Social Media as a Main Source of Customer Feedback – Alternative to Customer Satisfaction Surveys

Sharon Grubner Hasson
Computer Science
Johns Hopkins University
Laurel, MD, USA
sgrubne1@jhu.edu

John Piorkowski
Data Science
Johns Hopkins University
Laurel, MD, USA
jpiorko2@jhu.edu

Ian McCulloh
Accenture
Washington, DC, USA
ian.mcculloh@accenturefederal.com

***Abstract*—Customer satisfaction surveys, which have been the most common way of gauging customer feedback, involve high costs, require customer active participation, and typically involve low response rates. The tremendous growth of social media platforms such as Twitter provides businesses an opportunity to continuously gather and analyze customer feedback, with the goal of identifying and rectifying issues. This paper examines the alternative of replacing traditional customer satisfaction surveys with social media data. To evaluate this approach the following steps were taken, using customer feedback data extracted from Twitter: 1) Applying sentiment to each Tweet to compare the overall sentiment across different products and/or services. 2) Constructing a hashtag co-occurrence network to further optimize the customer feedback query process from Twitter. 3) Comparing customer feedback from survey responses with social media feedback, while considering content and added value. We find that social media provides advantages over traditional surveys.**

Keywords — Twitter, social media, survey, customers, feedback, machine learning, classifier

I. Introduction

Customer satisfaction surveys have been the traditional and most prevalent way of gaining customer feedback for years. Typically, these surveys are conducted through a third-party vendor that gathers the data and delivers analyzed results back to the surveyed company. These services involve added costs, require customer active participation and risk not reflecting specific points for improvements. Since surveys are generally conducted on an annual basis, customers are less likely to raise issues that occurred several months prior.

Twitter is a vast source of live information; customers from both academia and industry use this platform to share news, feedback and experiences with their followers via Tweets. While a significant portion of these Tweets may contain neutral information, there is an assumption that customers are more likely to turn to that readily-available platform when they are either highly satisfied or highly frustrated.

That alternative of gaining customer feedback through social media was evaluated using data gathered about a large global company based in the US, Illumina. Illumina is a biotechnology company that specializes in sequencing, genotyping, gene expression, and bioinformatics data analysis services. Illumina customer survey response rates normally stand around 5%, which does not sufficiently cover the breadth of customers' opinions and leads the company to seeking for alternatives. Similarly to other companies, Illumina has been growing its social media presence and as of May 2019 has approximately 65K Twitter followers.

This study aims to answer the question: does Twitter provide equivalent findings as surveys, at a lower cost and higher volume?

In order to answer this question, this study focuses on the following research goals:

- Analyze Tweets focusing on specific Illumina products and/or services and compare sentiment throughout.
- Discover optimal hashtags to track for any future social media customer feedback monitoring by building a hashtag co-occurrence network across the different offered services/products.
- Compare customer feedback from survey responses with social media feedback in terms of content and value.

These goals are studied using data extracted from Twitter as described in the Method section.

II. Background

Twitter has been one of the most active social media platforms in the world, enabling its users to share whatever they have on their minds, as long as it is 280-characters long. As of February 2019, Twitter has 126 million daily active users [1]. Being a very popular social media platform, Twitter has been the focus of various academic studies trying to take advantage of the vast amount of data it encompasses.

ASONAM '19, August 27–30, 2019, Vancouver, BC, Canada
© 2019 Association for Computing Machinery.
ACM ISBN 978-1-4503-6868-1/19/08…$15.00
https://doi.org/10.1145/3341161.3345642

Bajic and Lyons [2] studied the factors that influence overall use of social media for gathering user feedback for software development companies. While the study does not directly examine the use case of replacing customer surveys with social media data, it does inspect utilizing social media to gain dynamic feedback from customers. They examined factors such as company size, user community size, user types, project transparency, and number of social media channels, to determine which are the most prominent factors. Based on their analysis it seems that while the significance of these factors fluctuates based on the circumstances, social media has proven itself as a powerful and dynamic source of customer feedback.

