Sentiment Analysis from Facebook Comments using Automatic Coding in NVivo 11

Sameerchand Pudaruth[1], Sharmila Moheeputh[2], Narmeen Permessur[3] and Adeelah Chamroo[4]

[1]Department of ICT, Faculty of Information, Communication & Digital Technologies, University of Mauritius
s.pudaruth@uom.ac.mu
[2,3,4]Department of Computer Science and Engineering, Faculty of Engineering, University of Mauritius
sharmila.moheeputh@umail.uom.ac.mu, beebee.permessur@umail.uom.ac.mu,
bibi.chamroo3@umail.uom.ac.mu

KEYWORD	ABSTRACT
Sentiment Analysis; Emotions	*The number and size of social networks have grown significantly as years have passed. With its 1.5 billion active users, Facebook is by far the most popular social networks on the planet. From kindergarten kids to grandparents to teenagers, Facebook attracts users of all ages, religions, personalities and social status. Facebook users are sharing their personal information, their lifestyle, their precious moments and their feelings online. In this paper, we download a set of comments from the page 'Opposing Views' from Facebook. These were then categorised into either a positive comment or a negative comment using the auto code feature in NVivo 11. Comments where no positive or negative sentiments are found are considered to be neutral. Out of 626 comments, 29.6% were found to contain positive sentiments while 62.0% were found to contain negative sentiments. The outcome of this work can be used by businesses to assess public reviews about their products. This will help them understand what is working and what is not. Thus, they can improve their products and respond to customer demands sufficiently quickly.*

1. Introduction

Social networks have become an integral part in the lives of billions of people. The most popular one is Facebook which is a social network where people of all ages from different communities across the world are connected. It is a platform with over 1.5 billion active users (Statista, 2016). Activities on Facebook involve public conversations through profile pages or walls. Depending on the privacy settings, these conversations may be either public to all Internet users or are limited to only one or more preferred users. Through Facebook, people can express their lows and highs of everyday life, share life experiences, establish new friendships and keep in touch with old ones. All these activities are possible through messages, pokes, likes, posts, comments, images and videos. Simple or complex emotions are usually embedded in most of these conversations.

Sameerchand Pudaruth, Sharmila Moheeputh, Narmeen Permessur and Adeelah Chamroo
Sentiment Analysis from Facebook Comments using Automatic Coding in NVivo 11

ADCAIJ: Advances in Distributed Computing and Artifical Intelligence Journal
Regular Issue, Vol. 7 N. 1 (2018), 41-48
eISSN: 2255-2863 - http://adcaij.usal.es
Ediciones Universidad de Salamanca - CC BY NC DC

41

In psychology, emotions refers to a complex state of feeling resulting in physical and psychological changes that has a direct impact on human thought and behaviour. Emotions can be expressed with words like happy, sad, angry, depressed, love, hate and so on. In this paper, emotions in posts and comments on a Facebook page will be analysed so that it can be represented and understood in a meaningful way. There are several ways to perform the analysis of the data. Sentiment analysis allows us to track attitudes and feelings from posts and comments. Sentiment analysis has proved to be profitable to businesses where they can track new product perception, brand loyalty or reputation management. For example, people's views on a certain product can be analysed to see whether the product is being viewed positively or negatively. Consequently, feedbacks on products can be obtained and businessmen can use these feedback to improve their products. To perform sentiment analysis, the relevant text must be extracted from the web using an appropriate web scraping tool. Thereafter, the text needs to be analysed to find out whether it carries any sentiment. This is usually done by looking for a set of words from a set of lexicons or by using pre-trained classifiers like Support Vector Machines (SVM) or the Naïve Bayes.

This paper proceeds as follows. In the next section, we give an overview of existing works and how sentiment analysis have been used in various fields in order to understand human behaviour. Section III describes how the data was collected, stored and processed. The results are described in detail in section IV. Section V concludes the paper with a note on limitations and future works.

2. Related Works

According to the survey done by Sharef (2014) on the Scopus database, the number of publications on sentiment analysis has been rising by a significant amount year after year in the last decade. Although most of these studies target social networks such as Twitter, Facebook and MySpace, a number of interesting works have also been carried out on other types of datasets. Some of these works are described in this section.

