Business Intelligence with Social Media and Data Mining to Support Customer Satisfaction in Telecommunication Industry

Vitri Tundjungsari

Abstract**—** In today business, meeting customer satisfaction is a must. Customer satisfaction is related with various factors such as: hope, loyalty, and complaint. Previously, surveying customer opinion in experiencing the products or services is a main tool to collect and analyze the data. Nowadays, social media applications replace survey as promising media to support and enhance customer satisfaction. Social media provides easily an accessible platform for users to share information. As a result to this, the social media has been pervasively used all over the world 24 hours a day; thus, it has generated unprecedented amounts of social data. Mining social media has its potential to extract actionable patterns that can be beneficial for business, users, and consumers. This paper reviews the basics of data mining and social media, introduces several techniques which promising to be chosen to mining social media, and illustrates some applications of data mining, specifically in telecommunication industry, to support customer satisfaction and maintain customer relationship in telecommunication industry.

Keywords— customer, satisfaction, social media, data mining.

I. Introduction

WITHIN the last decade social media is widely used for various purposes, i.e.: entertainment, education, information and knowledge sharing. As a result of this, vast amounts of user-generated data exist and can be made available for data mining. Data mining of social media is used to discover new phenomena due to the use of social media and improve business intelligence to provide better services and in turn better customer satisfaction. There are a wide number of open source and commercial social media applications, e.g.: Facebook, Twitter, Google, YouTube. These applications are generally emerged as websites or web application components and serve as an interface to a cloud-based application that collects the data.

There are many potential benefits offered by social media applications, such as: information sharing provided by Facebook, media sharing (e.g.: video sharing by YouTube, sound sharing by Soundcloud, picture sharing by Flickr). Facebook and Twitter also provide reporting tools by supporting number of *like*, number of *seen*, or number of *Tweet feeds*; which are particularly useful for gaining an instant overview on a focused subset of activity, or for looking at recent changes either in real time or for a fixed time frame. However social media applications often do not support deeper insight for gaining deeper information into the behavior or needs concerns since these tools and services do not actually provide the data, nor do they provide any context for the summarized data being presented [17]. Mining the existing data within social media applications is promising to produce the required information regarding customer satisfaction purpose. However, social media data are vast, noisy, unstructured, and dynamic in nature, and thus mining the data is not easy to do. This paper discusses the potential of social media and data mining in exploring more information for customer satisfaction purposes. Section 2 of this paper introduces representative research problems of mining social media for customer satisfaction purposes and describes how social media can be very useful to increase customer satisfaction in telecommunication industry. Section 3 explains some data mining techniques that can be used for mining social media and illustrates using examples and figures. Finally, section 4 discusses the conclusion of the paper.

Vitri. Tundjungsari. Yarsi University, Faculty of Information Technology, Jl. Letjen Suprapto, Jakarta Pusat, Indonesia (corresponding author's phone/fax: 62214269301; e-mail: vibara11@gmail.com).

II. Mining SOcial Media

The usage of social media has created new culture and habit to society because social media provides easily an accessible platform for users to share information. This will in turn generate unprecedented amounts of social data. Having an abundant data is a potential way to explore and discover new information and knowledge to business by mining the existing data. Moreover, mining social media is very useful to extract actionable patterns that can be beneficial for business, users, and consumers.

Research in data mining has successfully produced numerous methods, tools, and algorithms for handling large amounts of data to solve real-world problems in many application domains, e.g.: data warehousing, bioinformatics, business intelligence, predictive analytics, and decision support systems. Data mining is a process of discovering useful or actionable knowledge in large-scale data [16]. Data mining is an integral part of many related fields including statistics, machine learning, pattern recognition, database systems, visualization, data warehouse, and information retrieval [6].

