Systematic literature review on customer emotions in social media

Prashanth Madhala[1], Jari Jussila[1,2], Heli Aramo-Immonen[3], Anu Suominen[1]
[1]Tampere University of Technology, Finland
[3]University of Jyväskylä, Finland
[2]Örebro University, Sweden
prashanth.madhala@student.tut.fi
jari.j.jussila@jyu.fi
heli.aramo-immonen@oru.se
anu.suominen@tut.fi

Abstract:
Customers are human beings who express their emotions openly on social media platforms. There is a wealth of social media data that companies can make use of to improve their business decision making and tailor their marketing strategies. In order to benefit from this, organizations need to apply computational methods, which can save time and effort rather than applying traditional consumer research approaches, such as surveys or interviews. The purpose of this study is to investigate existing computational studies on detecting consumer emotions from social media data. We conducted a systematic literature review on articles published in ScienceDirect, IEEE Explore, ACM Digital Library, and Emerald Insight from the period 2009-2017. The aim was to discover how social media data was extracted, how large datasets were used in detecting emotions, the type of computational methods used, and the accuracy of the results obtained from the existing studies. Most of the studies were focused on sentiment analysis and different machine learning algorithms. The computational methods were applied in business decision making and marketing functions. Practical scenarios included emotion detection in customer reviews and sentiment analysis of retail brands. Based on these studies, we have uncovered situations where the results of the analysis are either sufficiently accurate or supportive for decision making. We provide recommendations for organizations and managers on developing their resources to make use of different computational methods for emotion and sentiment detection. Finally, we present the limitations of these methods and provide recommendations for aligning future research studies toward big social data analytics on customer emotions.

Keywords: Social media, big data, emotions, consumer behavior, sentiment analysis

1. Introduction

The widespread diffusion of social media platforms, such as Facebook and Twitter, are driving the world-wide spread of the phenomenon called 'Big Data' (Fosso Wamba *et al.*, 2015). Moreover, social media has transformed the ways in which companies and customers interact. A great deal of social media interaction is emotionally loaded. Furthermore, competition is increasingly based on a brand's ability to inspire emotional experiences. (Jalonen and Jussila, 2016) According to a recent study, fully emotionally connected customers are 52% more valuable on average than just highly satisfied ones (Magids, Zorfas, and Leemon, 2015). In fact, for more than a decade, emotions have been recognized as being essential in marketing and consumer behavior studies (Laros and Steenkamp, 2005). These studies have highlighted the fact that, instead of rational decision making based on utilitarian product attributes and benefits, consumer decisions are "biased" by emotions (Jalonen and Jussila, 2016). There is a wealth of social media data on consumer emotions available for companies to collect and analyze, yet the sheer volume of the data makes its manual analysis difficult, time-consuming, and costly. However, social media analytics tools and methods can provide a cost-effective way to gather relevant data and the means of processing the data into knowledge, enabling more accurate and valuable marketing decisions (Jussila, Boedeker *et al.*, 2017)

In their systematic review and longitudinal case study, Fosso Wamba *et al.* (2015) studied the impacts that Big Data may bring about. They discovered that big data should be used for understanding customers better. Therefore, Fosso Wamba *et al*. (2015) called for future research in the big data domain, particularly focusing on developing explanatory and predictive theories that encompass the cross functional facets of the domain. In the literature, Big Social Data has been further distinguished from the broader category of big data. Big data has been, for instance, conceptualized as any data that is produced as a result of the quantification of the world that may include data from sensors, multiple industrial and domestic networks as well as financial markets, whereas big social data is produced as a result of the mediated communication practices of our everyday life, “whenever we go online, use our smartphone, use an app or make a purchase” (Coté, 2014). In recent years, several approaches to big social data analytics (Bravo-Marquez, Mendoza and Poblete, 2014; Batrinca and Treleaven, 2015; Bello-Orgaz, Jung and Camacho, 2016; Jussila, Menon, *et al.*, 2017) have been introduced that make use of computational methods for detecting and analyzing consumer emotions.

Criticism has been also leveled against current big data research. For instance, Boyd and Crawford (2012) question the objectivity and accuracy of the data, and challenge whether big data is always more suited to the research task than small data and whether it can preserve the contexts of what it is intended to describe. Bruns (2013) continues this by stating that the challenge is to improve the quality rather than (or at least the same time as) the quantity of the data gathered. Researchers must carefully consider the data they collect, and what the limitations of these sources are – rather than use the datasets that were the easiest to capture while still appearing to contain meaningful data at face value (Bruns, 2013). Bruns (2013) also points out that discussion of research choices in big data research should not be limited to data gathering approaches, for it is just as important that the further steps in processing the data are documented in detail in order to ensure the replicability of results.

Taking these recommendations into consideration, we conducted a systematic literature review on customer emotions in social media guided by the following research questions:

- RQ1. How is social media data on consumer emotions currently collected?
- RQ2. How are large datasets used in detecting emotions?
- RQ3. What types of computational methods are being used to detect consumer emotions?
- RQ4. What are the limitations and the accuracy of current computational methods in detecting consumer emotions?

