Sentiment Analysis Tools Should Take Account of the Number of Exclamation Marks!!!

Phoey Lee Teh
Department of Computing and Information Systems
Sunway University
Bandar Sunway, Malaysia
phoeyleet@sunway.edu.my

Paul Rayson
School of Computing and Communications,
Lancaster University, UK.
p.rayson@lancaster.ac.uk

Irina Pak
Department of Computing and Information Systems
Sunway University
lpak1992@gmail.com

Scott Piao
School of Computing and Communications,
Lancaster University, UK
s.piao@lancaster.ac.uk

ABSTRACT

There are various factors that affect the sentiment level expressed in textual comments. Capitalization of letters tends to mark something for attention and repeating of letters tends to strengthen the emotion. Emoticons are used to help visualize facial expressions which can affect understanding of text. In this paper, we show the effect of the number of exclamation marks used, via testing with twelve online sentiment tools. We present opinions gathered from 500 respondents towards "like" and "dislike" values, with a varying number of exclamation marks. Results show that only 20% of the online sentiment tools tested considered the number of exclamation marks in their returned scores. However, results from our human raters show that the more exclamation marks used for positive comments, the more they have higher "like" values than the same comments with fewer exclamations marks. Similarly, adding more exclamation marks for negative comments, results in a higher "dislike".

Categories and Subject Descriptors

D.3.2 [**Languages Classifications**]: Applicative (functional) languages – *specialized application language, exclamation marks.*

General Terms

Algorithms, Measurement, Design, Human Factors, Standardization, Languages, Theory.

Keywords

Sentiment, tools, Exclamation Marks, Comments, Emotions, Reviews.

1. INTRODUCTION

Sentiment analysis software reads text and uses an algorithm to produce an estimate of its subjective content and attitudes expressed therein [1]. Research on sentiment analysis has focused on identifying the valence of positive or negative orientation and recognizing the types of emotions and their strength or intensity. [2][3][4]. Sierdorfer et al [5] analyzed the text written in YouTube.

Gill et al. [6] classified eight emotions (surprise, joy anticipation, acceptance, sadness, disgust, anger, fear) from the content of Short Blog Texts. They related their finding to human emotion, but do not include capitals or exclamation marks in their emotional measurement.

Emotions are usually quantified with a word-based approach applying the results of predefined lexicon-based methods, such as LIWC by Pennebaker [7] and SentiStrength by Thelwall [1].

Bonny et al [8] have studied the effect of punctuation (e.g. position of punctuation in the sentence) used in Indonesian language to verify whether the type of statement either is an opinion statement or not an opinion statement. Liu at al. [9] have pinpointed that the usage of punctuation, has over-modifiers (e.g.: the word "too" big, will invert the meaning from positive to negative) which and the adverbs of degree in sentiment sentences. They proposed an approach to measure the effect of punctuation in Chinese texts. However, previous research on punctuation has not included elements of like and dislike values in their measurements, nor the intention of the punctuation used in the comments.

Naradhipa and Purwariantu [9] addressed the problem of sentiment classification though analyzing short text messages from social media. They built a system that pre-processes and eliminates the "repeated letters", where they concluded that "noooooo" should be treated as "no". They also assume that punctuation marks are misused and should be eliminated in their pre-processing. Ameur and Jamoussi [10] proposed a new technique to design positive and negative dictionaries based on Facebook comments. Their approach can predict positive or negative polarities considering symbols like emoticons, exclamations etc. Fu-liang and Gang [11] also proposed a new

improved method to analyze sentiment in micro blogs. They highlighted that users tend to express feelings though punctuation, emoticons, and symbols. They studied algorithms that can determine the grammatical structure of micro blogs including relationships with punctuation that can provide emotional value.