Sumit, et al. [3] explores a way of monitoring and analyzing customer feedback through social media platforms to resolve customer problems in real time. As part of their pipeline they repeatedly assign polarity to each given feedback by applying a proprietary sentiment classifier, which its details are not disclosed. While that study has made significant steps towards building their own set of tools, they do not detail their methodology that allows reproducing it for other purposes. Moreover, while the focus of their work was to get a live customer remediation solution, one of this paper's main goals is to enable finding general trends, topics, and in the future regions, that would need to be improved. Finally, this study offers added value by comparing data collected from social media to a sample gathered from a customer survey and examines if the latter is a viable alternative.

III. Method

Twitter data was extracted using Tweepy, a Python library that enables accessing data through Twitter API [4]. The used query's search terms included '@illumina', which is associated with Illumina's verified account, and '#illumina', which is often used by other Twitter users in context of the company. Data was extracted into two main batches. The first batch, which is smaller and extracted on April 24, 2019, was used for the following plotting steps in order to efficiently visualize the results. This batch included Twitter data from April 15 – 24, 2019. A second batch is composed of the original batch data in addition to data extracted on May 5, 2019, in an identical fashion. The data extracted on May 5, 2019 includes data from April 26 – May 5, 2019. Extracted data was then parsed and written into comma delimited files containing the following columns: Author, Location, Number of Followers, Number of Favorites, Retweet Count, Content, Hashtags, and Posted At.

The collected data was then processed and prepared for sentiment analysis. To achieve that, each row of the Content column, containing the Tweet's text, was cleaned from URLs and stop words. English stop words corpus was downloaded using Natural Language Toolkit (NLTK) Python library [5]. NLTK is a suite of tools and programs for statistical and symbolic natural language processing (NLP), that also offers over 50 corpora and lexical resources.

The same process of cleaning the data from URLs and stop words was repeated on a Sentiment140 corpus [6], which was used for training a sentiment classifier. The Sentiment140 dataset contains 497 rows, with the following columns: polarity of the Tweet (0=negative, 2=neutral, 4=positive), ID of the Tweet, date of the Tweet, the query, the user that tweeted, and the text of the Tweet. That resource was chosen since both its data and this study's data were extracted from Twitter social media platform and is therefore appropriate for a machine learning application. That corpus was used to build a classifier that identifies the polarity of Tweets, whether they are positive, neutral or negative. The columns used for the classifier training included the Tweet text and the determined polarity. That classifier was used to classify extracted Tweets' sentiment by calling the TextBlob library [7]. That library specializes in processing textual data by providing relatively simple APIs for common NLP tasks. Both preprocessed text and classifier are taken as TextBlob inputs to perform sentiment analysis. By default, without providing a classifier, TextBlob's sentiment analysis method is based on training data that is manually curated from commonly found words in product reviews. A list of products and services was assembled in order to construct a dictionary composed of products as keys, and total of sentiment categories associated with each Tweet. These dictionary values were then counted and processed before presenting them as a bar plot which shows the count of both positive and negative sentiment Tweets for each product and/or service of interest.

For the next phase of the study, extracted data was constructed in another dictionary, composed of services or products as keys, and a list of associated hashtags as values. The purpose of this phase is to discover optimal hashtags by building a hashtag co-occurrence network. In order to optimally visualize the network from the created dictionary in a manner that allows gaining insights, the plotted network was limited to the top 6 products and/or services of interest: NovaSeq, HiSeq, iSeq, FAS, TruSeq, and DRAGEN. NovaSeq, HiSeq and iSeq are Illumina sequencers, FAS (Field Application Scientist) is a direct customer facing role, TruSeq is a commonly used Illumina product line, and DRAGEN is the new comprehensive analysis solution offered by Illumina. The drawn network allows to intuitively visualize all the associated hashtags used for each product and/or service and observe which hastags overlap between the different categories. The main Python libraries utilized for both of these plots include pyplot by matplotlib [8] and networkX [9].

For the final phase of analysis, a sample of survey data in Excel was retrieved from Illumina Customer Experience team and processed as a more easily readable format. Out of the detailed Excel workbook, which contains customer information and therefore could not be publicly shared, one column was extracted for the purpose of this study, with customers answering the question: "What do you see as the most important thing to improve overall satisfaction?".