2. Methodology

According to Tranfield *et al*. (2003), the systematic literature review method aims at locating all the relevant studies without the biased view of the researcher. This is achieved by making explicit the values and assumptions that constitute a review. Unlike narrative reviews, systematic reviews adopt a replicable, scientific, and transparent process in order to minimize bias, furthermore offering an audit trail for reviewers.

We followed a seven-stage process model for conducting systematic literature reviews (Fink, 2005): 1) selecting research questions, 2) selecting the bibliographic or article database, 3) choosing search terms, 4) applying practical screening criteria, 5) applying methodological screening criteria, 6) performing the review, and 7) synthesizing the results.

First, we defined the research questions, which are described in the introduction section. Second, we selected the appropriate databases and the type of literature in which we were interested. In order to guarantee a comprehensive sample covering the most important data related to our research objectives, we used the following databases: ScienceDirect, Emerald Insight, IEEE Explore, and ACM Digital Library. The combination of the chosen databases was diverse, covering both business- and technology-related academic journal and conference articles. Third, we used as search terms “social media,” “emotions,” and “consumer,” as the aim was to understand especially emotions related to consumer behavior in social media. Therefore, we created two search strings: “social media emotions” and “social media emotions consumer.” All of their results are combined and documented in the findings. Fourth, we applied the practical screening criteria outlined in Table 1.

Table 1. Practical screening criteria.

Inclusion criteria	**Type**
Include only studies written in English	Publication language
Include studies focused on social media	Content
Include studies focused on emotions	Content
Include studies focused on consumer behavior	Content
Exclusion criteria	**Type**
Exclude duplicates	Content
Exclude articles with missing description of data extraction methods	Content

Fifth, as methodological screening criteria we included only empirical studies, excluding conceptual articles and studies that either introduce tools or perform experiments without collecting any social media data. Sixth, in reading the full papers we further excluded articles that talk about emotions but did not use any computational method to classify social media text into sentiment or emotion categories. Table 2 outlines the resulting articles.

3. Results

Table 2. Overview of the reviewed empirical studies.

No.	Author	Title	Publication
1	(Sun *et al.*, 2017)	Detecting users' anomalous emotion using social media for business intelligence	Journal of Computational Science
2	(Bernabé-Moreno *et al.*, 2015)	Emotional profiling of locations based on social media	Information Technology And Quantitative Management
3	(He *et al.*, 2015)	A novel social media competitive analytics framework with sentiment benchmarks	Information and Management
4	(Mostafa, 2013)	More than words: Social networks' text mining for consumer brand sentiments	Expert Systems with Applications
5	(Ghiassi, Skinner and Zimbra, 2013)	Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network	Expert Systems with Applications
6	(Kontopoulos *et al.*, 2013)	Ontology-based sentiment analysis of Twitter posts	Expert Systems with Applications
7	(Gao *et al.*, 2018)	Identifying competitors through comparative relation mining of online reviews in the restaurant industry	International Journal of Hospitality Management
8	(Li and Xu, 2014)	Text-based emotion classification using emotion cause extraction	Expert Systems with Applications
9	(Deng, Sinha and Zhao, 2017)	Adapting sentiment lexicons to domain-specific social media texts	Decision Support Systems
10	(Howells and Ertugan, 2017)	Applying fuzzy logic for sentiment analysis of social media network data in marketing	International Conference on Theory and Application of Soft Computing, Computing with Words and Perception
11	(Lee, 2017)	Social media analytics for enterprises: Typology, methods, and processes	Business Horizons
12	(Benthaus, Risius and Beck, 2016)	Social media management strategies for organizational impression management and their effect on public perception	Journal of Strategic Information Systems
13	(Wang *et al.*, 2016)	Fine-Grained Sentiment Analysis of Social Media with Emotion Sensing	Future Technologies Conference
14	(Zhao *et al.*, 2014)	PEARL: An Interactive Visual Analytic Tool for Understanding Personal Emotion Style Derived from Social Media	IEEE Conference on Visual Analytics Science and Technology
15	(Hong Keel Sul, Dennis and Yuan, 2014)	Trading on Twitter: The Financial Information Content of Emotion in Social Media	Hawaii International Conference on System Sciences
16	(Larsen *et al.*, 2016)	TV ratings vs. Social media Engagement Big Social Data Analytics of the Scandinavian TV Talk Show Skavlan	IEEE International Conference on Big Data
17	(Shukri *et al.*, 2015)	Twitter Sentiment Analysis: A Case Study in the Automotive Industry	IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies
18	(Sarakit *et al.*, 2015)	Classifying Emotion in Thai YouTube Comments	International Conference of Information and Communication Technology for Embedded Systems
19	(Bing, Chan and Ou, 2014)	Public Sentiment Analysis in Twitter Data for Prediction of A Company's Stock Price Movements	IEEE International Conference on e-Business Engineering
20	(Castellanos *et al.*, 2011)	LivePulse: Tapping Social Media for Sentiments in Real-Time	International conference companion on World Wide Web
21	(Xu *et al.*, 2017)	A New Chatbot for Customer Service on Social Media	Conference on Human Factors in Computing Systems
22	(Uren *et al.*, 2016)	Social media and sentiment in bioenergy consultation	International Journal of Energy Sector Management
23	(He *et al.*, 2017)	Application of social media analytics: a case of analyzing online hotel reviews	Online Information Review

