Understanding Personality through Social Media

Yilun Wang
Department of Computer Science
Stanford University
yilunw@stanford.edu

Abstract

In this paper, we study the relationship between language use on Twitter and personality traits. Specifically, we want to know how various linguistic features correlate with each personality trait and to what extent can we predict personality traits from language. We gather personality data from Myers-Briggs Type Indicator (MBTI) personality test which contains thinking, feeling, sensation, intuition, introversion, extroversion, judging and perceiving. Using the 90K users in our dataset, we collect most recent tweets from them and design three categories of feature, namely bag of n-grams, Twitter POS tags, and word vectors to explore the most related linguistic features for different personality traits. Analysis of these features provide insights of language use for different personalities. For instances, extroverts tend to use hashtag and phrases like ”so proud”, ”so excited”, and ”can’t wait”. People who like to use emoticon are more likely to be Sensing and Feeling personality type. Moreover, we investigate the predictive power of individual features and combined features in our analysis. With the concatenation of all the features we extracted, we can predict the personality traits with an average AUC of 0.661.

1 Introduction

Personality has been studies extensively in social science and psychology as it reflects the way people behave and react in online social media and in the society. Previous studies showed that personality significantly correlated with several real-world behaviors which makes it important in providing personalized services. For instance, it correlates with music taste, Extroverted people tend to like popular music, while open to experience people are more likely to enjoy unpopular one (Rawlings and Ciancarelli, 1997). Personality is also related to the formation of social relations (Selfhout et al., 2010), the pages that people like on Facebook (Kosinski et al., 2013), and the language that people use to communicate (Boyd et al., 2015).

People are increasingly using social media platforms, such as Twitter, Facebook, and Pinterest, to share their thoughts and opinions with their friends or people who are interested. Such scale of social media platforms provide us with a unprecedented opportunity to understanding psychological attributes on a large user base. In this paper, we want to analyze and predict personality by constructing a bridge between personality and language in popular social media such as Twitter, Facebook, and Pinterest. Specifically, we aim to find the linguistic features that distinguish people with different personality types and explore how these features can be explain by personality. Further, using the these features we want to understand the degree to which we can predict personality traits from social media language.

However, little research has touch upon understanding personality through social media because of a few reasons. First, language on social media has richer content that makes the typical linguistic analysis tool perform poorly. For example, Twitter, an online social networking service that enables users to send and read short 140-character messages called ”tweets”, contains many Twitter-specific language such as hashtag (#), at-mention (@), url, and emoticons. People tend to use shorten version of phrases on Twitter, for example, ”iono” means ”I don’t know”. Twitter poses additional challenges due to the conversational nature of the text, the lack of conventional orthography, and 140-character limit of each message (tweet). Also, collecting personality data is costly as nor-

Table 1: Examples of positive and negative tweets

Positive Tweets	@ProfCarol Just wondering, what's your type? I'm an **ENFJ** @whitneyhess that's an interesting test.. i got **ENTP** and it seems pretty accurate IMO @megfowler I'm **INTP** according to this http://similarminds.com/jung.html
Negative Tweets	I'll bet that Jeremiah @jowyang is an **ESTJ**. @mark **ENTJ** You should have known... http://typelogic.com/entj.html I love my wife. Even though she's **INFP**.

mally subjects need to take questionnaires with tens or hundreds of questions in order to estimate their personality traits. This questionnaire-based approach makes it difficult to scale up the personality test for a large amount of users.

In this paper, we try to solve the aforementioned problem by designing new richer linguistic analysis tools which can extract language feature in social media context and introducing a mechanism to automatically extract personality from text in social media. Using Twitter as a case study, we investigate the relationship between language features in tweets and personality traits, which leads to further experiment in predicting personality from language.

The contribution of our work is listed as follows:

- We compile a Twitter dataset with around 90,000 users by extracting and filtering all personality-related tweets on Twitter from 2006 to 2015. The dataset contains not only the personality types, but also the most recent tweets for all the 90,000 users, which creates enormous opportunities to study the relationship between tweets and personality.

- We design and implement three categories of linguistic features based on tweets, and further explore the correlations between each linguistic feature and each personality type. Several interesting findings can be observed here. For instance, extroverts tend to use hashtag and phrases like ”so proud”, ”so excited”, and ”can't wait”. People who like to use emoticon are more likely to be Sensing and Feeling personality type.

- We investigate the predictive powerful of the three categories of linguistic features we design by predicting each personality trait respectively. With the combination of all the three categories of features, we can predicting between introversion and extroversion with an AUC of 0.691 and an average AUC 0.661 of all personality traits .

