. ,The meanings of extracted topics are easily interpreted if the synonyms are assigned in common topics, such as "hungry","famished" and "starving" are assigned in a common topic about "hunger". Second, we use other information sources for the model to be supervised, where "supervised" means that the observed numeric data (or ordered categorical data) work as dependent variable, and the frequencies of topics work as covariates. In this article, we the model by combining two kinds of models and call it supervised–labeled LDA (SL–LDA) in the following.

Besides textual content, directly observed numerical data, i.e., satisfaction score, helpfulness numerator, and helpfulness denominator, are available in the data set. The topics in customer reviews can directly serve as covariates for the variation in satisfaction scores and helpful reference of product expectation. For example, if one product feature is especially satisfying customer needs, we expect that the topic about this feature co-occurs with a high satisfaction score in online customer reviews. On the contrary, topics about dissatisfying features should occur along with low satisfaction scores. As for the connection between topic assignment and product expectation, it is likely that a textual review is regarded as helpful by readers if it contains expected topics.

Let $y_d = \left(y_{SA,d}, y_{HN,d}, y_{HD,d}\right)$ denote the vector of three continuous dependent variables of customer satisfaction, helpfulness numerator and helpfulness denominator in document *d*. The satisfaction reflects consumer's current evaluation based on their past experience, and we assume that helpfulness measures imply their interest and expectation on the product when they purchase in the future,

Given the word-topic assignment z_i in document *d*, $\sum_{i=1}^{N_d} \#\left\{z_i = k\right\}$ is the number of words assigned to topic *k*, and it works as explanatory variables after logarithm transformation. We define the supervised model by multivariate regression:

$$y_d = X_d \gamma + \delta_d, \quad d = 1, ..., D, \tag{1}$$

where $X_d = \left[1, x_{1,d}\ x_{2,d}\ x_{3,d}\ x_{4,d}\ x_{5,d}\right]$, and we define the covariate

$$x_{k,d} = \log\left(\sum_{i=1}^{N_d} \#\{z_i = k\} + 1\right) \quad (2)$$

where N_d is the number of words of document d. In other words, in order to explain the variations in satisfaction score, helpfulness numerator, and helpfulness denominator, we quantize the topics about product features with the natural logarithm of topic sizes $\sum_{i=1}^{N_d} \#\{z_i = k\} + 1$. We assume the error term follows multivariate normal distribution, $\delta_d \sim N_3\left(\mathbf{0}, diag\left(\sigma_1^2, \sigma_2^2, \sigma_3^2\right)\right)$.

Figure 3. Graphical model

The graphical model of SL-LDA is shown in Figure 3. In contrast with standard LDA model, the information of y_d is included in the model corresponding to "supervised". The core of our modeling is to sample the probabilities with which the word w_i is assigned to topic k, and the probabilities are inferred from the following conditional posterior density :

$$p(z_i = k \mid z_{-i}, \theta, \phi_k, y_d, w_i,) \propto p(z_i = k \mid z_{-i}, \theta) p(w_i = v_\lambda^k \mid z_i = k, z_{-i}, \phi_k) p(y_d \mid z_i = k, z_{-i}, \gamma, \delta) \quad (3)$$

The detailed process of model estimation is shown in the appendix.

3.2 Transformation Matrix and Labeled Words

We demonstrate how to accomplish the labeled perceived topics by using labeled words in this section. After removing unnecessary words, we consider the total vocabulary which includes all words occurred in customer reviews as text data set. To make topics interpretable, some meaningful and frequently-used *labeled words* are chosen from the total vocabulary and fixed in respective perceived topics. At the same time, the remaining words which have no definite meaning are called as *non-labeled words* and assigned to topics with specific probabilities. For example, in Figure 4, given five topics about product features and total vocabulary of $\{v_1, v_2, v_3, v_4, v_5, v_6, v_7\}$, we assume that word v_1 works as the labeled word of Topic 1 and it is thus excluded from the other four topics. Other labeled words $v_2, v_3, v_4,$ and v_5 are specified as the labeled words for the remaining topics. The non-labeled words v_6 and v_7, can be assigned to any topic with certain probabilities; hence, the vocabulary of Topic 1 is $\{v_1, v_6, v_7\}$, which is the subset of total vocabulary of $\{v_1, v_2, v_3, v_4, v_5, v_6, v_7\}$. After establishing vocabulary for topic k, where k=1,2…5,according to LDA model we expect that customers express the topic k and choose the words from vocabulary of k with specific probabilities, and the chosen words follow Dirichlet distribution conditional on β^*_k.

Figure 4. Labeled and non-labeled words for respective topics

In online customer review data or other text analytics, we are faced with huge sized vocabulary. To build the vocabulary for topic k from total vocabulary, we imitate and simplify the methodology from labeled LDA (Ramage et al. 2009). We first generate the topic's *transformation matrix* $\Lambda^{(k)}$ ($V_k \times V$) conditional on $\{\lambda_1, \lambda_2 .. \lambda_{Vk}\}$, where $\{\lambda_1, \lambda_2 .. \lambda_{Vk}\}$ is the sequence number of probable words. Just like $\{v_1, v_6, v_7\}$ for Topic 1 in Figure 4, V_k=3 and V=7, consequently the choosing of sub-vocabulary is accomplished. The mechanism of $\Lambda^{(k)}$ is illustrated in the appendix.

The goal of "Labeled" is making the different vocabularies for each topic through $\Lambda^{(k)}$, in other words, we exclude other topics' labeled words from the objective topic's vocabulary, just as we exclude $\{v_2, v_3, v_4\}$ from the vocabulary topic 1 in Figure 4. In practical computation, some software packages, such as Python, R, and MATLAB, have a filter function to extract subvectors from the whole vector, so researchers can use these filter functions to avoid having to make the matrix calculation when V_k and V are especially large. Given the parameter vector β^*_k, we sample ϕ_k, a topic-vocabulary vector and obeys the Dirichlet distribution:

$$\phi_k \sim \mathrm{Dir}(\cdot \mid \beta^*_k) \tag{4}$$

Similar to collapsed Gibbs sampling of the labeled LDA model (Ramage et al. 2009), we sample the probabilities in which the topic assignment z_i of w_i is assigned to topic k according to the conditional posterior density:

$$p(z_i = k \mid z_{-i}, \alpha, \beta^*_k, w) \propto \frac{\left(n^{-i}_{d,k} + \alpha_k\right)}{n_d + \alpha^T \times 1} \frac{\left(n^{-i}_{k,w=v^k_\lambda} + \beta^*_{\lambda_{vk}}\right)}{n_k + \beta^{*T}_k \times 1} \tag{5}$$

where $w_i = v^k_\lambda$ means the word w_i is chosen from vocabulary $\{v^k_1, v^k_2, v^k_3, ..v^k_\lambda ..\}$ of topic k. The term $n^{-i}_{d,k}$ means, except for z_i in document d, the number of words assigned to topic k. The term $n^{-i}_{k,\ w=v^k_\lambda}$ means, given the topic assignment k, the number of times that v^k_λ is chosen. In addition, if w_i is the labeled word of topic k like v_1 in Topic 1, we ignore the equation (5) and assign w_i to topic k directly because no Dirichlet prior exists for the other topic assignments of labeled word w_i.