The reviewed empirical articles consist of 11 journal articles and 12 conference articles, with the earliest article from the year 2011 and the latest from 2017. As our aim was to understand data collection, dataset attributes and the types of emotions investigated in the studies, we compiled Table 3 that describes the social media platforms where the data was extracted, the data extraction method used to collect the data, and the emotion classification used in the study.

Table 3. Social media platforms, data extraction, dataset size and emotion classifications used in the studies.

No.	Social Media Platform	Data Extraction Method	Dataset Size	Emotion Classification
1	Sina weibo	Crawler	10275 microblogs on 100 users	Distinct sentiment – 3 categories, Discreet emotions – 5 categories
2	Twitter	Harvester	852319 interactions	Valence, Arousal, Dominance of emotions – 51 emotional states.
3	Twitter	API	Not described	Distinct sentiment – 2 categories
4	Twitter	Miner	3516 tweets	Distinct sentiment – 2 categories
5	Twitter	API	10345184 tweets	5 categories of sentiment
6	Twitter	API	Not described	Distinct sentiment – 2 categories
7	Dianping	Crawler	35872 reviews	Distinct sentiment – 2 categories
8	Sina weibo	Crawler	16845 microblogs	Discreet emotions – 6 categories
9	Twitter	Querying	743069 tweets	Distinct sentiment – 2 categories
10	Twitter	Open source intelligence	Not described	5 categories of sentiment
11	Yelp	Nvivo	50 customer postings(d)	Distinct sentiment – 4 categories
12	Twitter	API	17 million user generated tweets, 200000 sent by companies.	Distinct sentiment – 5 categories
13	Twitter	Crawler	Not described	Distinct sentiment – 3 categories, Emotions – 4 categories
14	Twitter	PEARL	Study 1: 10 participants, 308 labeled tweets. Study 2: 6 participants, 600 tweets per participant	Discreet emotions – 8 categories
15	Twitter	CRSP	2503385 tweets	Distinct sentiment – 2 categories
16	Facebook	SODATA	24000 comments	Distinct sentiment –3 categories, Emotions – 5 categories.
17	Twitter	API	3000 tweets	Distinct sentiment– 2 categories, Emotions – 7 categories
18	YouTube	API	Movies: 2771 comments Advertisement: 3077 comments	Discreet emotions – 6 categories
19	Twitter	Crawler	15 million tweets	Distinct sentiment – 5 categories
20	Twitter	API	Not described	Distinct sentiment – 3 categories
21	Twitter	API	703 study participants	Distinct emotions – 3 categories
22	Twitter	API	300 tweets	Distinct sentiment – 3 categories
23	TripAdvisor	API	11403 reviews	Distinct sentiment – 3 categories

Regarding the size of data, from as little as 50 postings to 17 million tweets have been analyzed in terms of customer emotion detection. The largest size of dataset was extracted from the Twitter platform. The datasets were extracted using APIs, crawlers, miners, harvesters, open source intelligence, and custom-built software (e.g., SODATA, PEARL, and VOZIQ).

A variety of different software tools were employed in the studies described in the articles. For instance, visual analytics was carried out by custom-built PEARL software in Zhao *et al.* (2014) and MATLAB was used for graphical representation of data in Sun *et al*. (2017). R was used for data pre-processing in several articles (He *et al.*, 2015, 2017; Shukri *et al.*, 2015). In addition, several R libraries were used in sentiment score calculation and visualization, such as twitteR, plyr, stringr, and ggplot2. Restaurant review pre-processing, element extraction, and web crawling were implement in Python using MySQL for data storage (Gao *et al.*, 2018). Similarly, Larsen et al. (2016) carried out an experiment using Python. However, some of the studies did not explicitly report the software and analysis tools used in the study.

Most of the studies categorized emotions into 2-5 different sentiment categories. Seven studies categorized sentiment into positive and negative categories. Six studies categorized sentiment into positive, neutral, and negative categories. Lee (2017) categorized emotions into four sentiment categories: very positive, moderately positive, very negative, and moderately negative. In three studies, emotions were categorized into five sentiment categories, e.g., strongly positive, mildly positive, neutral, mildly negative, and strongly negative (Ghiassi, Skinner, and Zimbra, 2013). The exceptions were Benthaus *et al*. (2016), who differentiated sentiment into average, positive, negative, ambivalent, and neutral and Bing *et al*. (2014), who classified sentiment into positive^{+}, positive, neutral, negative, negative^{-}.

