Journal of Theoretical and Applied Information Technology
20th December 2015. Vol.82. No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

SENTIMENT ANALYSIS BASED ON APPRAISAL THEORY FOR MARKETING INTELLIGENCE IN INDONESIA'S MOBILE PHONE MARKET

[1] ANDRY ALAMSYAH, [2] WIDITA RAHMAH, [3] HERRY IRAWAN
School of Economics and Business, Telkom University, Bandung, Indonesia
E-mail: [1]andrya@telkomuniversity.ac.id, [2] rahmahwidita@gmail.com, [3] herryir@telkomuniversity.ac.id

ABSTRACT

The growth of internet has prompt online marketing intelligence activities. Twitter becomes most popular social media in Indonesia and provides millions of text data. People share their opinion and emotions by their tweets. This phenomenon gives an opportunity to conduct a research for sentiment analysis. In this research we use appraisal theory that can analyze the emotions on the tweets according to basic descriptions on appraisal theory that include attitude, graduation and engagement towards products. We apply this theory in sentiment analysis using Bahasa Indonesia to determine and compare positive and negative sentiment of popular smartphone products in Indonesia, which is Lumia and Xperia. We perform calculation based on target and appraisal related, in the sentences level and the tweets level. The results of this research can be used as knowledge and consideration for business especially marketing intelligence field. Based on the calculations, Lumia have more positive results than Xperia.

Keywords: *Sentiment Analysis, Appraisal theory, Marketing Intelligence, Twitter, Lumia and Xperia*

1. INTRODUCTION

Internet has become a powerful platform that has changed the paradigm of communication and business, and a universal source of information that can be used by millions of people. *InternetWorldsStats* showed per 30 June 2014, the biggest internet population is in Asia region. Indonesia was on the fourth rank in Asia after China, India and Japan. As many as 71.2 million of internet users [1]. In Indonesia, one of social media's most widely used is *Twitter*. *Twitter* has penetration of about 11% of the Indonesian total population. There are many active social networks services such as *Facebook, WhatsApp, Facebook Messenger, Google +, LinkedIn, Instagram, Skype, Pinterest and Line* [2], but the openness of *Twitter* give us richer sources and varied resources of opinions and sentiments that can be about anything, including attitude towards product [3].

Sentiment analysis or opinion mining analyzes people's opinions, appraisals, attitudes, and emotions toward entities, individuals, issues, events, topics, and their attributes. Opinions are important because they are key influencers of our behaviors [4]. *Lumia* and *Xperia* are popular in Twitter, their accounts entered into the rank of fastest growth profile, based on *socialbakers.com* [5]. In 14 March 2014, Nokia has challenged Sony in terms of excellence cameras. At one time there are heat debates in Twitter, between official *Twitter* account *Nokia* and *Sony* to defend their own product [6]. Consumers use various types of online forums for social engagement including social network services such as *Facebook* and *Twitter*. Using the services, consumer involvement occurs in real time. Type of interaction offers the opportunity unprecedented for Marketing Intelligence [7]. We used sentiment analysis based on appraisal theory. Appraisal theory is a theory that describes the type of language used in communicating emotions and opinions [8]. The purpose of this research is to show that we can summarize market opinions from abundance data on the internet's, especially opinions about Sony *Xperia* and Nokia *Lumia* products.

2. LITERATURE REVIEW

2.1 Data Mining

Data mining is the process of discovering insightful, interesting, and novel patterns, as well as descriptive, understandable, and predictive models from large-scale data [9]. Data mining is an interdisciplinary field with contributions from many areas, such as statistics, machine learning, information retrieval, pattern recognition and bioinformatics. Data mining is widely used in many domains, such as retail, finance, telecommunication, social and many more [10].

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

2.2 Text Mining

Text mining is data mining of textual data, and now becoming a growing area with huge potential for business. Text mining requires a text and can be analyzed in electronic form. It involves filtering analysis on the meaning, summarizing the content through text processing, using keyword analysis, cluster themes and documents, obtain information with important words and phrases that are relevant. Text mining process is to analyze the words of data or non-numeric data and a document to define words that are attractive as a variable [11].

