Certain Investigations on Sentimental Analysis Architecture and Tools

R. P. Ramkumar, Sanjeeva Polepaka

Abstract*— The sentiment is defined as the feeling(s) about the review or comment. The Sentimental Analysis aims to determine the attitude of content or product for a period at a given moment. Later, these observations are categorized as negative, neutral, positive and sometimes no sentiment(s) at all. The review(s) or comment(s) on a concern product is beneficial for the companies to prioritize the issues, narrow down the problems to be solved and to explore the scenarios for success. This article deals with the study of sentimental analysis or opinion mining architecture and tools used for Sentimental Analysis for the naive users.

Keywords: Opinion Mining, Tweets, Opinion Polarities, Crawling, Sentimental Analysis, Twitter Statistics.

1. INTRODUCTION

In the technological era, others opinion about the product influences the decision-making process. One or two decades before to get the review about the product(s), we relied mostly on relatives and friends. However, in the modern world, opinion feedback from diverse people may be sought over the internet. Before buying the product, people use to look the website(s) for review about a particular product. Similarly, organizations use to get feedback about the services and the products for the customers. The identifying and extracting processes generate the subjective information using text analysis, natural language processing and computational linguistics is referred to as Sentimental Analysis or Opinion Mining. Nowadays, this type of sentimental analysis becomes more popular to monetize the products. The following shows the organization of the paper: Section 2 deals with related works about opinion mining. Usage of sentimental analysis and its architecture is discussed in Sections 3 and 4, respectively. Section 5 draws the conclusion and future work.

2. RELATED WORKS

The following section shows the related work regarding the opinion mining framework.

Meenambigai [4] presented a product based opinion mining to examine the nature of the product. The objective of this work was to categorize the opinion polarities (like a negative, neutral or positive) of the product. From the product's opinion statement, sentiment analysis was done and classified as objective, positive and negative. This method was tested on blog posts) from social media. There were three steps in determining the analysis. In step 1, all the product features were extracted. In step 2, simultaneously separation of text and emoticons and extraction of opinion words took place. The categorization of opinion words such as positive, neutral and negative words was done in step 3. Results concluded that people use emoticons to express their feelings in natural language text.

Vivekanandan and Josephine [5] constructed an automated framework called Review Opinion Mining (ROM) to determine the opinion about the online products. The ROM method has five steps namely (a) Data preprocessing, (b) Aspect extraction, (c) Identification of opinion, (d) Polarity identification and (e) Summarization of features. This method was tested on unstructured data to extract the viewer's opinion about the products. Though the ROM framework analyzed product reviews in a timely and efficient manner, it failed to adhere to appropriate algorithms for determining the summary report.

Khandelwal, Mishra and Mishra [6] implemented an analyzer to classify twitter comments as positive, negative and neutral. This workflow model initiates with tweet data identification followed by the data preprocessing and training data definition. Later, parsing of tweets was carried out by choosing an efficient algorithm. Subsequently, the training and evaluation follow to classify the comments as subjective and objective cases, with the test data. Authors concluded that machine learning algorithms such as Support Vector Machines (SUM) and Naïve Bayes were widely used to analyze the tweets even though the algorithm suffers limitations.

Kim and Kim [7] presented a case study on nuclear power using opinion mining on Twitter. There were four phases or stages in analyzing the tweets, namely (a) Crawling of tweets, (b) Text preprocessing, (c) Constructing the dictionary with sentiment words and (d) Predicting the feelings with tweets. Locoy spider is the crawling tool used to identify the terms "nuclear power" or simply the "nuclear" in the Korean language from 2009 to 2013. Among the five years of data, the first three years were used to construct the dictionary, whereas the later year's data were used for evaluation purpose. Results of the assessment method were compared with the human evaluator(s) results. Results concluded that this method has higher prediction accuracy on analyzing sentiments than the human evaluators.

Revised Version Manuscript Received on 22 February, 2019.

