Computers, Materials & Continua
DOI:10.32604/cmc.2022.017284
Article

A Netnographic-Based Semantic Analysis of Tweet Contents for Stress Management

Jari Jussila[1], Eman Alkhammash[2,*], Norah Saleh Alghamdi[3], Prashanth Madhala[4] and Mohammad Ayoub Khan[5]

[1]HAMK Smart Research Unit, Häme University of Applied Sciences, Hämeenlinna, 13100, Finland
[2]Department of Computer Science, College of Computers and Information Technology, Taif University, Taif, 21944, Saudi Arabia
[3]College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh, 11671, Saudi Arabia
[4]Faculty of Management and Business, Tampere University, Tampere, 33720, Finland
[5]College of Computing and Information Technologies, University of Bisha, Bisha, 67714, Saudi Arabia
[*]Corresponding Author: Eman Alkhammash. Email: eman.kms@tu.edu.sa
Received: 26 January 2021; Accepted: 19 May 2021

Abstract: Social media platforms provide new value for markets and research companies. This article explores the use of social media data to enhance customer value propositions. The case study involves a company that develops wearable Internet of Things (IoT) devices and services for stress management. Netnography and semantic annotation for recognizing and categorizing the context of tweets are conducted to gain a better understanding of users' stress management practices. The aim is to analyze the tweets about stress management practices and to identify the context from the tweets. Thereafter, we map the tweets on pleasure and arousal to elicit customer insights. We analyzed a case study of a marketing strategy on the Twitter platform. Participants in the marketing campaign shared photos and texts about their stress management practices. Machine learning techniques were used to evaluate and estimate the emotions and contexts of the tweets posted by the campaign participants. The computational semantic analysis of the tweets was compared to the text analysis of the tweets. The content analysis of only tweet images resulted in 96% accuracy in detecting tweet context, while that of the textual content of tweets yielded an accuracy of 91%. Semantic tagging by Ontotext was able to detect correct tweet context with an accuracy of 50%.

Keywords: Social media; stress; semantic analysis; Twitter; context recognition

1 Introduction

A massive volume of data is generated daily by social media, which allows companies to tap into the customers' world. The generated data are not equally valuable to companies. According to Borgman [1], the data must be large or small depending on their usefulness in helping us

This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

to understand problems and answer questions. Wang [2] argues that thick data are needed to rescue big data from resolution and context loss. Thick data can complement or refine big data when analyzing people's emotions, stories, and models by using qualitative, ethnographic or netnographic research methods [3]. As a strategy to optimize the size of the data, it should be small to be manageable while gathering enough information to draw rigorous conclusions. Latzko-Toth et al. [4] proposed a thickening approach to reduce the breadth of data while enhancing its depth.

The aim of this study is to analyze tweets about stress management practices to identify the context of the tweets. The tweets are mapped to the expressed emotions of tweet content on pleasure and arousal dimensions to draw new customer insights. This is important to understand the emotions expressed in tweets because interactions on social media are loaded to a great extent with emotions [5]. In studying emotions on social media, there is extant research analyzing the sentiment of social media interactions [6,7]. There are many benefits of collecting knowledge about consumers to have a better understanding of their dimensions of emotions, such as arousal and pleasure, as well as the context.

Social network analytical systems have a more efficient manner for collecting appropriate data that result in useful insights, thereby allowing reliable and valuable decisions [8,9]. The most common feature shared by these tools is sentiment analysis of social media conversations. The drawback of conventional sentiment analysis is as follows: "its inability to capture the deeper contents of the affective experience of the customer—simple positive versus negative categorization of affective experience is not able to reveal how the customer will truly act" [9]. Furthermore, the context of sentiment analysis is often lacking, making it difficult to determine whether it has positive or negative emotion. It can be argued that reducing the resolution of data to include only sentiment (pleasure-displeasure dimension) of social media conversations may dramatically decrease the potential value gained from the analysis.

Recently, many studies have shown that technology firms can utilize big data, which have been built to strengthen consumer relationships [10]. The present big data paradigm and studies have also been criticized. Kitchin et al. [11] pointed out that big data will not fully replace small data for all studies because big data usually cannot be useful for topics such as humanities and social sciences. Boyd et al. [12] discussed whether large data are more appropriate than small data in their respective studies to investigate data objectivity and quality.

This study employs a combination of small data and big data methods to explore how emotional expressions on pleasure-arousal dimensions are assessed. The context may be identified from the extracted data from Twitter. The objective is to create classifiers for the emotional context that can recognize the emotions that are expressed using the corpora of tweets.

The remaining sections of this paper are organized as follows: The related work and discussion are presented in Section 2, while the proposed research methodology is presented in Section 3; the results and discussion are presented in Section 4. Finally, Section 5 presents the conclusion and future work of our research.

2 Related Work

2.1 Identifying Emotional Expressions

Emotions are a main component of a human personality. Emotions are associated with behavior, physiology, experiences, feelings, perceptions, and conceptualizations that can emerge as a result of perceiving a certain situation [13]. The content on social media is filled with the

experiences and opinions of products and life [14]. When consumers want to buy new products or services, they often seek opinions from close friends and family members [15]. In the context of microblogs, consumers often reveal their perceptions about products in the form of feelings and reactions [16]. Sometimes, when buying new products, users also post reviews that are helpful for others [17]. Many managers are eager to understand these expressions of opinion to assess their company's products and services [18].

