liu_2017_an_investigation_of_brand_related_user_generated_content_on_twitter

Year

2017

Author(s)

Xia Liu and Alvin C. Burns and Yingjian Hou

Title

An Investigation of Brand-Related User-Generated Content on Twitter

Venue

Journal of Advertising

Topic labeling

Manual

Focus

Secondary

Type of contribution

Established approach

Underlying technique

Manual labeling (assisted by pre-defined label set)

Topic labeling parameters

\

Label generation

One of the critical tasks after running LDA is to give labels to these topics based on research requirements and domain knowledge.
Because our purpose here is to compare brands within and across industries, we opted to use a generic list of topic labels:

- product
- service
- promotion
- competitors
- news / trend
- shows / games
- price
- location

TABLE 3
Burger King Topic Proportion, Top Words, and Labels

Number	Proportion	Top 10 Words	Label
T0	.045	order (.080); drive (.040); wrong (.032); home (.031); time (.023); minutes (.020); wait (.016); delivery (.015); cream (.015); line (.015)	Customer service
T1	.022	menu (.101); kids (.086); meals (.048); meal (.045); soda (.035); drinks (.028); keyboard (.024); kid (.020); soft (.012); item (.012)	Products
T2	.029	McDonald's (.145); Wendy's (.066); Taco Bell (.054); shake (.045); red (.040); velvet (.040); Oreo (.039); KFC (.036); Subway (.030); Pizza Hut (.022)	Competitors
T3	.253	fries (.298); chicken (.223); back (.138); nuggets (.031); happy (.021); day (.019); life (.013); forever (.011); strips (.010); fry (.009)	Promotions

"After the researchers assign labels to those topics, it might be wise to employ the assistance of domain experts to verify the validity of problematic labels."
(No mention on whether it was actually done in this case)

Motivation

We […] categorized the 300 topics using these labels and averaged the incidences of each topic label within each industry.

TABLE 4
User-Generated Content Topics: Industry Similarities and Differences

Topics/Industry	Fast Food	Department Store	Footwear	Telecommunications	Electronics	Average
Product	47.9%	23.5%	55.9%	45.6%	70.6%	48.7% (17.2%)
Service	20.4%	29.6%	1.8%	28.5%	8.0%	17.6% (12.4%)
Promotion	15.7%	24.6%	15.3%	5.0%	5.8%	13.3% (8.1%)
Competitors	4.8%	5.4%	7.7%	5.7%	9.0%	6.5% (1.8%)
News/trends	6.7%	9.2%	9.2%	2.6%	5.1%	6.6% (2.8%)

Note. Standard deviation in parentheses.

Topic modeling

LDA

Topic modeling parameters

Nr of topics (k): 15 (15 topics for each of the 20 brands in the data set)

Nr. of topics

300

Label

Manually assigned pre-defined single word label: (product, service, promotion, competitors, news / trend, shows / games, price, location.)

Label selection

\

Label quality evaluation

\

Assessors

\

Domain

Paper: Social media analysis
Dataset: Social media (Twitter)

Problem statement

This article presents a framework that automatically derives latent brand topics and classifies brand sentiments.

It applies text mining with latent Dirichlet allocation (LDA) and sentiment analysis on 1.7 million unique tweets for 20 brands across five industries: fast food, department store, footwear, electronics, and telecommunications.
The framework is used to explore four brand-related questions on Twitter.

Corpus

Origin: Twitter
Nr. of documents: 1.7M
Details:

TABLE 2
Summary Tweet Information for Industries and Brands

Industry	Brands and Tweets			
Fast-food restaurant	McDonald's (318,003)	Burger King (122,075)	Wendy's (84,219)	KFC (70,533)
Department store	JCPenney (43,887)	Macy's (184,715)	Sears (33,005)	Kohl's (53,469)
Footwear	Nike (151,437)	New Balance (27,205)	Adidas (57,987)	Puma (23,427)
Electronics	LG (32,230)	Panasonic (4,286)	Samsung (14,857)	Sony (131,264)
Telecommunications	Comcast (261,914)	TWC (57,771)	Dish (35,603)	Cox (20,993)

Document

Text of a single Tweet

Pre-processing

- remove tweets created by Twitterbots
- filtered out hashtags and URLs
- tokenisation
- POS tagging
- Stop word removal

```
@article{liu_2017_an_investigation_of_brand_related_user_generated_content_on_t
witter,
  author = {Xia Liu and Alvin C. Burns and Yingjian Hou},
  date-added = {2023-03-31 20:24:22 +0200},
  date-modified = {2023-03-31 20:24:22 +0200},
  doi = {10.1080/00913367.2017.1297273},
  journal = {Journal of Advertising},
  month = {mar},
```

```
  number = {2},
  pages = {236--247},
  publisher = {Informa {UK} Limited},
  title = {An Investigation of Brand-Related User-Generated Content on
Twitter},
  url = {https://doi.org/10.1080%2F00913367.2017.1297273},
  volume = {46},
  year = 2017}
```

#Thesis/Papers/BS