In seven studies, emotions were categorized into distinct emotion categories, ranging from 4 to 51 distinct emotions. Bernabé-Moreno (2015) classified emotions into 51 categories based on Russel's (1980) circumplex model: ambitious, amorous, amused, apathetic, ashamed, attentive, bitter, conscientious, contemplative, convinced, courageous, dejected, desperate, despondent, determined, disappointed, discontented, disgusted, dissatisfied, distrustful, embarrassed, enthusiastic, envious, expectant, feeling_guilt, feeling_superior, friendly, hesitant, hopeful, impatient, impressed, indignant, insulted, interested, jealous, joyous, languid, light-hearted, loathing, longing, melancholic, passionate, peaceful, pensive, polite, serious, solemn, suspicious, taken_aback, uncomfortable, worried.

In some studies several emotion theories were used, but the most common theories included Ekman's (1992) basic emotions and Plutchik's (2001) emotion theory. Wang *et al*. (2016) categorized emotions into anxiety, sadness, anger, and others based on combined emotion studies (Ekman, 1992; Plutchik and Robert, 2001; Socher *et al.*, 2013; Chafale and Pimpalkar, 2014). Zhao *et al*. (2014) used Plutchik and Robert's (2001) categorical model for primary emotion pairs such as anger-fear, anticipation-surprise, joy-sadness, and trust-disgust; (Mehrabian, 1980; Russell, 1980) for dimensionality (valence, arousal, dominance); and (Davidson and Sharon Begley, 2012) for emotional styles on the above-mentioned emotions and dimensions. Shukri *et al*. (2015) employed (Strapparava and Valitutti, 2004; Wilson, Wiebe and Hoffmann, 2005) to categorize emotions into classes such as anger, disgust, fear, joy, sadness, surprise, and unknown. In the Li and Xu (2014) study, emotion classification was carried out based on basic emotions (Ekman and Friesen, 1971): happiness, anger, disgust, fear, sadness, and surprise, Larsen *et al*. (2016) and Sarakit *et al*. (2015) also employed Ekman's (1992) model of basic emotions. Sun *et al.* (2017) classified emotion into neutral, happy, surprised, sad, and angry, without referring to any emotion theories. In addition, Xu *et al*. (2017) classified emotions without referring to any emotion theories, first, into two categories: emotional and informational, and second, into three additional categories: appropriateness, empathy, and helpfulness.

The studies introduced a variety of approaches for detecting the above described emotions categories. We compiled the used computational methods, the evaluation and assessment of the methods, as well as, primary limitations outlined by the authors into Table 4.

Table 4. Limitations, computational methods, and their evaluation in reviewed empirical studies.

Article	Computational Methods	Evaluation of Computational Methods	Primary Limitations
1	Anomaly detection: single and multivariate Gaussian distribution	Accuracy of abnormal user emotion detection = 83.49%	Sparseness of micro-blog data of individual user, abnormal user reaction can only be detected in a week or in a month.
2	Multivariate kernel density function	Visual evaluation	Analyzed only geo-located social media users and not the entire population at the airport.
3	Modified Chi-square feature selection, modified N-Gram model	N-Gram correctly recognized 79% of the comments. The values of accuracy showed that the N-Gram correctly marked positive message	Not described