IV. Results and Analysis

The first generated plot [Figure 1] shows the associated sentiment for each service and/or product as a total count of positive and negative sentiment Tweets. The topic distribution of the analyzed Tweets was somewhat unexpected. Relatively general topics such as 'sequencers', 'genomics', 'NGS' and 'cancer' were expected to have a higher count, both of positive and negative nature, since they could be discussed in context that is indirectly related to Illumina's performance as a company. However, UMI (Unique Molecular Identifier) was discussed substantially more than the other tracked topics, both in positive and negative context, while it has not seemed to be the focus of Illumina's work behind the scenes. Alternatively, topics that represent groundbreaking work such as 'TSO500' (TruSight Oncology 500), are not evident to be popular in the extracted Tweet sample. These results prove

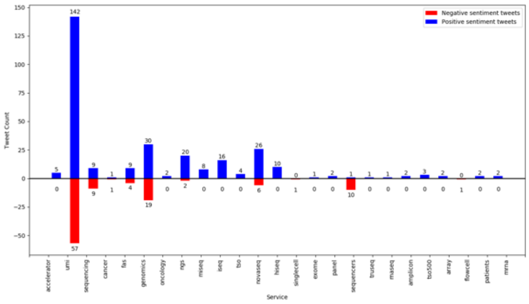

Figure 1: Sentiment category count associated with each service and/or product line

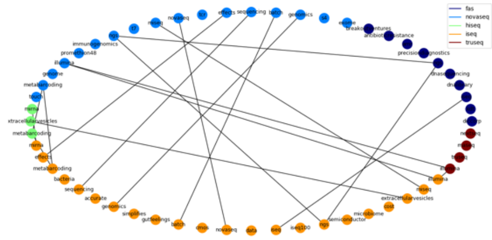

Figure 2: Hashtag Co-Occurrence Network

that the focus of discussion within the company does not always reflect the customers' main interests.

The second generated plot [Figure 2] shows the co-occurrence hashtag network of the extracted Tweets. That network provides a tool to further understand how customers are posting about Illumina and help design a more efficient query for any future efforts. Based on the constructed network it is evident that there is an overlap between different hashtags used to discuss Illumina's products and/or services. From the relatively small sample, the following hashtags were found to overlap: 'NovaSeq', 'MiSeq', 'MiRNA', 'metabarcoding', 'sequencing', 'batch', 'NGS', and 'Illumina'. Based on these findings, some hashtags that have not been considered before for customer feedback queries, were found to overlap across different topics. These hashtags include 'metabarcoding' and 'miRNA'. The same methodology could be applied for larger datasets in order to discover a more comprehensive list of hashtags that could be used to further improve social media feedback queries.

Extracted social media data includes types of customer feedback that would be valuable to a company such as Illumina and are not necessarily mentioned in official annual customer satisfaction surveys, since their associated events may have occurred a relatively long period of time beforehand. The following example shows a Tweet originating in London referring to a specific event with clear guidance how customer experience could be improved: "PS60k of #novaseq reagents stowed away in the nick of time. bit more dry ice next time please @illumina". Tweets could also include positive feedback from social media. For example, the following Tweet posted in Texas expresses overall high customer satisfaction, which is strengthened by a recent positive experience: "@illumina everything arrived in great shape and the next two days @illumina outstanding fas crew will do their thing and get all our systems up and running.". Social media could also include Tweets that express feedback and thoughts about a company's strategic move, such as an acquisition. The following Tweet from San Diego, California, expresses thoughts about an acquisition that is currently in progress and what they may think the next move should be: "@gbw42 @illumina @pacbio interesting idea. 10x is currently valued at $1.28b. with the premium they'd have to offer, it may be too expensive. i keep wondering if they'll buy someone in the clinical space.".

Customer satisfaction survey data includes some topics that were not stressed in the social media feedback sample data. For example, additional workshops, technical trainings and more environmentally friendly packaging seem to be the common themes raised by various customers. Having said that, a substantial portion of responses mentioned that lower costs would increase their overall satisfaction with the company, which could be said about almost any service provider. It is important to mention that a portion of the survey responses include general remarks that are difficult to translate into action items without additional information. For instance: "customer service", "customer support and reagents", "dealing with device failures", and "make the system more user friendly" are difficult to drive change.