3.3 Model Estimation

We make the SL-LDA model based on two aspects, i.e., one is "supervised" that the numerical dependent variable y_d works as reference information as shown in (3), and the other is "labeled" that we suppose the perceived labeled topics and labeled words strictly specified according to prior information about product features as shown in (5). The conditional probability density of topic assignment $z_i = k$ is given as

$$p(z_i = k \mid w, y, z_{-i}, \alpha, \beta^*, \gamma, \delta) \propto \frac{\left(n_{d,k}^{-i} + \alpha_k\right)}{n_d + \alpha^T \times 1} \frac{\left(n_{k,w=v_\lambda^k}^{-i} + \beta_{\lambda_{vk}}^*\right)}{n_k + \beta_k^{*T} \times 1} p\left(y_d \mid z_{d,i} = k, \gamma, \delta\right) \quad (6),$$

where i = 1, 2, 3 … , N.

Then, the MCMC process by Gibbs sampling is as follows:

(i) Initiate the topic assignment z of both labeled and non-labeled words.

(ii) For each document d = 1, 2 … D, and for each word w_i, i = 1, 2, … , N_d. If w_i is a non-labeled word, for topic k = 1, 2, … , K, assign the topic based on (6), else directly assign the labeled word w_i to a certain topic according to step (i).

(iii) Given the topic assignment Z, sample the following posterior mean of the Dirichlet distribution:

$$\theta_{d,k} = \frac{n_{d,k} + \alpha_k}{\sum_K \left(n_{d,k} + \alpha_k\right)} \qquad \phi_{k,w=v_\lambda^k} = \frac{n_{k,w=v_\lambda^k} + \beta_{\lambda_{vk}}^*}{\sum_{V_k} \left(n_{k,w=v_\lambda^k} + \beta_{\lambda_{vk}}^*\right)} \quad (7)$$

(iv) Sample $p(\gamma, \delta \mid z, y)$ with Bayesian linear regression and return to step (ii).

4. Empirical Application

Our research object is a brand of potato chips, whose data set has 564 customer reviews. This data consist of two parts, one of which is textual review and the other is observed numerical data, including Satisfaction Score, Helpful Numerators, and Helpful Denominators. First of all, we need to choose the labeled words for perceived topics mentioned in section 2.2 appropriately. We list the word occurrence in descending order, then remove irrelevant words such as “a,” “the,” and “and,” finally selecting the following labeled words for their respective perceived topics:

- **Flavor** and **Taste**: flavor, flavors, sweet, salt, taste, pepper, vinegar, tastes, tasty, delicious
- **Packing**: size, bag, box, bags
- **Healthy**: calories, calorie, fat, grease, greasy, healthy, health, unhealthy, nutrition, nutritious
- **Money**: price, money, cheap, expensive
- **Ingredient**: oil, flour, ingredient, ingredients, potato, sugar, starch, flakes

Our model assumes the perceived topics above, meanwhile the labeled words are fixed in their nominated perceived topics and never switched to other topics, and the remaining words in total vocabulary will be clustered to the five topics with specific probabilities.

4.1 Word Classification

After MCMC process of Gibbs sampling, which takes 4000 iterations, we illustrate the word classifications of the top 15 words in Table 2, where we can clearly see that the word classification of SL–LDA has better interpretation than Supervised–LDA (SLDA). Generally, all five topics of SL-LDA contains corresponding words. In contrast, the first four topics of SLDA contain heterogeneous words and are hardly interpreted, only the fifth topic has clear meaning about food ingredient. Note that the meanings of topics in SL–LDA have been presupposed when we appoint labeled words for the topics, so topic names are directly demonstrated in Table 2. In contrast, “Topics 1” through “Topic 5” serve as the topic names in Supervised–LDA since the meaning of most topics is difficult to interpret.

Table 2. Word Classification of SL–LDA and SLDA

The most non-labeled words, just like v_6 and v_7 in Figure 4, are clustered with related labeled words and assigned to corresponding topics:

- **Flavor** and **Taste**: bbq, favorite, best, better, eating, baked, original
- **Packing**: weight, single, small, per, ounce, individual, amount, filling, large, larger
- **Ingredient**: rice, corn, natural, cake, seasoning, heat, sunflower

The name of the topic “Money” was changed to “Money and Buying” because words such as *saving*, *subscribe*, *average*, *reviews*, *shipping*, *gives*, *sampling*, *alternatives*, and *extra* were assigned to this topic, and these words contain meanings about purchase behavior.

The topic “Health”, however, presents a problem that the clustered non-labeled words do not have significant meanings about health, maybe because there is really no other word about the health topics discussed by customers after 10 labeled words are selected for “Health”. In addition, many words in “Ingredient” have meanings close to that of “Health”.

We demonstrate the word count for each topic in Figure 5. Apparently, the most-discussed topic is “Flavor and Taste,” which means that most customers are willing to talk about that topic.

Figure 5. Number of words assigned to each topic

4.2 Supervised Learning Model: Multivariate Regression

Besides the word classification, our other research purpose is to predict the dependent variables of customer satisfaction, helpfulness numerator and helpfulness denominator through the interpretable perceived topics about product features. We use DIC as goodness of fit to compare the models: (i)"LDA + Multivariate Linear Regression", (ii) "SLDA" (Supervised LDA), (iii) "SL–LDA" (Supervised–labeled LDA), and (iv) "SL–LDA with Independent Linear Regression"in Table 3, which demonstrates that SL–LDA has lower DIC, and it means extracted topics in SL–LDA are quite reliable in interpreting the variation in customer satisfaction and helpfulness of product expectation.

Table 3. Model Comparison of DIC

"LDA + Multivariate Linear Regression" means a separate process that consists of LDA and multivariate linear regression. Specifically, after topics are assigned by a standard LDA model, we use multivariate linear regression with {topic1, topic2 … topic5} frequencies as the explanatory variables. "SLDA" means that topic assignment is done according to text data of customer reviews and it connects to numerical variables data in the form of multivariate regression where the {topic1, topic2 … topic5} frequencies defined in (2) serve as the explanatory variables.