Mining social media is very potential and useful for extracting more information and gaining deeper insight about customer. For example, it can be used to identify the influential customer in a social networking site, detect implicit or hidden groups in a social networking site, perceive customers opinion related to their product or service's satisfaction for proactive planning, develop recommendation systems to maintain existing customers and gaining new ones, or build and strengthen trust among customers or between customers and other stakeholders. In short, mining social media is a promising multidisciplinary area, thus researchers of different backgrounds can make important contributions that matter for social media research and development.

A. Providing Customer Satisfaction with Social Media

Customer satisfaction is a customer's reaction toward his/ her needs fulfillment [14]. According to *American Customer Satisfaction lndex* (ACSI), customer satisfaction is made based on economics structural model that connected satisfaction of various factors, such as: hope, loyalty, complaint [7]. Figure 1 shows the relationship among customer value, customer satisfaction, customer loyalty, service quality, and switching barriers. Customer value is defined as the whole assessment of customer's perception toward product's usefulness based on what he/ she takes and gives [8].

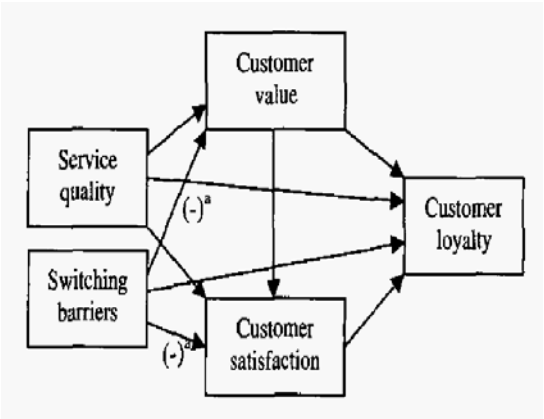

Fig 1. Customer Satisfaction Model [8]

The most common way to collect and measure customer's perception is surveys. However, nowadays social media applications provide useful approaches to measuring customer satisfaction without surveys, e.g.: to monitor what customers are saying about your business on websites like Twitter, Facebook, and Yelp. Yelp is a site that started in San Francisco but has migrated to other cities, and it gives customers a way to review businesses using 1-5 stars and provide detailed comments on their experience. The reviews often provide a wealth of useful details on what people liked and did not like. On the other hand, Twitter or Facebook might be a better bet but there is no overall numerical rating system like Yelp's 5 stars, so it is harder to get quantifiable data. It is undeniable that the words and comments in Twitter or Facebook are interesting, but it is harder to sort and categorize them into useful summaries. Those social media tools can provide a wealth of data that might be used to measure customer satisfaction without having to do surveys. Therefore, with today's internet technologies, unhappy customers and stakeholders can be much more visible than in the past, using social media tools. Thus, it has never been more important for any organization to focus on customer satisfaction strategy and techniques than in today's environment.

B. Data Mining for Customer Satisfaction Analysis in Telecommunication Industry

Telecommunications is one of the most data-intensive industries in the world, and a great opportunity exists for telecom managers to analyze the large amounts of data that have been collected in their network databases in order to improve the short-term and long-term operations of their organizations. Data mining is one highly effective tool to aid in this data analysis. Careful analysis using data mining can provide managers with the ability to optimize internal network operations and better manage external customer-facing activities such as churn and marketing. In terms of customer satisfaction, telecommunication data about customers such as call detail and customer information can be profitably data mined. Mining these data types can help you determine customer behavior and identify opportunities to support the goals of expansion of your customer base and reduction of customer churn.

Several data mining techniques have been used on traditional data customers, e.g.: regression analysis, Naïve Bayes (NB), Support Vector Machine (SVM), Neural Network (NN). Regression analysis is the most popular technique to predict customer satisfaction [5], and the best technique for data sample less than 1000 compare to any data mining techniques [13]. For example, Mihelis et al. (2001) use regression model to measure customer satisfaction; while [15] employ logistic regression to relate between customer satisfaction and customer retention's attributes. Author in [4] propose ITACS as a customer satisfaction application to analyze client's data from e-commerce and telecommunication industry. ITACS is used as a customer satisfaction solution that combines text classification and business intelligence, using NB and SVM. The research shows that SVM more accurate than NB.