2 Data

To realize the research about understanding personality from social media, we used Twitter as a case study and constructed a dataset with around 90,000 users with their personality traits and the most recent tweets.

2.1 Personality

Personality model used here, the Myers-Briggs Type Indicator (MBTI) (Myers, 1962), make the theory of psychological types understandable and useful in people's lives. The essence of the theory is that much seemingly random variation in the behavior is actually quite orderly and consistent, being due to basic differences in the ways individuals prefer to use their perception and judgment. MBTI captures four types of personality traits such as introversion (I) or extroversion (E), Sensing (S) or Intuition (N), Thinking (T) or Feeling (F), and Judging (J) or Perceiving (P). When you decide on your preference in each personality trait, you have your own personality type, which can be expressed as a code with four letters. The four personality traits constitute 16 distinctive personality types that have 16 personality codes ranging from “ISTJ”, “ISFJ”, to “ENFP”.

2.2 Users with personality data

These uppercase distinctive personality tokens give us a unique opportunity to extract personality types from users who talked about their MBTI personalities in social media. Using GNIP APIs from Twitter and Stanford Data Science Initiative[1], we have crawled all the users who have self-identified their personality types from 2006 to 2015 on Twitter, for example, “ESTJ” or “INFP”. In the end, We got 1.7 million tweets that contain the personality codes. Then, we need to pre-processing the

[1] https://gnip.com/

Table 2: Descriptive statistics of variables used.

Variable	Details
# of users	92,152
# of users with 20 tweets	89,548
# of tweets	17,469,989
# of tweets w/o personality codes	17,440,512
# of tokens	221,821,500

Table 3: Distribution of personality types in our dataset.

ISTJ: 3446	ISFJ: 3267	INFJ: 12885	INTJ: 12247
ISTP: 1874	ISFP: 2492	INFP: 11706	INTP: 7446
ESTP: 1132	ESFP: 2164	ENFP: 10400	ENTP: 4386
ESTJ: 2006	ESFJ: 2364	ENFJ: 6812	ENTJ: 4921

tweet data to make sure: 1. users are talking about their own personalities when writing about the personality codes; 2. the users are using English. The examples of positive labels and negative labels are listed in Table 1. We used simple heuristic rules by searching for prefixes such as ”I’m”, ”I got”, ”I have been a” of the personality codes. We didn’t apply complicated classification method because we want the rules we used to certain that the personality types are related to the users who wrote them and make sure a high precision in this process. Therefore, we retrieve 120K tweets out of all the 1.7M tweets with personality codes.

2.3 Tweets

Using standard Twitter API, we managed to crawl 200 most recent tweets of 92,152 among users who wrote the 120K tweets with personality codes (some users were banned or removed from Twitter after they posted about personality). In order to remove the obvious psychological signals in the tweets, we filtered out tweets with personality codes. The detail statistics of our dataset can be seen in Table 2.

We used the users with more than 20 tweets in our following analysis. The personality distribution of these users is shown in Table 3. The personality distribution in this dataset is skewed. However, the personality distribution of whole US population is also skewed.

3 Understanding personality

In this section, we aim to understand the relationship between MBTI personality of each user and the language in their tweets. In order to do so, we designed three categories of features: bag of n-grams, part-of-speech tags, word vectors. In the following subsections, we explain the motivations and the details of these features. Moreover, we show the relationship between personality traits and each feature in three categories which give us some interesting findings.

4 Bag of n-gram

n-gram models are a very important constituent in statistical language model, a probability distribution over sequences of words. Given such a sequence, say of length n, it assigns a probability $P(w_1, \ldots, w_n)$ to the whole sequence. The length n of the sequence of words decide the name of the model n-gram, for example, statistical language model based on sequence of one word, two words, and three words are corresponding to 1-gram (unigram), 2-gram (bigram), and 3-gram (trigram) models. Having a way to estimate the relative likelihood of different phrases is useful in many natural language processing applications. Language modeling is used in speech recognition, machine translation, part-of-speech tagging, parsing, handwriting recognition, information retrieval and other applications.

In this paper, we use the most frequent 1000 unigram, bigram, trigram words and phrases from the all the tweets in our dataset. For 1-gram word, we remove stop words that don’t provide useful information about the language. The reason we only use the most frequent 1000 words and phrases is that data sparsity is a major problem in building language models. Using the most frequent 1000 words and phrases can effectively reduce the size of feature sets while still keep the valuable signals in these linguistic features.

