DETC2013-12661

Fad or Here to Stay: Predicting Product Market Adoption and Longevity Using Large Scale, Social Media Data

Tuesday, August 6th , 2013

Suppawong Tuarob

Computer Science and Engineering

Conrad S. Tucker

Engineering Design and Industrial and Manufacturing Engineering

The Pennsylvania State University

{suppawong,ctucker4}@psu.edu

Overview

- Problem definition
- Social media vs Web-based information
- Use of social media in literature
- Main contributions
- Past literature
- Study 1: Predictive Power of Social Media on Product Demand
- Study 2: Predicting Product Longevity
- Study 3: Mining Product Features

Problem Definition

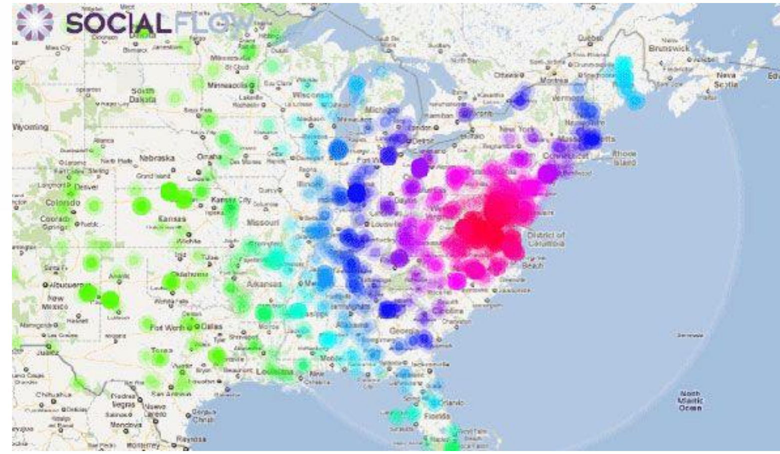
- Social media has recently been acknowledged as a powerful source of information
  - Large quantity
  - Available Instantly
- Can social media serve as a good indicator in product market domain?

Social Media vs Web-based information

Social Media	Web-based Information (Blogs, Reviews, Etc.)
Promptly available	Takes some amount of time for pre-publishing process including verifying content and proofreading.
Abundant	Limited amount of data
Contains opinions and rumor (Good for predictions)	Tend to be more factual

Use of Social Media in Literature

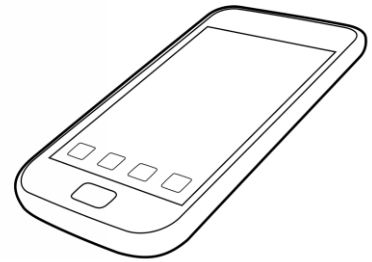
- Knowledge extracted from social media has proven valuable in various applications.
  - Earthquake warning [18]
  - Identifying emergency needs during disasters [6]
  - Predicting stock market trends [4]
  - Etc.



Selected Related Literature

- Product feature analysis
  - Lim et al.[13] proposed a Baysian network for modeling user preferences on product features.
  - Tucker and Kim [24] propose a machine learning based approach for mining product feature trends in the market from the time series of user preferences.
  - Wan et al. [25] define a feature as a set of connective attributes of edges or faces, such as convexity, concavity, and tangency.
- Mining information in social media
  - Asur et al.[2] successfully used tweets collected during a 3 month period to predict box office revenues.
  - Bollen et al.[4] showed that the changes of moods in Twitter correlate with the shifts in the Dow Jones Industrial Average that occur 3 to 4 days later.
  - Paul and Dredze[16] identify 14 unique ailments along with terms that represent the symptoms of each ailment from Twitter data using a topic modelling approach.

Fad or Here to Stay?



Can social media indicate consumer demand?

Study 1: Predicting product demand using sentiment in social media

Can a product stay competitive in the market?

Study 2: Quantifying product longevity

Why is a product a fad or a here-to-stay?