Other techniques use Recency Frequency Monetary (RFM) approach, such as: [9] and [10]. (Li et al. [9] employ Timely RFM (TRFM) as a method to measure customer loyalty. They develop TRFM using Adaptive Resonance Theory (ART) of Artificial Neural Network (ANN) to cluster customers based on their buying behavior by combining property of the product and buying period to calculate buying loyalty level and customers buying behavior. Moreover, they can analyze what kinds of products are needed by customers on different period of time. Liu and Shih [10] develop a product recommender model to increase customer satisfaction, using weighted RFM (WRFM) and K-nearest neighbors with preference-based collaborative filtering (KNN based CF

method) derived from Customer Life Value (CLV) and purchase preferences. This method results in top N recommendation. Other data mining techniques that can be used as product recommenders i.e.: association rule mining [1].

III. Proposed Techniques

Based on the discussion in section 1 and 2 above, this paper discusses several potential techniques to extract and mine data from social media application, i.e.: text mining, clustering, visualization. The explanations of each technique are presented as below:

- Text Mining: Text mining is an emerging technology that can be used to augment existing data in corporate databases by making unstructured text data available for analysis. One of the difficulties in getting started with text mining is acquiring the tools, i.e., the software for implementing text mining. Much of the software is expensive and/or difficult to use. For instance, some of the software requires purchase of expensive data mining suites. Other commercial software is more suited to large scale industrial strength applications such as cataloging and searching academic papers. One innovation that has made budget-friendly software available to text miners is open source software [3].
- Clustering: Cluster analysis divides data into meaningful or useful groups (clusters) [16]. If meaningful clusters are the goal, then the resulting clusters should capture the "natural" structure of the data. For example, cluster analysis has been used to group related documents for browsing, to find genes and proteins that have similar functionality, and to provide a grouping of spatial locations prone to earthquakes. However, in other cases, cluster analysis is only a useful starting point for other purposes, e.g., data compression or efficiently finding the nearest neighbors of points. Whether for understanding or utility, cluster analysis has long been used in a wide variety of fields: psychology and other social sciences, biology, statistics, pattern recognition, information retrieval, machine learning, and data mining [18].
- Visualization: Visualization seeks to harness the remarkable capabilities of the human visual system to aid cognition, through the use of computer-generated representations. Visualization techniques find applications in a wide range of areas, from the elucidation of structure in complex datasets to the creation of virtual environments to aid training in dangerous or challenging procedures. In science, visualization can play an important role in exploratory data analysis, where visual representations can help the scientist to build up an understanding of the content of their datasets, while in the specific arena of e-science, much effort is being devoted to the development of collaborative visualization techniques, which aid the interaction between scientists at different locations. Some types of representation used in visualization are computationally expensive to generate, so that one challenge for the visualization community is to respond to the rapidly expanding data volumes seen in many application areas. Another obvious barrier to the ready use of visualization tools in science is the wide variety of data formats used in different scientific disciplines, which often necessitates the translation of data before it can be visualized [11].

Fig 2. Sample of Text Mining using Visualization to identify Customer's Characters [19]

IV. Conclusion

The challenge comes in accessing that data from social media and transforming it into something that is usable and actionable. Generally, organizations want to use the social media data to understand the needs and behavior of their customers or specific targeted groups of individuals with respect to the organizations' current or future products or services. There are several data mining techniques that can be looked at to extract social media's data, e.g.: text mining, clustering, and visualization. Each has its useful aspects for different purposes to support good relationship and maintain customer relationship in telecommunication industry.

References

[1] C. Aggarwal, C. Procopiuc, L. J. Wolf, S. P. Yu, S. J. Park (1999), "Fast Algorithms for Projected Clustering," *In Proc. SIGMOD '99, Philadelphia PA*, 1999, ACM Digital Library.