A hybrid approach involving lexicons and machine learning techniques were used to extract sentiments from Facebook status messages in the context of e-learning by Ortigosa et al. (2014a). In a similar study, Ortigosa et al. (2014b) describes how personality traits can be extracted from Facebook data. They used a five-class model and machine learning classifiers to predict the personality of some 20000 Facebook users. The overall accuracy of their proposed system was found to be 62%.

Terrana et al. (2014) used sentiment analysis techniques to investigate the social relationships between Facebook users. In particular, the researchers were interested to learn in real-time what a particular user talks about, with whom he discusses the most, with whom he often agrees and with whom he is constantly in disagreement. The sentiments of Facebook users with respect to 187 politicians after the 2013 German federal elections were analysed in order to assess public opinions (Caton et al., 2015). Mihaltz et al. (2015) analysed 1.9 million political Facebook comments to determine past, present and future optimism. In a study on MySpace, it was found that female users tend to receive and give more positive comments than male users although there is no significant difference for negative comments (Thelwall et al., 2010).

Bae and Lee (2012) used sentiment analysis techniques to analyse the effect of influential real-world people on their twitter followers in order to understand the impact of their tweets on the emotions of their audience. Pearce et al. (2014) analysed tweets about the Intergovernmental Panel on Climate Change (IPCC) to discover how users responded to the post of others on climate change issues. They found that, in general, people tend to reply to people who hold the same views as them. Anwar Hridoy et al. (2015) used tweets in order to analyse the popularity of the iPhone 6 in seven different regions of the USA. Ravichandran et al. (2015) proposed a new approach called the bigram item response theory (BIRT) and showed that it is more effective in deriving sentiments from tweets compared to more traditional approaches.

Sonnier et al. (2011) used a proprietary web crawler technology and a proprietary sentiment extraction technology to assess whether positive or negative reviews from online communications have an effect on the sales of a physical good. They concluded that good and neutral reviews improve sale while bad reviews can do the opposite. Their study was limited to one firm only. An interesting work was done by Purao et al. (2012) in which they demonstrate that it is possible to extract sentiments from public documents in order to assess the progress of a project or to know the reasons for its failure.

In the context of online collaboration on the Wikipedia encyclopedia, Iosub et al. (2014) found that emotions in written text are strongly influenced by gender and status. Females tend to be more emotional while people of higher status display more positive emotions than negative ones. Marrese-Taylor et al. (2014) used opinion mining techniques in order to analyse reviews from tourists obtained from TripAdvisor (TripAdvisor, 2016). They were able to demonstrate that the majority of businesses in Lake District feel that their proposed system is helpful to them.

Dubreil et al. (2008) went one step further in their analysis of French blogs. They classified blog posts and comments into either an opinion, appreciation, acceptance, refusal and judgement. However, the process was not automated but was carried out by three trained annotators. Gopaldas (2014) believes in a hybrid approach to sentiment analysis. He explains via convincing examples why computational analysis of market sentiments is unreliable, inaccurate and often incomplete.

3. Methodology

To extract the posts and comments from Facebook, we used a qualitative research software called QSR NVivo 11 (QSR International, 2016). NVivo is a software which can be used to analyse unstructured data. Hilal and Al-abri (2013) described how to use NVivo in qualitative research. The add-on NCapture for NVivo was installed in our browser to download the required information from the public page '*Opposing Views*' on Facebook (Opposing Views, 2016).

It is a page where controversial issues are shared and discussed by Facebook users. Six hundred and twenty-six posts and comments were downloaded. The dataset was then exported to NVivo for further analysis. The following steps were performed in order to classify a comment as either positive, negative or neutral.

3.1. Data Clean-Up

Data extracted from Facebook comes out with a lot of meta-data such as the PostID, CommentID, the name of the person or entity making the comment, the actual text of the post or comment, the date and time the post or comment was made and the number of likes. For the purpose of this study, only the actual texts of the posts and comments field were used. Even this field contained some extra characters that had to be removed.

3.2. Tokenisation

Each post or comment is actually stored in a string. To enable further processing, this string must be split into individual words. Tokenisation is the process which splits a string into one or more words.

3.3. Stemming

Many words in English have different forms. For example, the words analyse, analyst, analysis and analysing have the common root of *analys*. Thus, all such words are stemmed to their root so that the search process is more complete.

