

Finding Optimists and Pessimists on Twitter

Xianzhi Ruan, Steven R. Wilson, and Rada Mihalcea

University of Michigan

Ann Arbor, MI

{rxianzhi, steverw, mihalcea}@umich.edu

Abstract

Optimism is linked to various personality factors as well as both psychological and physical health, but how does it relate to the way a person tweets? We analyze the online activity of a set of Twitter users in order to determine how well machine learning algorithms can detect a person's outlook on life by reading their tweets. A sample of tweets from each user is manually annotated in order to establish ground truth labels, and classifiers are trained to distinguish between optimistic and pessimistic users. Our results suggest that the words in people's tweets provide ample evidence to identify them as optimists, pessimists, or somewhere in between. Additionally, several applications of these trained models are explored.

1 Introduction

Optimists believe that future events are going to work out for the best; pessimists expect the worst (Carver et al., 2010). Research has shown that optimism is correlated with many positive life outcomes including improvements in life expectancy (Diener and Chan, 2011), physical health (Peterson and Bossio, 2001), and mental health (Achat et al., 2000). Previously, it was found that optimism and pessimism are differentiable but related: pessimism was principally associated with neuroticism and negative affect while optimism was primarily associated with extraversion and positive affect (Marshall et al., 1992). Another study found that optimism was correlated with personality factors, including extraversion, emotional stability, conscientiousness, and agreeableness (Sharpe et al., 2011).

It is clear that optimism relates to a wide variety of psychological and social variables, but how might optimism influence the way a person utilizes a social media platform? What features distinguish optimistic users from pessimistic ones? In order to answer these questions, we must first establish a means by which we can measure people's levels of optimism and pessimism. The Life Orientation Test (LOT) is commonly used to assess the degree to which a person is an optimist (Scheier and Carver, 1985). This short survey asks respondents to evaluate their own agreement with a number of short statements using a five-point scale. However, distributing such a survey over a large population requires both time and a form of incentive. Recent work has shown that open-ended text samples can be computationally analyzed to provide a more comprehensive view of a person's personal values than can be achieved using a more constrained, forced choice survey (Boyd et al., 2015). Furthermore, we know that language use is an independent and meaningful way of exploring personality (Pennebaker and King, 1999), and personality is correlated with optimism (Sharpe et al., 2011). Given a large enough text corpus, it may therefore be possible to build computational models that can automatically recognize optimism itself by looking at the words people use. The vast amount of publicly available social media data provides an excellent source of data that can be used to build models of users' psychological traits, as was done in previous studies that trained machines to predict aspects of personality from tweets (Golbeck et al., 2011; Sumner et al., 2012).

A tool that could identify optimists and pessimists by analyzing their text would aid in large scale studies of optimism among social media or other web users by providing a large number of subjects to analyze. This would open the door to

massive studies of the relationships between optimism, pessimism, and a range of online behaviors. On the other hand, an optimism classification system could help improve the social platform itself. For example, by learning more about the psychological traits of its users, Twitter could improve its “who to follow” suggestions so that they reflect people who have a similar outlook on life.

2 Data and Method

As a data source, we chose Twitter because it is widely used and is ripe with short pieces of text (i.e., tweets) containing people’s everyday thoughts, observations, and conversations. We used Twitter’s basic search function to look for users whose tweets include words and phrases that indicate that they might identify as optimistic or pessimistic. Looking for phrases such as “I am optimistic” within a user’s tweets to find potentially optimistic users, we identified 714 candidates. Finding pessimistic Twitter users proved more difficult because users would not usually tweet something negative such as “I am pessimistic” and present themselves in an unflattering way. We instead searched for keywords such as “hate,” “unfair,” and “disgust,” which may indicate a pessimistic nature. This led to 640 potential pessimists. For each user, we crawled their 2,000 most recent tweets (or all their tweets if the user had less than 2,000). In order to verify that the accounts identified were owned mostly by individual users (as opposed to organizations), we manually inspected a random sample of 50 accounts and found only one that appeared to be related to an organization.