In "SL–LDA," the frequencies of topic assignments about {"Flavor and Taste," "Packing," "Health," "Money and Buying," and "Ingredient"} work as the explanatory variables in step (iv) of MCMC iteration in section 3.3. The topic assignment is made on the basis of probabilities generated in equation (6).

"SL–LDA with Independent Linear Regression" means that we execute the SL–LDA model three times, with {"Flavor and Taste," "Packing," "Health," "Money and Buying," "Ingredient"} as the explanatory variables, while the Satisfaction Score, Helpfulness Numerator, and Helpfulness Denominator work as the response variables respectively in the linear regression. Furthermore, the "SL–LDA with Independent Linear Regression" will output three separate groups of word classification, regression estimation, and goodness of fit.

We use an HPD region to check the significance of estimated coefficients of intercept and five perceived topics. In Table 4, "SL–LDA" and "SL–LDA with Independent Linear Regression" both output seven significant coefficients (marked with *) compared with three significant coefficients in "LDA + Multivariate Linear Regression" and "SLDA". After comparing the DIC and HPD region of four models, we can say that the extracted topics of "SL–LDA" can effectively affect the dependent variables about customer satisfaction and helpfulness of product expectation, so SL-LDA is the best model in our research. Specifically, the estimates of SL-LDA shows that "Satisfaction Scores" is positively affected by the topics of "Flavor and Taste" and "Health," and negatively affected by "Money and Buying" and "Ingredient."

Table 4. Estimates of Gamma in Linear Regression (0.1 HPD region test)

In the following section, we will have discussion about estimated coefficients of SL-LDA and topic proportions for all satisfaction levels.

5. Discussion and Managerial Implications

According to the estimated coefficients of SL-LDA, "Satisfaction Score" is significantly affected by four topics. Specifically, dissatisfied customers are more likely to talk about "Flavor and Taste" and "Health," while satisfied customers usually talk about "Money and Buying" and "Ingredient." We suggest that the manufacturer of these potato chips improve the attributes of flavor, taste, and health, such as raising awareness and advertising their nutritive function, so that the needs of dissatisfied customers will be fulfilled. In addition, reviews including the topic "Health" are regarded as helpful by readers of online customer reviews, while the topic "Ingredient" weakens helpfulness. We deduce that novice customers highly expect the healthiness information from reviews, whereas they pay less attention to the reviews about ingredient issues, possibly because these readers are impatient to read the terminology. Consequently, retailers need to improve the healthy attributes of their products and advertise healthy attributes to meet novice customers' expectations.

We simulate the topic assignment for each word in customer reviews through SL-LDA, so that these customer reviews can be regarded as the mixture comprised with topics of {"Flavor and Taste," "Packing," "Health," "Money and Buying," "Ingredient"}, and accordingly we make the bar chart of topic composition of Figure 6 and Figure 7.

Figure 6 shows the total and average number of words assigned to respective topics in each satisfaction level. In the left figure of total word number, we can clearly see that satisfied customers are more willing to write online reviews, and the topic “Flavor and Taste” makes up the most words. In the right figure about average word number (total number of words divided by the number of reviews for the respective satisfaction score level), customers who gave four scores wrote more words in their reviews, and “Health,” “Ingredient,” and “Packing” receive proportionally more when satisfaction scores increase.

Figure 6. Total and average word number

Figure 7 illustrates the helpfulness degree of product expectation voted by its readers in different satisfaction levels. In terms of total quantity (left figure), the most helpful reviews are made by customers who give 5 stars to goods. However, individually (right graph), reviews made by customers who give just 1 and 2 stars are thought to be more helpful, it means that readers are more likely to search for expected information from the negative reviews made by dissatisfied customer. Furthermore, readers care more about “Flavor and Taste,” “Health,” and “Ingredient” when mentioned by dissatisfied customers.

Figure 7. Total and average helpfulness numerator

6. Conclusion

We introduced a supervised labeled LDA (SL-LDA) modelcombined the labelled LDA to extract interpretable perceived topics with the supervied LDA model for exploring the structure of predicting satisfaction of experienced customers and expectation of novice customers. In order to distinguish meanings among latent topics in online customer reviews, priori labeled words related with product features are assigned to respective topics. Accordingly, we connect the product features with customer satisfaction as past demand and expectation as future demand through supervised learning.

According to word classification in section 4.1 and predictive performance in section 4.2, SL-LDA model performs best compared with other models including "LDA + Multivariate Linear Regression", "SLDA", and "SL–LDA with Independent Linear Regression.". The research method through which we create perceived topics strictly according to the prior knowledge about product features is quite effective for improving the interpretability of topics and goodness of fit for structural model to predict satisfaction and expectation.

In the empirical study applied to a product of potato chips, we find that "Flavor and Taste," "Health," and "Ingredient" mentioned by dissatisfied customers are likely recognised as helpfulness by review readers, who will click the "Yes" button if they think the reviews are helpful for their product expectation, and this public expectation should be a important reference about word-of-mouth improvement for marketers. "Flavor and Taste" and "Health" negatively affect satisfaction, so retailers should track the details of the two product features to find what is complained about by dissatisfied customers.

Two characteristics of online customer reviews help us to employ the prior information and enhance the predictive fit of our model. One is that customers discuss product features with concise textual reviews without deep semanteme such as slang and adage. The other one is the simple and convenient extraction of features from single product, thus the perceived topics about product features are easily artificially presupposed, otherwise the complex features of heterogeneous products will cause great difficulties for choosing perceived topics. In managerial application, as long as the retailers take advantage of the prior information and presuppose perceived topics properly, the SL-LDA model should make good predictive fit and help retailers making textual analytics accurately.

Future research is needed in two aspects about marketing improvement. One is the improvement of product feature analytics through evolutive LDA models or other linguistic methods, such as employing word sequence, linguistic structure and word nature as more informative “labeled”. Another aspect is using the topic model to predict more “supervised” objectives such as detecting fake or deceptive reviews, in which meaningless reviews and high score rating are made after false purchase behavior in P2P e-shop.

References

Andersen BV, Hyldig G (2015) Food satisfaction: Integrating feelings before, during and after food intake. *Food Qual. Prefer.* 43:126–134.

Bickart B, Schindler RM (2001) Internet forums as influential sources of consumer information. *JIM* 15(3):31–40.

Blei DM, Ng, AY, Jordan MI (2003) Latent Dirichlet Allocation. *J. Mach. Learn. Res.* 3(January):993–1022.

