Singapore Management University
Institutional Knowledge at Singapore Management University

Research Collection School Of Information Systems | School of Information Systems

11-2016

Spiteful, one-off, and kind: Predicting customer feedback behavior on Twitter

Agus SULISTYA
PT Telekomunikasi Indonesia

Abhishek SHARMA
PT Telekomunikasi Indonesia

David LO
Singapore Management University, davidlo@smu.edu.sg

DOI: https://doi.org/10.1007/978-3-319-47874-6_26

Follow this and additional works at: https://ink.library.smu.edu.sg/sis_research

Part of the Databases and Information Systems Commons, and the Social Media Commons

Citation

SULISTYA, Agus; SHARMA, Abhishek; and David LO. Spiteful, one-off, and kind: Predicting customer feedback behavior on Twitter. (2016). *Proceedings of 8th International Conference on Social Informatics (SocInfo 2016): Seattle, USA,*. 10047, 368-381. Research Collection School Of Information Systems.
Available at: https://ink.library.smu.edu.sg/sis_research/3614

This Conference Proceeding Article is brought to you for free and open access by the School of Information Systems at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Research Collection School Of Information Systems by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email libIR@smu.edu.sg.

Spiteful, One-Off, and Kind: Predicting Customer Feedback Behavior on Twitter

Agus Sulistya[1,2], Abhishek Sharma[2], and David Lo[2]

[1] Human Capital Center, PT Telekomunikasi Indonesia, Indonesia
[2] School of Information Systems, Singapore Management University, Singapore
{aguss.2014, abhisheksh.2014, davidlo}@smu.edu.sg

Abstract. Social media provides a convenient way for customers to express their feedback to companies. Identifying different types of customers based on their feedback behavior can help companies to maintain their customers. In this paper, we use a machine learning approach to predict a customer's feedback behavior based on her first feedback tweet. First, we identify a few categories of customers based on their feedback frequency and the sentiment of the feedback. We identify three main categories: spiteful, one-off, and kind. Next, we build a model to predict the category of a customer given her first feedback. We use profile and content features extracted from Twitter. We experiment with different algorithms to create a prediction model. Our study shows that the model is able to predict different types of customers and perform better than a baseline approach in terms of precision, recall, and F-measure.

Keywords: Social Media, Customer Relationship Management, Machine Learning

1 Introduction

The use of social media in the customer relationship context has gained popularity nowadays. A report by VB Insight [1] reveals that modern consumers complain about brands 879 million times a year on Facebook, Twitter, and other social media portals. About 10 percent of those consumers make a complaint on social media every day. With this extensive use of social media by customers, opportunities arise for companies to engage with their customers and be aware of the issues that they face. For example, a customer can complain on social media after experiencing a failure of service; this complaint notifies the company and prompts it to take necessary actions to prevent further damage to the company's reputation and customer base. Therefore, it is important for the company to continuously monitor the voices of their customers, which refer to an activity called *social listening and monitoring.*

It would be interesting to be able to predict different types of customer feedback behavior. Such prediction can help a company to formulate a suitable strategy to manage and improve customer satisfaction and retention. For example, some users may complain many times to a company's Twitter account if the users are not given sufficient attention in a short period of time, others may complain only once, and yet others may express their thanks after a good service has been rendered by a company. From the company's side, multiple complaints are considered as something that should be avoided, since it can affect the company's reputation. Having an ability to predict this type of customer would allow the company to take preventive action before the user spreads negative opinion about the company in social media. A company may also want to provide good reasons for the third category of customers to publicize good service and improve the company's reputation.

In this study, we try to address the aforementioned prediction problem by employing a two-stage machine learning algorithms. In the first stage, our approach clusters social media users into several categories based on their feedback frequency and sentiment polarity. We identify three categories of users: spiteful (i.e., the user complains many times in social media), one-off (i.e., the user only provides negative feedback once), and kind (i.e., the user provides positive feedback). In the second stage, our approach builds a prediction model that can assign a user into one of the three categories based on his/her first feedback. We experiment with different supervised machine learning algorithms (i.e., Naive Bayes, Logistic Regression, and Random Forest), to build an automated prediction model.

As a case study, we use an internal data from a state-owned telecommunication company in Indonesia to evaluate the effectiveness of our proposed approach. The company named Telkom extensively uses social media such as Facebook and Twitter, to interact with its customers. To facilitate social listening, the company has set up a dedicated unit to actively monitor customer feedback. Our work extends the current social listening platform that the company has by adding some predictive capabilities. Under 10-fold cross validation, our experiments show that our proposed approach can predict customer feedback behavior categories with a weighted precision, recall, and F-measure of up to 0.797 (Random Forest), 0.881 (Naive Bayes), and 0.800 (Random Forest) respectively. Our approach outperforms a baseline that randomly assign categories to customers based on the distribution of customer feedback behavior categories in a training data.

Extracting knowledge from microblogs has been one of active research areas. We believe that this study would be important towards the development of techniques that make use of social media data to improve product and service quality. Specifically, our contributions are as follows:

1. We propose a new problem of predicting different types of customer feedback behavior on Twitter.
2. We use a clustering algorithm to identify different types of customer feedback behavior.
3. We propose a set of features, i.e. content features and profile features, that can be used to predict customer feedback behavior by leveraging a supervised machine learning algorithm to build a prediction model.

4. We have evaluated our proposed approaches on a dataset containing 11,809 tweets. Our proposed approaches can achieve reasonable precision, recall and F-measure which are higher than those of a baseline approach.

The structure of the remainder of this paper is as follows. In Section 2, we describe social listening activities in a company used as our case study and data analysis techniques that we leverage for this work. We describe how we cluster customers to create several categories in Section 3. In Section 4, we explain our approach which extract features from customer Twitter accounts and their corresponding tweets and use them to build a prediction model to predict customer categories based on their first feedback tweet. We describe our experiments which evaluate the prediction accuracy of our approach in Section 5. Related work is presented in Section 6. We finally conclude and mention future work in Section 7.

2 Preliminaries

2.1 Social Listening at Telkom

In this paper, we experiment with a dataset collected and annotated by a state-owned telecommunication service provider in Indonesia, namely Telkom[3]. The company serves tens of millions of customers throughout Indonesia, offering a wide range of products including broadband internet connections, cable TV, and land line telephone connections.

Telkom has set up a system that actively monitors what customers say on social media, and handles each issues raised by forwarding the problem to a back-room unit. To monitor customer voices, the company uses tools provided by Brand24[4] and BrandFibres[5]. The first tool is used to crawl any contents containing keywords related to the company's product from different platforms, including Facebook, Twitter, blog posts, and news media. These crawled records are then filtered by removing irrelevant posts. The filtering process requires manual work performed by several social media analysts. The analysts use a second tool called BrandFibres dashboard. Using this tool, they evaluate each post, and then assign a sentiment score to each post. They give scores ranging from "+5" (very positive feedback) to "-5" (very negative feedback). The analysts also assign a post into one of the 8 different categories shown in Table 1. Note that a tweet can be assigned to more than one category, and an analyst will assign a sentiment score for every category that applies to a tweet.

Figure 1 shows an example of a customer complaint on Twitter. In the figure, the tweet mention a company's account (@telkomcare). The tweet also mentions other users (@detikcom and @telkompromo). The first one is an online news media account and the latter is the company's other account that focuses on disseminating the company's promotional events and deals.

[3] `http://www.telkom.co.id/en/tentang-telkom`
[4] `http://www.brand24.com/`
[5] `http://www.brandfibres.org/`

Fig. 1. A tweet posted by a customer to a company's customer care channel

This study analyzes data consisting of tweets collected and annotated by Telkom for a 3 month period from June-August 2015. In total, there are 12,634 posts. We consider only the posts that have been collected from Twitter, which results in about 11,809 posts (or tweets) constituting about 93.4% of the total posts.These tweets are those that mention the official company's customer care account on Twitter, namely @telkomcare. For the tweets in our dataset, we extract distinct twitter users who posted them, resulting in 6,031 distinct users. We use this set of users as the input to our clustering and prediction tasks described in the next two sections. The data provided by Telkom did not include the profile of these 6,031 Twitter users. To get these profiles, we call the standard Twitter API using Tweepy[6] Python module.

2.2 Handling Imbalanced data

Imbalanced data problem typically refers to a classification problem where the classes are not represented equally. For customer feedback, typically we would see more negative feedback rather than positive ones. One way to deal with imbalanced data is by using sampling methods, which modifies the distribution of the original training samples to obtain a relatively balanced data. There are two types of sampling methods: oversampling and undersampling [11]. Oversampling is conducted by adding more samples to the minority class, while undersampling is done by creating a subset of the majority class. One popular oversampling algorithm to handle imbalanced data is SMOTE (Synthetic Minority Over-sampling Technique) [6]. This oversampling algorithm creates synthetic samples from the minority class instead of creating copies. SMOTE works by finding the k nearest neighbors of each sample in the minority class. Next, artificial samples are then generated along the line of some or all of the k nearest neighbors, depending on the amount of oversampling required.

3 Clustering Customers

In the first stage of our work, we cluster customers in our dataset (i.e., the 6,031 users described in Section 2.1) into several categories based on their feedback frequency and the sentiment polarity of these feedback. Figure 2 shows our overall approach to cluster customers.

[6] `http://www.tweepy.org/`

Table 1. Customer Feedback Categories

Category	**Description**
Quality Evaluation	Tweets related to general quality of a product or service (for example, slow or unstable internet connection).
Offer Evaluation	Tweets providing feedback to a product offering (such as an ongoing promotion of a certain product)
Activity Disturbance	Tweets reporting specific disturbance in a user's activity while using a product (for example, trouble when browsing or downloading).
Invoice Related	Tweets reporting issues related to product or service invoicing (such as reports of incorrect billing).
Customer Service Quality	Tweets reporting issues related to quality of customer service (such as quality of customer service agents, and how the company handles current problems experienced by the customer)
Actions	Tweets related to action taken by the customer (such as comparing product provided by the company with other competitors).
Social Media	Tweets related to social media interactions between company and customers.
Others	Tweets about other issues related to the company and subsidiaries.

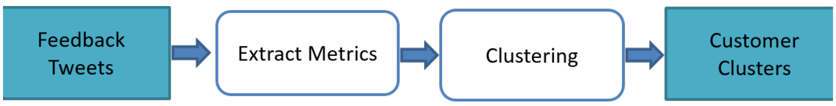

Fig. 2. Our approach for clustering customers

We represent each customer as a set of metrics: NumOfFeedback, NumOfPosFeedback and NumOfNegFeedback. These metrics are listed and defined in Table 2. Next, based on this representation, we cluster the users together. To cluster the users, we use Expectation-Maximization (E-M) algorithm. E-M algorithm assigns a probability distribution to each instance which indicates the probability of it belonging to each of the clusters. A previous study conducted by Meilă and Heckerman [13] has found that the E-M algorithm often performs better than other clustering methods such as k-means and model-based hierarchical agglomerative clustering.

We use the implementations of E-M Algorithm in Weka [10]. We do not initiate number of cluster and let the E-M algorithm decides the best number of clusters. All parameters are set into Weka default setting.

Table 2. Metrics used for clustering users

Features	Description
NumOfFeedback	Number of feedback tweets generated by a user
NumOfPosFeedback	Number of feedback tweets that are of positive sentiment polarity.
NumOfNegFeedback	Number of feedback tweets that are of negative sentiment polarity.

Table 3 shows the results of the E-M clustering algorithm. We verify the result by manually investigating the properties of each cluster. Based on this manual investigation, we conclude general properties for each group as shown in the fourth column of the table.

Table 3. Clusters of users based on their tweets mentioning the company

Cluster	Count	Percentage	Observed Properties
0	1235	20.48 %	post one or two times, with at least one positive feedback
1	152	2.52 %	post more than 2 tweets, with more than two possitive feedback
2	481	7.98 %	post at least 4 tweets, with majority of negative feedback
3	82	1.36 %	post at least 9 tweets, with majority of negative feedback
4	2837	47.04 %	post only one tweet with negative feedback
5	1244	20.63 %	post 2 or 3 times with majority of negative feedback

Note that there are similarities among these clusters. Cluster 4 represents the majority of customers who only provide one negative feedback, without posting further tweets. Clusters 2, 3 and 5 correspond to customers who post more than one tweet with negative sentiment. These customers are typically the group of customers that may damage a company's reputation if they are not managed well. The other two groups (clusters 0 and 1) are groups of customers that post at least one positive feedback such as thanking the company for its good service. These customers can improve the company's reputation. Based on this observation, we decide to group the clusters further into three groups based on how the customers complain or behave. This new groups are shown in Table 4. We will use these three groups as class labels for the second stage of our approach that predicts customer feedback behavior.

Table 4. Three main categories of customers

Class	Cluster	Percentage
Kind	0,1	23.0%
One-Off	4	47.0%
Spiteful	2,3,5	30.0%

4 Predicting Customer Categories

In the second stage, our approach builds a prediction model that can assign a customer into one of the three categories based on their first feedback tweet. With our prediction model, a company would be able to know the category of a customer early and take necessary actions. Our approach first extracts a number of features that characterize a customer and his/her first feedback tweet. Features of customers belonging to the three categories are then used to train a prediction model that can differentiate each category. The following subsections explain features used and our approach to build the prediction model.

4.1 Feature Engineering

We use two types of features: profile features (i.e., features that we extract from a customer's Twitter profile) and content features (i.e., features that we extract from a customer's first feedback tweet).

Profile Features Twitter provides several information about its user which include the user's number of followers, number of followee, etc. We consider five profile features to infer customer categories. These five features are described in detail below.

- **TweetCount** This feature is the number of tweets or re-tweets generated by a user. This metric represents a user's level of activity on Twitter.
- **FollowerCount** This feature is the number of followers that a user has. If A follows B on Twitter, all B's tweets would be propagated to A. This feature is a basic measure of a user's popularity on Twitter.
- **FolloweeCount** This feature is the number of people a user follows. It represents the user's level of interest on others and correlates to the number of tweets that the user would receive daily.
- **FavCount** Twitter users may express their liking of a tweet by marking the tweet as a favorite. This feature is the number of the tweets that a particular user has favorited. A higher value of this metric indicates that this user often gives positive feedback to others and may indicate his/her level of agreeableness.
- **ListCount** A Twitter user can create lists of other Twitter users whom he/she follow. Each of this list typically contains related Twitter users who belong to a particular topic or interest (e.g., a list of friends, co-workers,

celebrities, athletes, etc.). This feature is the number of lists that a user creates. We use this feature to capture another aspect of a user's level of activity on Twitter.

Content Features Content features characterize a Twitter user's first feedback tweet. The tweets in the dataset that we use has been annotated by Telkom's social media analysts (see Section 2.1). We leverage the annotations and use them as content features. We use a total of eight content features; each of them corresponds to one of the eight possible categories of tweets listed in Table 1. The value of each of these eight features is the sentiment polarity score that is assigned manually by Telkom's social media analyst.

4.2 Methodology

In general, our methodology contains of two phases: a model building phase and a prediction phase, as shown in Figure 3. In the model building phase, our goal is to build a prediction model based on a training set of customers along with their profiles, first feedback tweets and category labels. In the prediction phase, this model is used to predict the category of a new customer based on his/her profile and first feedback tweet.

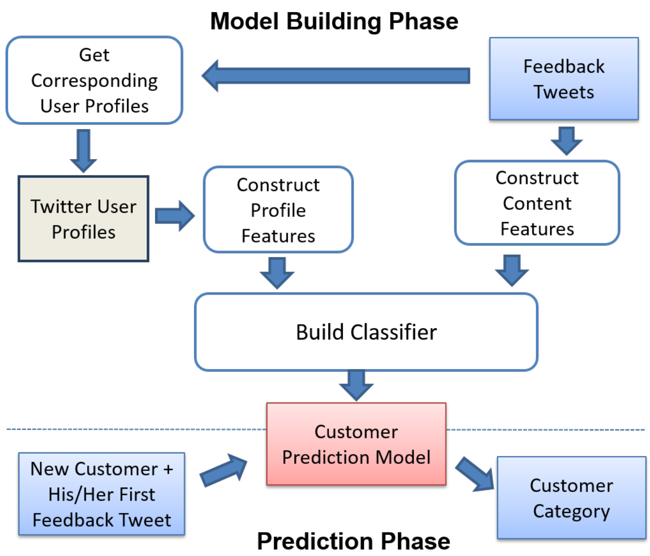

Fig. 3. Our approach for predicting customer categories

In the model building phase, we first extract profile and content features from customers in the training data. Next, for the profile features (i.e., *TweetCount,* *FollowerCount, FolloweeCount, FavCount, ListCount*), since the variation of the feature values is high, we normalize them to have values between 0 and 1. However, we do not normalize the content features, since we want to preserve the actual sentiment polarity scores and the variation of these scores is not high. After the features are extracted, we apply SMOTE (described in Section 2.2) to handle imbalanced data. Finally, we use a classification algorithm to build a prediction model.

We explore three classification algorithms, namely Logistic Regression, Naive Bayes and Random Forest. These algorithms are widely used in data mining research such as in [3, 5, 20].

In the prediction phase, we extract values of profile and content features for a new customer whose category is to be inferred. These feature values are extracted from the new customer's Twitter profile and his/her first feedback tweet. Next, we apply the prediction model that we have learned in the model building phase on the new customer's feature values. This model will output a prediction, which is one of the three categories listed in Table 4.

5 Experiments and Results

5.1 Dataset and Experiment Setting

There are 6,031 distinct Twitter users in our Telkom dataset. However, we could not collect profile features of a number of them. This is the case since not all Twitter accounts are public. Among the 6,031 users, we are able to get 5,813 user profiles. This represents 96.39% of distinct users in our dataset. For each of these users, we identify his/her tweet that will be used as input to the prediction task. We consider the earliest feedback tweet that is posted during the observation period (i.e., June-August 2015) as such tweet.

We use the implementations of Logistic Regression, Naive Bayes and Random Forest in Weka [10]. We apply SMOTE filter for all of the three variants. All parameters are set into Weka default settings. We also perform 10-fold cross validation to investigate the effectiveness of our approach.

As a baseline, we use an approach which we refer to as *WeightedRandomPicker*. This baseline picks one of the three categories randomly based on the percentage of customers of each category in our dataset (see table 4). For example, given a new customer, *WeightedRandomPicker* predicts that the customer belongs to Class 1 (Kind) with a probability of 0.10, Class 2 (One-Off) with a probability of 0.47 or Class 3 (Spiteful) with a probability of 0.30.

5.2 Evaluation Metrics

As yardsticks to measure the effectiveness of our approach and the baseline, we use precision, recall, and F-measure. These metrics are common metrics that have been widely used in many past studies such as [17, 12, 8, 21, 18].

These metrics are calculated based on four possible outcomes of a Twitter user in an evaluation set: True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN). For example, in case of predicting spiteful customer, TP is when a spiteful customer is correctly predicted as such; FP is when a non spiteful customer is wrongly predicted as a spiteful customer; FN is when a spiteful customer is wrongly predicted as a non spiteful customer; TN is when a non-spiteful customer is correctly predicted as non-spiteful customer.

Since we deal with multi-class classification, we also calculate weighted precision, weighted recall and weighted F-measure. We use the following formula to calculate weighted F-measure (similarly for weighted precision and recall):

$$WeightedFM = \frac{\sum_c FM(x) \times x))}{n} \tag{1}$$

In the above equation, c is total number of classes (in our case: 3), FM(x) is the F-measure score for class x, x is total number of data instances that belong to a particular class, and n is the total number of instances in the dataset.

5.3 Research Questions and Results

RQ1: How well does our approach perform in predicting different categories of customers?

Approach: In this research question, we investigate three variants of our approach which uses three supervised classification algorithms (Logistic Regression, Naive Bayes and Random-forest), and compare its performance (measured in terms of precision, recall, and F-measure) with that of the *WeightedRandomPicker* baseline.

Results: Table 5 shows the results of our experiments. From the table, we can see that the three variants of our approach consistently outperform the *WeightedRandomPicker* baseline. Among the three classes, determining spiteful customers (C3) is the hardest problem. In this case, Logistic Regression performs the worst when compared to the other two supervised algorithms. But still, it outperforms the baseline by 13% in terms of weighted F-measure. Meanwhile, determining kind customers (C1) is the easiest task, and all three variants of our approach outperform the baseline by more than 53% in terms of F-measure.

RQ2: How effective is the oversampling strategy to improve classification accuracy?

Approach: We apply SMOTE to handle imbalanced class problem. In this research question, we compare results obtained by our approach when SMOTE is used and when it is not used.

Results: Table 6 shows results that our approach achieves when SMOTE is turned off and on. We can note that by applying SMOTE the effectiveness of our approach (measured in terms of weighted F-measure) can be improved by 33-44%. This result gives evidence that handling imbalanced data by using minority-class oversampling improves the accuracy of the constructed prediction model.

Table 5. Effectiveness of various variants of our approach which uses different underlying classification algorithms to predict customer categories. C1: Kind customers, C2: One-Off customers, C3: Spiteful customers.

Algorithm	**Metrics**	**C1(Kind)**	**C2(One-Off)**	**C3(Spiteful)**	**Weighted**
Logistic Regression	precision	0.969	0.732	0.574	0.738
Logistic Regression	recall	0.769	0.971	0.372	0.744
Logistic Regression	F-measure	0.858	0.835	0.451	0.724
Random Forest	precision	0.973	0.754	0.697	0.787
Random Forest	recall	0.819	0.926	0.667	0.824
Random Forest	F-measure	0.890	0.832	0.682	0.800
Naive Bayes	precision	0.881	0.733	0.525	0.704
Naive Bayes	recall	0.767	0.925	0.897	0.881
Naive Bayes	F-measure	0.820	0.818	0.663	0.772
Baseline	precision	0.452	0.474	0.296	0.415
Baseline	recall	0.212	0.474	0.366	0.382
Baseline	F-measure	0.288	0.472	0.322	0.385

Table 6. Weighted F-Measure of our approach with and without SMOTE

Algorithm	**No SMOTE**	**With SMOTE**	**Improvement**
Logistic Regression	0.542	0.724	33.58 %
Random Forest	0.591	0.800	35.33 %
Naive Bayes	0.534	0.772	44.49 %

5.4 Discussion

Our experiments show that among the three variants of supervised classification algorithm, Random Forest performs the best with F-Measure of 0.890, 0.832 and 0.682 for predicting kind, one-off, and spiteful customers respectively. This finding is consistent with previous study by Caruana et.al [5] which observed that random forests tended to perform well across different settings. Even for the variant that uses the most basic machine learning approach among the three (i.e., Naive Bayes), the prediction performance is 30% better than that of *WeightedRandomPicker*. These results highlight a promising potential of applying machine learning techniques to identify different categories of customer based on their first feedback tweets.

Our prediction model relies on sentiment polarity of customer feedback. To ensure the correctness of sentiment polarity of user's tweet, we decided to use labeled/annotated data that Telkom provides and none of the authors are involved in the labeling/annotation process.

A limitation of our study is the sample used in the case study. We have evaluated the effectiveness of our approach to infer customer categories from tweets that mention one company in Indonesia. In the future, we plan to address this limitation by considering a larger set of tweets collected over a longer period of time. We also plan to experiment with other companies situated in different countries.

6 Related Work

Inferring User's Attributes and Behavior. Pennacchiotti et al. presented an approach to infer the values of a Twitter user's hidden attributes such as political orientation or ethnicity by analyzing observable information such as the user behavior, network structure and the linguistic content of the user's Twitter feed [16]. Another work by On et al. studies interactions in email network [15]. They use internal company dataset, and build a model to predict email reply order. However, our work differs with previously mentioned works since we use different sets of features taken from user profile and feedback contents. We also focus on a different problem, namely the prediction of customer category based on feedback tweets.

Handling imbalanced data. Van et al. [19] and Huang et al. [11] have investigated the effectiveness of oversampling strategies to handle issues with imbalanced datasets. Our findings in this work further demonstrate the value of using an oversampling method to deal with imbalanced dataset. In RQ2 (see Section 5.3), we show that oversampling substantially improves the prediction accuracy of our customer category prediction approach.

Social Listening Framework. Bhatia et al. develop a system that automatically monitors social network platforms, analyzes data from the platforms, and triggers events that lead to corrective actions [4]. Ajmera et al. analyze posts and messages in social network platforms and identify posts relevant to an enterprise [2]. Einwiller et al. examined the complaining behavior and complaint management on Social Media, focusing primarily on how companies manage the complaints [9]. Millard et al. found that customers engage with brands not only to complain but also to complement [14]. Chen et al. introduce a brand-specific intelligent filters on Twitter which is called CrowdE using a common crowd-enabled process [7]. Our work highlights another framework that has been implemented and currently used by a large telecommunication company using customized commercial tools. Our work extends the company's social listening framework with a capability to predict customer categories.

7 Conclusion and Future Work

In this study, we propose a method to predict customer categories (i.e., kind, one-off, and spiteful) given a customer's profile and first feedback tweet. Our approach extracts a set of profile features and content features and use these to build a prediction model using a classification algorithm. To demonstrate the accuracy of our approach, we evaluate our approach using a real dataset of labeled tweets mentioning an official account of a large telecommunication company in Indonesia. We evaluate our approach by using common evaluation metrics in data mining research (i.e., precision, recall, and F-measure), and compare its performance with that of a weighted random picker baseline. Our experiment results show that three variants of our approach that uses different underlying classification algorithms can substantially outperform the baseline. Our approach can benefit companies to improve their customer service strategies to deal with different categories of customers. For the company in our case study, our approach extends the current capability of their social media listening system by adding a prediction functionality.

In the future, we plan to evaluate our proposed approach on more dataset. To improve the accuracy of our approach further, we plan to extract more features to better characterize different categories of users. We also plan to investigate more advanced classification algorithms.

Acknowledgments. This research is supported by the National Research Foundation, Prime Minister's Office, Singapore under its International Research Centres in Singapore Funding Initiative, and PT Telekomunikasi Indonesia (Telkom).

References

1. VentureBeat Report social media (2014), `http://venturebeat.com/2014/12/12/social-media-we-complain-879-million`
2. Ajmera, J., Ahn, H.i., Nagarajan, M., Verma, A., Contractor, D., Dill, S., Denesuk, M.: A crm system for social media: challenges and experiences. In: Proceedings of the 22nd international conference on World Wide Web. pp. 49–58. International World Wide Web Conferences Steering Committee (2013)
3. Arapakis, I., Cambazoglu, B.B., Lalmas, M.: On the feasibility of predicting news popularity at cold start. In: International Conference on Social Informatics. pp. 290–299. Springer (2014)
4. Bhatia, S., Li, J., Peng, W., Sun, T.: Monitoring and analyzing customer feedback through social media platforms for identifying and remedying customer problems. In: Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. pp. 1147–1154. ACM (2013)
5. Caruana, R., Niculescu-Mizil, A.: An empirical comparison of supervised learning algorithms. In: Proceedings of the 23rd international conference on Machine learning. pp. 161–168. ACM (2006)
6. Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P.: Smote: synthetic minority over-sampling technique. Journal of artificial intelligence research pp. 321–357 (2002)
7. Chen, J., Cypher, A., Drews, C., Nichols, J.: Crowde: Filtering tweets for direct customer engagements. In: ICWSM. Citeseer (2013)
8. Dev, H., Ali, M.E., Mahmud, J., Sen, T., Basak, M., Paul, R.: A real-time crowd-powered testbed for content assessment of potential social media posts. In: International Conference on Social Informatics. pp. 136–152. Springer (2015)
9. Einwiller, S.A., Steilen, S.: Handling complaints on social network sites–an analysis of complaints and complaint responses on facebook and twitter pages of large us companies. Public Relations Review 41(2), 195–204 (2015)

10. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The weka data mining software: an update. ACM SIGKDD explorations newsletter 11(1), 10–18 (2009)
11. Huang, P.J.: Classication of imbalanced data using synthetic over-sampling techniques (2015)
12. Kedia, D., Gupta, V.: Identifying suggestions for improvement of product features from online product reviews. In: Social Informatics: 7th International Conference, SocInfo 2015, Beijing, China, December 9-12, 2015, Proceedings. vol. 9471, p. 112. Springer (2015)
13. Meilă, M., Heckerman, D.: An experimental comparison of several clustering and initialization methods. In: Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence. pp. 386–395. Morgan Kaufmann Publishers Inc. (1998)
14. Millard, N.J.: Serving the social customer: How to look good on the social dance floor. In: HCI in Business, pp. 165–174. Springer (2015)
15. On, B.W., Lim, E.P., Jiang, J., Purandare, A., Teow, L.N.: Mining interaction behaviors for email reply order prediction. In: Advances in Social Networks Analysis and Mining (ASONAM), 2010 International Conference on. pp. 306–310. IEEE (2010)
16. Pennacchiotti, M., Popescu, A.M.: A machine learning approach to twitter user classification. ICWSM 11(1), 281–288 (2011)
17. Rangnani, S., Devi, V.S., Murty, M.N.: Autoregressive model for users retweeting profiles. In: International Conference on Social Informatics. pp. 178–193. Springer (2015)
18. Tian, Y., Lo, D., Xia, X., Sun, C.: Automated prediction of bug report priority using multi-factor analysis. Empirical Software Engineering 20(5), 1354–1383 (2015)
19. Van Hulse, J., Khoshgoftaar, T.M., Napolitano, A.: Experimental perspectives on learning from imbalanced data. In: Proceedings of the 24th international conference on Machine learning. pp. 935–942. ACM (2007)
20. Van Vlasselaer, V., Eliassi-Rad, T., Akoglu, L., Snoeck, M., Baesens, B.: Afraid: Fraud detection via active inference in time-evolving social networks. In: Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015. pp. 659–666. ACM (2015)
21. Xia, X., Lo, D., Shihab, E., Wang, X., Yang, X.: Elblocker: Predicting blocking bugs with ensemble imbalance learning. Information and Software Technology 61, 93–106 (2015)