Using the collected data set, which we expected would be more representative of optimistic or pessimistic nature than the norm based on the content of their tweets, we selected a fraction of the users to create a ground truth set for our task. We used Amazon Mechanical Turk (MTurk)¹ to obtain human annotations for a subset of our corpus. We randomly selected 500 users who were retrieved by the optimistic queries and 500 users found when searching for pessimists. For each user, we randomly selected 15 tweets for a total of 15,000 tweets to be labeled on a scale of -3 (very pessimistic) to 3 (very optimistic) by five independent annotators. Before labeling began, we provided clear definitions of optimism and pessimism

¹<http://www.mturk.com>

to the annotators.

In order to pick the tweets from each user that had a stronger emotional signal, we took advantage of the “positive emotions” and “negative emotions” word categories included in the Linguistic Inquiry and Word Count Tool (Pennebaker et al., 2001).² If any of the original 15 tweets did not contain at least one word from either category, the tweet was removed and a new tweet was chosen at random to replace it. This process was repeated until we had a set of 15 tweets per user without skewing that user’s true distribution of positive and negative tweets.

During the MTurk annotation, to identify workers who were quickly selecting options without even reading the tweets, we added a “check” question that asked the workers to choose a specific value for that question. All the workers who did not correctly answer this “check” question were removed from the annotation. When a worker’s annotations had to be thrown out, the tweets were put back onto MTurk for reannotation. Additionally, we compared the scores of each annotator with the average score and removed workers who deviated significantly from the others. The final agreement (Krippendorff’s alpha) between the five annotators was measured at 0.731, assuming an interval scale.

For each individual tweet, we assigned a label of “optimistic,” “pessimistic,” or “neutral”. Any tweet with an average score greater than one (slightly optimistic or higher in the annotation task) was considered an “optimistic” tweet, and those with an average score less than one (slightly pessimistic or lower) were given the “pessimistic” class label. The tweets with average MTurk annotation scores between -1 and 1 were considered to be “neutral.”

We also assigned a class label to each user. To accomplish this, we calculated the average of the assigned scores, sorted the Twitter users by their level of optimism, and considered the top 25% of users as optimists, the bottom 25% as pessimists, and the remaining ones as neutral.

Before moving on, we decided to investigate the online behaviors and attributes of the optimistic and pessimistic users in our new data set. A summary of some of the differences between the two groups is shown in Table 1. Interestingly the optimists have more followers and are following more

²The 2007 version of LIWC was used

| | Optimistic | Pessimistic |
|----------------------|------------|-------------|
| mean followers | 6238 | 1840 |
| mean following | 1898 | 1156 |
| mean tweets | 12156 | 28190 |
| median followers | 972 | 572 |
| median following | 718 | 443 |
| median tweets | 5687 | 17451 |
| std. dev of tweets | 17952 | 37851 |
| min number of tweets | 108 | 28 |
| max number of tweets | 99344 | 231220 |
| tweet rate | 10.24 | 19.184 |
| favorite count | 4314.34 | 8761.2 |
| listed count | 112.62 | 10.23 |

Table 1: Statistics for the most extreme 100 optimistic & 100 pessimistic users.

other users than the more pessimistic users. On the other hand, the pessimists tend to tweet much more frequently, with a mean and median number of tweets both more than twice as large as the optimist group. This is not just a factor of the pessimists having been around longer to build up a history of tweets- we also compute the “tweet rate” for each user by dividing their total number of tweets by the total number of days since the activation of their Twitter account. Looking at this variable, we see that the average number of tweets per day is much higher for the pessimists. Optimists are also included in more lists, while pessimists choose to label things as a “favorite” more often.

In order to build computational models to differentiate between the optimistic and pessimistic users, we use five different methods from the scikit-learn python library³: Naive Bayes (NB), Nearest Neighbor (NN), Decision Tree (DT), Random Forest Classifier (RFC), Gradient Boosting Classifier (GBC) and Stochastic Gradient Descent (SGD). The default parameters are used for each. The preprocessing method was the same for all different classifiers: the text was preprocessed by removing mentions (@), web links, and the phrase RT. We also used the Emoji unicode tables to replace all Emoji unicodes to their corresponding meanings (e.g., “<smiling-face>”). We tried performing classification both with and without removing stopwords to see what the effect was. For all different classifiers, we tested with different settings: with and without stopwords; and adding a user’s profile information as additional features or not.