		and negative messages at a success rate of 82.42%.	
4	Hu and Liu Lexicon	T-mobile brand has highest negative sentiment of 72%, DHL has highest positive sentiment of 60%	Analysis does not reveal the reason behind consumers' expression of sentiment, meaning that it fails to identify the sentiment topic.
5	Supervised machine learning algorithm - DAN2, Support vector machine (SVM).	Accuracy for strong positive emotions using DAN2 – 96%. Accuracy for mildly negative and strongly negative emotions – 89.9% and 95.1%	Analysis was performed only on a single corpus (single brand).
6	Ontology based semantically enabled system (SEM), Ontology support (ONT); Custom built system without ontology learning techniques (CUS).	Calculates sentiment score for tweets: (a) SEM outperforms ONT with regards to recall ratios, (b) SEM, ONT perform better than CUS.	Presence of high ratio of advertising tweets could contribute to the distortion of results.
7	Network construction: Single-line graph, Dichotomic-line graph, and Multi-line graph. Sentiment strength dictionary.	Proportion of 5 star rating – R_{20} (56.8%) Service - R_2 (9.1/10) Environment - R_2 (8.9/10),R_{20} (8.9/10) Price - R_2 (89/100) Taste - R_{20} (9.0/10) Restaurant no 20 is better in 4/7 categories.	Dataset used is limited to 50 restaurants.
8	Support vector machine (SVM), Support vector regression (SVR), Emotion cause extraction.	Higher precision for happiness, anger, fear, and surprise emotions. Increase in statistical f-score percentage in happiness, anger, disgust, surprise (1.6%, 2%, 9.2%, and 0.8%) respectively.	Emotion causes derived from sad posts are not effective when classifying them from other posts.
9	Lexicon that learns sentiment words generated by corpus. General Inquirer (GI), MPQA subjectivity lexicon (MPQA), Opinion lexicon (OL), SentiWordNet (SWN). Domain specific: Loughran and McDonald (LM).	F-measure - Experimental results – 80.31%, Class level results - 81.05%, political tweets results – 85.46% Proposed algorithm - combined lexicon with part-of-speech tagging excluding SWN (Combined4_ARVN) outperforms in F-measure by almost 10% .	Higher false-positive rate is obtained when many words are added to the lexicon which have no sentiment and higher false-negative rate when fewer words are added during which sentiment words are missed. Therefore, choosing the right threshold is important.
10	Fuzzy logic analysis	Social bot account follower rate = 27% of the user population, Information gathered from 12% of the user population for customer relationship management.	Not described
11	Sentiment analysis, Descriptive statistics	Stage 1: evaluation of reviews on sentiment polarity Stage 2: Statistics performed on six factors, where customer satisfaction evident on high correlation between restaurant's food, price, and atmosphere.	Small dataset size.
12	SentiStrength 2, Naïve bayes, Data analysis	User generated positive tweets increased by 4.3% and negative tweets decreases by 3.5%, successful deployment of social media management tools.	Analysis was based on companies that have success with social media management tools and unconscious interviewer bias.
13	Fuzzy rule inference technique, multi-source lexicon combination.	Real-time analysis of tweets/data and emotions of tweets are visualized with geographic information. Negative tweets are broken down into finer emotions for real-time crisis management.	Not described
14	NRC, ANEW, WordNet, Clustering-based mood detection	In the proposed tool, anger mood accuracy is estimated to be 0.81 (Cohen's Kappa). The tool can help a	A small reliability issue in the PEARL analytic engine is that it should be evaluated with a

		user visualize his/her own emotional resilience, emotional style.	large number of users or using a broader population.
15	Word analysis strategy (sentiment analysis), Harvard-IV dictionary, regression using adjusted R squared value, correlation coefficient.	Positive correlation between emotional valence and same day stock returns. More impact from users with more followers (median: 177)	Not described
16	No mention of any method exclusively. Only general mention of sentiment analysis.	No correlation between activity on Facebook and viewing figures. Controversial views expressed are received with negative reactions which may be an opportunity for brand exposure and attention. However, it may affect the brand in the long run.	The time interval used may have been too short to study the significance of the relationship between the viewing figures and Facebook activity .
17	Naïve Bayes (NB), Wiebe's polarity lexicon, Strapparava emotions lexicon.	Positive polarity of AUDI brand (83%), which was higher when compared to other brands.	Not described
18	LexTo: Thai Lexeme Tokenizer, Lexitron dictionary, Support vector machine (SVM), Multinomial naïve bayes (MNB), J48 decision tree, Weka. 10-fold cross validation	TF - In movie category, MNB has the best accuracy of 84.48%. In advertisement category, SVM has the best accuracy of 76.14%. For the TF-IDF, SVM performs better for both movie and advertisement categories with 82.28% and 72.41%, respectively.	Accuracy is affected by the presence of ambiguous comments.
19	SentiWordNet 3.0, Naïve Bayes (NB), Support vector machine (SVM), C4.5	Proposed algorithm accuracy for sentiment detection in IT industry 76.12% , Media – 73.78% performs better than other classifiers.	Sometimes, tweets contain no meaning which can hamper the accuracy of the model and the best prediction accuracy is obtained only after a 3-day period.
20	General sentiment lexicon, WordNet, Language heuristics	Successful classification of positive and negative sentiments from tweets.	Not described
21	Long short-term memory networks (LTSM), Sequence-to-sequence learning, Word embeddings	Deep learning performs better than information retrieval method in emotional context.	Performance decrease (from emotional requests to informational requests) of information retrieval and deep learning systems.
22	Entity-level contextual sentiment extraction	Negative emotions are mainly targeted by opposition groups.	Limitations due to the analysis being performed on a small dataset.
23	Natural language pre-processing, text mining (cluster analysis, categorization, concept maps) and sentiment analysis.	Customers, both satisfied and unsatisfied, share a common interest such as location, food, rooms, service quality, and staff.	Not described