Study 3: Identifying notable product features

Methodology

- Predicting product demand
  - Using social media to predict product sales
  - Analyzing buzzing effects on social media
- Predicting product longevity
  - Using social media to quantify the ability to stay competitive in the market for a particular product
- Extracting notable product features
  - Using social media to identify strong/weak/controversial features of a particular product

Case Study: Twitter and Smartphones

- Social Media: Twitter
  - Publicly and massively available
  - ~ 800 million tweets in the United States during the period of 19 months, from March 2011 to September 2012
- Study Product: Smartphones
  - Obtained a smartphone database from GSMArena.com consisting of 2,547 smartphone models manufactured by 33 different companies

Study 1: Predictive Power of Social Media on Product Demand

- How does collective sentiment in social media reflect the buying decision of the consumers?
- Can social media be used to predict the product sales?
- Is social media a good place to spread product rumors (buzzes)?

Sentiment Analysis In Social Media

- Subjective messages tend to express disappointment or satisfaction relating to a particular product/product feature.
  - Absolutely loving my new iPhone 4 (p.s. I wrote this tweet with #siri lol)
  - I hate the fact that my iPhone 4 home button is intermittently unresponsive.
- We used ***SentiStrength***** algorithm to quantify emotion level in social media text. Then compute ***Emotional Strength*** score using:

$$Emotion\ Strength(ES) = NegativeScore - PositiveScore$$

- A social text is then classified into one of the 3 categories based on the sign of Emotion Strength score (i.e. positive (+ve) /neutral (0ve)/negative (-ve))

*Mike Thelwall, Kevan Buckley, Georgios Paltoglou, Di Cai, and Arvid Kappas. Sentiment in short strength detection informal text. J. Am. Soc. Inf. Sci. Technol., 61(12):2544–2558, December 2010.

Study 1: Prediction Methodology

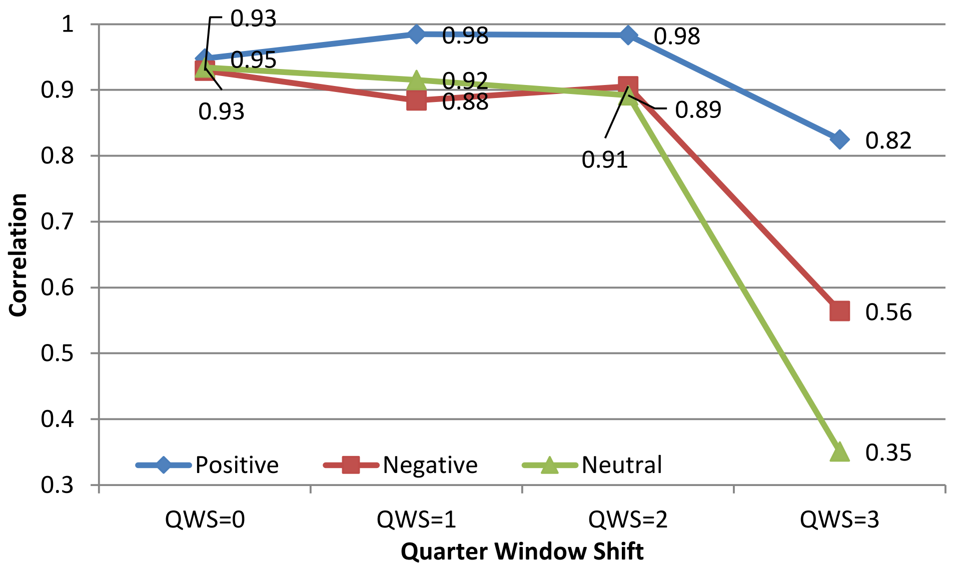
- Use collected tweets in the past to predict current actual product sales.

<table>
<tr><th rowspan="2">QWS</th><th colspan="3">2010</th><th colspan="9">2011</th></tr>
<tr><th>Oct</th><th>Nov</th><th>Dec</th><th>Jan</th><th>Feb</th><th>Mar</th><th>Apr</th><th>May</th><th>Jun</th><th>Jul</th><th>Aug</th><th>Sep</th></tr>
<tr><td>0</td><td colspan="3"></td><td colspan="3">Q1</td><td colspan="3">Q2</td><td colspan="3">Q3</td></tr>
<tr><td>1</td><td colspan="2"></td><td colspan="3">Q1</td><td colspan="3">Q2</td><td colspan="3">Q3</td><td>Q4</td></tr>
<tr><td>2</td><td></td><td colspan="3">Q1</td><td colspan="3">Q2</td><td colspan="3">Q3</td><td colspan="2">Q4</td></tr>
<tr><td>3</td><td colspan="3">Q1</td><td colspan="3">Q2</td><td colspan="3">Q3</td><td colspan="3">Q4</td></tr>
</table>

An illustration of the shifts of quarter windows with various Quarter Window Shift (QWS) values.