The results from Madhala et al. [19] showed that the emotion theories from Ekman [20] and Plutchik [21] are the most common theories upon which several studies have been conducted with regard to emotions in social media. In addition, various other theoretical models have been introduced for identifying emotional expressions from social media, namely, Plutchik's wheel of emotions, the theory of basic emotions, the consumption emotion set, and the hierarchical consumer emotion model.

Zhao et al. [22] proposed a visual analytics tool (PEARL) that supports an interactive user interface, an emotional profile overview, mood words, and raw tweets. It also has features such as the important data points, the action menu, the time window of tweets, the interactive legend for filtering data, and other visual tool tips. The authors relied on the inputs from 10 participants, and each dimension was rated on a scale of range −4 to +4, where −4 represents the most negative rate, +4 represents the most positive rate and 0 represents a neutral rate. The labeled tweets (n = 308) were categorized into 60 mood segments. To evaluate the performance of PEARL, a study was conducted to judge the emotional style.

To detect the emotional expressions from Facebook posts, Zimmerman et al. [23] utilized an intelligent method to build a test variant of an automatic 'Feelings Meter' for uncovering feelings from the text. In the study, the valence and arousal dimensions were used for mapping feelings. The feelings were taken from Scherer [24] and Russell [25]. In the final analysis, 44 'Facebook Feelings' were manually selected. The scale was used to score feelings between −0.5 and +0.5. On the valence scale, feelings such as 'sad' and 'disappointed' were rated with high scores of 0.488 and 0.464, respectively. Feelings such as 'wonderful' and 'excited' were rated with low scores of −0.427 and −0.444, respectively. Alternatively, on the arousal scale, the 'excited' and 'angry' words were scored with the highest values, 0.431 and 0.329, respectively. The 'tired' and 'sleepy' words scored −0.489 and −0.487, respectively. In the context of marketing analytics, Jussila et al. [9] proposed a tool to detect an affective phenomenon in the pleasure, arousal and dominance (PAD) dimensions that can replace the traditional measurement system for consumer satisfaction, such as a survey.

Wang et al. [26] categorized text sentiment into four types: positive, negative, neutral and mixed sentiment. In addition, the underlying emotions from the text were also identified. These emotions were adapted from the following emotions studies: Ekman [20], Plutchik [21], Socher et al. [27], Chafale et al. [28]; the methodology used was adaptive fuzzy similarity-based classification for classifying text into sentiment. Some of the emotions studied were anxiety, sadness, anger, satisfaction, excitement, and happiness.

Menon et al. [29] introduced text analytics for Facebook data such as posts and comments. The methodology contains several elements, of which NLTK was used for extracting, preprocessing, analyzing the frequency of word occurrence, the classification of text, and collocation analysis. In this study, two companies were selected for testing the tool where the Facebook posts were classified into different emotions, such as anger, fear, sadness, happiness, disgust and surprise. For this purpose, a consumer psychological domain model [30] and basic emotions were used. Other

studies [31], Sarakit et al. [32] and Larsen et al. [33] also used Ekman's emotions to classify social media data.

Shukri et al. [34] used Wilson et al. [35] polarity lexicon for classifying polarity, while Strapparava et al. [36] used an emotion lexicon for classifying emotions. The case study was related to polarity and emotion classification with regard to the automation industry. Tweets (n = 3000) on different brands of car, such as Mercedes, Audi and BMW, were retrieved using Twitter API and R Language.

Some of the investigations did not use any concepts of emotions for classifying texts. For instance, Xu et al. [37] classified texts into information or emotions and then into further classes such as appropriateness, empathy, and helpfulness. Similarly, Sun et al. [38] used five emotions, namely, neutral, happy, surprised, sad, and angry. Neither of these studies referred to any emotional theories.

2.2 Semantic Tagging

Semantic tagging is the process of associating a given text with meaningful tags. Semantic tagging plays an important role in several applications, including linked-data generation, alignment of ontologies, information retrieval and interoperability. It is possible to perform semantic tagging manually, semimanually, and automatically [39]. In manual tagging, human "annotators" perform annotation. This kind of annotation is considered expensive and may result in unreliable tagging [40]. Semimanual tagging requires human intervention at some points, which can overcome the limitations of manual tagging. Automatic tagging performs annotation in an automatic way.

Several approaches have been proposed for the semantic tagging of texts. Gadelrab et al. [41] proposed a tagging model based on a nonnegative matrix factorization technique to extract topics from texts. The approach makes use of lexical semantic correlation to capture semantics from text and performs well compared to state-of-the-art models. Some semantic tagging approaches have been developed according to the domain requirements. For the health smart home (HSH) domain, Wollschlaeger et al. [42] proposed a framework-based ontology for semantic annotation that consists of taxonomy creation methods.

Other semantic tagging approaches have been developed based on the language of the texts. Albukhitan et al. [43] developed a prototype for Arabic documents using deep learning methods. Word embedding language models are used to link nonphrases with the ontology classes and the verb phrases that are matched with the relationships of the ontology.

The locations in social networks have also attracted the attention of many researchers. Li et al. [44] proposed an approach called the similar user pattern (SUP) for tagging the locations in social networks. The approach considers the similarities of the check-in activities made by the users. The approach uses the semantic tags of the locations and the algorithm based on extreme learning machine models.