2.3 Sentiment Analysis

Sentiment analysis, also called opinion mining, is the field of study that analyzes people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services, organizations, individuals, issues, events, topics, and their attributes. It represents a large problem space. Sentiment analysis automated is needed because the average human reader will have difficulty identifying relevant sites, extracting and summarizing the opinions in them [12].

2.4 Appraisal Theory

Appraisal Theory describes how authors use language to communicate their engagement with others [8]. Appraisal Theory shifts sentiment classification further and considers the appraisal expression – a basic grammatical unit by which an opinion expressed, Appraisal Theory as a basic description has basic attributes [13]. The following attributes are:

a. Attitude, expresses the current state of a person at the time she or he wrote a text. It has multiple subcategories: affect, which represents the feelings of the author such as happy, sad, or others
b. Appreciation, which talks about the opinion that a person has about the inner or outer qualities of an object such as ugly, beautiful, shy, and judgment of an object, which describes the behavior of somebody in a social context such as heroic, feeble-minded.
c. Engagement determines the position of text proposal. It reflects probability or possibility such as in words perhaps or seems, in most cases.

2.5 Marketing Intelligence

Marketing intelligence represents a continuous process of understanding, analyzing, and assessing a firm's internal and external environments associated with customers, competitors, and markets and then using the acquired information and knowledge to support the firm's marketing-related decisions. Marketing intelligence provides a road map of current and future trends in customers' preferences and needs, new market and segmentation opportunities, and major shifts in marketing and distribution in order to improve the firm's marketing planning, implementation, and control [14].

2.6 User Generated Content (UGC)

UGC denotes any form of content such as blogs, wikis, discussion forums, posts, chats, tweets, podcasting, pins, digital images, video, audio, and other forms of media that was created by users of an online system or service, often made available via social media web sites [15].

3. METHODOLOGY

In this research, we use data mining method to conduct Knowledge Discovery in Database (KDD). KDD is the nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data [4]. We use social network services *Twitter* as our data source. The tweet mining is in Bahasa Indonesia which is contained basic description on appraisals theory (attitude, graduation, engagement) and target (*Lumia, Xperia* and/or its feature). We analyze tweets data using entity and aspect level in sentiment analysis, this level is looking at language constructs such as documents, paragraph, sentences, clauses or phrases) and directly looks at the opinion itself. The opinion consists of positive or negative sentiment and a target of an opinion [12]. We crawl data using *Twitter* Application Programming Interface (API) functionality. The data crawling duration is from 1 January 2015 until 30 June 2015.

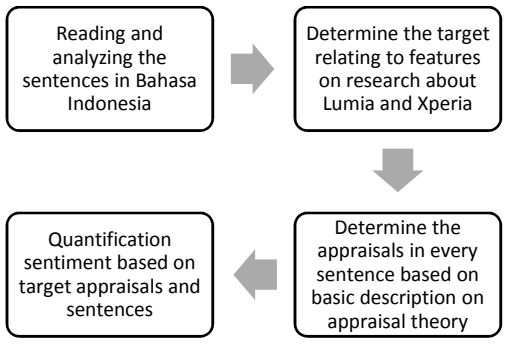

Figure 1: The research workflow

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

We construct range scoring library to support Bahasa Indonesia structures and grammars. We show small part of the library in table 1 below:

Table 1: Range scoring library

Attitude		*Graduation*				*Engagement*
A1 Value = +2	A2 Value = -2	B1= A1 + Words Value = +3	C1 = Words + A1 Value = +1	D1 = B1 + Emote Value = +4	D4 = A2 + Emote Value = -3	E
				D2= B1 + Emote Value = +4	D5 = B2 + Emote Value = -4	
		B2 = A2 + Words Value = -3	C2 = Words + A2 Value = -1	D3 = Emote Value = +1	D6 = Emote Value = -1	
Senang/Happy	*Kecewa*/Disappoint	*Sangat*/Very	*Lumayan*/Pretty fair	:)	"-___-"	*Sebelumnya*/Before
Suka/Like	*Nyesel*/Regret	*Terlalu*/Too	*Agak*/Rather	:D	:(	*Sesudahnya*/After
Bagus/Good	*Jelek*/Bad	*Benar-benar*/Really	*Rada*/Rather (Informal form)	^___^	!	*Mungkin*/Maybe