Study 1: Result Highlight

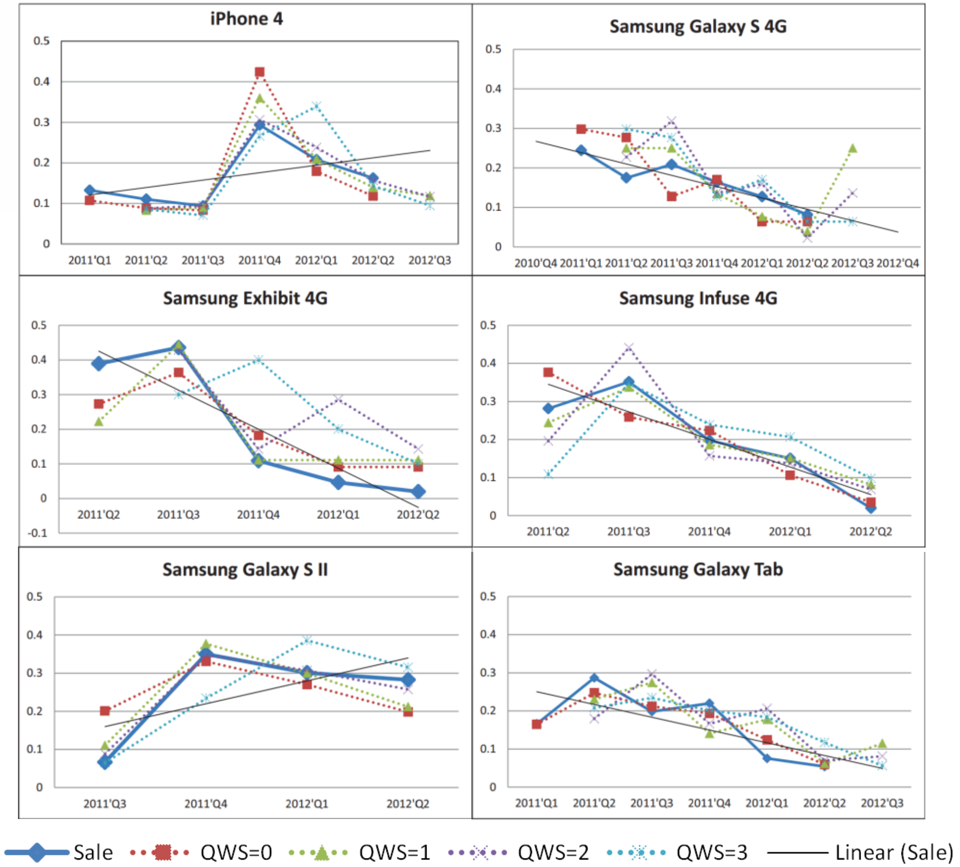
- How does collective sentiment in tweets reflect the buying decision of the consumers?



Correlations of the iPhone 4 sale prediction using different emotion type of tweets.

Study 1: Result Highlight

- Can tweets be used to predict the product sales?
  - The ***+ve*** tweets continue to be a good predictor for up to 2 months based on the average correlation.
  - The predictability of tweets on product sales varies significantly across different smartphone models.



Sale prediction of different smartphone models. X axis represents Quarter Window Shift (QWS). Y axis represents the normalized values of actual sale and the quantity of ***+ve*** tweets at each QWS.

Model	Correlation				MAPE				MSE			
	QWS=0	QWS=1	QWS=2	QWS=3	QWS=0	QWS=1	QWS=2	QWS=3	QWS=0	QWS=1	QWS=2	QWS=3
iPhone 4	0.9476	**0.9845**	0.9834	0.8249	0.2268	0.1335	**0.0867**	0.3896	0.0035	0.0012	0.0003	0.0182
Samsung Galaxy S 4G	0.7561	0.9203	**0.9362**	0.8306	0.3240	0.3459	0.3933	0.3615	0.0040	0.0025	0.0040	0.0047
Samsung Galaxy S II	0.6890	0.9150	**0.9912**	0.7966	0.6187	0.2491	0.1122	0.1883	0.0066	0.0019	0.0003	0.0054
Samsung Exhibit	**0.9663**	0.8708	0.1423	-0.1163	1.1386	1.2971	2.5452	2.2725	0.0062	0.0081	0.0451	0.0570
Samsung Infuse 4G	0.8493	**0.9812**	0.8720	0.6319	0.3750	0.8176	0.7652	1.1512	0.0029	0.0010	0.0031	0.0028
Samsung Galaxy Tab	**0.9579**	0.6304	0.4007	0.7418	0.1784	0.4887	0.6289	0.6303	0.0008	0.0051	0.0083	0.0047
Average	0.8610	**0.8837**	0.7215	0.5122	0.4769	0.5553	0.7553	0.8322	0.0040	0.0033	0.0102	0.0155

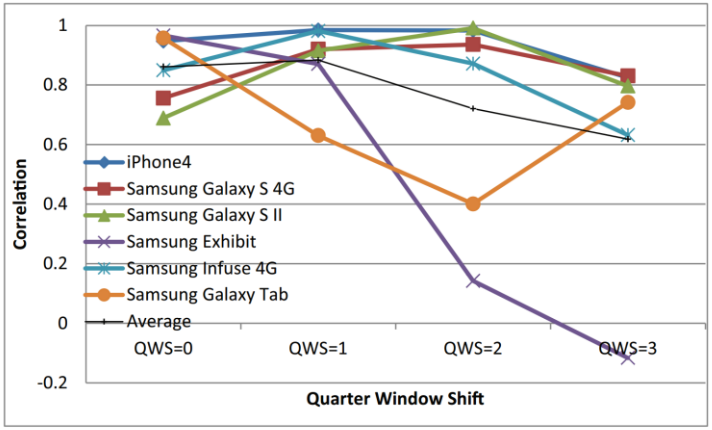
Correlation, MAPE, and MSE of the predictions with QWS varying from 0 to 3 of each model.

Study 1: Result Highlight

- Is Twitter a good place to spread product rumors (buzzes)?
  - Oftentimes, buzzes or short-term rumors are observed before a product is even launched.
  - Popular products (e.g. iPhone4) tend to have a lasting and powerful buzz before the products are even launched.
  - Some products (e.g. Samsung Exhibit) may draw attention from the consumers just a short time before they are released.

Model	Release Date	3 Month Hits	1st Qtr Sale
iPhone 4	June 24, 2010	1,220,000	6810
Samsung Galaxy S 4G	February 1, 2011	910,000	354
Samsung Galaxy S II	April 1, 2011	840,000	275
Samsung Exhibit	June 1, 2011	21,100	118
Samsung Infuse 4G	May 15, 2011	42,300	288
Samsung Galaxy Tab 4G	January 6, 2011	4,240	77

The number of hits returned by the Google search engine, using each smartphone names as search queries, during 3 month period before each model is released.



Correlation of the sale prediction at each QWS

Study 2: Predicting Product Longevity

- Aims to investigate whether social media data can be used to quantify the ***longevity***, the ability to stay competitive in the market, of a particular product.
  - $S = \{s_1, s_2, \dots, sn\}$ be the set of products
  - $Positive(si)/Negative(si)/Neutral(si)$ are set of ***+ve/+ve/0ve*** tweets corresponding to product s_i.

$$Polarity(s_i) = \frac{|Positive(s_i)|}{|Positive(s_i) \bigcup Negative(s_i)|}$$

$$Subjectivity(s_i) = \frac{|Positive(s_i) \bigcup Negative(s_i)|}{|Positive(s_i) \bigcup Negative(s_i) \bigcup Neutral(s_i)|}$$