R. P. Ram Kumar, Professor, Department of Computer Science and Engineering, Malla Reddy Engineering College (Autonomous), Maisammaguda, Secunderabad, India.(Email: rprkvishnu@gmail.com)

Sanjeeva Polepaka, Associate Professor, Department of Computer Science and Engineering, Malla Reddy Engineering College (Autonomous), Maisammaguda, Secunderabad, India. (e-mail:psanjesus@gmail.com)

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

3. MEASURE THE MATTERS USING SENTIMENTAL ANALYSIS

According to Katie Delahaye Paine [1], metrics such as likes, shares, onsite engagements, comments and inbound links are very helpful in identifying the sentiment of the people or the user engaged with the content. By just having the metric counts only gives the false sense or bad branding about the particular product. Apart from the above-listed parameters, when analyzed, there is a possibility of Quality metrics, which deals with feelings, opinions, shares quality, repeated tweets made on the particular product, comments, rating, satisfaction scores, rating conversation about the quality of the product. Figure 1 shows such a typical example chart of sentimental analysis depicted from [1] and Table 1 shows some of the tools used to track and crack the user's sentiments.

	#s	Month 2	#s	Metrics	Chng
Likes	2,000	Likes	4,000	Like Growth	100%
Posts	100	Posts	125	Post Growth	25%
Comments	200	Comments	300	Comment Growth	50%
Comments-per-Post	2	Comments-per-Post	2.4	CPP Growth	20%
Comments-per-Like	0.1	Comments-per-Like	0.075	CPL Growth	-25%

Fig. A

Platform	Objective	Metric	Goal	Alternate Metric
Facebook	Customer Engagement	Av. #Comments/Post	10	Av. # Shares/Week
Twitter	General Awareness	Av. New Followers/Post	5	Av. #RTs/Post
LinkedIn	Thought Leadership	# Best Answers	20	# InRecommendations
Youtube	Sales/Lead Generation	# Leads or Sales/View	1%	Likes/Views
Google+	Customer Service	# Hangouts/Week	3	NetPromoter Score
Pinterest	General Awareness	# Likes/Pin	10%	# Repins
Slideshare	Sales/Lead Generation	# Leads or Sales/View	2%	# Downloads
iTunes	Thought Leadership	# Downloads/Month	500	Ratings
Quora	Thought Leadership	# Best Answers	10	Referring Traffic
Blog	General Awareness	# Unique Visitors/Month	1000	Comments/Post

Figure 1 Sentimental analysis chart (Courtesy [1])

Table 1 Practical tool for Sentimental Analysis

S. No	Name of the Tool	Details
1	Meltwater	Used to uncover insight into targeted audience
2	Google Alerts	Very simple method to track the "content marketing" for regular updates
3	People Browser	A useful tool to evaluate the competitors, industries and brands to explore the exact status regarding before, during and after marketing campaigns.
4	Google Analytics	A perfect tool to discover the prejudiced subscribers and buyers, such as annotations, custom reports, web designs, etc.
5	Hootsuite	Free and subscription-based options of this tools allow to measure and manage the social media networks data directly.
6	Tweetstat	Free graphical tools to explore Twitter statistics.
7	Facebook Insights	This tool is used to extract the overall likes, fans, along with the user activities, the sum of fresh likes and unlikes, tab views, page views, media consumption, referrers and lot more.
8	Pagelever	This tool can measure the activities in the Facebook, such as consumed content, shared content on Facebook.
9	Social Mention	Similar to Google Alerts, this tool is useful in tracking the keywords in bookmarks, blogs, events, question and answers, comments, audio and even in videos.
10	Marketing Grader	This tool is used to grade the whole marketing funnel with more than 30 metrics including, blog posts, tweets, Facebook updates, the number of visitors and much more.

Table 2 shows the pinnacle software for text mining, text analysis and analytics along with the proprietary solutions [2].

S. No.	Name of the Software	S. No.	Name of the Software
1	Abzooba	31	Loop Cognitive Computing Platform
2	Ai-one	32	Luminoso
3	AlchemyAPI	33	MeaningCloud
4	Angoss Text Analytics	34	Medallia
5	Ascribe Forest Rim's Texual ETL	35	Megaputer
6	Attensity	36	muText Mu Sigma
7	AUTINDEX	37	NetOwl
8	Averbis	38	OpenText
9	AYLIEN	39	Oracle Endeca
10	Basis Technology	40	Oracle Social Cloud - Collective Intellect
11	Bitext	41	Pingar
12	Brainspace Discovery	42	Provalis Research
13	Buzzlogix	43	Rapid Miner
14	Clarabridge	44	Rocket Text Analytics
15	Content Analyst	45	SAP Text Analytics
16	Datumbox	46	Saplo
17	DiscoverText	47	SAS Text Analytics
18	Etuma	48	Semantria
19	Expert System	49	SIFT
20	General Sentiment	50	Smartlogic
21	Google Cloud Prediction	51	StatSoft
22	HP Autonomy	52	Synapsify
23	IBM Text Analytics	53	Sysomos
24	Indico	54	SYSTRAN
25	Intellexer	55	Taste Analytics
26	Kanjoya	56	Text2data
27	Language Computer Corp[oration	57	Thomson Reuters Open Calais
28	Lexalytics Text Analytics	58	Twinword
29	LingPipe	59	Verint Systems
30	LinguaSys	60	VisualText