Furthermore, emerging research has been conducted on annotations to detect the context of tweets. Alkhammash et al. [45] used Ontotext to tag microcontents over tweets to understand the similarities and to improve the social services of smart cities. Ontotext is a tool used for semantic tagging where it extracts and recognizes entities such as people, places, organizations, things, date, and software. Ontotext also identifies relations between the concepts and the related entities in the domain knowledge. The recognized concepts are stored in the semantic graph database called GraphDB.

2.3 Detecting Context

Zhu et al. [46] introduced a technique that extracts, filters, and semantically categorizes "tips" from Foursquare service for the area of San Francisco. The text includes a timestamp of the posts and the location information. Authors have listed 3343 texts describing habits in the various classes, including homes, society, relaxation, food, drink, consumer purchases, sports, exercise and entertainment, occupational activities, transportation, personal cares, and training.

Zhu et al. [47] studied an automated method to estimate the amount of time a person spends using Twitter. Using "CrowdFlower" as a total task, 6004 were manually labeled and classified.

Beber et al. [48], using people trajectories, introduced a new approach for recognizing different activities conducted at one venue to classify participating individuals in the group activities of social media. For classifying the events, the ATUS dataset was utilized. Foursquare was used to collect information about where and how many tweeters were located. Earlier studies have employed an established taxonomy to categorize activities and for crowdsourcing [49].

3 Research Methodology

3.1 Use Case

We chose a use case study approach to explore social media data to enhance customer value propositions [50]. The context of the case study is a health technology company that develops wearable Internet of Things (IoT) devices and services for stress management, making use of social media data about users' stress management practices. Data collection was designed to benefit from the small data and big data approaches. The study of the use case helps to accurately discover people's feelings since a circumplex model has been employed. In the company's product, the arousal dimension is calculated by a smart wearable sensor designed to assess the electrodermal activity. Additionally, the mobile application can record valence dimensions with the diary function of the application [51]. Business entities are genuinely interested in researching and analyzing social media topics that are relevant to alleviating stress.

A campaign for Twitter marketing was planned in the university-industry partnership where the followers of the business were encouraged to share their ways to alleviate the stress and disruption. The call for participant tweets was mainly targeted to Finnish language audiences. The call for participation included the following message: "September brings challenges to everyday life. There is a hustle but also excitement! What is your best stress management practice?"

3.2 Data Gathering, Processing and Analysis

We used Twitter API and Python for data collection using Python "Twitter 1.18.0". Afterwards, forty-seven tweets with photos were processed and analyzed. Fig. 1 shows the dataset, APIs, and algorithms used by the researchers. Second, the tweets were annotated for arousal and valence, and activity was annotated using manual intervention.

To analyze the images, an image processing application programming interface (API) was used that has computer vision (CV) capabilities. The CV API was applied to detect text on images. Only a single image tag is returned for each image along with the title. English version of tweets is performed using Googletrans using PyPI, which is a public domain library available to translate tweets from Finnish to English (Step 3 in Fig. 1). For semantic tagging, we applied Ontotext tagging on the tweets. Ontotext identifies people, location, organization, etc. We noticed that a large number of words were tagged by Ontotext as a key phrase. The key phrase identifies general

"important" words and phrases from a given text. Tab. 1 shows an example of a tweet tagged using Ontotext.

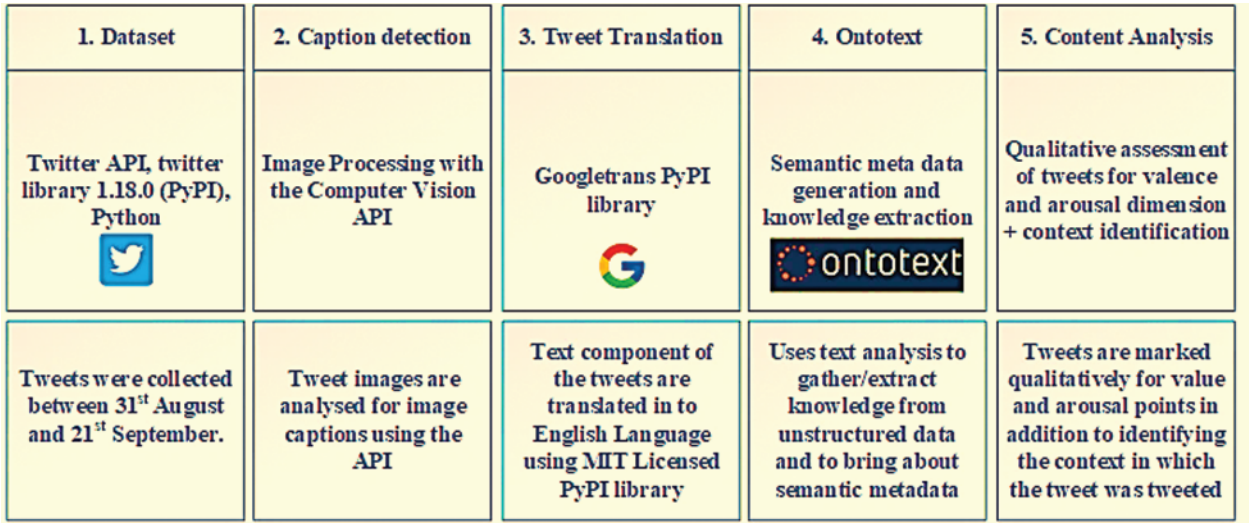

Figure 1: Analysis of tweet content

Table 1: Example of a tweet tagged using Ontotext

Tweet	Key phrase	Organization
"September brings many challenges for everyday life. There is a hurry but also enthusiasm! What is your best stress management tool? Take part in the Moodmetric Twitter campaign-you can win the @ Moodmetric smart ring". Campaign instructions: https://t.co/Qaxfig2dC4#moodmetricstressinhallintakeino, https://t.co/sfUNyirjjF	Challenges, everyday life, hurry, enthusiasm, stress management tool, part	Twitter