Bag of n-grams model is used here to represents linguistic features of users based on n-gram models. By combining all the tweets for each user, we count number of occurrence for each frequent unigrams, bigrams, and trigrams and normalize by the number of unigrams, bigrams, and trigrams for each users. In this way, we construct a vector representation (1000 dimensions for unigram, bigram, and trigram respectively) for each user using language model.

The bag of n-grams vector representations can be used to measure the correlations between n-gram words and phrases, and the personality traits.

For example, in Figure 1(a) and Figure 1(b), we plot the top related unigram words for Thinking and Feeling. Interesting, the most correlated unigram for thinking is the question mark (?) while "love", "happy", "thank", "beautiful" are highly correlated with Feeling. We have the same plots for bigram phrases, and introversion or extroversion. From these bigrams, we can see introverts tend to complaint ("my god", "holy shit") and refuse ("I don't", "I can't"), while extroverts are more energetic ("so proud", "can't wait", "so excited").

These results prove our assumption that language varies for people with different language. Even with simple count of probabilities for unigrams, bigrams, and trigrams, we can see signals from sequences of words that reflect the corresponding personalty traits.

5 Part-of-speech tag

One of the most fundamental tasks in linguistic analysis and natural language processing is part-of-speech (POS) tagging, a basic form of syntactic analysis which has countless applications. Most POS taggers are trained from treebanks in the newswire domain, such as the Wall Street Journal corpus of the Penn Treebank. Tagging performance degrades on out-of-domain data, and Twitter poses additional challenges due to the conversational nature of the text, the lack of conventional orthography, and 140- character limit of each message (tweet).

Here we adopt a POS tagger which has over 90% accuracy from (Gimpel et al., 2011). Unlike, traditional POS tagger based on Penn Treebank, this Twitter POS tagger has 25 types of distinctive tags with some Twitter-specific tags such as hashtag, at-mention, discourse marker, URL, and emoticon. The description of tags is shown in Table 4

Similar with bag of n-gram feature, we compute the distribution of each POS tag for users in our dataset. Then, we compute the Pearson correlations between each POS tag and each personality trait as in Table 4. There are many interesting insights in Table 4. For example, in nominal and nominal+verbal category, common noun is a good indicator for personality. People who use common nouns more often in their language tend to be in Extroversion, Intuition, Thinking, or Judging type. Introverted people use more pronouns but less common nouns.

Interjection, which includes "lol", "haha", "FTW", "yea", is more likely to be used by people who are in Sensing and Perceiving type. Emoticon is more likely to be used by people who are in Sensing and Feeling type while numbers are more likely to be used by people who are in Sensing and Thinking type. Also, extroverted people are more likely to use hashtags. Those seemingly random online behaviors, such as use of hashtags, emoticons, and nouns can be somehow explained by the psychological traits of the users.

6 Word Vector

Distributed representations of words in a vector space help learning algorithms to achieve better performance in natural language processing tasks by grouping similar words. The word representation idea has since been applied to statistical language modeling with considerable success (Bengio et al., 2003). Recently, Mikolov et al. introduced the Skip-gram model, an efficient method for learning dense vector representations of words from large amounts of unstructured text data (Mikolov et al., 2013). The application of word representation includes automatic speech recognition and machine translation, and a wide range of NLP tasks.

In this paper, we want to explore the relationship between vector representations of words in tweets and the personality traits. In order to improve the generalization of the word vectors we use, we trained the model based on an external Twitter dataset which has hundreds of millions of tweets. This gave us word vectors of 2,334,564 words and each word has a 500 dimension distributed representation of its semantic meanings.

