Customer Feedback-Text Analytics

Santhi Bhanavam

15/06/2022

## Abstract

As a result of the complexity that surrounds the holistic nature of customer experience feedback, it has become challenging to measure the customer’s perceptions concerning the services provided to them or goods delivered to them. In addition, the technological advancements and the change in techniques through which organizations collect explicit customer feedback have resulted in an increase in the generation of large volumes of unstructured textual data and hence making it difficult to analyze and gain useful information from this data manually. Therefore, an approach of text mining has been useful as it helps in the automatic extraction of information from textual data. However, in terms of accuracy and depth analysis of the customer feedback, the text mining technique has performed below expectations as identified in related literature. Therefore, this project has implemented a linguistics-based text mining model that incorporates essential elements of customer experience feedback. The text mining model illustrated a higher accuracy level. The project applies the text analytics technique in R that will help in analyzing the customer feedback from a woman’s E-commerce shop to help in solving problems that the business has been experiencing in understanding how satisfied their customers are with the clothing products they deliver to them. The project will generate information from the textual data from the feedback from customers and help to categorize them as either neutral, positive, or negative. The results of this project illustrated the satisfaction and loyalty of the customers for the Woman's eCommerce clothing business. The implementation of this project will help in handling any size of textual data for text analytics.

## Introduction

The technique of collecting and analyzing customer feedback is essential as it provides the businesses with the ability to gain insights into their customer's preferences and act accordingly to any positive and negative reviews. In recent years, there has been the usage of multiple communication channels through which customers utilize to provide feedback concerning any service or product. This has made it tedious for businesses to have a reliable and efficient process through which they can analyze all that information and gain useful insights. With technological advancement, the choice of available channels that organizations use to obtain feedback from their customers has expanded. Most of this is collected in the form of unstructured data, and this has made it challenging for businesses to uncover any meaning from this data. This brings in the need to have automatic models that help with textual analytics in the most efficient and effective way, as discussed in this project. The project is divided into different sections. These sections include a literature review, data, methodology, results, implications, and conclusion.

## Literature Review

The study by Lin & He (2009) concentrated on the research of customer feedback and text mining that was more on extracting information that was related to sentiments and opinions. The authors indicated that at different levels, it is possible to analyze the opinions that are found in the feedback from customers (Li & He, 2009). According to Owens et al. (2009), they described text mining is a technological development that has the potential that is highly commercial. According to a study by Ur-Rahman & Harding (2011), most of the company's information is mostly found in text documents. Ludwig et al. (2013) depicted the benefits that are associated with the process of automating the process of analyzing qualitative customer feedback data, which is essential in informing business strategy.

It has been identified that various techniques have been applied to provide analysis of the customer feedback. However, unlike other studies, this project aims to implement a text analytics model that will be essential in categorizing the feedback into positive, neutral, and negative feedback.

## Data

## X Clothing.ID Age Title

## 1 0 767 33

## 2 1 1080 34

## 3 2 1077 60 Some major design flaws

## 4 3 1049 50 My favorite buy!

## 5 4 847 47 Flattering shirt

## 6 5 1080 49 Not for the very petite

## Review.Text

## 1 Absolutely wonderful - silky and sexy and comfortable

## 2 Love this dress! It's so pretty. I happened to find it in a store, and i'm glad i did bc i never would have ordered it online bc it's petite. i bought a petite and am 5'8". i love the length on me- hits just a little below the knee. would definitely be a true midi on someone who is truly petite.

## 3 I had such high hopes for this dress and really wanted it to work for me. i initially ordered the petite small (my usual size) but i found this to be outrageously small. so small in fact that i could not zip it up! i reordered it in petite medium, which was just ok. overall, the top half was comfortable and fit nicely, but the bottom half had a very tight under layer and several somewhat cheap (net) over layers. imo, a major design flaw was the net over layer sewn directly into the zipper - it c

## 4 I love, love, love this jumpsuit. it's fun, flirty, and fabulous! every time i wear it, i get nothing but great compliments!

## 5 This shirt is very flattering to all due to the adjustable front tie. it is the perfect length to wear with leggings and it is sleeveless so it pairs well with any cardigan. love this shirt!!!