³<http://scikit-learn.org/stable/>

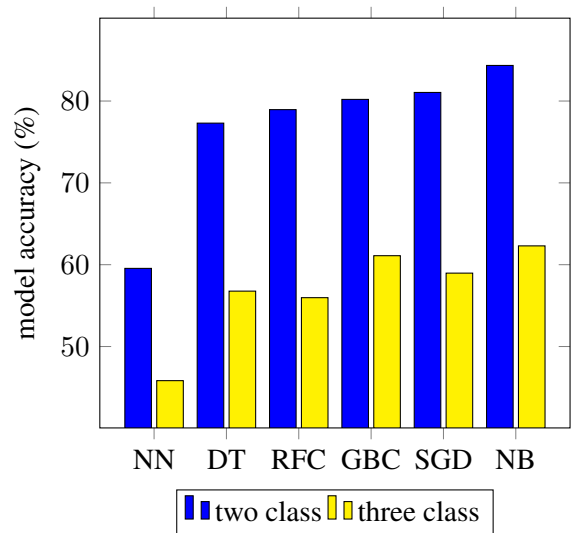


Figure 1: Tweet level classification accuracy for the two-way and three-way classification problems.

3 Results

We first evaluate the ability of our classifiers to distinguish between optimistic and pessimistic tweets (two-way classification) or among optimistic, pessimistic, and neutral tweets (three-way classification). We randomly selected 1,000 tweets from each class. Figure 1 shows the ten-fold cross validation results obtained using the six classifiers. During each classification, we made sure that tweets from the same user were not shared between the training and testing folds. In both cases, the best setting was using the Naive Bayes classifier and not including profile information as features. Stopword removal had no noticeable effect. Note that the majority baseline is a score of 50% in the two-class case, while it is 33% in the three-class case.

For additional insight, Table 2 shows some of the top features for the optimistic and pessimistic class, sorted by the probability of the feature given the class. We can see that, as one might expect, the useful words for detecting optimists are generally very positive, while the pessimistic features are negative and sprinkled with profanity. Since we formulated the problem as a three-way classification, it is reasonable that some words may have high scores for both optimistic and pessimistic classes. These words distinguish optimism/pessimism from the neutral class.

We perform our next evaluation at the user level, which means that we consider all tweets from

| Optimism | Pessimism |
|-----------------------|------------------------|
| love, so, that, be | fuck, that, not, so |
| have, good, am, this | like, am, do, hate |
| on, your, not, day | have, be, this, just |
| like, just, do, will | up, life, on, shit |
| can, get, what, at | no, people, can, what |
| great, make, up, much | feel, your, about, I'm |
| best, we, if, go | go, know, get, even |
| was, from, thing, out | want, at, was, off |
| look, thank, know, he | out, kill, if, done |

Table 2: Most discriminating features collected by the Naive Bayes classifier for the three-class tweet-level prediction setting.

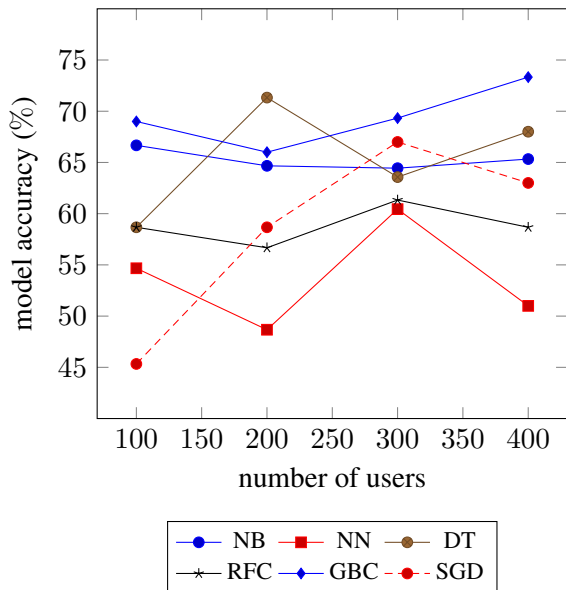


Figure 2: Accuracy on the user-level prediction task for different data sizes.