In this example, the word "Twitter" is tagged as location, whereas the words "challenges", "everyday life", "hurry", "enthusiasm", "stress management tool", and "part" are tagged as key phrases.

Using Twitter content review (step 5), the Twitter post is divided into different tasks depending on the context. The emotional valence and arousal values were established from the previously mentioned related studies conducted by Zimmerman et al. From this description, the behavior is qualitatively mapped under the dimensions of valence and arousal. According to the quantitative and qualitative evaluations, the picture has a high level of arousal and negative valence. Evaluating the text and image part of the tweets at the same time was made possible.

4 Results and Analysis

The results of the content and semantic analysis of the tweets are presented in Tab. 2, which shows the number of correct and incorrect tweets using two different content analysis methods and a semantic analysis of tweets.

Table 2: Results from content and semantic analysis of tweets

Classification	Content analysis		Semantic analysis
	Text (%)	Images (%)	Text (%)
Correct	91	96	50
Incorrect	9	4	50

In this case study, the content analysis of only tweet images was found to be the best method in detecting the tweet context with 96% accuracy. The second best method was the content analysis of the tweet textual content, which yielded 91% accuracy. Further analysis revealed that this was the case because in two tweets, the context was illustrated only in an image and not written in text. The semantic analysis of the tweet context using Ontotext was found to be the third best method. Semantic tagging by Ontotext was able to detect the correct tweet context with 50% accuracy. However, the Ontotext was able to tag only 30% of the tweets.

The analysis of the text component of tweets was mapped, and the visualization of the obtained results is shown in Fig. 2.

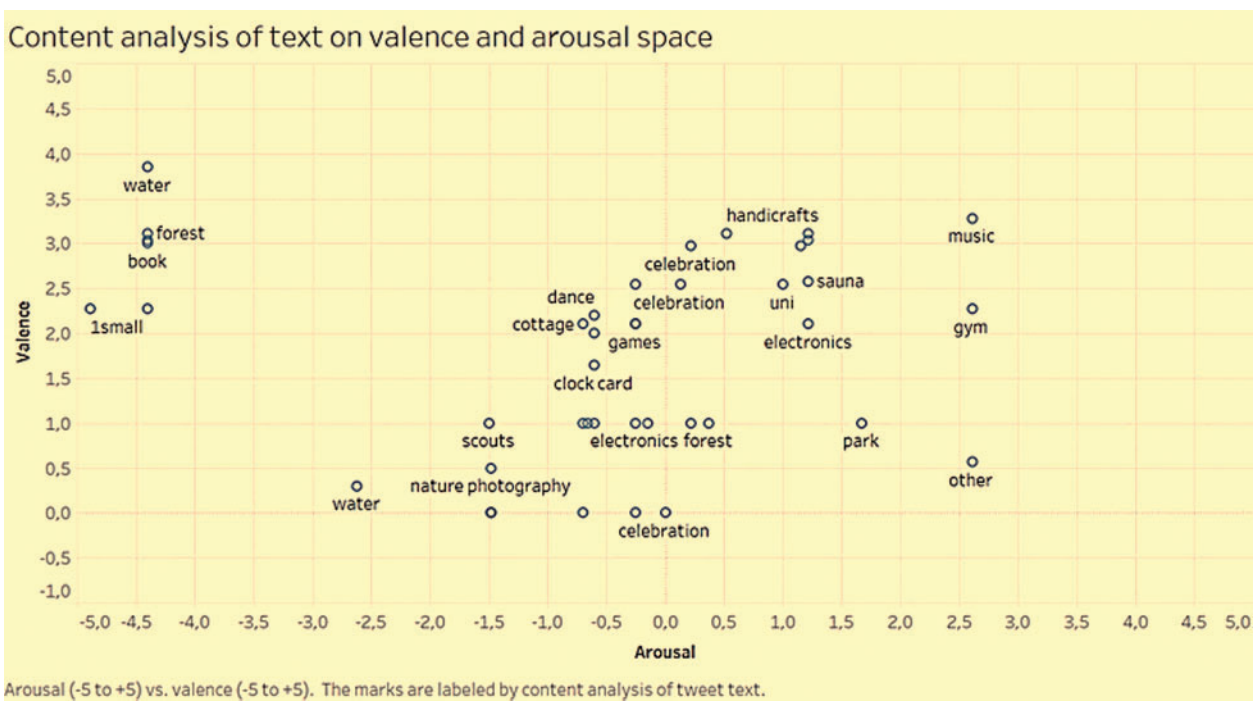

Figure 2: Content analysis of behaviors produced by emotions

Fig. 2 shows that 'music', 'gym', and 'park' received high arousal values. However, 'water', 'nature photography', and 'scouts' received low arousal values. On the other hand, 'water', 'handicrafts', 'forest', 'music', 'celebration', and 'book' received high valence ratings. Activities such as 'nature photography', 'celebration', 'scouts', 'clock card', 'forest', and 'park' received low valence ratings. Similarly, content analysis of the tweet images was performed, and the results are displayed in Fig. 3.