## 6 I love tracy reese dresses, but this one is not for the very petite. i am just under 5 feet tall and usually wear a 0p in this brand. this dress was very pretty out of the package but its a lot of dress. the skirt is long and very full so it overwhelmed my small frame. not a stranger to alterations, shortening and narrowing the skirt would take away from the embellishment of the garment. i love the color and the idea of the style but it just did not work on me. i returned this dress.

## Rating Recommended.IND Positive.Feedback.Count Division.Name Department.Name

## 1 4 1 0 Initmates Intimate

## 2 5 1 4 General Dresses

## 3 3 0 0 General Dresses

## 4 5 1 0 General Petite Bottoms

## 5 5 1 6 General Tops

## 6 2 0 4 General Dresses

## Class.Name

## 1 Intimates

## 2 Dresses

## 3 Dresses

## 4 Pants

## 5 Blouses

## 6 Dresses

Create a corpus for the Review.Text column from the imported dataset.

feedback\_corpus<- Corpus(VectorSource(feedback$Review.Text))

Convert all words in the created corpus to lower case

feedback\_corpus <- tm\_map(feedback\_corpus, tolower)

## Warning in tm\_map.SimpleCorpus(feedback\_corpus, tolower): transformation drops

## documents

Remove punctuation from the corpus

feedback\_corpus <- tm\_map(feedback\_corpus, removeNumbers)

## Warning in tm\_map.SimpleCorpus(feedback\_corpus, removeNumbers): transformation

## drops documents

Remove all stopwords in English from the corpus.

feedback\_corpus <- tm\_map(feedback\_corpus,removeWords , stopwords("english"))

## Warning in tm\_map.SimpleCorpus(feedback\_corpus, removeWords,

## stopwords("english")): transformation drops documents

Also, specific stopwords in the Corpus have been removed.

feedback\_corpus <- tm\_map(feedback\_corpus, removeWords,c("also", "get","like", "company", "made", "can", "im", "dress", "just", "i"))

## Warning in tm\_map.SimpleCorpus(feedback\_corpus, removeWords, c("also", "get", :

## transformation drops documents

Through the use of SnowballC package, the documents will be stemmed.

feedback\_corpus <- tm\_map(feedback\_corpus , stemDocument)

## Warning in tm\_map.SimpleCorpus(feedback\_corpus, stemDocument): transformation

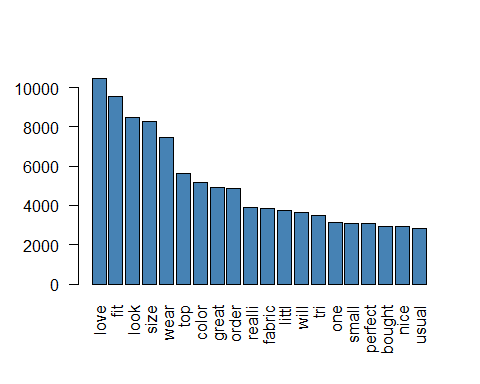
## drops documents

## Methodology

Once the corpus is cleaned, then we create a Document Term Matrix from the preprocessed corpus.

Next, we perform a visualization of the top 20 most frequent words in the corpus.

barplot(feedback\_freq\_term[1:20] , col ="steel blue" , las=2)



Create a worldcloud for the 50 most commonly used words through the use of the wordcloud package.

feedback\_freq\_word <- data.frame(term = names(feedback\_freq\_term) , counts = feedback\_freq\_term)

#Creating wordcloud

wordcloud(feedback\_freq\_word$term , feedback\_freq\_word$counts ,max.words=50 , colors ="red" )

 Next, create a comparison wordcloud for positive and negative feedback concerning the products.

feedback\_Pos <- subset(feedback$Review.Text , feedback$Recommended.IND==1)

feedback\_Neg <- subset(feedback$Review.Text , feedback$Recommended.IND==0)

termsPositive <- paste(feedback\_Pos , collapse =" ")

termsNegative <- paste(feedback\_Neg , collapse =" ")

#Positive and negative feedback combined

terms\_comb <- c(termsPositive, termsNegative)

#Create a corpus for combined terms

corpus\_comb <- VCorpus(VectorSource(terms\_comb))