[2] K. Ehrlich, I. Carboni, *Inside Social Network Analysis*. (2009). [online] Available: http://ppr.cs.dal.ca/sraza/files/social%20pdf.

[3] L. Francis, M. Flynn. (2010), *Text Mining Handbook-Casualty Actuarial Society*. [online] Available:

www.casact.org/pubs/forum/10spforum.
[4] S. Godbole, S. Roy, "Text Classification, Business Intelligence, and Interactivity: Automating C-Sat Analysis for Services Industry," *In Proc. KDD'08*, Las Vegas, Nevada, USA, August 24–27, 2008, ACM Digital Library.
[5] J. Hadden, A. Tiwari, R. Roy, D. Ruta, "Computer Assisted Customer Churn Management: State-Of-The-Art and Future Trends," *Computers & Operations Research*, October 2007, vol. 34, no. 10, pp. 2902-2917.
[6] J. Han, M. Kamber, J. Pei, *Data Mining: Concepts and Techniques*, Morgan Kaufmann: San Francisco, 2011.
[7] C. Fornell, "A National Customer Satisfaction Barometer: The Swedish Experience", *Journal of Marketing*, 1992, vol. 55, no. 1, p. 121.
[8] Z. Huan, X. Xu, H. Li, "The Impact of Service Quality, Satisfaction, Value and Switching Barrier on Customer Loyalty in Chinese Mobile Telecommunication Industry," *IEEE Digital Library*, 2005.
[9] L. Li, F. Lee, W. Liu, "Timely Product Recommendation Based on RFM Method," *In Proc. The International Conference on Business and Information (***BAI 2006***)*, Singapore, 12-14 July 2006.
[10] D. Liu, Y. Shih. Y, "Hybrid approaches to product recommendation based on customer lifetime value and purchase preferences," *The Journal of Systems and Software*, 2005, vol. 77, pp. 181–191.
[11] T. M. Mann. (2002), *Visualization of Search Result from the WWW*. [online] Available: citeseerx.ist.psu.edu
[12] G. Mihelis, E. Grigoroudis, Y. Siskos, Y. Politis, Y. Malandrakis, "Customer Satisfaction Measurement in the Private Bank Sector", *European Journal of operational Research*, 2001, vol. 130, pp. 347-360.
[13] I. Ograjensek, "Use of Customer Data Analysis in Continuous Quality Improvement of Service Processes", *In Proceedings of the Seventh Young Statisticians Meeting*, Metodološki zvezki, 2003, vol. 21, Ljubljana: FDV.
[14] S. Oliver, "A model for the future of electronic commerce", *Information Management & Computer Security*, 1997, Vol. 5, No.5, pp.166-9.
[15] R.T. Rust, A. J. Zahorik, "Customer Satisfaction, Customer Retention, and Market Share", *Journal of retailing*, 1993, vol. 69, no. 2, pp. 193-215.
[16] P. N. Tan, M. Steinbach, V. Kumar, *Introduction to Data Mining*. Pearson AddisonWesley: Boston, 2006.
[17] K. Thiel, T. Kotter, M. Berthold, R. Silipo, P. Winters, "Creating Usable Customer Intelligence from Social Media Data: Network Analytics meets Text Mining", *KNIME.com Report*, 2012.
[18] http://cs.umn.edu
[19] http://www.smashingmagazine.com/2009/09/11/25-useful-data-visualization-and-infographics-resources/

Vitri Tundjungsari was bornt in Yogyakarta, Indonesia. Member of IEEE in 2009, PhD degree in Computer Science from Gadjah Mada University Yogyakarta, Indonesia; Master degree in Information Technology from University of East London, London, UK; Bachelor degree in Civil Engineering from Bandung Institute of Technology, Bandung, Indonesia. Having research interest in Decision Support Systems, Social Media, e-Participation, e-Government, and Public Decision Making.