a user as a single document. The classification is performed using a randomly selected set of 100, 200, 300, and 400 users from the annotated set (each set adds 100 new users to the previous group). In each case, the 25% users with highest annotation score are considered the optimistic group, 25% users with lowest annotation score as pessimist group, and the other 50% of users is the neutral group. The results of the ten-fold cross validation are shown in Figure 2. In this setting, the Gradient Boosting Classifier usually outperforms the others and achieves an accuracy of 73.33% on the 400 user data set.

We also sought to discover how accurate the classifiers would be if the objective was simply to identify the top N optimists or pessimists. For example, if we wanted to find the 10 users out of a group with the greatest number of optimistic tweets, how accurately could this be done? To

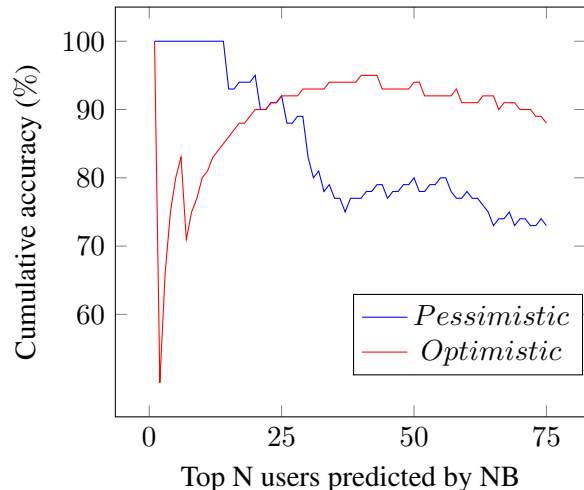


Figure 3: Cumulative classification accuracy on the top N users sorted by Naive Bayes predicted probabilities

| City | Neutral Users | Optimists | Pessimists |
|-------------|---------------|-----------|------------|
| Chicago | 20.65% | 39.27% | 40.08% |
| Los Angeles | 14.55% | 31.87% | 53.58% |
| New York | 20.74% | 40.63% | 38.63% |

Table 3: Predicted optimism & pessimism in three major cities

carry out this analysis, we sorted the users by the probabilities that they belonged to either the optimistic class or the pessimistic class as predicted by a Naive Bayes classifier (Figure 3). Then, we compute the accuracy for the top N optimists and pessimists. As we can see, it is possible to predict the most pessimistic 14 users with perfect accuracy. On the other hand, some of the most likely optimistic users actually belonged to another class based on the ground truth labels. With a larger number of users to classify, it becomes easier to correctly label optimists than pessimists.

4 Applications

What kinds of things can we learn with a tool for classifying optimists and pessimists? First, we look at groups of users from three major cities in the United States: Chicago, Los Angeles, and New York. We found users who listed their location as one of these three cities (494 users from Chicago, 433 from Los Angeles, 480 from New York), then collected 2,000 tweets from each user. Using our best models from the user-level experiments, we obtain predictions for the optimism/pessimism of the users. The breakdown of predicted optimists, pessimists, and neutral users is listed in Table 3.

Chicago and New York are fairly balanced with roughly 40% of people falling into each category (leaving 20% as neutral). However, pessimists were predicted much more often than optimists in Los Angeles.

For a second sample application, we went to the official twitter accounts of six presidential candidates: Hillary Clinton, Donald Trump, Marco Rubio, Martin O'Malley, Bernie Sanders and Ben Carson. We randomly picked approximately 500 followers of each of the candidates and predicted the optimism/pessimism of them (Table 4).⁴ While these scores are only estimates, we see that O'Malley's followers tend to be the users who posted a greater number of optimistic tweets, while the users who tweeted lots of pessimistic tweets are those keeping up-to-date with Rubio's campaign. Overall, we see that most of the followers of these candidates are optimistic.

| Candidate | Neutral Users | Optimists | Pessimists |
|-----------|---------------|-----------|------------|
| Clinton | 31.52% | 49.22% | 19.26% |
| Trump | 36.20% | 39.46% | 24.32% |
| Rubio | 30.00% | 49.41% | 20.59% |
| O'Malley | 29.22% | 64.51% | 6.26% |
| Sanders | 44.62% | 44.42% | 10.96% |
| Carson | 31.54% | 49.90% | 18.56% |

Table 4: Predicted optimism & pessimism of those following some of the candidates for the 2016 Presidential election.