Blei DM, McAuliffe J (2007) Supervised topic models. *Adv. Neural Inf. Process Syst.* 21:121–128.

Büschken J, Allenby GM (2016) Sentence-based text analysis for customer reviews. *Marketing Sci* 35(6):953–975.

Chen Y, Xie J (2008) Online consumer review: Word-of-mouth as a new element of marketing communication mix. *Management Sci*. 54(3):477–491.

Chevalier J, Mayzlin D (2006) The effect of word of mouth on sales: Online book reviews. *J. Mark. Res*. 43(August):345–354.

Chi CG, Chua BL, Othman M, Ab Karim S (2013) Investigating the structural relationships between food image, food satisfaction, culinary quality and behavioral intentions: The case of Malaysia. *Int. J. Hospit. Tourism Admin.* 14(2):99–120.

Chintagunta PK, Gopinath S, Venkataraman S (2010) The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets. *Marketing Sci.* 29(5):944–957.

Decker R, Trusov M (2010) Estimating aggregate consumer preferences from online product reviews. *Int. J. Res. Mark.* 27(4):293–307.

Fay S, Xie J (2008) Probabilistic goods: A creative way of selling products and services. *Marketing Sci.* 27(2):674–690.

Durairaj Maheswaran, Joan. Meyers Levy(1990) The influence of message framing and issue involvement. *J. Marketing Res*. 27 (3) : 361-367

Griffin A, Hauser JR (1993) The voice of the customer. *Marketing Sci.* 12(1):1–27.

Küster I, Vila N (2017) Health/nutrition food claims and low-fat food purchase: Projected personality influence in young consumers. *J. Funct. Foods* 38:66–76.

Lee T, Bradlow E (2011) Automated marketing research using online customer reviews. *J. Marketing Res*. 48(5):881–894.

Liu Y (2006) Word of mouth for movies: Its dynamics and impact on box office revenue. *J. Mark.* 70(July):74–89.

Mason MC, Nassivera F (2013) A conceptualization of the relationships between quality, satisfaction, behavioral intention, and awareness of a festival. *J. Hospit. Market Manag.* 22(2):162–182.

Netzer O, Feldman R, Goldenberg J, Fresko M (2012) Mine your own business: Market-structure surveillance through text mining. *Marketing Sci*. 31(3):521–543.

Paul M. Herr, Frank R. Kardes, John Kim(1991) Effects of Word-of-Mouth and Product-Attribute Information on Persuasion: An Accessibility-Diagnosticity Perspective. The *J. Consumer Res*. 17(4):454-462.

Ramage D, Hall D, Nallapati R, Manning CD (2009) Labeled LDA: A supervised topic model for credit attribution in multi-labeled corpora. *Proc. Conf. Empir. Methods Nat. Lang. Process* 1:248–256.

Wilson EJ, Sherrell DL (1993) Sources effects in communication and persuasion research: A meta-analysis of effect size. *Marketing Sci*. 21(Spring):101–112.

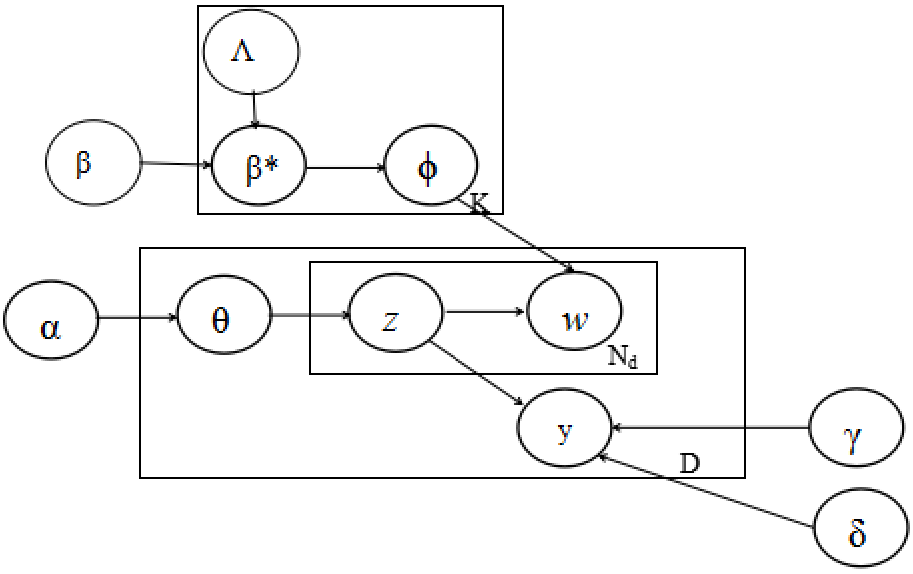
Xie J, Gestner E (2007) Service escape: Profiting from customer cancellations. *Marketing Sci*. 26(1):18–30.

Jo Y, Oh A (2011) Aspect and sentiment unification model for online review analysis. *The 4th ACM International Conference on Web Search and Data Mining* February:815–824.

Zhao Y, Yang S, Narayan V (2013) Modeling consumer learning from online product reviews. *Marketing Sci.* 32(1):153–169.

Appendix

A.1 Graphic Model of SL–LDA



Let $k \in \{1, . . . , K\}$ and $d \in \{1, . . . , D\}$, where K and D mean the number of topics and documents (customer reviews), respectively, and N_d stands for the number of words in document d. α is a K-dimensional vector acting as the Dirichlet prior for each row of θ, where θ is the $D{\times}K$ dimensional document-topic matrix defining mixing proportion. Denote $\beta = (\beta_1, \beta_2 ... \beta_V)$ and $\beta^*_k = (\beta_{\lambda_{1,k}}, \beta_{\lambda_{2,k}} ... \beta_{\lambda_{vk}} ... \beta_{\lambda_{VK}})$ as V and V_k dimensional vectors, respectively, where V is the size of the total vocabulary. V_k is the size of vocabulary $\{v_1^k, v_2^k, ... v_\lambda^k, ... v_{Vk}^k,\}$ for topic k, and $V_k < V$. Λ_k is defined as the transformation matrix for the reducing dimension of topic k. $\phi = \{\phi_1, \phi_2, ... \phi_k ... \phi_K\}$, where ϕ_k means the V_k dimensional topic-vocabulary vector. The term w = (w_1, w_2, … w_{Nd}) is the occurred words of document d. z = (z_1, z_2, … z_{Nd}), z_i is the word-topic assignment of word w_i.

A.2 Transformation Matrix of Labeled LDA

In Figure 4, the specific form of $\Lambda^{(k)}$ is as under.