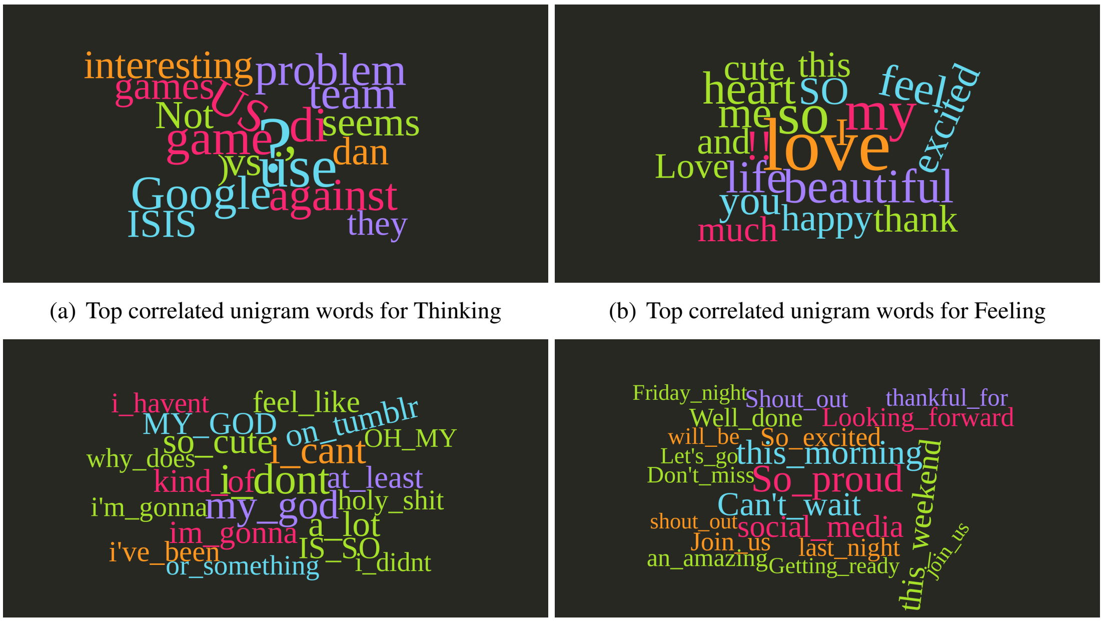
Aiming to predict the personalty traits, we need to compute a general representations of all tweets from a user based the word vectors. After removing stop words in all the tweets, we use two approaches to compute the textual representations of users listed as below.

- **Average word vectors.** We average all the vectors of all the word that is available in the tweets of a user to represent the vector representations of that user.

- **Weighted average word vectors.** Instead of average all the vectors of all the word equally, we weighted average the vectors of the words

Table 4: The description of POS tags and their correlations with each personality trait (*** $p < 0.001$). Negative values mean that POS tags are positively correlated with I, S, T, or J, while positive values mean that POS tags are positively correlated with E, N, F or P. The top 3 positively correlated and negatively correlated POS tags are bold for each personality trait.

Tag	Description	I/E	S/N	T/F	J/P
	Nominal, Nominal+Verbal				
N	common noun (NN, NNS)	**0.0474*** **	**0.0535*** **	**-0.0509*** **	**-0.0586*** **
O	pronoun (personal/WH; not possessive)	**-0.0715*** **	-0.0191***	**0.1014*** **	0.0350***
ˆ	proper noun (NNP, NNPS)	-0.0014	-0.0193***	**-0.0788*** **	**0.0468*** **
S	nominal + possessive	0.0179***	0.0205***	0.0064	-0.0406***
Z	proper noun + possessive	0.0260***	0.0202***	-0.0156***	-0.0357***
	Other open-class words				
V	verb incl. copula, auxiliaries (V*, MD)	-0.0299***	0.0081	0.0534***	-0.0142***
A	adjective (J*)	0.0037	**0.0440*** **	0.0340***	-0.0388***
R	adverb (R*, WRB)	**-0.0767*** **	-0.0121***	0.0702***	0.0145***
!	interjection (UH)	-0.0458***	**-0.0592*** **	0.0274***	**0.0639*** **
	Other closed-class words				
D	determiner (WDT, DT, WP$, PRP$)	0.0204***	0.0416***	0.0401***	-0.0519***
P	pre- or postposition, or subordinating conjunction (IN, TO)	**0.0541*** **	**0.0492*** **	0.0128	**-0.0761*** **
&	coordinating conjunction (CC)	-0.0381***	0.0146***	**0.0763*** **	-0.0090
T	verb particle (RP)	0.0197***	-0.0168***	0.0306***	-0.0204***
X	existential there, predeterminers (EX, PDT)	-0.0208***	0.0228***	0.0243***	-0.0108
	Twitter/online-specific				
#	hashtag (indicates topic/category for tweet)	**0.0912*** **	-0.0305***	0.0225***	-0.0252***
@	at-mention (indicates another user as a recipient of a tweet)	0.0082	-0.0080	-0.0143***	0.0092
˜	discourse marker, indications of continuation of a message across multiple tweets	-0.0324***	0.0034	0.0130***	0.0276***
U	URL or email address	0.0432***	0.0005	-0.0038	0.0036
E	emoticon	-0.0139***	**-0.0546*** **	**0.0744*** **	0.0025
	Miscellaneous				
$	numeral (CD)	0.0283***	**-0.0502*** **	**-0.0461*** **	-0.0031
,	punctuation	0.0455***	0.0365***	-0.0377***	**-0.0741*** **
G	other abbreviations, foreign words, possessive endings, symbols, garbage (FW, POS, SYM, LS)	-0.0423***	-0.0197***	-0.0297***	**0.0518*** **
	Other Compounds				
L	nominal + verbal; verbal + nominal	**-0.0808*** **	-0.0268***	0.0632***	0.0417***
M	proper noun + verbal	0.0042	-0.0064	0.0028	0.0035
Y	X + verbal	-0.0256***	0.0002	0.0046	0.0172***