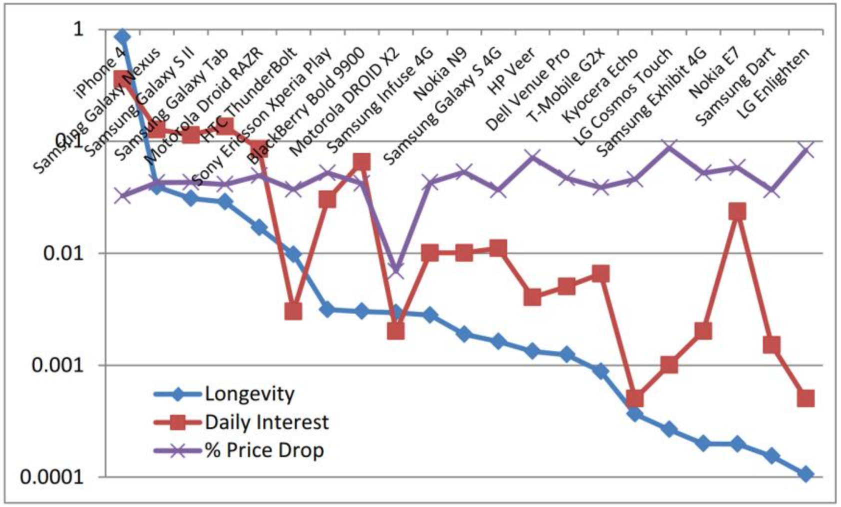
$$Favorability(s_i) = \frac{|Positive(s_i) \bigcup Neutral(s_i)|}{\sum_{s \in S} |Positive(s) \bigcup Negative(s) \bigcup Neutral(s)|}$$

$$Longevity(s_i) = Polarity(s_i) \times Subjectivity(s_i) \times Favorability(s_i)$$

Returns a real number between 0 and 1, and is served as a comparable score, instead of an absolute score.

Study 2: Result Highlight

- The Longevity scores are computed for the 21 smartphone models.
- The scores are compared with the GSMArea.com's Daily Interest rates and the percentages of price drops.
- A high correlation of 0.8743 is observed between the Longevity scores and the GSMArea Daily Interest rates.
- However, the correlation between the Longevity scores and the percentage price drops is weak (-0.2113).
  - There are multiple reasons to drop prices!



Comparison between the Longevity score vs GSMArea Daily Interest and % Price Drop for each sample smartphone model.

Study 3: Mining Product Features

- Comments that express **positive** or **negative** sentiment about a product usually infer information about the product features.
- The feature extraction problem is transformed into the *term ranking* problem.
- Strong/Weak/Controversial features of product s_i are extracted from $Positive(s_i)$/$Negative(s_i)$/$Positive(s_i) \cup Negative(s_i)$ respectively.

FaceTime Iss Amazing:) #iPhone4

I hate the samsung galaxy s2 homescreen it keeps not workingUghh

The Term Ranking Algorithm

- A document ***d*** is a bag of terms $\{t_1,t_2,\ldots,t_n\}$.
- Given a set of documents ***D*** = $\{d_1, d_2,\ldots, d_m\}$ and subset **θ** $\subseteq$ ***D***:

STEP1 The set of all distinct terms T are extracted from D.
STEP2 For each term $t \in T$, compute $P(t|\theta,D,T,M)$, the likelihood of the term t given θ,D,T, and M (where M, Calculation Method, can be either TFIDF or LDA).
STEP3 Rank the terms by their likelihood.
STEP4 Top K terms are returned.

- The positive/negative features of the product s are the first K ranked terms, where ***D*** = Positive(s) ∪ Negative(s) ∪ Neutral(s) and ***q*** = Positive(s)/Negative(s) respectively.
- The controversial features are the first ***K*** terms ranked by the controversial scores defined as:

$$Controversial(t) = P(t|Postive(s),D,T,M) \times P(t|Negative(s),D,T,M)$$

TFIDF-based P(t|q,θ,T,*TFIDF*)

- TFIDF reflects how important a term is to a document in a corpus.
- TFIDF has two components: the term frequency (TF) and the inverse document frequency (IDF):

$$tf(t,d) = \sqrt{count(t,d)}$$

$$idf(t,D) = \sqrt{log\left(\frac{|D|}{|d \in D; t \in d|}\right)}$$

$$tfidf(t,d,D) = TF(t,d) \cdot IDF(t,D)$$

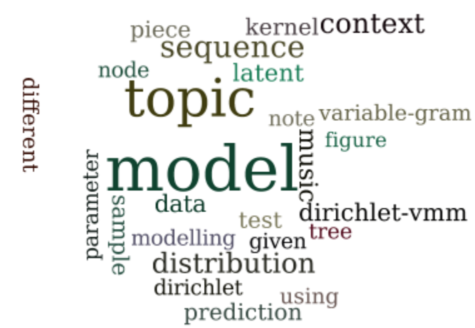
- Finally, **P(t|q,θ,T,TFIDF)** is computed by normalizing **tfidf(t,d,D)**.