For each row $i \in \{1,2...V_k\}$ and column $j \in \{1,2...V\}$:

$$\Lambda_{i,j}^{(k)} = \begin{cases} 1 & \text{if } \lambda_i = j \\ 0 & \text{otherwise} \end{cases} \tag{A1}$$

Hence, the parameter vector of Dirichlet prior β is transformed to a lower dimensional vector β^*_k.

$$\beta_k^* = \Lambda^{(k)}\beta$$

$$\beta_k^* = \begin{pmatrix} \beta_{\lambda_1} \\ \beta_{\lambda_2} \\ \beta_{\lambda_3} \\ \cdots \\ \beta_{\lambda_{Vk}} \end{pmatrix} = \Lambda^{(k)} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \cdots \\ \beta_V \end{pmatrix} \tag{A2}$$

Just as in the example in Figure 4, the sequence number of probable words for Topic 1 is $\{\lambda_1, \lambda_2, \lambda_3\} = \{1, 6, 7\}$; in other words, $\{v_1, v_6, v_7\}$ are chosen as potential words of topic k, and the dimension reduction process is as follows:

$$\beta_1^* = \begin{pmatrix} \beta_1 \\ \beta_6 \\ \beta_7 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \\ \beta_7 \end{pmatrix} \tag{A3}$$

A.3 Supervised–LDA

To build the conditional likelihood $p(y \mid z, \gamma, \delta)$, multivariate linear regression is applied as under.

For the document $d = 1, 2, ..., D$,

$$y_d = X_d \gamma + \delta_d,$$

$$\underbrace{\begin{bmatrix} y_{SA,d} & y_{HN,d} & y_{HD,d} \end{bmatrix}}_{y_d} = \underbrace{\begin{bmatrix} 1 & x_{1,d} & x_{2,d} & x_{3,d} & x_{4,d} & x_{5,d} \end{bmatrix}}_{X_d} \underbrace{\begin{bmatrix} \gamma_{SA,0} & \gamma_{HN,0} & \gamma_{HD,0} \\ \gamma_{SA,1} & \gamma_{HN,1} & \gamma_{HD,1} \\ \gamma_{SA,2} & \gamma_{HN,2} & \gamma_{HD,2} \\ \gamma_{SA,3} & \gamma_{HN,3} & \gamma_{HD,3} \\ \gamma_{SA,4} & \gamma_{HN,4} & \gamma_{HD,4} \\ \gamma_{SA,5} & \gamma_{HN,5} & \gamma_{HD,5} \end{bmatrix}}_{\gamma} + \underbrace{\begin{bmatrix} \delta_{SAd} & \delta_{HNd} & \delta_{HDd} \end{bmatrix}}_{\delta_d} \quad \text{(A3)}$$

A.4 Bayesian Inference of SL–LDA

Our purpose is to obtain the following conditional posterior:

$$p(z_i = k \mid z_{-i}, \theta, \phi_k, y_d, w_i,) \propto \underbrace{p(z_i = k \mid z_{-i}, \theta) p(w_i = v_\lambda^k \mid z_i = k, z_{-i}, \phi_k)}_{part1} \underbrace{p(y_d \mid z_i = k, z_{-i}, \gamma, \delta)}_{part2} \quad \text{(A4)}$$

Part 1 is about the text data of labeled LDA, and Part 2 is about the numerical data of supervised–LDA. First, we build the joint distribution according to our graphic model:

$$p(w, z, \theta, \phi, y \mid \alpha, \beta^*, \gamma, \delta) = \underbrace{p(\theta \mid \alpha) \cdot p(z \mid \theta) \cdot p(\phi \mid \beta^*) \cdot p(w \mid z, \phi_z)}_{part1} \cdot \underbrace{p(y \mid z, \gamma, \delta)}_{part2} \quad \text{(A5)}$$

Identical with a collapsed Gibbs sampling in a general LDA model, we can bypass the parameters of Φ and θ in Part 1 and make the specific form of joint distribution as follows:

$$\begin{aligned} & p(z, w, y \mid \alpha, \beta^*, \gamma, \delta) \\ & = \int\int p(w, z, \theta, \phi, y \mid \alpha, \beta^*, \gamma, \delta) d\theta \, d\phi \\ & = \int p(\theta \mid \alpha) \cdot p(z \mid \theta) d\theta \cdot \int p(\phi \mid \beta^*) p(w \mid z, \phi_z) d\phi \;\; p(y \mid z, \gamma, \delta) \\ & = \underbrace{\prod_{d=1}^{D} \frac{\Gamma\left(\sum_k \alpha_k\right)}{\prod_k \Gamma(\alpha_k)} \cdot \frac{\prod_k \Gamma(n_{d,k} + \alpha_k)}{\Gamma\left(\sum_k n_{d,k} + \alpha_k\right)} \prod_{k=1}^{K} \frac{\Gamma\left(\sum_w \beta^*_{w=v_\lambda^k}\right)}{\prod_w \Gamma\left(\beta^*_{w=v_\lambda^k}\right)} \cdot \frac{\prod_w \Gamma\left(n_{k,w=v_\lambda^k} + \beta^*_{\lambda_{vk}}\right)}{\Gamma\left(\sum_w n_{k,w=v_\lambda^k} + \beta^*_{\lambda_{vk}}\right)}}_{part\,1} \underbrace{\prod_{d=1}^{D} p(y_d \mid z_d, \gamma, \delta)}_{part\,2} \end{aligned} \quad \text{(A6)}$$

Because Φ and θ both disappear in the end of (A6), the collapsed Gibbs sampling is available in Part 1. Part 2 is the conditional likelihood of (A3), in which y_d obeys multivariate normal distribution. Given the above joint distribution function, we iterate conditional density of z_i conditional on $\{z_1, z_2, \ldots z_{i-1}, z_{i+1}, \ldots z_N,\}$ in collapsed Gibbs sampling, where i = 1, 2, 3 … N:

$$\begin{aligned}
& p\left(z_i = k \mid w, y, z_{-i}, \alpha, \beta^*, \gamma, \delta\right) \\
& = \frac{p\left(z_i = k, z_{-i}, w, y \mid \alpha, \beta^*, \gamma, \delta\right)}{p\left(z_{-i}, w, y \mid \alpha, \beta^*, \gamma, \delta\right)} \\
& \propto p\left(z_i = k, z_{-i}, w, y \mid \alpha, \beta^*, \gamma, \delta\right) \\
& \propto p\left(z_i = k, z_{-i} \mid \alpha\right) p\left(w_i = v_\lambda^k, w_{-i} \mid z_i = k, z_{-i}, \beta_k^*\right) p\left(y \mid z_i = k, z_{-i}, \gamma, \delta\right) \\
& \propto \underbrace{\frac{\left(n_{d,k}^{-i} + \alpha_k\right)}{n_d + \alpha^T \times 1} \frac{\left(n_{k,w=v_\lambda^k}^{-i} + \beta_{\lambda_{\mathrm{vk}}}^*\right)}{n_k + \beta_k^{*T} \times 1}}_{part\,1} \cdot \underbrace{p\left(y_d \mid z_{d,i}, \gamma, \delta\right)}_{part\,2}
\end{aligned} \tag{A7}$$