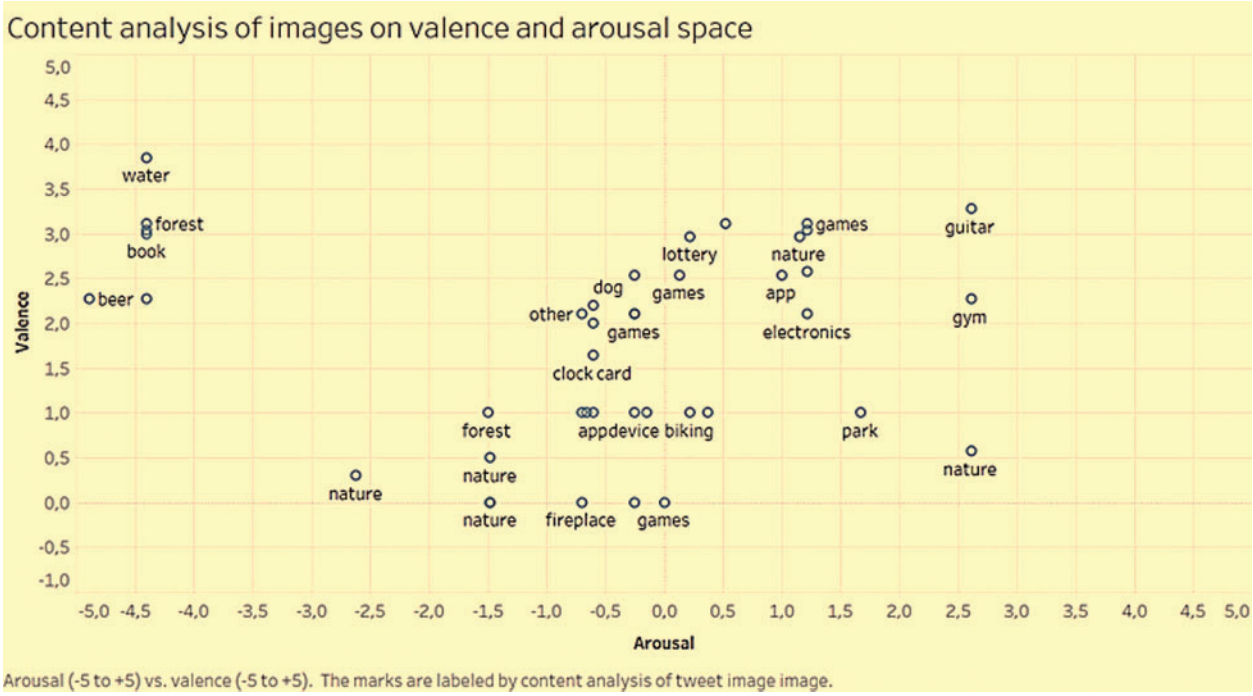

Figure 3: Content analysis of activities mapped into arousal and valence space

From Fig. 3, it can be understood that 'beer', 'water', 'book', and 'forest' received the lowest arousal values and 'guitar', 'gym', and 'nature' received the highest arousal values. On the other hand, 'nature', 'fireplace', and 'games' received lower valence ratings. However, 'water', 'forest', 'games', 'lottery', and 'guitar' received high valence scores.

The semantic analysis of the text component of the tweets was performed using Ontotext. A similar methodology was applied where the results of context detection of the Ontotext were mapped and visualized on manually coded valence, as shown in Fig. 4.

Ontotext was able to correctly detect a few of the most positive contexts of tweets, including sauna, (nature) photography and clock (card). However, it missed many of the contexts that users found the most positive in terms of relaxation or stress management, such as forest, nature, water, books, games, guitar, dance, gym, and handicrafts. Ontotext did not recognize any contexts that were classified as neutral or negative in terms of valance.