$$P(t|\theta,D,T,TFIDF) = \frac{tfidf(t,d,D)}{\sum_{\tau \in T} tfidf(\tau,d,D)}$$

LDA-based $\mathbf{P(t|q,\theta,T,\mathit{LDA})}$

- Latent Dirichlet Allocation (LDA)* is a generative model that allows a document to be represented by a mixture of topics.
- Mathematically, the LDA model is described as following:

$$P(t_i|d) = \sum_{j=1}^{|Z|} P(t_i|z_i = j) \cdot P(z_i = j|d)$$



  - $P(t_i|d)$ is the prob. of term t_i being in document *d*.
  - z_i is the latent (hidden) topic
  - $P(t_i|z_i = j)$ is the prob. of term t_i being in topic *j*
  - $P(z_i=j|d)$ is the prob. of picking a term from document *d*.
  - $|Z|$ is the total number of topics (need to specify this)

* David M. Blei, Andrew Y. Ng, and Michael I. Jordan. Latent dirichlet allocation. J. Mach. Learn. Res., 3:993–1022, March 2003.

LDA-based $\mathbf{P(t|q,\theta,T,}\boldsymbol{LDA}\mathbf{)}$

- 5 topics are modeled from $\boldsymbol{\theta}$, using Stanford Topic Modeling Toolbox with 3,000 iterations and the collapsed variational Bayes approximation to the LDA objective.
- Two topics whose numbers of feature terms within the first 30 terms ranked by ***P(t |z)*** are chosen.
- The term distribution of the two chosen topics are averaged to represent the new term distribution of the merged topic ***z****:

$$P(t|\theta, D, T, LDA) = P(t|z^*)$$

Study 3: Result Highlight

Features extracted from tweets related to each selected smartphone model using the TFIDF based feature extraction algorithm												
Features	iPhone 4			Samsung Galaxy S II			Motorola Droid RAZR			Sony Ericsson Xperia Play		
	Strong	Weak	Controversial	Strong	Weak	Controversial	Strong	Weak	Controversial	Strong	Weak	Controversial
1	case	phone	case	touch-screen	touch-screen	touch-screen	battery-life	touch-screen	android	game	network	game
2	camera	case	face-time	update	update	update	commercial	update	battery-life	play	play	play
3	face-time	face-time	camera	ics	screen	screen	update	screen	commercial	control	game	network
4	app	battery-life	battery-life	battery-life	video	battery-life	screen	video	update	controller	video	game
5	screen	camera	app	screen	note	sensation	ics	sensation	screen	playstation	control	control
6	life	screen	screen	sensation	sensation	ics	app	battery-life	cream	commercial	picture	commercial
7	battery-life	app	texting	support	battery-life	contract	keyboard	case	ics	freebie	style	picture
8	price	life	update	message	case	camera	message	ics	app	bootloader	data	battery-life
9	update	update	video	upgrade	ics	internet	picture	carrier	price	battery-life	emulator	playstation
10	video	video	price	camera	function	app	price	function	release	carrier	cpu	integration

Features extracted from tweets related to each selected smartphone model using the LDA based feature extraction algorithm												
Features	iPhone 4			Samsung Galaxy S II			Motorola Droid RAZR			Sony Ericsson Xperia Play		
	Strong	Weak	Controversial	Strong	Weak	Controversial	Strong	Weak	Controversial	Strong	Weak	Controversial
1	camera	battery-life	battery-life	touch-screen	touch-screen	touch-screen	battery-life	keys	picture	game	game	game
2	battery-life	face-time	face-time	update	function	email	screen	price	price	battery-life	accessories	video
3	screen	app	camera	battery-life	email	video	picture	browser	browser	control	video	commercial
4	app	video	app	screen	video	bootloader	android	bootloader	webpage	fun	battery-life	control
5	price	jailbreak	video	ics	bootloader	photo	glass	warranty	life	hardware	commercial	gaming
6	music	wifi	update	sensation	photo	texting	app	microphone	music	performance	style	battery-life
7	face-time	bug	voice-control	display	gallery	price	camera	delay	update	experience	control	baseball
8	message	charge	wifi	video	button	jelly bean	keyboard	bloatware	screen	wifi	app	hardware
9	voice-control	location	screen	app	texting	app	network	fixes	touch-screen	video	size	experience
10	case	touch-screen	case	picture	price	network	noise	email	android	controller	carrier	controller

- LDA based approach yields better performance with the average Precision@50 of 0.47/0.27/0.36 for the strong/weak/controversial features respectively.