In majority of the studies, marketing and business intelligence were the context for detecting emotions. In Sun et al. (2017), the main context was business intelligence, where user emotions were detected to help companies make better decisions with respect to product image and also to help the government monitor public opinion. Geographically tailored marketing advertisements and messages based on the emotional profile of the location were the main focus in Bernabé-Moreno (2015). He et al. (2015) focused on customer opinions on competitors' product and service offerings for improved decision making and marketing intelligence. Xu et al. (2017) applied emotion detection in customer service chatbots. In Mostafa (2013), social media data was used to redesign marketing and advertising campaigns. Ghiassi et al. (2013) similarly focused on marketing and public relations. Gao et al. (2018) focused on competitor environment analysis, market positioning, and development of service strategy. Howells and Ertugan (2017) study was directed at customer relationship management, marketing, and retention of customers. Lee (2017) presented a 2x2 typology of enterprise social media analytics. Benthaus et al. (2016) evaluated the use of social media management tools to help firms influence public perception besides marketing and customer engagement activities. In Wang et al. (2016), the focus was on customer behavior and consumer preference context. Zhao et al. (2014) focused on providing individualized customer care. Larsen et al. (2016) and Shukri et al. (2015) both studied customer emotions for assisting marketing activities. In Bing et al. (2014) and Hong et al. (2014), the focus of emotion detection was on financial markets, and the impact on stock prices and stock returns, for example.

4. Discussion and Conclusions

In our systematic literature review we first sought answers to three research questions, i.e., how social media data on consumer emotions is currently collected, how large datasets are used in detecting emotions, and what type of computational methods are being used to detect consumer emotions. We discovered that social media data on consumer emotions was collected by eight types of methods, most commonly using APIs and web crawling. The datasets used in detecting emotions ranged from small data (e.g., 50 customer postings) to what can be considered as big data in terms of volume (e.g., 17 million tweets). We found that a large variety of computational methods were being used in the process of detecting consumer emotions. Some of the studies combined several computational methods to achieve more accurate results (Sarakit *et al.*, 2015; Deng, Sinha and Zhao, 2017).

The answers to our last research question *"What are the limitations and the accuracy of current computational methods in detecting consumer emotions?"* revealed a fundamental weakness of the current research on detection of sentiment or emotions from social media, in particular regarding the limitations and evaluation of computational methods. First of all, a major deficiency in the research is that less than half of the studies report the accuracy of the computational method used. Furthermore, in those studies that report the accuracy, it ranged between 73% and 95%. Additionally, context seems to affect the accuracy of the computational methods. For example, in the study of Sarakit *et al*. (2015), Support Vector Machine yielded the best result with an accuracy of 76% in the advertisement genre, whereas the multinomial naïve Bayes yielded the best result with an accuracy of 84% in the music video genre. Therefore, a custom-built tool trained with a particular dataset, e.g., data collected from restaurant reviews, may not be transferable directly to a different context; at minimum the accuracy should be estimated in the new context. This finding affirms the view of Boyd and Crawford (2012) that objectivity and the accuracy of the data is the challenge of big data research. Therefore, our results indicate that, in addition to data gathering approaches, researchers should also focus on the processing of the data, as well as on its documentation in detail to ensure the replicability of results, as suggested by Bruns (2013).

To sum up our results regarding practical implications: managers should understand that the results of the detection of sentiment or emotions from social media may not provide accurate results to support decision making unless implemented rigorously, which is important for marketing and business intelligence purposes. Thus, there should be more research on the accuracy of various computational methods as well as the effect of context on accuracy and transferability of the results to another context. What was missing from most studies is how these methods can be linked to business performance. Research was also found lacking on the perspective, how managers perceive the value of social media emotion detection, and overall how be believable they find social media data in the first place.

References

Batrinca, B. and Treleaven, P. C. (2015) 'Social media analytics: a survey of techniques, tools and platforms', *AI & SOCIETY*, 30(1), pp. 89–116. doi: 10.1007/s00146-014-0549-4.
Bello-Orgaz, G., Jung, J. and Camacho, D. (2016) 'Social big data: Recent achievements and new challenges', *Information Fusion*, 28, pp. 44–59.
Benthaus, J., Risius, M. and Beck, R. (2016) 'Social media management strategies for organizational impression management and their effect on public perception', *The Journal of Strategic Information Systems*. North-Holland, 25(2), pp. 127–139. doi: 10.1016/J.JSIS.2015.12.001.
Bernabé-Moreno, J. *et al.* (2015) 'Emotional Profiling of Locations Based on Social Media', *Procedia Computer Science*. Elsevier, 55, pp. 960–969. doi: 10.1016/J.PROCS.2015.07.107.
Bing, L., Chan, K. C. C. and Ou, C. (2014) 'Public Sentiment Analysis in Twitter Data for Prediction of a Company's Stock Price Movements', in *2014 IEEE 11th International Conference on e-Business Engineering*. IEEE, pp. 232–239. doi: 10.1109/ICEBE.2014.47.