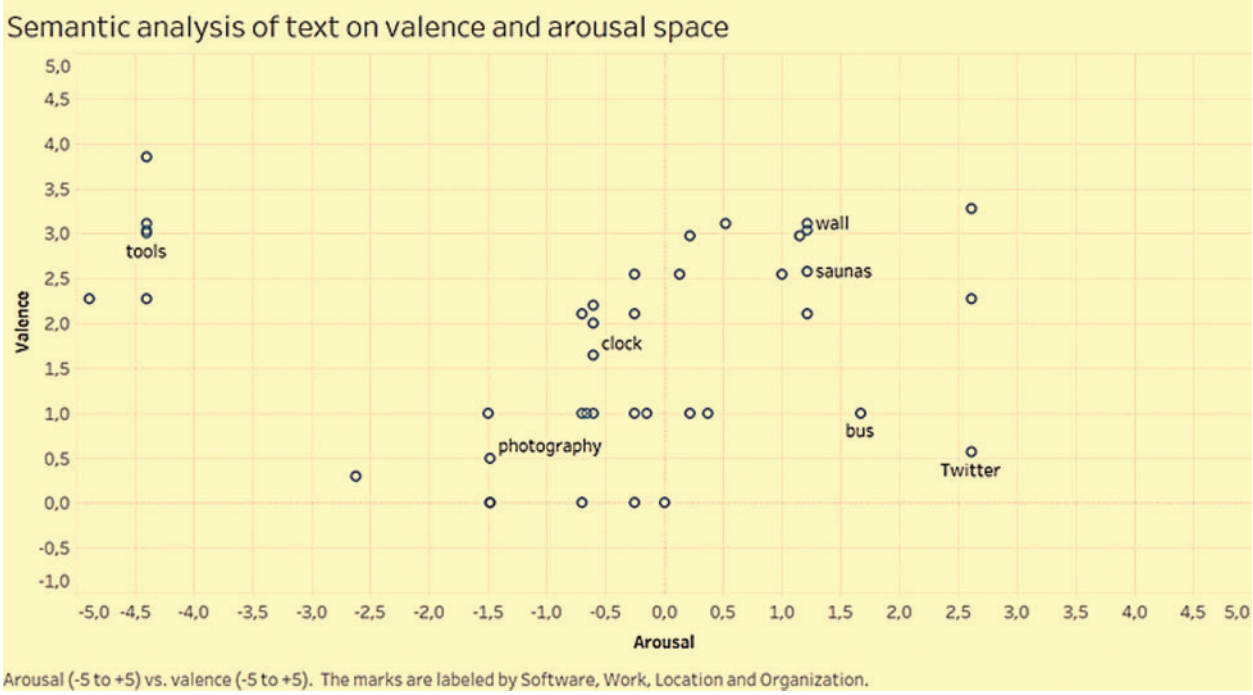

Figure 4: Semantic analysis of tweet contents plotted on valence and arousal space

5 Conclusion

The objective of this investigation was to explore the content analysis and the semantic analysis of the context of tweets for stress management. The analysis may help to develop value propositions and elicit new customer insights. Although the dataset in the study was small, the results of the study clearly indicate that increasing only the volume of social media data does not lead to valuable insights to develop value propositions. The data were limited to one language; however, making use of the Googletrans PyPI library, this method can be applied for various languages, including the Arabic language. The findings support previous research [2] that acknowledges small data in the qualitative dimension. The content analysis of only tweet images resulted in 96% accuracy to detect the tweet context, while for the tweets of the textual content, it yielded 91% accuracy. Semantic tagging by Ontotext was able to correctly detect the tweet's context with an accuracy of 50%. Future research is needed to see whether the results of this case study hold for larger data sets. Alternatively, big data methods may develop to such a degree that context can be accurately detected from social media content. In light of this study, however, it remains doubtful that such big data development methods will happen in the near future that can leave humans out of the loop from detecting context or activities from social media.

Acknowledgement: The authors would like to acknowledge the support of Taif University Researchers Supporting Project number (TURSP-2020/292), Taif University, Taif, Saudi Arabia. This research was funded by the Deanship of Scientific Research at Princess Nourah bint Abdulrahman University through the fast-track Research Funding Program.

Funding Statement: This work was supported by Taif University Researchers Supporting Project number (TURSP-2020/292), Taif University, Taif, Saudi Arabia. This research was funded by the Deanship of Scientific Research at Princess Nourah bint Abdulrahman University through the fast-track Research Funding Program.

Conflicts of Interest: The case study investigates the social media data of a health tech company in which the first author is a shareholder. The first author provided information about the case study context and participated in the review of related work. The social media data are, however, publicly accessible and objectively verifiable. The data gathering, processing and analysis in the study was conducted by researchers with no conflicts of interest.

References

[1] C. Borgman, "What are data?," in *Big Data, Little Data, No Data: Scholarship in the Networked World*. USA: MIT Press, pp. 17–29, 2015.

[2] T. Wang, "Big data needs thick data, ethnography matters," in *Ethnomining and the Combination of Qualitative and Quantitative Data*. Berkeley, California, United States: Ethnography Matters. [Online]. Available: http://ethnographymatters.net/blog/2013/05/13/big-data-needs-thick-data/.

[3] R. V. Kozinets, "Investigating: Five steps to social media data colleciton," in *Netnography: The Essential Guide to Qualitative Social Media Research*, 3rd ed., California, USA: SAGE Publications, pp. 190–211, 2019.

[4] G. Latzko-Toth, C. Bonneau and M. Millette, "Small data, thick data: Thickening strategies for trace-based social media research," in *The SAGE Handbook of Social Media Research Methods*, L. Sloan, A. Quan-Haase (Eds.), 1st ed., California, USA: SAGE Publications, pp. 199–214, 2017.

[5] J. Jussila and P. Madhala, "Cognitive computing approaches for human activity tecognition from tweets—A case study of twitter marketing campaign," in *Research & Innovation Forum 2019. RIIFORUM 2019. Springer Proceedings in Complexity*, A. Visvizi, M. Lytras (Eds.), Cham, Rome, Italy: Springer, pp. 153–170, 2019.

[6] M. Thelwall, K. Buckley and G. Paltoglou, "Sentiment strength detection for the social web," *Journal of the American Society for Information Science and Technology*, vol. 63, no. 1, pp. 163–173, 2012.