Then, the conditional posterior probability is

$$\begin{aligned}
p\left(z_i = k \mid w, y, z_{-i}, \alpha, \beta^*, \gamma, \delta\right) &= \frac{p\left(z_i = k \mid w, y, z_{-i}, \alpha, \beta^*, \gamma, \delta\right)}{\sum_{j=1}^{K} p\left(z_i = j \mid w, y, z_{-i}, \alpha, \beta^*, \gamma, \delta\right)} \\
&= \frac{\frac{\left(n_{d,k}^{-i} + \alpha_k\right)}{n_d + \alpha^T \times 1} \frac{\left(n_{k,w=v_\lambda^k}^{-i} + \beta_{\lambda_{\mathrm{vk}}}{}^*\right)}{n_k + \beta_k^{*T} \times 1} \cdot p\left(y_d \mid z_i = k, \gamma, \delta\right)}{\sum_{j=1}^{K} \frac{\left(n_{d,j}^{-i} + \alpha_j\right)}{n_d + \alpha^T \times 1} \frac{\left(n_{j,w=v_\lambda^j}^{-i} + \beta_{\lambda_{\mathrm{vj}}}{}^*\right)}{n_j + \beta_j^{*T} \times 1} \cdot p\left(y_d \mid z_i = j, \gamma, \delta\right)}
\end{aligned} \tag{A8}$$

Finally, we sample the topic assignment z_i according to the posterior conditional probability for each probable topic in the MCMC process.

Table 1. Word Frequency of Dataset

Word	Frequency	Word	Frequency	Word	Frequency
chips	760	regular	117	original	63
flavor	353	try	112	texture	62
like	345	vinegar	108	something	60
potato	300	eat	107	time	59
taste	281	tried	106	crunchy	59
popchips	243	much	106	product	57
bag	232	salty	99	lime	57
good	232	favorite	93	light	56
great	212	best	93	box	56
love	209	get	91	bought	53
sweet	196	healthy	84	weight	53
salt	194	better	82	chili	53
chip	170	eating	81	crunch	52
really	137	pop	80	tasty	52
fat	136	little	80	definitely	52
snack	130	buy	75	fried	52
calories	127	baked	71	case	51
bags	123	low	70	cheddar	51
flavors	119	pepper	69	rice	49
bbq	119	find	65	food	49

Table 2. Word Classification of SL–LDA and SLDA

SL-LDA				
Flavor and Taste	**Packing**	**Health**	**Money and Buying**	**Ingredients**
chips	bag	calories	price	potato
flavor	bags	fat	save	rice
like	box	healthy	money	ingredients
taste	snack	chip	after	oil
good	chips	low	subscribe	potatoes
salt	size	calorie	expensive	flour
popchips	popchips	greasy	reading	corn
love	weight	snack	average	natural
great	great	chips	reviews	starch
really	single	crunch	thembr	ingredient
flavors	love	crunchy	helps	cake
tried	eat	grams	thicker	list
much	small	sodium	shipping	flakes
try	per	great	decided	sugar
bbq	ounce	salty	seems	seasoning
SLDA				
Topic 1	**Topic 2**	**Topic 3**	**Topic 4**	**Topic 5**
snack	sour	chips	weight	flour
calories	cream	flavor	watchers	rice
makes	onion	like	points	ingredients
perfect	cheddar	taste	chips	corn
bag	after	potato	portion	oil
popchips	original	good	love	starch
relatively	garlic	love	bags	ingredient
afternoon	reading	popchips	great	natural
gives	parmesan	bag	bag	flakes
each	idea	great	single	sugar
mine	reviews	salt	low	safflower
shipments	thicker	chip	serve	sodium
compared	helps	really	pop	seasoning
boxes	preference	tried	control	list
tasty	bagbr	flavors	popchips	preservatives

Table 3.Model Comparison of DIC

Model	DIC
LDA + Multivariate Linear Regression	2707.46
SLDA	2611.38
SL–LDA	186.66
SL–LDA with Independent Linear Regression	1559.92 + 1605.84 + 1603.81

Table 4. Estimates of Gamma in Linear Regression (0.1 HPD region test)

Explanatory Variable	**Response Variable**	**LDA + Multivariate Linear Regression**	**SLDA**	**SL-LDA**	**SL-LDA with Independent Linear Regression**
Intercept	Satisfaction Scores	−0.104081	−0.16475	0.17948	0.17365
γ1		***−0.082277****	***−0.45349****	***−0.35696****	***−0.17229****
γ2		0.025689	−0.27023	−0.26295	***0.18504****
γ3		−0.047175	0.14893	***−0.16791****	***0.25785****
γ4		0.016622	−0.0499	***0.12293****	−0.03308
γ5		0.202006	−0.09601	***0.11739****	−0.06994
Intercept	Helpfulness Numerator	−0.465789	−0.02932	***0.17587****	***−0.33906****
γ1		***−0.013201****	−0.0463	0.0261	***0.11597****
γ2		0.152468	−0.0601	−0.02763	0.03422
γ3		0.048836	0.04158	***0.26881****	−0.03415
γ4		0.007903	***0.14977****	−0.02411	−0.01212
γ5		0.096529	0.09486	***−0.10661****	0.05451
Intercept	Helpfulness Denominator	−0.296761	***0.23187****	−0.05071	−0.2654
γ1		0.030989	0.10702	−0.01633	***0.11932****
γ2		0.101681	0.0237	0.05048	−0.02704
γ3		***0.039018****	−0.06036	−0.07703	***−0.10624****
γ4		−0.047418	0.04784	0.05149	0.0375
γ5		0.021323	0.04359	0.06881	0.06569