The algorithm we use is modified algorithm from [13] adapted for Bahasa Indonesia and informal language used in *Twitter* data.

```
appraisalvalue == 0
    for target : sentence.target
        for appraisal : target.appraisal
            if appraisal.category == attitude or graduation or engagement
                if appraisal.category == attitude
                    if appraisal.category.value == A1 or A2
                        targetscore +2
                    else
                        targetscore -2
                elseif appraisal.category == attitude and graduation
                    if appraisal.category.value == A1 + Words
                        targetscore +3
                    elseif appraisal.category.value == A2 + Words
                        targetscore -3
                    elseif appraisal.category.value == A1 + Emote
                        targetscore +3
                    elseif appraisal.category.value == A2 + Emote
                        targetscore -3
                    elseif appraisal.category.value == B1 + Emote
                        targetscore +4
                    elseif appraisal.category.value == B2 + Emote
                        targetscore -4
                    elseif appraisal.category.value == Words + A1
                        targetscore +1
                    elseif appraisal.category.value == Words + A2
                        targetscore -1
                    end if
                        elseif appraisal.category == graduation
                            if appraisal.category.value == D3
                            targetscore +1
                            if appraisal.category.value == D6
                            targetscore -1
                            end if
                elseif appraisal.category == attitude and engagement
                    if appraisal.category.value == A1 + E
                    targetscore +2
                appraisal.category.value == A2 + E
                    targetscore -2
                    end if
                end if
            else
                targetscore == 0
```

Figure 2: Sentiment analysis based on appraisal theory algorithm

Example of tweet processing:

Mau Reparasi **Xperia z1** lama dan mahal sekali ternyata ya T1 A A
Layanan @sonyindonesia makin mengecewakan saja T2 A
Jadi **nyesel** rasanya pakai Sony A T3

Description

C: Category (Sentences in Bahasa Indonesia)
T: Target / research object
A: Appraisal

T1: Xperia Z1
T2: @sonyindonesia
T3: Sony

We split sentences on the tweets, and analyze them according to the algorithm into category

C1: *mau reparasi Xperia z1 lama dan mahal sekali ternyata yaa*
C2: *layanan* @sonyindonesia *makin mengecewakan saja*
C3: *jadi nyesel rasanya pakai sony*

C1: T1 related with *Appraisal* (A) = *Lama* (Long), *Mahal Sekali* (Very Expansive)
C2: T2 related with *Appraisal* (A) = *Makin Mengecewakan* (More Disappointed)
C3: T3 related with *Appraisal* (A) = *Nyesel* (Regret)

Targetscore (T1) = -2-3 = -5
Targetscore (T2) = -3
Targetscore (T3) = -2
Total TargetScore = -10 (negative)

Journal of Theoretical and Applied Information Technology
20th December 2015. Vol.82. No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4. RESULTS AND DISCUSSION

By using sentiment analysis based appraisal theory algorithm, we calculate the amount of target score on each tweet. The total target score of all tweets is shown in figure 3.

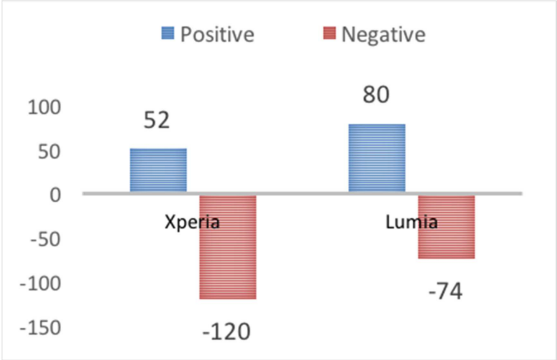

Figure 3. Sentiment results of Lumia and Xperia based on total target score

We also classify the product sentiment based on on details opinion in each tweets to one of following class: products/features, services, brand and price. We provide this approach in order to support marketing intelligence effort, which is to get details opinions about specific aspects of the products. By far the classification process is done manually, because of the complexity doing automation process in Bahasa Indonesia without proper corpus available. The details results of *Xperia* is in figure 4 and *Lumia* is in figure 5 below:

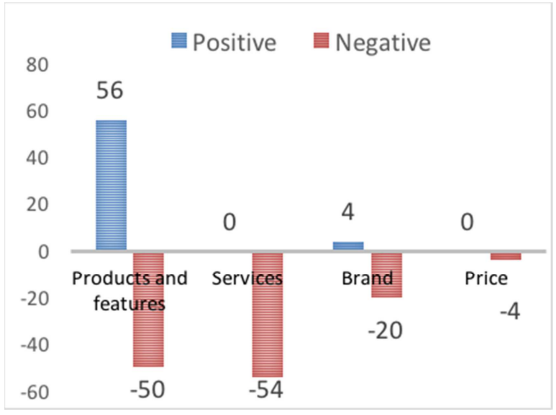

Figure 4: Sentiment results of Xperia details opinion

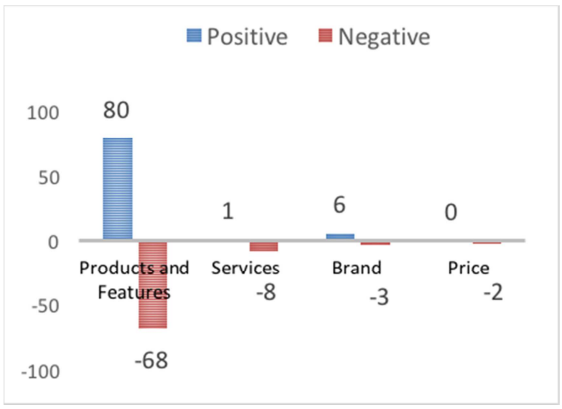

Figure 5: Sentiment results of Lumia details opinion

In addition to the calculation, we also identify the most often used words as appraisal when people gives their opinions about *Lumia* and *Xperia* products. The result is in figure 6 and figure 7.

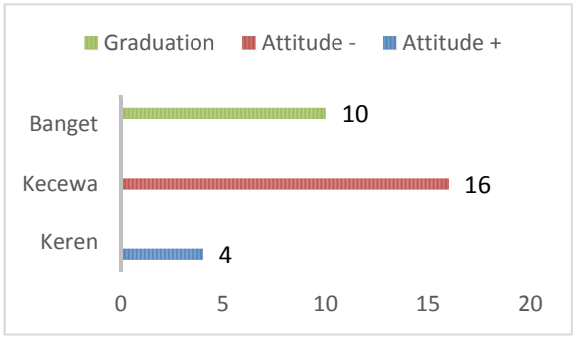

Figure 6 The most often used words in appraisal about Xperia product

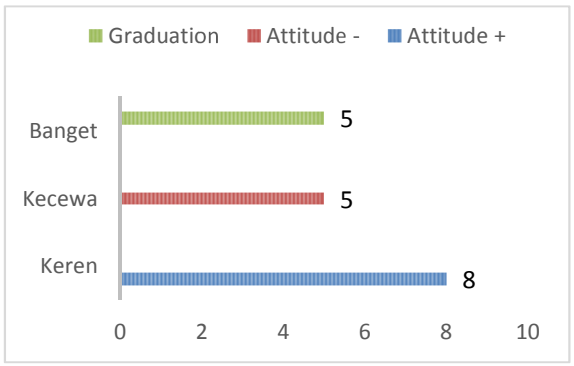

Figure 7 The most often used words in appraisal about Lumia product

In Bahasa Indonesia "*Banget*" (informal form in Bahasa) means "Very", "*Kecewa*" means "Disappoint" and "*Keren*" means "Cool". After we calculate the results on every tweet refers to total

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

targets core and based on target that often appear, we found a knowledge for marketing intelligence considerations using sentiment analysis based appraisal theory.