[7] F. Bravo-Marquez, M. Mendoza and B. Poblete, "Meta-level sentiment models for big social data analysis," *Knowledge-Based Systems*, vol. 69, no. 2, pp. 86–99, 2014.

[8] R. Vatrapu, "Understanding social business," in *Emerging Dimensions of Technology Management*, K. Akhilesh (Ed.), India: Springer, pp. 147–158, 2013.

[9] J. Jussila, M. Boedeker, H. Jalonen and N. Helander, "Social media analytics empowering customer experience insight," in *Strategic Innovative Marketing. Springer Proceedings in Business and Economics*, A. Kavoura, D. Sakas, P. Tomaras (Eds.), Cham: Springer, pp. 25–30, 2017.

[10] S. F. Wamba, S. Akter, A. Edwards, G. Chopin and D. Gnanzou, "How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study," *International Journal of Production Economics*, vol. 165, no. 2, pp. 234–246, 2015.

[11] R. Kitchin, "Big data, new epistemologies and paradigm shifts," *Big Data & Society*, vol. 1, no. 1, pp. 1–12, 2014.

[12] D. Boyd and K. Crawford, "Critical questions for big data: Provocations for a cultural, technological and scholarly phenomenon," *Information, Communication & Society*, vol. 15, no. 5, pp. 662–679, 2012.

[13] A. Ortony, G. Clore and A. Collins, "Factors affecting the intensity of emotions," in *The Cognitive Structure of Emotions*. Cambridge: Cambridge University Press, pp. 59–84, 1988.

[14] Y. Li and T. Li, "Deriving market intelligence from microblogs," *Decision Support Systems*, vol. 55, no. 1, pp. 206–217, 2013.

[15] M. Kumar, "Expanding the boundaries of your research using social media: Stand-up and be counted," *IETE Technical Review*, vol. 31, no. 4, pp. 255–257, 2014.

[16] G. Petz, M. Karpowicz, H. Fürschuß, A. Auinger, V. Stříteský *et al.,* "Computational approaches for mining user's opinions on the web 2.0," *Information Processing & Management*, vol. 50, no. 6, pp. 510–519, 2105.

[17] M. Rathan, H. Vishwanath, K. Venugopal and L. Patnaik, "Consumer insight mining: Aspect based twitter opinion mining of mobile phone reviews," *Applied Soft Computing*, vol. 68, no. 1, pp. 765–773, 2018.

[18] H. Wu, W. Harris, Y. Gongjun, A. Vasudeva and S. Jiancheng, "A novel social media competitive analytics framework with sentiment benchmarks," *Information & Management*, vol. 52, no. 7, pp. 801–812, 2015.

[19] P. Madhala, J. Jussila and H. Aramo-Immonen, "Systematic literature review on customer emotions in social media," in *Proc. of the 5th European Conf. on Social*, Media, Limerick, Ireland, 2018.

[20] P. Ekman, "An argument for basic emotions," *Cognition and Emotion*, vol. 6, no. 3–4, pp. 169–200, 1992.

[21] R. Plutchik, "The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice," *American Scientist*, vol. 89, no. 4, pp. 344–350, 2001.

[22] J. Zhao, L. Gou, F. Wang and M. Zhou, "PEARL: An interactive visual analytic tool for understanding personal emotion style derived from social media," in *IEEE Conf. on Visual Analytics Science and Technology*, Paris, pp. 203–212, 2014.

[23] C. Zimmerman, M. Stein, D. Hardt and R. Vatrapu, "Emergence of things felt: Harnessing the semantic space of facebook feeling tags," in *Proc. of the Thirty Sixth Int. Conf. on Information Systems*, Fort Worth, Texas, 2015.

[24] K. Scherer, "What are emotions? And how can they be measured?," *Social Science Information*, vol. 44, no. 4, pp. 695–729, 2005.

[25] J. Russell, "Pancultural aspects of the human conceptual organization of emotions," *Journal of Personality and Social Psychology*, vol. 45, no. 6, pp. 1281–1288, 1983.

[26] Z. Wang, C. Chong, S. Ho, Y. Yang, L. Lan *et al.,* "Fine-grained sentiment analysis of social media with emotion sensing," in *Future Technologies Conf.*, San Francisco, CA, pp. 1361–1364, 2016.

[27] R. Socher, A. Perelygin, J. Wu, J. Chuang, C. Manning *et al.,* "Recursive deep models for semantic compositionality over a sentiment treebank," in *Proc. of the 2013 Conf. on Empirical Methods in Natural Language Processing*, Seattle, Washington, USA, pp. 1631–1642, 2013.

[28] D. Chafale and A. Pimpalkar, "Review on developing corpora for sentiment analysis using Plutchik's wheel of emotions with fuzzy logic," *International Journal of Engineering & TechnologyInternational*, vol. 1, no. 2, pp. 1–8, 2014.

[29] K. Menon, H. Kärkkäinen, J. Jussila, J. Huhtamäki, R. Mukkamala *et al.,* "Analysing the role of crowdfunding in entrepreneurial ecosystems: A social media event study of two competing product launches," *International Journal of Entrepreneurship and Small Business*, vol. 33, no. 4, pp. 575–606, 2018.

[30] R. Lavidge and G. A. Steiner, "A model for predictive measurements of advertising effectiveness," *Journal of Marketing*, vol. 25, no. 6, pp. 59–62, 1961.