3.4. Query Augmentation

In NVivo, we also have the facility to look for synonyms of the keywords found in the posts/comments. For example, if the search term is fear, words such as fright, awe, dread, etc. It is also able to deal with similar words like pick, picks, picked and picking.

3.5. Classification

The posts and comments are finally tagged into positive and negative emotions or neutral by making use of the auto code feature in NVivo 11. The auto coder has built-in lexicons for positive and negative sentiments.

Examples of positive words are: happy, smile, hope, etc., and examples of negative words are: sad, fear, hate, shame, regret, anger, etc. Words which are not positive or negative are considered to be neutral.

4. Experiments & Results

We have used NVivo 11 for the analysis of the comments. NVivo 11 has a feature for the automatic tagging of sentiments to text. Sentiments can be coded as moderately positive, very positive, moderately negative and very negative. NVivo maintains separate lexicons for each of these categories. Furthermore, word modifiers like very, more or somewhat can change the class of that emotion.

Table 1: Sample of Positive Emotions.

Word	Similar Words
love	love, loved, lovely, loves, loving
please	please, pleased, pleases, pleasing
trust	trust, trusted, trusting, trusts
wish	wish, wished, wishes, wishing
proud	proud, proudly
laugh	laugh, laughed, laughing, laughter
light	bright, light, lights
promise	promise, promised, promises, promising

Table 1 shows a sample list of words that are considered to carry positive emotions. NVivo almost maintains a list of words that are similar to the keyword. For example, the word promise can occurs in several forms like promised, promises and promising. All these will be stemmed or normalised to the word promise and will count as occurences of that emotion.

Table 2: Sample of Negative Emotions.

Word	Similar Words
sad	sad, sadly, sadness
badly	bad, badly
shame	shame, shameful, shamefully, shaming
hate	hate, hated, hateful, hatefulness, hates
damn	damn, damned, damning
blaming	blame, blamed, blames, blaming
fear	fear, feared, fearful, fearing, fears
awful	awful, awfully

Table 2 shows a sample list of words that are considered to carry negative emotions. Besides maintaining a list of words that are similar to the keyword, NVivo uses uses synonyms extracted from Wordnet (Wordnet, 2016) to further augment the capabilitie of their sentiment auto-coder. For example, words like awe and fright which are synonyms of the word fear will also be considered as occurences of that emotion. Figure 1 and Figure 2 show the most frequent positive and negative words that have been used in the comments.

Figure 1: Most Frequent Words Describing Positive Emotions.

Figure 2: Most Frequent Words Describing Negative Emotions.

Sameerchand Pudaruth, Sharmila Moheeputh,
Narmeen Permessur and Adeelah Chamroo
Sentiment Analysis from Facebook Comments
using Automatic Coding in NVivo 11

ADCAIJ: Advances in Distributed Computing
and Artifical Intelligence Journal
Regular Issue, Vol. 7 N. 1 (2018), 41-48
eISSN: 2255-2863 - http://adcaij.usal.es
Ediciones Universidad de Salamanca - CC BY NC DC

45

Table 3: Sample Results after Automatic Coding.

Comment	Emotion
I hope he suffers a lot before he dies	moderately negative
Wish him death.	moderately negative
A bunch of over privileged rich, students, that have no wisdom, or a desire to get it.	moderately negative / moderately positive
We have a Canadian Goose couple who comes to our pond every year and makes their nest. We love watching the goslings grow.	moderately positive
I love happy ending. ....ba ha ha ha	moderately positive
I wish I could like this post about a hundred times	neutral
Read & learn please.	neutral
Bueatiful :)	neutral
all the students involved should have been arrested and expelled from school!	very negative
I love all of our technology, but I knew it bite us.... sooner or later!	very negative / moderately positive
I purchased one of those SMART tvs ...and pretty much proud to say I'm still dumb in using it!! Thank God !!	very negative / very positive
I went to Berkeley in the 60s when things were tough on campus. Yet, I never missed a class.	very positive
Whereas I'm saying "You're failing because you're not focusing on what's important, and you're giving up too easily."	very positive