Figure 1. Layout form of Amazon food review

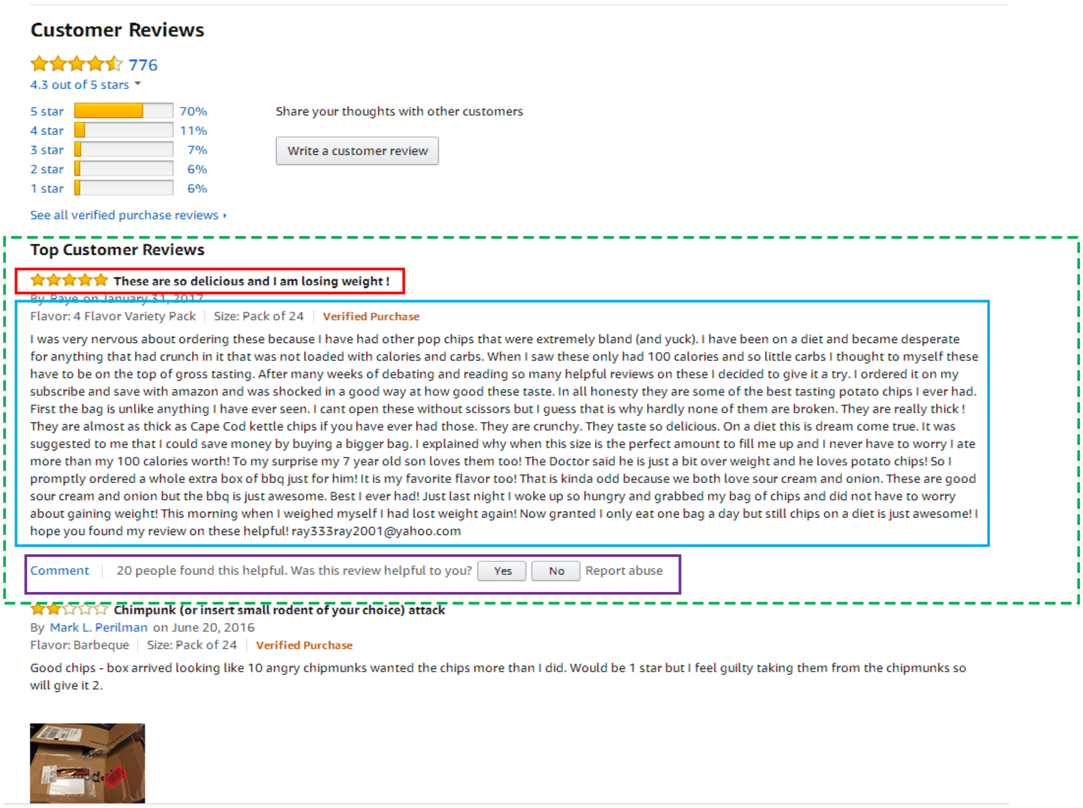


Figure 2. Reviews toward Product Features and Helpfulness Recognition

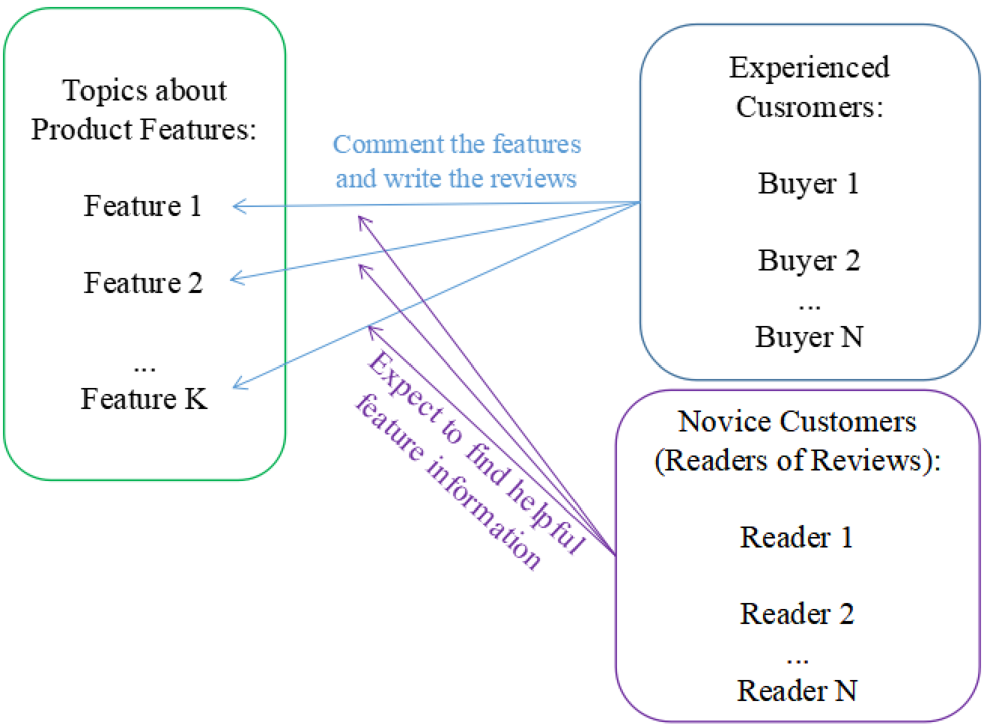


Figure 3. Graphical model

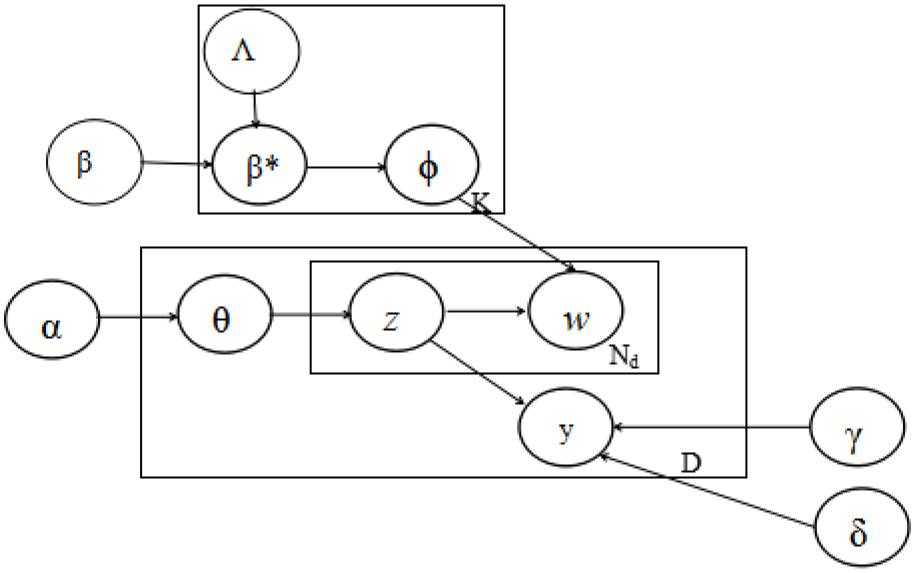


Figure 4. Labeled and non-labeled words for respective topics

$$
\begin{array}{r l}
 & \begin{array}{ccccccc} v_1 & v_2 & v_3 & v_4 & v_5 & v_6 & v_7 \end{array} \\
\begin{array}{r} \textit{Topic } 1,\ \beta_1^* \\ \textit{Topic } 2,\ \beta_2^* \\ \textit{Topic } 3,\ \beta_3^* \\ \textit{Topic } 4,\ \beta_4^* \\ \textit{Topic } 5,\ \beta_5^* \end{array} &
\begin{bmatrix}
\beta_1 & \times & \times & \times & \times & \beta_6 & \beta_7 \\