(a) Top correlated unigram words for Thinking

(b) Top correlated unigram words for Feeling

(c) Top correlated bigram phrases for Introversion

(d) Top correlated bigram phrases for Extroversion

Figure 1: Correlations between n-gram words and phrases, and the personality traits

that is available in the tweets of a user according to the TF-IDF values. The weighted vector representation is then used to represent the vector representations of that user.

7 Predicting Personality

In this section we discuss our models aimed at predicting MBTI personality traits based on the features discussed in the previous sections. We develop different prediction models using both individual features and combined features. For combined features, we concatenate features within or across categories to test the predictive power of our models. Also, results presented here were obtained using a Logistic Regression model with 10-fold cross-validation; other machine learning models, such as Random Forest and SVM, produced similar results. Since the distribution of people with different personalities is skewed, as shown in Table 3, the prediction performance is reported in terms of area under the receiver operating characteristic curve (AUC) which reflects the probability of correctly classifying two randomly selected users from the two classes of dichotomous variables (i.e. extrovert and introvert).

The two approaches we proposed to learn the vector representation of users based on word vectors turn out to have similar performance in predicting personalities traits. Therefore, in the following experiment, we use average word vectors as word-vector based feature.

7.1 Prediction Accuracy

7.1.1 Performance of Individual Features and Feature Sets

Prediction accuracies achieved using both individual features and combined feature sets are presented in Table 5.

The highest accuracy among individual features (AUC=0.651) was achieved for 500 dimension vector representations of users based on the word vector. This is a remarkable performance given that those word vectors were unsupervisedly trained on a external Twitter dataset Likes and most of words are only weakly correlated with the personality traits. This result nicely illustrates the potential of predictions based on social media in general and languages in. Even the relatively weak signal aggregated over many observations (e.g. many words) results in a good prediction performance.

Unigram (AUC=0.597), bigram (AUC=0.590), and trigram (AUC=0.586) were also reasonably predictive of the personality traits. Combining those n-gram features boosts the AUC to 0.607. Beside, it is worth noting that the length of the n-gram feature also affect the prediction results. As trigram contains a sequences of three words, it introduces more noise into the feature and therefore

Table 5: Accuracy of predicting personality traits in terms of AUC

Individual Features & Feature Sets					
Feature	**I/E**	**S/N**	**T/F**	**J/P**	**Average AUC**
Word vector	0.679	0.643	0.673	0.608	0.651
Bag of n-grams	0.631	0.588	0.621	0.588	0.607
Unigram	0.617	0.581	0.609	0.582	0.597
Bigram	0.609	0.569	0.607	0.573	0.590
Trigram	0.613	0.567	0.593	0.570	0.586
POS tag	0.593	0.575	0.603	0.569	0.585
Combined Features					
Feature	**I/E**	**S/N**	**T/F**	**J/P**	**Average AUC**
POS + n-grams	0.628	0.607	0.633	0.596	0.616
POS + n-grams + word vector	0.691	0.653	0.680	0.619	0.661

preforms worse in the prediction comparing to bigram and unigram.

POS tags give the lowest predictive accuracy (AUC=0.585) among the three feature categories. POS tags convert all the tweets of a users into a distribution of 25 tags, much useful information might be lost during this process which leads to lowest but moderate performance given the random baseline is AUC=0.5.

Another interesting observation is that language is good at predicting I/E and T/F than S/N and J/P. This result is also showed in our previous analysis (Table 4, Figure 1)

7.1.2 Performance of Combined Features

In practice, features belonging to separate categories are often combined to maximize the model performance. The benefits of combining multiple features can be clearly seen in the results included at the bottom of Table 5. The model combining POS tag, Bag of n-grams, and word vector turns out to be very predictive of the participation ($AUC = 0.661$). Removing word vector from the feature set significantly decreases the predictive power (AUC=0.616) proving that the word vector is a very import feature that contains richer information about the personality traits than other two features (as shown in performance of individual features).