2. RELATED WORK

Recent studies have started to place an emphasis on punctuation to show that emoticons are becoming popular and effective in expressing mood and emotions [12] [13]. Jang et al. [14] expressed that exclamation marks can increase or decrease the strength expressed, however, the number of exclamation marks is not counted in their paper. A comment with a repeated letter "I looooooove it" may be interpreted as expressing higher positive sentiment than "I love it", as Kalman and Gerglein 2010 [15] claim that repeating a letter imitates spoken verbal speech. A study by Vandergriff [16] shows that the number of exclamation points delivers high effect and/or explicability. People tend to use emoticons in computer communication for representing mood or feelings as nonverbal face-to-face communication [17] .

Table 1 compares score ranges for sentiment polarity from six tools. Thelwall's [1] SentiStrength reads texts and uses an algorithm to produce an estimation of its sentiment. His tools measure texts from -5 to +5, in a scale of 10 points (ignoring zero), results with separate positive and negative polarities.

TextSentiment Analyzer is a professional text sentiment analysis tool that provides a score for emotional level from inputted text (via the website). Its trained model using twitter aggregated text and produces results as "positive" or "negative". It gives "Positive" and "Negative" results in numerical confidence in the range of 0.1-1 for positive outcome and 0.1 - 1 for negative outcome. For instance, a phrase of "I do not like it" is negative with confidence 0.674345.

Table 1. Measurement Ranges for Sentiment Scores

No	Author/Source	Name of Tools	Measurement Domain
1	Mike Thelwall [1]	SentiStrength	5 (Strongly negative) to 5 (Strongly positive)
2	Mashape	TextSentiment Analyzer	0.100000-1 for positive 0.100000-1 for negative
3	NLTK	Sentiment Analysis with Python NLTK Text Classification	0.1-0.9 for positive 0.1-0.9 for negative Total polarity is 1
4	Sentitweey from SenticNet Sentic Team	Sentitweety SenticNet	Positive 0.00-100 Negative 0.00-100 Total polarity is 100
5	Sentiment analysis online beta	Sentiment Analysis Engine	Neutral = -0.25 to 0.25 Bad = Between -0.25 and -0.75 Very Bad = Less than or equal to -0.75 Good = Between 0.25 and 0.75 Very Good = Greater than or equal to 0.75
6	Dr. Daniel Soper	2.1 Sentiment Analyzer	0 to -100 (very negative) 0 to +100 (very positive)

Sentiment Analysis [18], is part of Python NLTK Text Classification for processing linguistic data [19]. It measures the positive polarity in the range 0.1 to 0.9 and negative polarity in the range 0.1 to 0.9 with its total adding to 1. For example, a phrase "I do not like it" is positive overall with the results of positive: 0.7 and negative: 0.3.

SenticTweety [20] is a publicly available online tool that is able to measure sentiment values based on Twitter posts and is part of SenticNet. It determines the level of polarity as a percentage and allows deeper and multi-faceted analysis of natural language opinions [21]. SenticTweety provides overall polarity in percentages. The phrase "I do not like it" is given 60% for negative polarity and 40% for positive polarity.

The Sentiment Analysis Engine [22] is online free tool which uses a five point scale consisting of 1) Neutral, 2) Bad, 3) Very Bad, 4) Good and 5) Very Good. The range for "Neutral" is -0.25 to 0.25, for "Bad" is between -0.25 and -0.75, for "Very Bad" is less than or equal to -0.75, those two referring to negative sentiment outcome. Good is between 0.25 and 0.75 and Very Good is greater than or equal to 0.75.

Finally, Sentiment Analyzer [23] is a free tool that can analyze English texts and produce sentiment outcome in the range of -100 to +100. It was developed by Soper [24]. The phrase " I do not like it" gives the result of -100.

3. METHODOLOGY

Our experiments occurred in three distinct phases described in the following subsections.

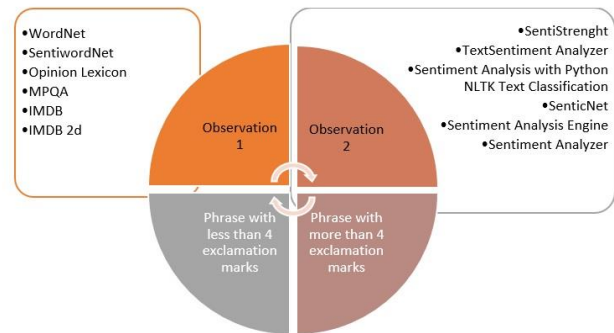## 3.1 Comparisons with Number of Exclamation Marks

Figure 1 Experimental setup for comparing sentiment tools

Observations were performed in two groups. Observation 1 was performed by comparing the nine tools in Christopher Potts' online interface. Then observation 2 compared the other six tools listed in Table 1. These two groups of tools were tested to observe the variance of outcomes with respect to the number of exclamation marks.

3.2 Corpus Collection of product review comments

1041 comments were collected relating to 10 different types of products in the following categories: 1) Beauty and Health 2)

Camera 3) Computer 4) Consumer Electronics 5) Fashion 6) Home appliance 7) Jewellery and Watch 8) Mobiles and Tables 9) Sport goods 10) Toys and Kids. Comments were analysed with the UCREL Wmatrix system [25] to calculate word frequency occurrence. The variety of sentiment expressions were categorized based on the observation of these comments. For more details, see [26].

3.3 Analysis of fine-grained sentiment categories

We carried out a survey of 30 comments with a 7 point Likert Scale ranging from 'strongly dislike' to 'strongly like'. This was used to obtain human ratings for the level of expression across variants such as capitalization and punctuation. Five hundred returned questionnaires were then analyzed with across the seven varieties. The number of Exclamation marks on both hypothesis of less than 2 exclamation marks "!!" and more than 2 exclamation marks "!!!!" of positive and negative comments has shown significant value in results in our hypothesis [26]. For the results presented below, we have exclusively focused on the variation in punctuation marks and on the number of exclamation marks in particular.

4. RESULT
4.1 Variations test on Exclamation Marks

Table 2 shows the result from the first observation tested on the online tool by Potts [27]. It provides text scores based on 6 sentiment lexicons and tools, which are: 1) WordNet, 2) SentiwordNet, 3) Opinion Lexicon, 4) MPQA, 5) IMDB, 6) IMDB 2d

Table 2 Scores from Potts [28] online tools from Observation 1

Tools	Score for "I love it !!"	Score for "I love it !!!!"
WordNet	0	0
SentiwordNet	0.25	0.25
Opinion Lexicon	1	1
MPQA	2	2
IMDB	+ 0.24	+ 0.24
IMDB 2d	+ 0.5	+ 0.5

In our comparison, none of sentiment lexicons shows difference in score when number of exclamation marks is used. Six tools from first observation with the Christopher Pott tool [27] has shown no score changes in the variant of exclamation marks used.

Table 3 shows the second observation of "exclamation marks" test on online tools such as 1) SentiStrenght, 2) TextSentiment Analyzer, 3)Sentiment Analysis with Python NLTK Text Classification, 4) SenticNet, 5) Sentiment Analysis Engine and 6) Sentiment Analyzer.

Table 3 Comparison of results of the six online tools from Obeservation 2

Tools	Score for "I love it !!"	Score for "I love it !!!!"
SentiStrenght	+ 4 -1	+ 4 -1
TextSentiment Analyzer	+0.809237	+0.809237
Sentiment Analysis with Python NLTK Text Classification	Overall is positive + 0.7 - 0.3	Overall is positive + 0.8 - 0.2
SenticNet	+ 88% - 12%	+ 86% - 14%
Sentiment Analysis Engine	+0.625	+ 0.625
Sentiment Analyzer	+100	+100

SentiStrength gives same results for both variants of exclamations marks used. TextSentiment Analyzer does not include value on the number of exclamation marks in its outcome. It gave same polarity of +0.809237 for both tests. Others results shown no significant on its sentiment value for difference in number of exclamations used, which include Sentiment Analysis Engine and Sentiment Analyzer.