\times & \beta_2 & \times & \times & \times & \beta_6 & \beta_7 \\
\times & \times & \beta_3 & \times & \times & \beta_6 & \beta_7 \\
\times & \times & \times & \beta_4 & \times & \beta_6 & \beta_7 \\
\times & \times & \times & \times & \beta_5 & \beta_6 & \beta_7
\end{bmatrix}
\end{array}
$$

Figure 5. Number of words assigned to each topic

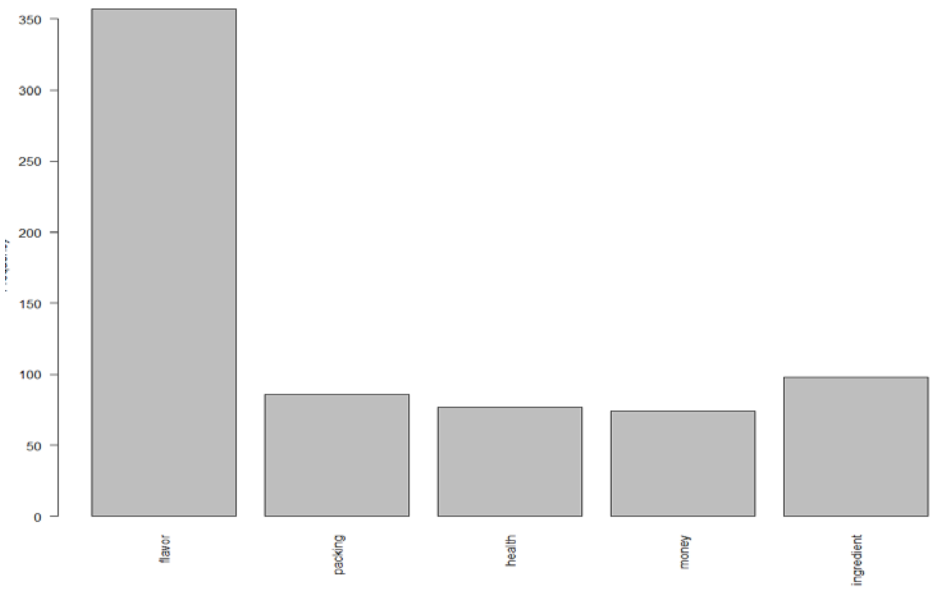
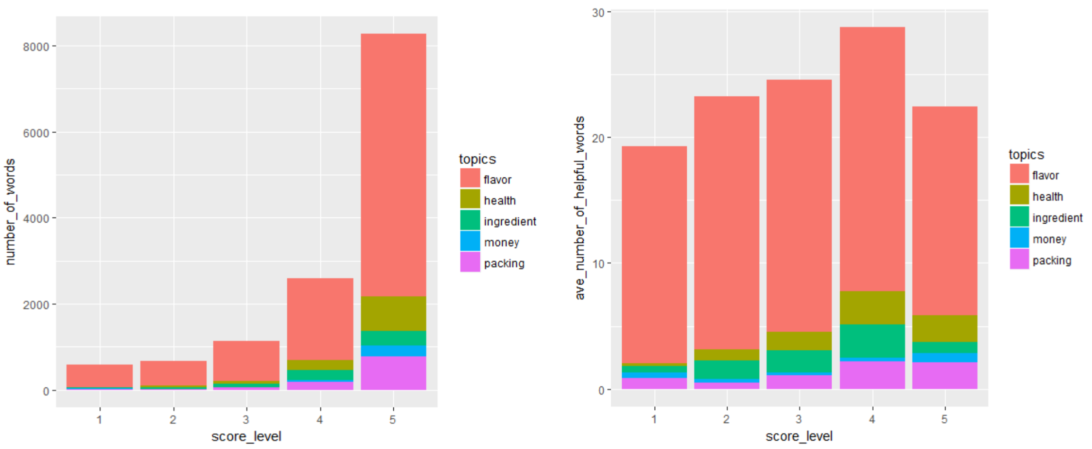
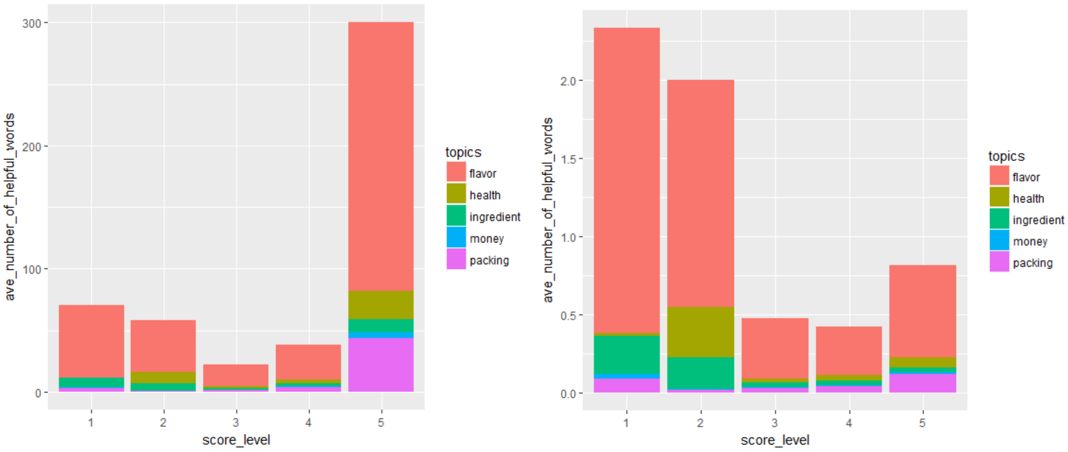


Figure 6. Total and average word number



Horizontal axis： satisfaction score. Vertical axis: number of words

Figure 7. Total and verage helpfulness numerator.



Horizontal axis： satisfaction score. Vertical axis: number of helpfulness numerator.