Conclusion

- A sequence of studies involving using social media such as tweets to extract meaningful information about products is presented.
- The first study involves investigating the predictability of tweets on selected 6 smartphone models.
  - The predictive ability varies across products, leading to a further study on the buzzing effect of social media.
  - The study shows that tweets can be a good medium for spreading rumors about products.
- The second study attempts to quantify the ability to stay in the market of individual products using their corresponding tweets.
  - A high correlation between the results from the proposed mathematical model and the today's current interest rates from end users is observed.
- The third study employs the information retrieval techniques (Term Frequency-Inverse Document Frequency and Latent Dirichlet Allocation) to extract notable strong/weak/controversial features from individual products.

References

- [1] Arthur Asuncion, Max Welling, Padhraic Smyth, and Yee Whye Teh. On smoothing and inference for topic models. In Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence, UAI '09, pages 27–34, Arlington, Virginia, United States, 2009. AUAI Press.
- [2] Sitaram Asur and Bernardo A. Huberman. Predicting the future with social media. In Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology - Volume 01, WIIAT '10, pages 492–499, Washington, DC, USA, 2010. IEEE Computer Society.
- [3] David M. Blei, Andrew Y. Ng, and Michael I. Jordan. Latent dirichlet allocation. J. Mach. Learn. Res., 3:993– 1022, March 2003.
- [4] J. Bollen, H. Mao, and X. Zeng. Twitter mood predicts the stock market. Journal of Computational Science, 2(1):1– 8, 2011.
- [5] J. Boyle, M. Jessup, J. Crilly, D. Green, J. Lind, M. Wallis, P. Miller, and G. Fitzgerald. Predicting emergency department admissions. Emergency Medicine Journal, 29(5):358–365, 2012.
- [6] C. Caragea, N. McNeese, A. Jaiswal, G. Traylor, H.W. Kim, P. Mitra, D. Wu, A.H. Tapia, L. Giles, B.J. Jansen, et al. Classifying text messages for the haiti earthquake. In Proceedings of the 8th International Conference on Information Systems for Crisis Response and Management (ISCRAM2011), 2011.
- [7] Nigel Collier and Son Doan. Syndromic classification of twitter messages. CoRR, abs/1110.3094, 2011.
- [8] Kushal Dave, Steve Lawrence, and David M. Pennock. Mining the peanut gallery: opinion extraction and semantic classification of product reviews. In Proceedings of the 12th international conference on World Wide Web, WWW '03, pages 519–528, New York, NY, USA, 2003. ACM.
- [9] Kevin Gimpel, Nathan Schneider, Brendan O'Connor, Dipanjan Das, Daniel Mills, Jacob Eisenstein, Michael Heilman, Dani Yogatama, Jeffrey Flanigan, and Noah A. Smith. Part-of-speech tagging for twitter: annotation, features, and experiments. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers -Volume 2, HLT '11, pages 42–47, Stroudsburg, PA, USA, 2011. Association for Computational Linguistics.
- [10] Saurabh Kataria, Prasenjit Mitra, and Sumit Bhatia. Utilizing context in generative bayesian models for linked corpus. In AAAI'10, pages –1–1, 2010.
- [11] Ralf Krestel, Peter Fankhauser, and Wolfgang Nejdl. Latent dirichlet allocation for tag recommendation. In Proceedings of the third ACM conference on Recommender systems, RecSys '09, pages 61–68, New York, NY, USA, 2009. ACM.
- [12] Onur Kucuktunc, B. Barla Cambazoglu, Ingmar Weber, and Hakan Ferhatosmanoglu. A large-scale sentiment analysis for yahoo! answers. In Proceedings of the fifth ACM international conference on Web search and data mining, WSDM '12, pages 633–642, New York, NY, USA, 2012. ACM.