Other than these, only NLTK and SenticNet has shown a slightly different in its result. Both NLTK and SenticNet changed the polarity with the variants. The nature polarity measurement from SenticNet is based on the number of tweet posted at the "certain period of time". Thus, result has shown less positive even though it is used with the positive comment and more exclamation. The Sentiment Analysis with Python NLTK Text Classification shows a slight increase (+0.1) in positive score when the number of exclamation marks is more than two.

4.2 Results of "like" and "dislike" value from 500 respondents on positives and negatives comments used with different number of exclamation marks

Turning now to the results from the 500 human raters, Table 4 and Figure 2 shows the comparison of mean "like" and "dislike" values for positive comments. A common trend can clearly be seen. Note that a score of 4 means neutral and 'original' means the same plain text with no exclamation marks. Positive comments with zero or two exclamation marks have lower "like" values than those with four exclamation marks. Positive comments with two exclamation marks tend to sit in the middle of scores for those with zero and four exclamation marks. Our conclusion from this is that the more exclamation marks in a comment, the more highly rated it is. Comment "Are you kidding me?" was picked from data collected as positive but the nature of comment might be also negative or sarcastic. That might course low "like" value rate for that comment.

Table 4 Comparison of means for comment with different number of exclamation marks

Comment	(!!)	(!!!!!)	Original
I love it	5.35	5.82	4.77
I like it	5.04	5.61	4.78
I am very happy	4.87	5.64	4.73
I am glad	5.1	5.5	4.63
I am big fan	4.73	5.41	4.76
My favourite	4.91	5.49	4.87
Hours of fun	4.8	5.45	4.4
Very satisfied	4.52	5.47	4.75
I prefer it	5.08	5.29	4.33
Really enjoy	5	5.59	4.55
I recommend it	5.12	5.46	4.45
Exceed expectations	4.85	5.56	4.73
I will continue taking this brand	3.42	5.51	4.7
Are you kidding me?	4	3.87	4.36
No need to say more	4.76	4.44	3.95

Table 5 Comparison of means for comment with different number of exclamation marks

Comment	(!!)	(!!!!!)	Original
Some serious abuse	3.24	2.66	3.57
Very disappointed	3.03	2.39	3.72
I don't care	3.34	2.68	3.51
I did hit it well	3.81	3.29	3.71
I hate it	2.99	2.44	3.27
It is really annoying	3.04	2.37	3.79
I boot it	3.65	3.18	3.81
Too much trouble	3.18	2.59	3.64
Totally fierce	3.27	2.75	3.59
I have to worry	3.68	3.27	3.88
I can afford it	4.99	5.42	4.6
What a lie	2.98	2.37	3.22
Don't come here to shop	3	2.45	3.34
Fine until it breaks	3.29	2.69	3.75
Never, ever, never	2.89	2.38	3.08

Figure 2 Like and Dislike Values of Positive comments

Positive comments

Figure 3 Like and Dislike Values of Negative comments
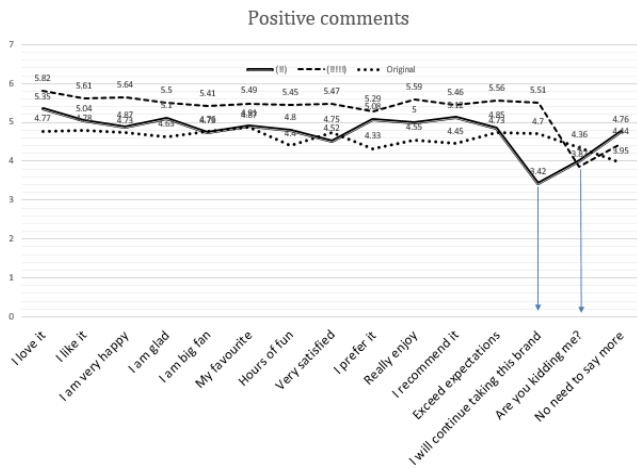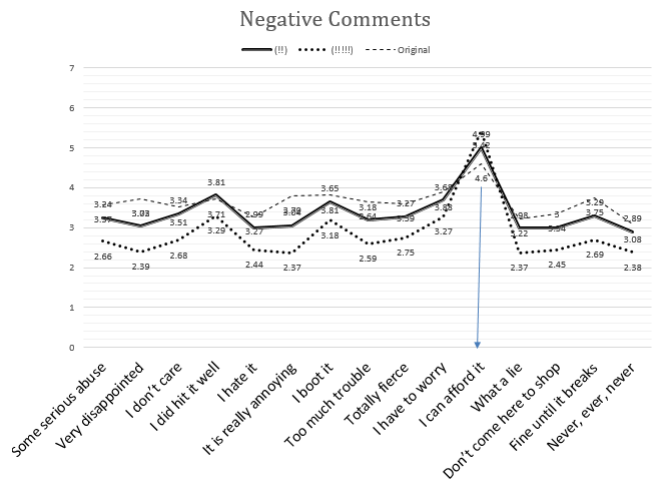