8 Related Work

Social media sites are now the most popular destination for Internet users, providing social scientists with a great opportunity to understand online behavior. There are a growing number of research papers related to social media, a small number of which focus on Analyzing and predicting personality that has recently attracted more and more attention in research community. Golbeck et al. used tweets on Twitter to extract Linguistic Inquiry and Word Count (LIWC) features, MRC language features, Twitter use, structural, and sentiment features to predict the personality traits (Golbeck et al., 2011). Sumner et al. explored the extent to which it is possible to determine antisocial personality traits based on Twitter use (Sumner et al., 2012). Quercia et al. studied the relationship between personality traits and five types of Twitter users: listeners, popular, highly-read, and two types of influentials (Quercia et al., 2011). These studies, however, were mostly based a small samples of hundreds or thousands of users that could be prone to bias in the dataset.

Schwartz et al. analyzed words, phrases, and topic instances collected from the Facebook messages of 75,000 volunteers, who also took standard personality tests, and found striking variations in language with personality, gender, and age (Schwartz et al., 2013). However, this work only looked at the general language use instead of online-specific language use which ignores an important part of online behaviors. Also, this work has not investigate the relationship between word vector features and personalities which are proved to be the best predictive feature in our analysis.

9 Conclusion & Discussion

In this paper, we study the relationship between human language on Twitter and personality traits. Specifically, we want to know how linguistic features correlate with each personality trait and to what extent can we predict personality traits from language. We gather personality data from Myers-Briggs personality test which contains thinking, feeling, sensation, intuition, introversion, extroversion, judging and perceiving. Also, we collect 200 most recent tweets from users with personality values. We design three categories of feature, namely bag of n-grams, Twitter POS tags, and word vectors. Analysis of these features provide insights of language use for different personalities. For instances, extroverts tend to use hashtag and phrases like "so proud", "so excited", and "can't wait". Moreover, we investigate the predictive power of individual features and combined features in our analysis. With the concatenation of all the features we extracted, we can predict the personality traits with an average AUC of 0.661.

Social media is one of the most frequent destination for internet users. Inferring the personality traits of users in social media not only helps us understand their online behaviors, but also gives us the information to provide better personalized services and improve the product. However, predicting personality can also lead to privacy issue that expose the psychological details of online users to the public. These reasons make this study more important that it tries to understand personality traits from social media and explores the degree to which we can predict personality traits simply using language on social media.

Acknowledgments

This project is a joint work with Dr. David Jurgens. The author would like to thank Dr. David Jurgens and TAs for valuable comments and discussions.

References

Yoshua Bengio, Réjean Ducharme, Pascal Vincent, and Christian Janvin. 2003. A neural probabilistic language model. *The Journal of Machine Learning Research*, 3:1137–1155.

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*.

Kevin Gimpel, Nathan Schneider, Brendan O'Connor, Dipanjan Das, Daniel Mills, Jacob Eisenstein, Michael Heilman, Dani Yogatama, Jeffrey Flanigan, and Noah A Smith. 2011. Part-of-speech tagging for twitter: Annotation, features, and experiments. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers-Volume 2*, pages 42–47. Association for Computational Linguistics.

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 Inernational Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on*, pages 149–156. IEEE.

Michal Kosinski, David Stillwell, and Thore Graepel. 2013. Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, 110(15):5802–5805.

Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Isabel Briggs Myers. 1962. The myers-briggs type indicator: Manual (1962).

Daniele Quercia, Michal Kosinski, David Stillwell, and Jon Crowcroft. 2011. Our twitter profiles, our selves: Predicting personality with twitter. In *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on*, pages 180–185. IEEE.

David Rawlings and Vera Ciancarelli. 1997. Music preference and the five-factor model of the neo personality inventory. *Psychology of Music*, 25(2):120–132.

H Andrew Schwartz, Johannes C Eichstaedt, Margaret L Kern, Lukasz Dziurzynski, Stephanie M Ramones, Megha Agrawal, Achal Shah, Michal Kosinski, David Stillwell, Martin EP Seligman, et al. 2013. Personality, gender, and age in the language of social media: The open-vocabulary approach. *PloS one*, 8(9):e73791.

Maarten Selfhout, William Burk, Susan Branje, Jaap Denissen, Marcel Van Aken, and Wim Meeus. 2010. Emerging late adolescent friendship networks and big five personality traits: A social network approach. *Journal of personality*, 78(2):509–538.

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.