Out of 626 comments, 215 have been coded as very negative, 173 as moderately negative, 110 as moderately positive and 75 as very positive. Comments which are not coded into these four categories are considered to be neutral. The auto code feature in NVivo does not attempt to classify a whole comment as either positive or negative, instead it looks at words in isolation. This is why we will notice from Table 3 that some comments are tagged as both moderately positive and moderately positive or both very negative positive and very negative. These results are shown graphically in Figure 3 below.
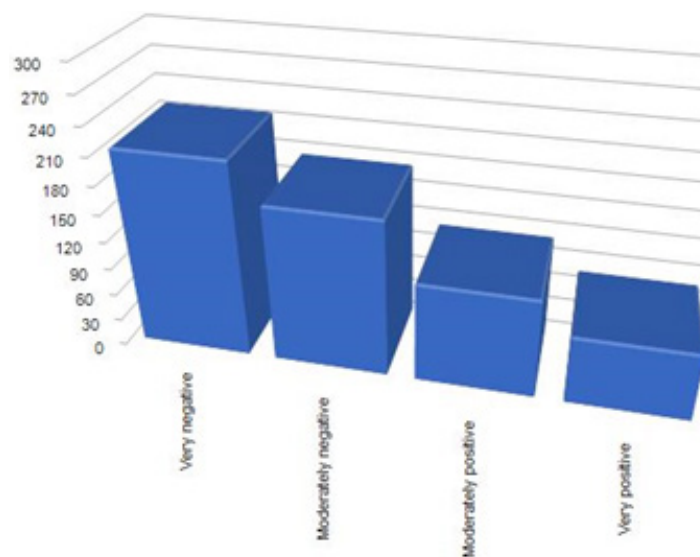

Figure 3: Summarised Results after Automatic Coding.

The sentiment auto coder in NVivo 11 would be useful for researchers who are interested to get a quick and dirty view of their data. In only a few minutes, it it possible to load the data, select the sentiment auto coder and get the results in the form of tables or charts. However, the main drawback of this auto coder is that it does not attempt the classify the whole comment into a specific emotion. It classifies only words and context is not taken into consideration. Furthermore, it cannot deal with slang, sarcasm, double negatives, idioms, etc.

5. Conclusions

In this paper, we have shown how QSR NVivo 11 can be used to extract and analyse posts and comments from a Facebook page. Our dataset is currently limited to only text data, that is, emoticons, images, audio files and videos were not taken into consideration. The auto code facility in NVivo 11 was used to tag the comments with the appropriate emotion. A comment can be tagged with both positive and negative sentiments. Stemmed words and synonyms are also used in the comparison process. For the page '*Opposing Views*', we have seen that the percentage of negative comments is more than twice the number of positive comments. Our aim in this paper is not to show that people tend to express more negative views than positive views but rather it should be considered more as a description of the use of sentiment analysis to extract human expressions on social media. Businesses can use sentiment analysis to understand the voice of the market, improve their brand management strategies, gain competitive advantage and develop new improved products. In the future, we intend to repeat the same experiment on several public pages on Facebook and then compare the results. We also intend to compare the NVivo auto coder with other sentiment classification tools.

6. References

Anwar Hridoy, S. A., Ekram, M. T., Islam, M. S., Ahmed, F. and Rahman, R. M., (2015). Localized Twitter Opinion Mining using Sentiment Analysis. *Decision Analytics*, 2(8). doi:10.1186/s40165-015-0016-4.

Bae, Y. and Lee, H., (2012). Sentiment Analysis of Twitter Audiences: Measuring the Positive or Negative influence of Popular Twitterers. *Journal of the American Society for Information Science and Technology*, 63(12), pp. 2521-2535.

Caton, S., Hall, M. and Weinhardt, C. (2015). How do Politicians use Facebook? An Applied Social Observatory. *Big Data & Society*, July-December 2015, pp. 1-18.

Dubreil, E., Vernier, M., Monceaux, L. and Daille, B., (2008). Annotating Opinion – Evaluation of Blogs. Workshop on Sentiment Analysis: Emotion, Metaphor, Ontology and Terminology (EMOT 2008), pp. 124, Marrakech, Morocco..

Gopaldas, A. (2014). Marketplace Sentiments. *Journal of Consumer Research,* 41(4), pp. 995-1014.