[31] W. Li and H. Xu, "Text-based emotion classification using emotion cause extraction," *Expert Systems with Applications*, vol. 41, no. 4, pp. 1742–1749, 2014.

[32] P. Sarakit, T. Theeramunkong, C. Haruechaiyasak and M. Okumura, "Classifying emotion in Thai youtube comments," in *6th Int. Conf. of Information and Communication Technology for Embedded Systems*, Hua Hin, Thailand, pp. 1–5, 2015.

[33] H. Larsen, J. Forsberg, S. Hemstad, R. Mukkamala, A. Hussain *et al.,* "TV ratings *vs*. social media engagement: Big social data analytics of the Scandinavian TV talk show Skavlan," in *2016 IEEE Int. Conf. on Big Data*, Washington, DC, pp. 3849–3858, 2016.

[34] S. Shukri, R. Yaghi, I. Aljarah and H. Alsawalqah, "Twitter sentiment analysis: A case study in the automotive industry," in *2015 IEEE Jordan Conf. on Applied Electrical Engineering and Computing Technologies*, Amman, Jordan, pp. 1–5, 2015.

[35] T. Wilson, J. Wiebe and P. Hoffmann, "Recognizing contextual polarity in phrase-level sentiment analysis," in *Proc. of the Conf. on Human Language Technology and Empirical Methods in Natural Language Processing*, USA, pp. 347–354, 2005.

[36] C. Strapparava and A. Valitutti, "WordNet-Affect: An affective extension of WordNet," in *Proc. of the Int. Conf. on Language Resources and Evaluation*, Lisbon, Portugal, pp. 1083–1086, 2004.

[37] A. Xu, Z. Liu, Y. Guo, V. Sinha and R. Akkiraju, "A new chatbot for customer service on social media," in *Proc. of the 2017 CHI Conf. on Human Factors in Computing Systems*, New York, NY, USA, pp. 3506–3510, 2017.

[38] X. Sun, C. Zhang, G. Li, D. Sun, F. Ren *et al.,* "Detecting users' anomalous emotion using social media for business intelligence," *Journal of Computational Science*, vol. 25, pp. 193–200, 2018.

[39] K. Bontcheva and H. Cunningham, "Semantic annotations and retrieval: Manual, semiautomatic, and automatic generation," in *Handbook of Semantic Web Technologies*, J. Domingue, D. Fensel, J. Hendler (Eds.), Berlin, Heidelberg: Springer. 2011.

[40] P. Bayerl, H. Lüngen, U. Gut and K. Paul, "Methodology for reliable schema development and evaluation of manual annotations," in *Proc. of the Workshop on Knowledge Markup and Semantic Annotation at the Second Int. Conf. on Knowledge Capture*, Florida, USA, pp. 17–23, 2003.

[41] F. Gadelrab, M. Haggag and R. Sadek, "Novel semantic tagging detection algorithms based non-negative matrix factorization," *SN Applied Sciences*, vol. 2, no. 1, pp. 1–18, 2020.

[42] B. Wollschlaeger, E. Eichenberg and K. Kabitzsch, "Explain yourself: A semantic annotation framework to facilitate tagging of semantic information in health smart homes," in *Proc. of the 13th Int. Joint Conf. on Biomedical Engineering Systems and Technologies*, Valletta, Malta, vol. 5, pp. 133–144, 2020.

[43] S. Albukhitan, A. Alnazer and T. Helmy, "Framework of semantic annotation of Arabic document using deep learning," *Procedia Computer Science*, vol. 170, no. 3, pp. 989–994, 2020.

[44] Y. Li, X. Zhao and Z. Zhang, "Annotating semantic tags of locations in location-based social networks," *Geoinformatica*, vol. 24, no. 1, pp. 133–152, 2020.

[45] E. Alkhammash, J. Jussila, M. Lytras and A. Visvizi, "Annotation of smart cities twitter micro-contents for enhanced citizen's engagement," *IEEE Access*, vol. 7, pp. 116267–116276, 2019.

[46] Z. Zhu, U. Blanke, A. Calatroni and G. Tröster, "Prior knowledge of human activities from social data," in *Proc. of the 2013 Int. Symp. on Wearable Computers*, New York, USA, Association for Computing Machinery, pp. 141–142, 2013.

[47] Z. Zhu, U. Blanke, A. Calatroni and G. Tröster, "Human activity recognition using social media data," in *Proc. of the 12th Int. Conf. on Mobile and Ubiquitous Multimedia*, New York, USA, Association for Computing Machinery, pp. 1–10, 2013.

[48] M. Beber, C. Ferrero, R. Fileto and V. Bogorny, "Individual and group activity recognition in moving object trajectories," *Journal of Information and Data Management*, vol. 8, no. 1, pp. 50–66, 2017.

[49] K. Shelley, "Developing the American time use survey activity classification system," *Monthly Labor Review*, vol. 128, pp. 3–15, 2005.

[50] N. Siggelkow, "Persuation with case studies," *Academy of Management Journal*, vol. 50, no. 1, pp. 20–24, 2007.

[51] J. Jussila, N. Venho, H. Salonius, J. Moilanen, J. Liukkonen *et al.,* "Towards ecosystem for research and development of electrodermal activity applications," in *Proc. of the 22nd Int. Academic Mindtrek Conf.*, New York, USA, pp. 79–87, 2018.