

ibrahim_2019_a_text_analytics_approach_for_online_retaili ng_service_improvement_evidence_from_twitter

Year

2019

Author(s)

Noor Farizah Ibrahim and Xiaojun Wang

Title

A text analytics approach for online retailing service improvement: Evidence from Twitter

Venue

Decision Support Systems

Topic labeling

Manual

Focus

Secondary

Type of contribution

Novel approach

Underlying technique

Manual labeling

Topic labeling parameters

\

Label generation

Literature Analysis of Online Retailing

This study integrates six common constructs from the literature as the main themes. These constructs were selected because they were often discussed in the literature on online retailing and were significant and relevant to this study.

A comprehensive list of themes and factors was created for further analysis.

Theme	Factor	Description	Reference
Delivery	Delivery timeliness	Timely delivery. Receive product according to the promised time	(Cheung, Lee, and Rabjohn 2008; Luo et al. 2012)
	Accuracy	Receive product at the desired location and be billed the correct amount	(Collier and Bienstock 2006; King, Racherla, and Bush 2014)
	Flexibility	Advantage to choose preferred delivery time	(Francis and White 2002; Jiang, Yang, and Jun 2013)
	Product condition	Receive products in good condition	(Collier and Bienstock 2006; Francis 2009)
	Reliability	Reliable delivery	(Ahn, Ryu, and Han 2004)
	Shipping cost	Charges and shipping process	(Jiang and Rosenbloom 2005)
	Tracking	Tracking option	(Jiang and Rosenbloom 2005)
Customer service	Customer support	Follow-up service post-purchase or complaints handling by service personnel	(Francis 2007, 2009)
	Competence	Professional customer service staff perceived as competent	(Ahn et al. 2004; Francis 2009; Francis and White 2002)
	Comprehensiveness	Clear and complete policies and procedures	(Cheung and Thadani 2012; Yoo, Sanders, and Moon 2013)

Etc...

Labeling

The topics were manually labelled into themes or categories based on the combination of human judgement with reference to the literature on online retailing.

We assigned the extracted topics into eight categories, of which six (delivery, customer service, product, security, website, and transaction) were from the comprehensive literature review made for online retailing work.

If the extracted topic could not be associated with any of the six factors, we labelled the terms that can best explain the combination of these keywords.

This labelling and mapping process were based on the researchers' interpretation of the literal meaning of the most frequent keywords and their logical connections with online retailing related themes/factors identified.

Furthermore, time-related keywords (e.g., time, day, today) were included in the topic keywords list.

Subsequently, after the identification, a candidate topic name was further tested via a logical connection to other top 10 words within the same topic based on human judgement.

If a connection was found, the name was confirmed; otherwise, the naming process was

restarted.

To verify the interpretation of topic labels, we also referred to the original tweets that contained these keywords to check whether the topic labels reflected the original meaning of tweets.

Motivation

Listing the importance of factors extracted from the dataset (and used as labels) with the weights measured by the word distribution for each factor.

Topic modeling

LDA

Topic modeling parameters

Nr of topics (k): 72

Nr. of topics

72

Label

Single or multi-word label extracted from one of six factors (delivery, customer service, product, security, website, and transaction) or simply using an arbitrary term (as a fallback mechanism)

Label selection

The labelling process was first conducted by two researchers independently. Then the researchers compared their results and confirmed the topic name by consensus.

Topic No.	Weight	Topic Keywords	Word cloud	Factor		
				Researcher A	Researcher B	Final
39	0.938	Order, delivery, day, today, time, delivered, service, customer, email, and days		Delivery timeliness	Delivery timeliness	Delivery timeliness
32	0.925	Home, week, online, give, price, put, win, delivery, pay, and offer		Delivery timeliness	Flexibility	Flexibility
63	0.811	Disappointed, customers, thought, arrived, box, lot, response, part, shop, and longer		Responsiveness	Responsiveness	Responsiveness
16	0.758	Shopping, year, night, kids, voucher, care, pls, company, friend, and products		Prior experience	Prior experience	Prior experience
58	0.599	Day, big, open, bag, slot, tomorrow, amazing, bags, eat, and due		Sentiment	Flexibility	Flexibility
37	0.532	Store, Christmas, food, today, great, good, time, bought, staff, and shop		Product experience	Brand experience	Store-related issues
11	0.480	Ruined, guarantee, changing, terrible, entire, stated, difficult, load, moon, and complained		Responsiveness	Responsiveness	Responsiveness
27	0.472	Online, poor, waiting, collect, car, milk, people, manager, morning, and turkey		Reliability	Delivery timeliness	Reliability
12	0.303	Prime, account, Christmas, package, book, shipping, app, seller, waiting and tracking		Delivery timeliness	Transaction system	Transaction system
29	0.297	Kindly, suggestion, addressed, handbag, dey, apologize, apt, Michael, regency and diary		Willingness to recommend	Willingness to recommend	Willingness to recommend

The accumulated weight indicates the distribution of words that put the highest weight in that topic

Label quality evaluation

/

Assessors

Two researchers.