boyd, danah and Crawford, K. (2012) 'CRITICAL QUESTIONS FOR BIG DATA', *Information, Communication & Society*. Routledge , 15(5), pp. 662–679. doi: 10.1080/1369118X.2012.678878.
Bravo-Marquez, F., Mendoza, M. and Poblete, B. (2014) 'Meta-level sentiment models for big social data analysis', *Knowledge-Based Systems*, 69, pp. 86–99. Available at: https://www.sciencedirect.com/science/article/pii/S0950705114002068 (Accessed: 6 February 2018).
Bruns, A. (2013) 'Faster than the speed of print: Reconciling "big data"social media analysis and academic scholarship', *First Monday*, 18(10). Available at: https://www.firstmonday.dk/ojs/index.php/fm/article/view/4879 (Accessed: 15 February 2018).
Castellanos, M. *et al.* (2011) 'LivePulse', in *Proceedings of the 20th international conference companion on World wide web - WWW '11*. New York, New York, USA: ACM Press, p. 193. doi: 10.1145/1963192.1963287.
Chafale, D. and Pimpalkar, A. (2014) 'Review on Developing Corpora for Sentiment Analysis Using Plutchik's Wheel of Emotions with Fuzzy Logic', *International Journal of Computer Sciences and Engineering*.
Coté, M. (2014) 'The Materiality of Big Social Data', *Cultural Studies Review*, 20(1), pp. 121–149.
Davidson, R. J. and Sharon Begley, W. (2012) 'The Emotional Life of Your Brain: How Its Unique Patterns Affect the Way You Think, Feel, and Live-and How You Can Change Them.' Available at: http://testrain.info/download/Richard J. Davidson - The Emotional Life of Your Brain.pdf (Accessed: 13 February 2018).
Deng, S., Sinha, A. P. and Zhao, H. (2017) 'Adapting sentiment lexicons to domain-specific social media texts', *Decision Support Systems*. North-Holland, 94, pp. 65–76. doi: 10.1016/J.DSS.2016.11.001.
Ekman, P. (1992) 'An argument for basic emotions', *Cognition and Emotion*. Taylor & Francis Group , 6(3–4), pp. 169–200. doi: 10.1080/02699939208411068.
Ekman, P. and Friesen, W. V (1971) 'Constants across cultures in the face and emotion.', *Journal of personality and social psychology*, 17(2), pp. 124–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/5542557
Fink, A. (2005) *Conducting research literature reviews: From the internet to paper*.
Fosso Wamba, S. *et al.* (2015) 'How "big data" can make big impact: Findings from a systematic review and a longitudinal case study', *International Journal of Production Economics*. Elsevier, 165, pp. 234–246. doi: 10.1016/j.ijpe.2014.12.031.
Gao, S. *et al.* (2018) 'Identifying competitors through comparative relation mining of online reviews in the restaurant industry', *International Journal of Hospitality Management*. Pergamon, 71, pp. 19–32. doi: 10.1016/J.IJHM.2017.09.004.
Ghiassi, M., Skinner, J. and Zimbra, D. (2013) 'Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network', *Expert Systems with Applications*. Pergamon, 40(16), pp. 6266–6282. doi: 10.1016/J.ESWA.2013.05.057.
He, W. *et al.* (2015) 'A novel social media competitive analytics framework with sentiment benchmarks', *Information & Management*. North-Holland, 52(7), pp. 801–812. doi: 10.1016/J.IM.2015.04.006.
He, W. *et al.* (2017) 'Application of social media analytics: a case of analyzing online hotel reviews', *Online Information Review*. Emerald Publishing Limited , 41(7), pp. 921–935. doi: 10.1108/OIR-07-2016-0201.
Hong Keel Sul, Dennis, A. R. and Yuan, L. I. (2014) 'Trading on Twitter: The Financial Information Content of Emotion in Social Media', in *2014 47th Hawaii International Conference on System Sciences*. IEEE, pp. 806–815. doi: 10.1109/HICSS.2014.107.
Howells, K. and Ertugan, A. (2017) 'Applying fuzzy logic for sentiment analysis of social media network data in marketing', *Procedia Computer Science*. Elsevier, 120, pp. 664–670. doi: 10.1016/J.PROCS.2017.11.293.
Jalonen, H. and Jussila, J. (2016) 'Developing a Conceptual Model for the Relationship Between Social Media Behavior, Negative Consumer Emotions and Brand Disloyalty', in, pp. 134–145. doi: 10.1007/978-3-319-45234-0_13.
Jussila, J., Boedeker, M., *et al.* (2017) 'Social Media Analytics Empowering Customer Experience Insight', in Kavoura, A., Sakas, D., and Tomaras, P. (eds) *Strategic Innovative Marketing*. Springer, Cham, pp. 25–30. doi: 10.1007/978-3-319-56288-9_4.
Jussila, J., Menon, K., *et al.* (2017) 'Who is who in Big Social Data? A Bibliographic Network Analysis Study', in *Proceedings of The 4th European Conference on Social Media*. Vilnius: Academic Conferences and publishing limited, pp. 161–169.
Kontopoulos, E. *et al.* (2013) 'Ontology-based sentiment analysis of twitter posts', *Expert Systems with Applications*. Pergamon, 40(10), pp. 4065–4074. doi: 10.1016/J.ESWA.2013.01.001.
Laros, F. and Steenkamp, J. (2005) 'Emotions in consumer behavior: a hierarchical approach', *Journal of business Research*, 58(10), pp. 1437–1445.
Larsen, H. H. *et al.* (2016) 'TV ratings vs. social media engagement: Big social data analytics of the Scandinavian TV talk show Skavlan', in *2016 IEEE International Conference on Big Data (Big Data)*. IEEE, pp. 3849–3858. doi:

10.1109/BigData.2016.7841058.
Lee, I. (2017) ‘Social media analytics for enterprises: Typology, methods, and processes’, *Business Horizons*. Elsevier. doi: 10.1016/J.BUSHOR.2017.11.002.
Li, W. and Xu, H. (2014) ‘Text-based emotion classification using emotion cause extraction’, *Expert Systems with Applications*. Pergamon, 41(4), pp. 1742–1749. doi: 10.1016/J.ESWA.2013.08.073.
Magids, S., Zorfas, A. and Leemon, D. (2015) ‘The new science of customer emotions’, *Harvard Business Review*, November 2. Available at: http://sproutresearch.com.au/wp-content/uploads/2015/11/HBR-The-New-Science-of-Customer-Emotions.pdf
Mehrabian, A. (1980) *Basic dimensions for a general psychological theory : implications for personality, social, environmental, and developmental studies*. Oelgeschlager, Gunn & Hain. Available at: https://books.google.fi/books/about/Basic_Dimensions_for_a_General_Psycholog.html?id=1RR9AAAAMAAJ&redir_esc=y (Accessed: 13 February 2018).
Mostafa, M. M. (2013) ‘More than words: Social networks’ text mining for consumer brand sentiments’, *Expert Systems with Applications*. Pergamon, 40(10), pp. 4241–4251. doi: 10.1016/J.ESWA.2013.01.019.
Plutchik, R. and Robert (2001) ‘The Nature of Emotions’, *American Scientist*, 89(4), p. 344. doi: 10.1511/2001.4.344.
Russell, J. A. (1980) ‘A circumplex model of affect.’, *Journal of Personality and Social Psychology*, 39(6), pp. 1161–1178. doi: 10.1037/h0077714.
Sarakit, P. *et al.* (2015) ‘Classifying emotion in Thai youtube comments’, in *2015 6th International Conference of Information and Communication Technology for Embedded Systems (IC-ICTES)*. IEEE, pp. 1–5. doi: 10.1109/ICTEmSys.2015.7110808.
Shukri, S. E. *et al.* (2015) ‘Twitter sentiment analysis: A case study in the automotive industry’, in *2015 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT)*. IEEE, pp. 1–5. doi: 10.1109/AEECT.2015.7360594.
Socher, R. *et al.* (2013) ‘Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank’. Available at: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.383.1327 (Accessed: 13 February 2018).
Strapparava, C. and Valitutti, A. (2004) ‘WordNet-Affect: an Affective Extension of WordNet’, in *Proceedings of the International Conference on Language Resources and Evaluation*. Available at: http://www.lrec-conf.org/proceedings/lrec2004/pdf/369.pdf (Accessed: 13 February 2018).
Sun, X. *et al.* (2017) ‘Detecting users’ anomalous emotion using social media for business intelligence’, *Journal of Computational Science*. Elsevier. doi: 10.1016/J.JOCS.2017.05.029.
Tranfield, D., Denyer, D. and Smart, P. (2003) ‘Towards a methodology for developing evidence-informed management knowledge by means of systematic review *’, *British Journal of Management*, 14, pp. 207–222. doi: 10.1111/1467-8551.00375.
Uren, V. *et al.* (2016) ‘Social media and sentiment in bioenergy consultation’, *International Journal of Energy Sector Management*. Emerald Group Publishing Limited , 10(1), pp. 87–98. doi: 10.1108/IJESM-11-2014-0007.
Wang, Z. *et al.* (2016) ‘Fine-grained sentiment analysis of social media with emotion sensing’, in *2016 Future Technologies Conference (FTC)*. IEEE, pp. 1361–1364. doi: 10.1109/FTC.2016.7821783.
Wilson, T., Wiebe, J. and Hoffmann, P. (2005) ‘Recognizing contextual polarity in phrase-level sentiment analysis’, in *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing - HLT ’05*. Morristown, NJ, USA: Association for Computational Linguistics, pp. 347–354. doi: 10.3115/1220575.1220619.
Xu, A. *et al.* (2017) ‘A New Chatbot for Customer Service on Social Media’, in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI ’17*. New York, New York, USA: ACM Press, pp. 3506–3510. doi: 10.1145/3025453.3025496.
Zhao, J. *et al.* (2014) ‘PEARL: An interactive visual analytic tool for understanding personal emotion style derived from social media’, in *2014 IEEE Conference on Visual Analytics Science and Technology (VAST)*. IEEE, pp. 203–212. doi: 10.1109/VAST.2014.7042496.