Negative Comments

Table 5 and Figure 3 illustrates the "like" and "dislike" mean values for negative comments. Again, we can observe from the questionnaire results that there is a clear trend as the more exclamation marks in a comment then the more negative it is rated. We can see that comments with more than 4 exclamation marks have the lowest result on the likert scale. While comments with two exclamation marks tend to be less negative, but more negative than comments without any exclamation marks. The trend is pretty consistent across all types of comments. Comment "I can afford it" was picked from data collected as negative comment but in another cases the same comment can be positive, it might be the reason for higher "like" value.

5. CONCLUSION

As can clearly be seen from our human rating study, the number of exclamation marks influences the sentiment value of informal text in social media. However, most of the automatic tools for sentiment analysis do not detect the difference in number of exclamation marks. In our experiments, we have tested Potts' online tools and others for determining sentiment scores for comments including "!!!!!!" and "!!" in order to see if the number of exclamation marks changes the reported sentiment value but in general there is little to no effect on the automatic scores.

In particular, we tested "I love it!!" and "I love it!!!!!" in order to track the difference in sentiment results when the same comment has between two and six explanation marks. Only two tools out of six identified any difference between "!!" and "!!!!!!".

However, from our large-scale manual rating study, the value attributed to expressions varies along with the number of explanation marks. If a comment is positive, then the value assigned manually to an expression becomes more positive when the comment contains more exclamation marks. We observe a similar effect for negative comments: the rating value assigned manually is more negative with six explanation marks, than with two or zero. In conclusion, automatic sentiment analysis tools should always include a weighting for the number of exclamation marks used. In future work, we will explore the same effect for other punctuation and textual features.

6. REFERENCES

[1] M. Thelwall, "Heart and soul: Sentiment strength detection in the social web with sentistrength," *Cyberemotions*, vol. 5, 2013.

[2] S. Aman and S. Szpakowicz, "Identifying expressions of emotion in text," *Text, Speech and Dialogue*, vol. 4629, pp. 196–205, 2007.

[3] C. Battaglino, C. Bosco, E. Cambria, and R. Damiano, "Emotion and Sentiment in Social and Expressive Media," *Citeseer*, pp. 1–226, 2013.

[4] P. Shaver, J. Schwartz, D. Kirson, and C. O'Connor, "Emotion knowledge: further exploration of a prototype approach.," *J. Pers. Soc. Psychol.*, vol. 52, no. 6, pp. 1061–1086, 1987.

[5] S. Siersdorfer, S. Chelaru, and V. Augusta, "How useful are your comments?: analyzing and predicting youtube comments and comment ratings," *Proc. 19th Int. Conf. World wide web*, vol. 15, pp. 891–900, 2010.

[6] A. J. Gill, R. M. French, D. Gergle, and J. Oberlander, "The language of emotion in short blog texts," *Proc. 2008 Conf. Comput. Support. Coop. Work*, pp. 8–11, 2008.

[7] J. W. Pennebaker, "What is LIWC?," 2015 DOI= http://www.liwc.net/.

[8] I. Bonny, T. Panggabean, and A. Purwarianti, "Opinion Sentence Classification in Indonesian Review Document," *Telecommun. Syst. Serv. Appl. (TSSA), 2012 7th Int. Conf.*, pp. 175–180, 2012.