Domain

Domain (paper): Online retailing

Domain (corpus): Social media (Twitter)

Problem statement

The purpose of this study is to identify the customers' primary topics of concern regarding online retail brands that are shared among Twitter users.

This study collects tweets associated with five leading UK online retailers covering the period from Black Friday to Christmas and New Year's sales.

We use a combination of text analytical approaches including topic modelling, sentiment analysis, and network analysis to analyse the tweets.

Through the analysis, we identify that delivery, product and customer service are among the most-discussed topics on Twitter. We also highlight the areas that receive the most negative customer sentiments such as delivery and customer service.

Corpus

Origin: Twitter

Nr. of documents: 386,379

Details: The dataset represents the tweets associated with the brands from 20 November 2016 to 20 January 2017, the period covering Black Friday, Christmas, Boxing Day, and New Year's sales events in the UK.

Table 1: Dataset

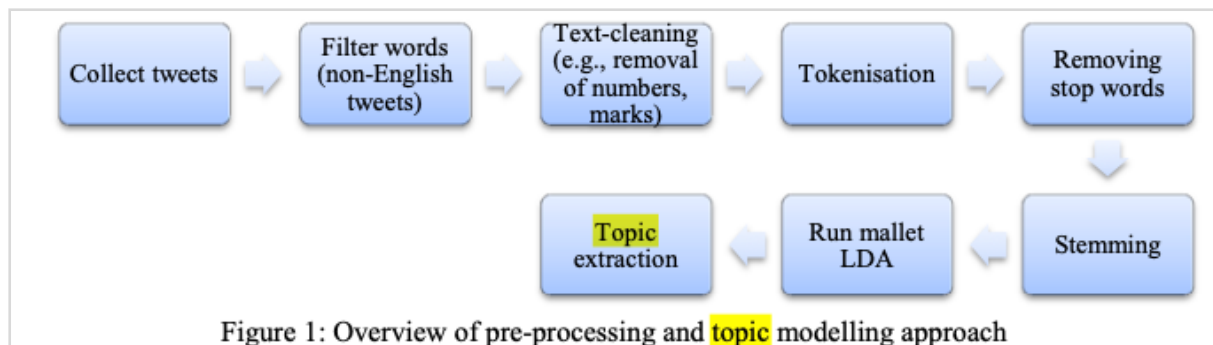
Start Time (GMT)	End Time (GMT)	Amazon # of tweets	Argos # of tweets	Asda # of tweets	John Lewis # of tweets	Tesco # of tweets
Nov 20, 2016 00:00	Jan 20, 2017 23:59	186,885	31,509	60,139	25,991	81,855

Document

Single tweet collected using a customary keyword– [brand name] lang:en since:[start date] until:[end date]

Five leading brands, Amazon UK, Argos, Asda, John Lewis, and Tesco

Pre-processing



@article{ibrahim_2019_a_text_analytics_approach_for_online_retailing_service_improvement_evidence_from_twitter,

abstract = {The purpose of this study is to identify the customers' primary topics of concern regarding online retail brands that are shared among Twitter users. This study collects tweets associated with five leading UK online retailers covering the period from Black Friday to Christmas and New Year's sales. We use a combination of text analytical approaches including topic modelling, sentiment analysis, and network analysis to analyse the tweets. Through the analysis, we identify that delivery, product and customer service are among the most-discussed topics on Twitter. We also highlight the areas that receive the most negative customer sentiments such as delivery and customer service. Interestingly, we also identify emerging topics such as online engagement and in-store experience that are not captured by the existing literature on online retailing. Through a network analysis, we underscore the relationships among those important topics. This study derives insights on how well the leading online retail brands are performing and how their products and services are perceived by their customers. These insights can help businesses understand customers better and enable them to convert the information into meaningful knowledge to improve their business performance. The study offers a novel approach of transforming social media data into useful knowledge about online retailing. The incorporation of three analytical approaches offers insights for researchers to understand the hidden content behind the large collections of unstructured bodies of text, and this information can be used to improve online retailing services and reach out to customers.},

author = {Noor Farizah Ibrahim and Xiaojun Wang},

date-added = {2023-03-09 18:32:38 +0100},

date-modified = {2023-03-09 18:32:38 +0100},

```
doi = {https://doi.org/10.1016/j.dss.2019.03.002},  
issn = {0167-9236},  
journal = {Decision Support Systems},  
keywords = {Online retailing, Social media research, Text analytics, Topic  
modelling, Sentiment analysis},  
pages = {37-50},  
title = {A text analytics approach for online retailing service improvement:  
Evidence from Twitter},  
url = {https://www.sciencedirect.com/science/article/pii/S0167923619300405},  
volume = {121},  
year = {2019}}
```

#Thesis/Papers/Initial