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

4. THE ARCHITECTURE OF SENTIMENTAL ANALYSIS PROCESS

Opinion plays a significant role in the technological world. However, in electronic commerce, the opinion on a product is vital. Further, customers are anticipated to acquire the magnificent goods based on the reviews from the customers who had already bought those products, without any interactions. The opinion mining process comprises of three phases, namely, (1) Opinion retrieval, (2) Classification of opinion, and (3) Views summarization. Figure 2 shows the architecture of opinion mining depicted from [3]. In the first step of opinion mining, the opinion retrieval phase deals with the collection of reviews from the websites/ blogs/databases. The second phase, classification of opinion is categorized as Supervised and unsupervised methods and is responsible for classifying the opinion as negative and positive classes. The third stage, summarization of opinion ought to highlight the review results of the opinion.

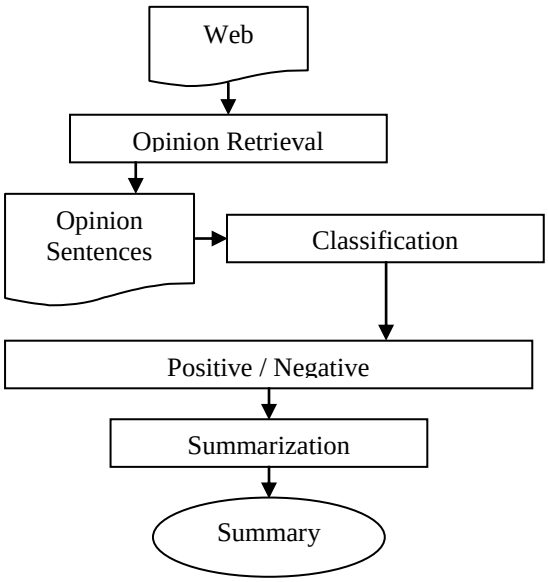

Figure 2. The architecture of Opinion Mining / Sentiment Analysis

5. CONCLUSION

There is no doubt that, Sentimental Analysis made us move from traditional market surveys and the research to online media monitoring surveys. Sentimental Analysis not only promoted business strategies for large investors but also for the number of small start-up companies. Because industries want to know and explore the customers perceive and their competitors. Henceforth, the technical challenges and practical needs of the opinion mining will be likely for years in all the domains.

REFERENCES

1. https://www.iprospect.com/en/ca/blog/10-sentiment-analysis-tools-track-social-marketing-success/
2. http://www.predictiveanalyticstoday.com/top-software-for-text-analysis-text-mining-text-analytics/
3. Parashar and Sharma, "A Literature Review on Architecture, Classification Technique and Challenges of Sentiment Analysis", International Journal of Engineering Research & Technology, Vol. 5, Issue 5, pp. 124-127, 2016.
4. Meenambigai, "An Efficient Surveillances of Products Based on Opinion Mining", International Journal of Innovative Research in Computer and Communication Engineering, Vol. 2, Issue 8, pp. 5261-5265, August 2014.
5. Vivekanandan and Helen Josephine, "ROM - Review Opinion Mining a Novelized Framework", International Journal of Computer Sciences and Engineering, Volume-2, Issue-11, pp. 86-89, November 2014.
6. Khandelwal, Mishra and V. K. Mishra, "A Survey on Subjective Sentiment Analysis from Twitter Corpus", IJSRSET, Volume 2, Issue 2, pp. 1198-1200, 2016.
7. DongSung Kim and Jong Woo Kim, "Public Opinion Mining on Social Media: A Case Study of Twitter Opinion on Nuclear Power", Advanced Science and Technology Letters, Vol. 51, pp. 224-228, 2014.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication