Towards identifying leading indicators of smoking cessation attempts from social media

Aron Culotta
Department of Computer Science
Illinois Institute of Technology
Chicago, IL 60616
culotta@cs.iit.edu

***Abstract*—Understanding when and why people choose to attempt smoking cessation can provide guidance for targeted intervention and public health campaign strategies. Online social networks provide a real-time, open ended data source with which to explore these questions. This paper presents preliminary methods and results to identify social media messages that are leading indicators of smoking cessation attempts.**

I. Introduction

While cigarette smoking in the U.S. is in decline, it is still the leading cause of preventable death, accounting for one of every five deaths [1]. Many applied and theoretical public health approaches aim to understand the psychological, behavioral, and environmental factors that influence a smoker to attempt to quit [2], [3]. These questions can be difficult to answer by traditional methods, as they rely on honest, self-reported survey data that often suffer from recall bias.

Emerging methods that analyze online social networks may potentially provide novel insights into these issues due to the real-time, open-ended nature of the data. Recent work has analyzed online media to identify characteristics of successful cessation attempts [4], [5] and categorize utterances during cessation attempts [6]. However, little is understood about contributing factors in a person's life that lead to a cessation attempt, for example, changing homes, careers, or relationships. Understanding these issues can help public health researchers individualize cessation attempts and better target public health campaigns [7].

In this paper, we briefly describe our preliminary, work in progress to identify leading indicators of smoking cessation attempts from Twitter data. Our approach first identifies smoking cessation attempts in a user's Twitter feed, then analyzes the tweets in the time period just prior to the cessation attempt. By comparing these recent tweets with both historical tweets from that user, and contemporary tweets from smokers who have not attempted to quit, we aim to identify linguistic patterns that are characteristic of smokers who are likely to attempt to quit smoking in the near future.

II. Methods and Preliminary Results

A. Smoking Cessation Classification

We begin with a large set of smoking-related tweets from the Twitter Firehose, spanning January 1, 2014 - June 30, 2015. These were collected by using a complex, manually refined set of queries (e.g., *cigs, ciggies, cigarettes, cigaretes, boges*, etc.).

TABLE I
Accuracy of the smoking cessation classifier.

Class	Precision	Recall	F1	Count
positive class	.80	.77	.78	296
negative class	.99	.99	.99	9704

TABLE II
Terms most predictive of smoking cessation tweets, with examples.

term	example
quit	I'm about to quit smoking after this cigarette!
quitting	Last weekend of smoking cigarettes quitting monday!
month	1 month free of Cigarettes
days	I haven't smoked a cig in 3 days #quitcoldturkey
without	3 weeks without smoking cigarettes

To identify which of these original tweets express a desire or intent to quit smoking, we performed standard supervised classification. We annotated 10k smoking-related tweets from February 2015 according to whether they contain an explicit indication that the user intends to quit smoking or is in the process of quitting smoking. These may be future tense ("Im quitting cigarettes") or past tense ("Two weeks not one cigarette #winning"). These represent a small fraction of all smoking related tweets — in this sample of 10K, only 296 were annotated as positive examples. We then trained a logistic regression classifier to identify smoking cessation tweets. We represent each tweet as a bag of words (unigrams and bigrams), removing terms occurring in fewer than two tweets. To mitigate the extreme class imbalance, we weight instances in the learning objective function inversely proportional to their frequency in the training data (thus, positive examples receive over 30x the weight of negative instances).

Table I shows the precision, recall, and F1 measures for each class, using 5-fold cross-validation in the 10k labeled tweets. Table II lists the top five coefficients for the positive class (cessation attempt), along with an example tweet containing each term for context. We observe that, despite the class imbalance, the classifier can fairly accurately identify cessation attempts, with an F1 score of .78.

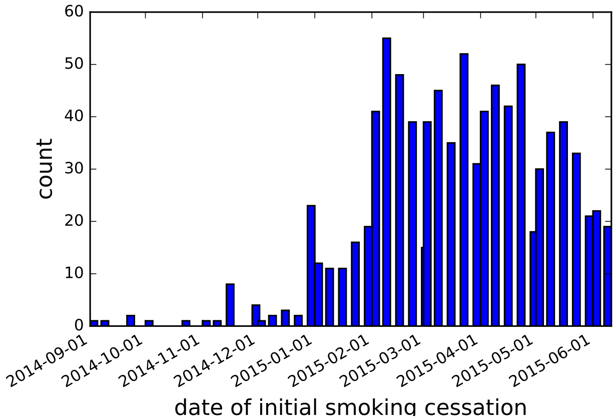

Fig. 1. The frequency of the initial smoking cessation date by week for each of the 918 users.

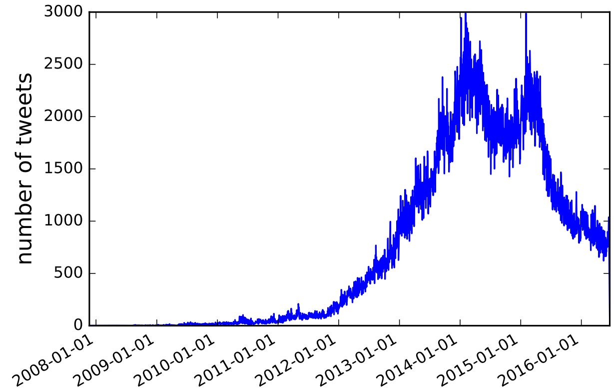

Fig. 2. Tweets by day for 2,237,472 tweets from 918 users who quit smoking in the first half of 2015.

B. Identifying Smoking Cessation Date

We next used the cessation classifier from the previous section to categorize all smoking-related tweets from February 2015 through June 2015 in our original data. For each tweet classified as positive, we used regular expressions to determine whether the tweet mentions the number of days since the user has stopped smoking (e.g., "3 days smoke free," "I haven't smoked in 3 days," "I stopped smoking today"). Based on the extracted information and the date of the tweet, we can then determine the day that the user first stopped smoking. To further reduce noise, we removed retweets, tweets containing URLs, and tweets from users with fewer than 50 or more than 10K total tweets. This resulted in tweets from 3,720 unique users for which we had high confidence in the precise day that the user quit smoking.

From these users, we sampled 918 users whose quit days occurred between September 2014 and June 2015. Figure 1 displays a histogram of the estimated date of cessation for the 918 users, binned by week. As expected, the cessation date is most commonly in the range of the cessation tweets identified from February-June; however, some tweets list larger day values when marking anniversaries (e.g., "100 days smoke free"), leading to earlier cessation dates. Thus, while January 1 was not included in the original set of identified cessation tweets, there is a spike on that day in Figure 1, due to New Year's resolutions.

For each of these 918 users, we then collected up to 3,200 of their most recent tweets (most of which are not related to smoking). Figure 2 displays the number of tweets per day in the resulting 2,236,472 tweets (roughly 2,400 tweets per user). Clearly, the users in this sample are extremely active Twitter users (at least, they were in 2014-2015), and caution must be applied when attempting to generalize conclusions to the larger population. (In future work, we will estimate the socio-economic status of our sample to assess generalizability.)

C. Identifying Leading Indicators of Cessation

We next design a classification task to identify leading indicators of a user's attempt to quit smoking. To do so, we define the 90 days prior to the user's first day of smoking cessation as the *cessation precursor window*. Our conjecture is that the user's online activity prior to a smoking cessation attempt differs from the user's activity at earlier points in time. For example, a user may enter a new relationship, start a new job, or move to a new home. The goal is to identify linguistic evidence of such changes.

We construct a training set by first concatenating all tweets from a user's cessation precursor window. We represent these tweets with a single bag of words (unigrams and bigrams). Such examples are considered positive training instances, as they should contain indicators of an impending cessation attempt. For each positive example, we construct a corresponding negative example from the user's tweets *a year prior* to the cessation attempt. For example, if a user quits smoking on January 1, 2015, we construct a positive instance using tweets from October 2, 2014 through December 31, 2014. We then construct a negative instance using tweets from October 2, 2013 through December 31, 2013. In this way, we can reduce cyclical effects in the classifier (e.g., we would like to discover more than the fact that people tend to quit smoking after Christmas).

We construct an additional set of negative examples by identifying smokers who have not attempted to quit smoking and are active on Twitter from September 2014-June 2015. This control group allows us to reduce additional temporal effects of one-time events (e.g., we do not want our classifier to determine that a sporting event in February 2015 is a leading indicator of smoking cessation). To do so, we manually identify tweets indicating that the user is a smoker (e.g., "coffee and cigarettes are my only escaaaape"). We identify 90 smokers from February 2015, collect their most recent 3,200 tweets, then restrict those tweets to those between September 2014 and June 2015 (the same range as the positive examples).

This results in 58,300 tweets. In this way, we construct a negative set of tweets that are from the same time period as the positive examples (the cessation precursor window), but are tweets from smokers rather than those who are attempting to quit smoking.

With this training data, we then fit a classifier to distinguish between these three groups: (1) **positive** (cessation precursor window), (2) **negative - year prior** (the past tweets of users who attempted to quit), and (3) **negative - contemporary** (tweets from smokers who have not attempted to quit). We use a multi-class logistic regression classifier, using the same features as the original cessation classifier. Table III displays the accuracy by class, and Table IV lists five of the terms most predictive of the positive class, with a sample tweet for context.

We observe that the classifier is again fairly accurate at identifying the positive class, with a final F1 score of .72. It is somewhat surprising that the classifier can distinguish between **positive** and **negative - year prior**, given that these tweets come from the same 918 users (and so we would expect them to have similar terms overall). The accuracy for **negative - contemporary** is rather low (F1 of .38), which is in large part due to class imbalance. (We used class weighting as in the cessation classifier, which helped somewhat.)

The top terms in Table IV provide some insight into the data. Recall that these terms are predictive of the months prior to a cessation attempt, and not for the same time period one year prior. These results suggest that relationship changes play a role in cessation attempts. (Note that "bae" is an abbreviation of "babe," a common term of endearment.) The top terms all contain references to romantic relationships and loved ones. Such life transition events are well-summarized by tweets like this one: "Had an amazing weekend with [*redacted*] now it's onwards and upwards for me! New job, better people to walk into my life from now please." One possibility is that entering a new relationship encourages one to change health behaviors, including smoking cessation, though further investigation is needed to understand this.

We also observe a common phenomenon of using retweeting to subtly reveal personal events. For example, in Table IV, the user in the fourth example retweets a message about a partner's behavior, presumably because the retweeter identifies with the feelings of the original poster. Similar retweets are used to indicate the start of a new relationship (e.g., "RT @[*redacted*]: When you start catching feelings for someone").

Other types of tweets we observed include indicators of college completion ("6 weeks of college left"), moving ("I'm moving to London"), and starting careers ("I start a 'big girl' job tomorrow!"). Further work is required to automatically identify such tweets and assess their significance.

TABLE III
ACCURACY OF THE CESSATION LEADING INDICATOR CLASSIFIER.

Class	Precision	Recall	F1	Count
positive	.71	.73	.72	918
negative - year prior	.73	.73	.73	918
negative - contemporary	.47	.31	.38	90

TABLE IV
TERMS MOST PREDICTIVE OF THE CESSATION PRECURSOR WINDOW, WITH EXAMPLES.

term	**example**
bae	Waking up at 6 am to take bae to work
goals	Be with a man who will wake you up, just to remind you to put ya scarf/bonnet on. #RelationshipGoals
valentine's	Niagara Falls for Valentine's Day maybe
when bae	RT [*redacted*]: When bae acting crazy for no reason
netflix	All I want right now is cuddles and netflix

III. Conclusion and Future Work

In this short paper, we have provide preliminary evidence that Twitter may provide early indicators of a user's attempt to quit smoking.

In ongoing work, we will increase our sample size to enable us to more reliable identify leading indicators, and we will investigate alternatives to controlling for the many potential temporal and demographic variables that may confound this analysis.

Acknowledgments

This research was funded in part by support the National Science Foundation under grant #IIS-1526674. Any opinions, findings and conclusions or recommendations expressed in this material are the authors' and do not necessarily reflect those of the sponsor.

References

[1] U. D. of Health, H. Services *et al.*, "The health consequences of smoking50 years of progress: a report of the surgeon general," *Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health*, vol. 17, 2014.

[2] A. Bandura, "Self-efficacy: The exercise of control," 1997.

[3] J. O. Prochaska and W. F. Velicer, "The transtheoretical model of health behavior change," *American journal of health promotion*, vol. 12, no. 1, pp. 38–48, 1997.

[4] E. L. Murnane and S. Counts, "Unraveling abstinence and relapse: Smoking cessation reflected in social media," in *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*, ser. CHI '14. New York, NY, USA: ACM, 2014, pp. 1345–1354.

[5] A. Tamersoy, M. De Choudhury, and D. H. Chau, "Characterizing smoking and drinking abstinence from social media," in *Proceedings of the 26th ACM Conference on Hypertext & Social Media*. ACM, 2015, pp. 139–148.

[6] G. Wadley, W. Smith, B. Ploderer, J. Pearce, S. Webber, M. Whooley, and R. Borland, "What people talk about when they talk about quitting," in *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: the Future of Design*. ACM, 2014, pp. 388–391.

[7] J. Paay, J. Kjeldskov, M. B. Skov, L. Lichon, and S. Rasmussen, "Understanding individual differences for tailored smoking cessation apps," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 2015, pp. 1699–1708.