Create a document Term Matrix for the positive and negative terms combined

options(warn = -1)

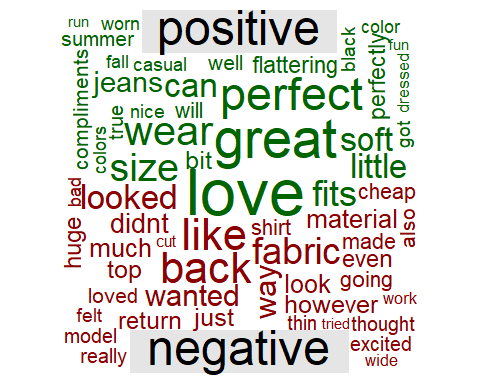
comparison.cloud(

tdm\_comb\_m,

max.words = 100,

colors = c("darkgreen", "darkred")

)

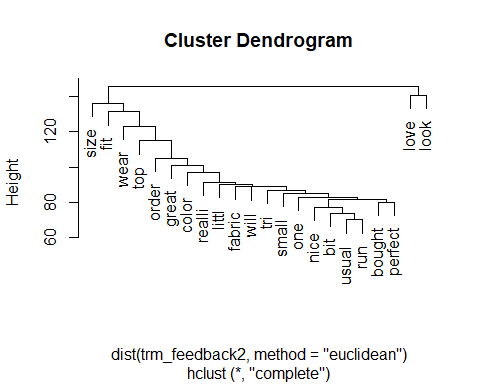
 Identify words used together through word clustering.

trm\_feedback2 <- removeSparseTerms(trm\_feedback, sparse=0.9)

p\_hc <- hclust(d = dist(trm\_feedback2, method = "euclidean"), method = "complete")

# Plot a dendrogram

plot(p\_hc)



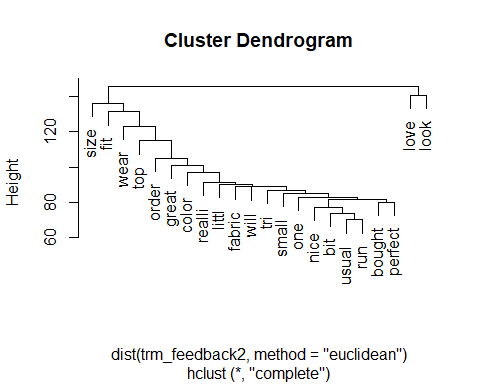
**Results**

The barplot illustrated that the most commonly used words by the customers in their feedback are love, fit, and size. From the word cloud, it is clear that there is size, look, love, and fit are the most commonly used words in the feedback.

From the comparison wordcloud created, the positive feedback and negative feedback have been depicted. From this worldcloud, it is clear that love, great, perfect, and comfortable are the most used positive words from the customer's feedback. On the other hand, fabric, back, like, and wanted are the most used words for negative feedback.



Using word clustering, we were able to identify that love and look, bought and perfect, and usual and run mostly used together in the customers' feedback.



The analysis also shows that 82% of the customers provide positive feedback and they would recommend the products from the store to other people, and only 17% would not recommend them.

##

## 0 1

## 0.1776377 0.8223623

## Conclusion

This paper has illustrated the process of reading data into the R environment, data cleaning, and performing data transformation in R., The paper has also demonstrated how to create worldcloud and identify word associations for feedback from customers. The paper has also helped in identifying what might be the reason for positive and negative feedback from customers regarding the products.

# References

Ur-Rahman, N., & Harding, J. A. (2012). Textual data mining for industrial knowledge management and text classification: A business oriented approach. Expert Systems with Applications, 39(5), 4729-4739.

Lin, C., & He, Y. (2009, November). Joint sentiment/topic model for sentiment analysis. In Proceedings of the 18th ACM conference on information and knowledge management (pp. 375-384).

Ludwig, S., De Ruyter, K., Friedman, M., Brüggen, E. C., Wetzels, M., & Pfann, G. (2013). More than words: The influence of affective content and linguistic style matches in online reviews on conversion rates. Journal of marketing, 77(1), 87-103.

Owens, L., Brown, M., Burnes, S., & Schmidt, P. (2009). Text Analytics Takes Business Insight to New Depths. Forrester Report.