5 Conclusions and Future Work

We have shown that we can use Twitter to collect a data set⁵ of optimistic and pessimistic users, and predict the most (top 25%) optimistic/pessimistic users with greater than 70% accuracy. The optimistic users on Twitter tended to have more social connections, but tweet less often than the pessimists. In the future, we hope to explore the social effects of optimism, such as the degree to which optimistic users follow one another and whether or not optimistic comments receive more "favorites" and retweets. Finally, we would like to compare the optimism and pessimism scores that our model predicts with those received when taking the LOT in order to compare the text-based analysis with a widely used tool for measuring optimism.

⁴data collected late December, 2015

⁵The data set introduced in this paper is available at <http://lit.eecs.umich.edu/research/downloads>.

Acknowledgements

We would like to thank Seong Ju Park, Tian Bao, and Yihan Li for their assistance in the initial project that led to this work. This material is based in part upon work supported by the National Science Foundation award #1344257 and by grant #48503 from the John Templeton Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation or the John Templeton Foundation.

References

- Helen Achat, Ichiro Kawachi, Avron Spiro, Deborah A DeMolles, and David Sparrow. 2000. Optimism and depression as predictors of physical and mental health functioning: the normative aging study. *Annals of Behavioral Medicine*, 22(2):127–130.
- Ryan L Boyd, Steven R Wilson, James W Pennebaker, Michal Kosinski, David J Stillwell, and Rada Mihalcea. 2015. Values in words: Using language to evaluate and understand personal values. In *Ninth International AAAI Conference on Web and Social Media*.
- Charles S Carver, Michael F Scheier, and Suzanne C Segerstrom. 2010. Optimism. *Clinical psychology review*, 30(7):879–889.
- Ed Diener and Micaela Y Chan. 2011. Happy people live longer: Subjective well-being contributes to health and longevity. *Applied Psychology: Health and Well-Being*, 3(1):1–43.
- Jennifer Golbeck, Cristina Robles, Michon Edmondson, and Karen Turner. 2011. Predicting personality from twitter. In *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (Social-Com), 2011 IEEE Third International Conference on*, pages 149–156. IEEE.
- Grant N Marshall, Camille B Wortman, Jeffrey W Kusulas, Linda K Hervig, and Ross R Vickers Jr. 1992. Distinguishing optimism from pessimism: Relations to fundamental dimensions of mood and personality. *Journal of personality and social psychology*, 62(6):1067.
- James W Pennebaker and Laura A King. 1999. Linguistic styles: language use as an individual difference. *Journal of personality and social psychology*, 77(6):1296.
- James W Pennebaker, Martha E Francis, and Roger J Booth. 2001. Linguistic inquiry and word count: Liwc 2001. *Mahway: Lawrence Erlbaum Associates*, 71:2001.

- Christopher Peterson and Lisa M. Bossio. 2001. Optimism and physical well-being. In *Optimism & pessimism: Implications for theory, research, and practice.*, pages 127–145. American Psychological Association (APA).
- Michael F Scheier and Charles S Carver. 1985. Optimism, coping, and health: assessment and implications of generalized outcome expectancies. *Health psychology*, 4(3):219.
- J Patrick Sharpe, Nicholas R Martin, and Kelly A Roth. 2011. Optimism and the big five factors of personality: Beyond neuroticism and extraversion. *Personality and Individual Differences*, 51(8):946–951.
- Chris Sumner, Alison Byers, Rachel Boochever, and Gregory J Park. 2012. Predicting dark triad personality traits from twitter usage and a linguistic analysis of tweets. In *Machine Learning and Applications (ICMLA), 2012 11th International Conference on*, volume 2, pages 386–393. IEEE.