Hilal, A. H. and Alabri, S. S. (2013). Using Nvivo for Data Analysis in Qualitative Research. *International Interdisciplinary Journal of Education*, 2(2), pp. 181-186.

Iosub, D., Laniado, D., Castillo, C., Morell, M. F. and Kaltenbrunner, A., (2014). Emotions under Discussion: Gender, Status and Communication in Online Collaboration. *PLoS ONE* 9(8): e104880. doi:10.1371/journal.pone.0104880.

Marrese-Taylor, E., Velasquez, J. D. and Bravo-Marquez, F., (2014). A Novel Deterministic Approach for Aspect-Based Opinion Mining in Tourism Product Review. *Expert Systems with Applications*, 41(17), pp. 7764-7775.

Mihaltz, M., Varadi, T., Cserto, I., Fulop, E., Polya, T. and Kovago, P., (2015). Beyond Sentiment: Social Psychological Analysis of Political Facebook Comments in Hungary. In: Proceedings of the 6th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis (WASSA 2015), pp. 127-133, Lisbon, Portugal.

OpposingViews, 2016. Opposing Views [online]. Available from: https://www.facebook.com/opposingviews [Accessed 15 January 2016].

Sameerchand Pudaruth, Sharmila Moheeputh,
Narmeen Permessur and Adeelah Chamroo
Sentiment Analysis from Facebook Comments
using Automatic Coding in NVivo 11

ADCAIJ: Advances in Distributed Computing
and Artifical Intelligence Journal
Regular Issue, Vol. 7 N. 1 (2018), 41-48
eISSN: 2255-2863 - http://adcaij.usal.es
Ediciones Universidad de Salamanca - CC BY NC DC

47

Pearce, W., Holmberg, K., Hellsten, I. and Nerlich, B., (2014). Climate Change on Twitter: Topics, Communities and Conversations about the 2013 IPCC Working Group 1 Report. *PLoS ONE* 9(4): e94785. doi:10.1371/journal.pone.0094785.

Purao, S., Desouza, K. C. and Becker, J., (2012). Investigating Failures in Large-Scale Public Sector Projects with Sentiment Analysis. *e-Service Journal*, 8(2), pp. 84-105.

Ravichandran, M., Kulanthaivel, G. and Chellatamilan, T., (2015). Intelligent Topical Sentiment Analysis for the Classification of E-Learners and their Topics of Interest. *The Scientific World Journal*. http://dx.doi.org/10.1155/2015/617358.

Sharef, N. (2014). A review of Sentiment Analysis Approaches in Big Data Era. In: *Proceedings of the Malaysian National Conference on Databases*, MaNCoD 2014, pp. 7-12, Serdang, Malaysia.

Sonnier, G., McAlister, L. and Rutz, O. J., (2011). A Dynamic Model of the Effect of Online Communications on Firm Sales. *Marketing Science*, 30(4), pp.702-716.

Statista, 2016. The Statistics Portal [online]. Available from: http://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/ [Accessed 15 January 2016].

Terrana, D., Augello, A., and Pilato, G., (2014). Facebook User Relationships Analysis based on Sentiment Classification. In: *Proceedings of the 2014 IEEE International Conference on Semantic Computing*.

Thelwall, M., Wilkinson D. and Uppal, S. (2010). Data Mining Emotion in Social Network Communication: Gender Differences in MySpace. *Journal of the American Society for Information Science & Technology*, 61(1), pp. 190-199.

TripAdvisor, 2016. TripAdvisor [online]. Available from: http://www.tripadvisor.com/ [Accessed 15 January 2016].

WordNet, 2016. WordNet: A Lexical Database for English [online]. Available from: https://wordnet.princeton.edu/ [Accessed 15 January 2016].

Sameerchand Pudaruth, Sharmila Moheeputh,
Narmeen Permessur and Adeelah Chamroo
Sentiment Analysis from Facebook Comments
using Automatic Coding in NVivo 11

ADCAIJ: Advances in Distributed Computing
and Artifical Intelligence Journal
Regular Issue, Vol. 7 N. 1 (2018), 41-48
eISSN: 2255-2863 - http://adcaij.usal.es
Ediciones Universidad de Salamanca - CC BY NC DC

48