V. Conclusion

Social media has the potential of having similar findings as surveys, when collecting data overtime, while offering additional, more dynamic information. Social media feedback can reflect customer reactions to Illumina's related news and product releases, which is not information generally found in surveys. Moreover, social media platforms like Twitter have a higher exposure to customers, including lab technicians and scientists, who may not always get targeted by survey requests.

Gathering feedback from social media allows getting more dynamic reactions to specific topics of interest. For instance, Bio-Rad's acquisition process has been on the news and social media for the past few months, yet Illumina would not be able to gauge any feedback before the next annual survey is released. Aiming for more frequent surveys is also not a viable alternative. Trying to conduct more frequent surveys might result in even lower response rates, as customers may find those tedious, while gathering feedback from social media does not require them to take any action.

Finally, since Illumina operates through Global Channel Partners (GCP) in some parts of the world, there is a significant number of customers who do not receive these survey requests. Gathering feedback from social media will enable to track these customers as well.

In both cases of customer feedback data, both from social media and from traditional surveys, there are issues of noise and usability of information, at different proportions. Since customer satisfaction surveys involve much higher costs and require customers to specifically dedicate their time, responses were expected to have higher focus and relevance. The most significant difference between data collected from social media and customer surveys is that the former could be performed continuously and potentially at a lower cost on an automated routinely manner, while the latter offers more concentrated feedback collected on an annual basis. The data collection for this study was completed on April 24 and May 5, 2019, and therefore represents the Tweets available through the Twitter API on these days. In order to derive additional valuable feedback, the same process described in the Methodology section would be automated and repeated on a weekly basis.

Finally, examining social media provided new insights on customer satisfaction that did not appear in traditional customer surveys.

VI. Future Work

With the additional metadata available with each Tweet, more thorough and sophisticated analysis could be done by including fields such as: author, location, number of followers, number of retweets, etc. Geographic location could help customer experience analysis to dynamically alert which regions should get special attention by customer support. A deeper analysis of authors could help distinguish normal customer from frequent complainers. Moreover, the analysis could be extended to multiple languages. While this study focused on parsing posts written in English, Twitter is a global platform with an enormous potential of growth.

References

[1] H. Shaban (2019), “Twitter Reveals Its Daily Active User Numbers For The First Time”, The Washington Post, https://www.washingtonpost.com/technology/2019/02/07/twitter-reveals-its-daily-active-user-numbers-first-time/

[2] F. Bajic and K. Lyons (2011), “Leveraging Social Media to Gather User Feedback for Software Development”, Proceedings of the 2Nd International Workshop on Web 2.0 for Software Engineering, DOI: http://doi.acm.org/10.1145/1984701.1984702

[3] B. Sumit, L. Jingxuan, S. Tong (2013), “Monitoring and Analyzing Customer Feedback Through Social Media Platforms for Identifying and Remedying Customer Problems”, IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, pp 1147-1154, DOI: 10.1145/2492517.2500287

[4] J. Roesslein (2009), “Tweepy. An Easy-To-Use Python Library for Accessing Twitter API”, http://www.tweepy.org/

[5] E. Loper, S. Bird (2002), “NLTK: The Nautral Language Toolkit”, https://arxiv.org/abs/cs/0205028

[6] A. Go, R. Bhayani, L. Huang (2009), “Twitter Sentiment Classification using Distant Supervision”, Stanford University, https://cs.stanford.edu/people/alecmgo/papers/TwitterDistantSupervision09.pdf

[7] S. Loria, P. Keen, M. Honnibal, R. Yankovsky, D. Karesh (2014), “TextBlob: simplified text processing”. https://textblob. readthedocs.org/en/dev/index.html.

[8] J. F. Hunter (2007), “Matplotlib: A 2D Graphics Environment”, Computing in Science & Engineering 9, 90. https://doi.org/10.1109/MCSE.2007.55

[9] A. Hagberg, P. Swart, D.S. Chult (2008), “Exploring Network Structure, Dynamics, and Function Using Networkx”, U.S Department of Energy, Office of Scientific and Technical Information. https://www.osti.gov/biblio/960616