[9] A. R. Naradhipa and A. Purwarianti, "Sentiment classification for Indonesian message in social media," *Proc. - Int. Conf. Cloud Comput. Soc. Netw. 2012 Cloud Comput. Soc. Netw. Smart Product. Soc. ICCCSN 2012*, 2012.

[10] H. Ameur and S. Jamoussi, "Dynamic Construction of Dictionaries for Sentiment Classification," *2013 IEEE 13th Int. Conf. Data Min. Work.*, pp. 896–903, 2013.

[11] G. Fu-liang and Z. Gang, "Research on Micro-blog Sentiment Orientation Analysis Based on Improved Dependency Parsing," *Consum. Electron. Commun. Networks (CECNet), 2013 3rd Int. Conf.*, pp. 546–550, 2013.

[12] Y. Urabe, R. Rzepka, and K. Araki, "Emoticon Recommendation System for Effective Communication," *Proc. 2013 IEEE/ACM Int. Conf. Adv. Soc. Networks Anal. Min.*, pp. 1460–1461, 2013.

[13] M. Thelwall, K. Buckley, G. Paltoglou, and D. Cai, "Sentiment Strength Detection in Short Informal Text," *Am. Soc. Informational Sci. Technol.*, vol. 61, no. 12, pp. 2544–2558, 2010.

[14] H. J. Jang, J. Sim, Y. Lee, and O. Kwon, "Deep sentiment analysis: Mining the causality between personality-value- attitude for analyzing business ads in social media," *Expert Syst. Appl.*, vol. 40, no. 18, pp. 7492–7503, 2013.

[15] Y. M. Kalman and D. Gergle, "052 Cmc Cues Enrich Lean Online Communication: the Case of Letter and Punctuation Mark Repetitions," *MCIS 2010 Proc.*, p. 45, 2010.

[16] I. Vandergriff, "Emotive communication online: A contextual analysis of computer-mediated communication (CMC) cues," *J. Pragmat.*, vol. 51, pp. 1–12, 2013.

[17] A. Garrison, D. Remley, P. Thomas, and E. Wierszewski, "Conventional Faces: Emoticons in Instant Messaging Discourse," *Comput. Compos.*, vol. 28, no. 2, pp. 112–125, 2011.

[18] NLTK, "Sentiment Analysis with Python NLTK Text Classification," 2015. DOI= http://text-processing.com/demo/sentiment/.

[19] M. Lobur, A. Romanyuk, and M. Romanyshyn, "Using NLTK for educational and scientific purposes," in *CAD Systems in Microelectronics (CADSM), 2011 11th International Conference The Experience of Designing and Application of*, 2011, pp. 426–428.

[20] SentiNet, "SenticTweety," 2015. DOI= http://tweety.sentic.net/.

[21] E. Cambria, D. Olsher, and D. Rajagopal, "SenticNet 3: a common and common-sense knowledge base for cognition-driven sentiment analysis," *AAAI Publ. Twenty-Eighth AAAI Conf. Artif. Intell.*, pp. 1515–1521, 2014.

[22] Online Sentiment analysis, "Sentiment Analysis Engine," 2015. DOI= http://www.sentimentanalysisonline.com/.

[23] D. Soper, "Free Sentiment Analyzer," 2006. DOI= http://www.danielsoper.com/sentimentanalysis/#.

[24] S. Daniel, "Free Sentiment Analyzer," 2015. DOI= http://www.danielsoper.com/sentimentanalysis/.

[25] P. Rayson, "From key words to key semantic domains," *Int. J. Corpus Linguist.*, vol. 13, no. 4, pp. 519–549, Jan. 2008.

[26] P. L. Teh, P. Rayson, I. Pak, and S. Piao, "Exploring fine-grained sentiment values in online product reviews," *2015 IEEE Conf. Open Syst.*, pp. 108–112, 2015.

[27] C. Potts, "Text scoring," 2011. DOI= http://sentiment.christopherpotts.net/textscores/.