The total results from calculations shows that *Lumia* is generally superior to *Xperia*. *Lumia* has total target score of +6 (positive) while *Xperia* has total score of -68 (negative). *Xperia* considered to have poor services by having the most negative sentiment score of -54. *Lumia* has the most negative sentiment about the product with score of -68. In general, *Lumia* has slightly better product/features than *Xperia.*

The methodology we proposed can mine market sentiment towards products, in this case is smartphone market. With the user generated content that share customer's experience, it can be a fundamentals evaluation to other potential buyers and marketers. This activity supports marketing intelligence efforts on the internet [16].

5. CONCLUSION

The sentiment analysis based on appraisal theory can support marketing intelligence effort in collecting market opinions about our and competitor's products. With the availability of accessible social network services such as *Twitter* to anyone means more data generated by users. Collecting these opinions are considered much cheaper and faster comparing collecting data using legacy methodology of marketing intelligence such as sending out questionnaires, interviewing peoples, or any other means.

The sentiment analysis based on appraisal theory algorithm that we propose can be used to distinguish appraisal in Bahasa Indonesia and informal language used in social network services such as *Twitter*. We construct range scoring library in figure 1 as complement of our algorithm.

We also introduce product details opinion classification such as products / features, services, brand and price. This approach is useful to get details insight about product strong and weakness. With this result companies can increase their quality of services, products and brands. This approach provides information for company to understand their products, services, competitor's strength and weakness faster than before.

In the future, we suggest the automation process of target and appraisal classification and also product details opinion classification. This automation process is important to support large-scale stream data from social network services and getting real time insight that support marketing intelligence and decision making.

REFERENCES:

[1] IWS, *Internet User in The World* [Online], Available at : http://www.internetworldstats.com/asia.htm , 2014.

[2] Statista.com, *Twitter penetration in selected Asia Pacific countries as of 4th quarter 2014* [Online], Available at: http://www.statista.com/statistics/254803/twitter-penetration-in-selected-apac-countries/ , 2015.

[3] A. Kumar, T.M. Sebastian, *"Sentiment Analysis on Twitter"*. IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 4, No 3, July 2012. Page 372.

[4] B. Liu, *Web Data Mining Exploring Hyperlinks, Contents, and Usage Data*. New York: Springer, 2011.

[5] Socialbakers, *Twitter Statistics in Indonesia - Phone* [Online], Available at: http://www.socialbakers.com/statistics/twitter/profiles/indonesia/brands/electronics/phone/page-2 / , 2015.

[6] Gadgets.ndtv, *Nokia challenges Sony to a camera duel, Lumia 1020 versus Xperia Z2* [Online], Available at: http://gadgets.ndtv.com/internet/news/nokia-challenges-sony-to-a-camera-duel-lumia-1020-versus-xperia-z2-497430 ,2015.

[7] M. Rambocas, J. Gama, *Marketing Research: The Role of Sentiment Analysis*, Retrieved from http://wps.fep.up.pt/. Page 1, 2013.

[8] J. Read, J. Carrol, *Annotating Expressions of Appraisal in English*, Springer, 2010.

[9] M.J. Zaki, W. Meira Jr, *Data Mining and Analysis: Fundamental Concepts and Algorithms*, Cambridge University Press, 2014

[10] Y. Zhao, *R and Data Mining: Examples and Case Studies*. Elsevier. 2013.

[11] D. Olson, S. David, Y. Shi, *Introduction to Business Data Mining*, New York: Mc-Graw-Hill, 2007.

[12] B. Liu, *Sentiment Analysis and Opinion Mining*. Toronto: Morgan & Claypool Publishers, 2012.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[13] P. Korenek, M. Šimko, *Sentiment Analysis on Microblog Utilizing Appraisal Theory*, Springer, 2013.

[14] C. Chern, A. Lee, J.T. Wei, *Introduction to The Special Issue on Data Analytics for Marketing Intelligence*. Springer, 2014.

[15] M.F. Moenz, J. Li, T. Chua, *Mining User Generated Content*, Boca Raton: CRC Press, 2014.

[16] P. Kotler, K. Keller, *Marketing Management*. New Jersey : Prentice Hall, 2012