- [13] Soon Chong Johnson Lim, Ying Liu, and Han Tong Loh. An exploratory study of ontology-based platform analysis under user preference uncertainty. In Proc. ASME 2012 Int. Design Engineering Technical Conf. Computers and Information in Engineering Conf. (IDETC/CIE2012), 2012.
- [14] B. OConnor, R. Balasubramanyan, B.R. Routledge, and N.A. Smith. From tweets to polls: Linking text sentiment to public opinion time series. In Proceedings of the International AAAI Conference on Weblogs and Social Media, pages 122–129, 2010.
- [15] M.J. Paul and M. Dredze. You are what you tweet: Analyzing twitter for public health. In Fifth International AAAI Conference on Weblogs and Social Media (ICWSM 2011), 2011.
- [16] M.J. Paul and M. Dredze. A model for mining public health topics from twitter. HEALTH, 11:16–6, 2012.
- [17] Ana-Maria Popescu and Oren Etzioni. Extracting product features and opinions from reviews. In Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing, HLT '05, pages 339–346, Stroudsburg, PA, USA, 2005. Association for Computational Linguistics.
- [18] Takeshi Sakaki, Makoto Okazaki, and Yutaka Matsuo. Earthquake shakes twitter users: real-time event detection by social sensors. In Proceedings of the 19th international conference on World wide web, WWW '10, pages 851–860, New York, NY, USA, 2010. ACM.
- [19] Mike Thelwall, Kevan Buckley, and Georgios Paltoglou. Sentiment in twitter events. J. Am. Soc. Inf. Sci. Technol., pages 406–418, 2011.
- [20] Mike Thelwall, Kevan Buckley, Georgios Paltoglou, Di Cai, and Arvid Kappas. Sentiment in short strength detection informal text. J. Am. Soc. Inf. Sci. Technol., 61(12):2544–2558, December 2010.
- [21] Suppawong Tuarob, Line C. Pouchard, and C. Lee Giles. Automatic tag recommendation for metadata annotation using probabilistic topic modeling. In Proceedings of the 13th ACM/IEEE-CS joint conference on Digital Libraries, JCDL '13, 2013.
- [22] Suppawong Tuarob, Line C Pouchard, Natasha Noy, Jeffery S Horsburgh, and Giri Palanisamy. Onemercury: Towards automatic annotation of environmental science metadata. In Proceedings of the 2nd International Workshop on Linked Science 2012, LISC '12, 2012.
- [23] C. Tucker and H. Kim. Predicting emerging product design trend by mining publicly available customer review data. In Proceedings of the 18th International Conference on Engineering Design (ICED11), Vol. 6, pages 43–52, 2011.
- [24] C.S. Tucker and H.M. Kim. Trend mining for predictive product design. Transactions of the ASME-R-Journal of Mechanical Design, 133(11):111008, 2011.
- [25] Sha Wan, Yunbao Huang, Qifu Wang, Liping Chen, , and Yuhang Sun. A new approach to generic design feature recognition by detecting the hint of topology variation. In Proc. ASME 2012 Int. Design Engineering Technical Conf. Computers and Information in Engineering Conf. (IDETC/CIE2012), 2012.
- [26] Ingmar Weber, Antti Ukkonen, and Aris Gionis. Answers, not links: extracting tips from yahoo! answers to address how-to web queries. In Proceedings of the fifth ACM international conference on Web search and data mining, WSDM '12, pages 613–622, New York, NY, USA, 2012. ACM.
- [27] Bing Xu, Tie-Jun Zhao, De-Quan Zheng, and Shan-Yu Wang. Product features mining based on conditional random fields model. In Machine Learning and Cybernetics (ICMLC), 2010 International Conference on, volume 6, pages 3353 –3357, july 2010.
- [28] X. Zhang, H. Fuehres, and P. Gloor. Predicting asset value through twitter buzz. Advances in Collective Intelligence 2011, pages 23–34, 2012.
- [29] Xiao Zhang and Prasenjit Mitra. Learning topical transition probabilities in click through data with regression models. In Procceedings of the 13th International Workshop on the Web and Databases, WebDB '10, pages 11:1– 11:6, New York, NY, USA, 2010. ACM.