Personality Detection

Group: 52

Problem statement

To automatically classify the personality traits of the author based on his/her posts in social media.

It is evident that there is a strong correlation between user's personality and the user's status text on the social media. Human personality can be summarized by five personality traits as mentioned below.

Extrovert vs. Introvert (sociable, assertive, playful vs. aloof, reserved, shy)

Emotional stability vs.Neuroticism (calm, unemotional vs. insecure, anxious)

Agreeableness vs. Disagreeable (friendly, cooperative vs. antagonistic, fault finding)

Conscientiousness vs unconscientiousness (self-disciplined, organized vs. inefficient, careless)

Openness to experience (intellectual, insightful vs. shallow, unimaginative)

Hence we represent human personality as a vector of the above five traits with boolean values (i.e yes/no)

Introduction

In recent years the interest of the scientific community towards Personality Recognition has grown incredibly, since there are many applications that can take advantage of personality recognition, including social network analysis, social computing, recommendation systems, deception detection, authorship attribution, sentiment analysis/opinion mining, and others.

The most difficult thing here is collecting the facebook data and how to classify the person. Most of the facebook users are not ready to give away their posts. But thanks to mypersonality.org, they have collected data from various people and they provide two two gold standard labeled datasets: Essays and MyPersonality. Essays is a large dataset of streamofconsciousness texts (about 2400, one for each user), collected between 1997 and 2004 and labeled with personality classes. This dataset can be used for identifying the feature set of the personality traits. MyPersonality consists of about 10000 Facebook status updates of 250 users, plus Facebook network properties (including network size, betweenness centrality, density and transitivity) labeled wxith personality. This dataset can be divided into training data and test data.

As said before, the main goal of the project is to automatically classify the personality traits of the author based on his/her posts in social media. Personality traits are commonly described using five dimensions (known as the Big Five), namely extraversion, neuroticism (the opposite of emotional stability), agreeable-ness, conscientiousness, and openness to experience. Since more than one trait can be present in the same user, for each trait we train a binary classifier that separates the users displaying the trait from those who do not. We use a variety of features as input for the classifiers, including features related to the text that users use in their status updates, features about the user's social network, time-related factors, etc.

Approach

Our goal is to predict these traits for a given user,

where we identify a user with his set of available status updates (treated together as one text per user when extracting linguistic features), their time stamps, and his social network properties.
Feature selection for each of the traits: Various features were extracted and used. The feature set of the trait is heavily accountable for the correctness of the classifier we build.Social Network features

- **Social network features:**

7 features related to the social network of the user: (1) network size, (2) betweenness,(3) nbetweenness, (4) density, (5) brokerage, (6) nbrokerage, and (7) transitivity.

- **Time-related features**

7 features related to the time of the status updates (we assume that all the times are based on one time zone): (1) frequency of status updates per day, (2) number of statuses posted between 6-12 am, (3) number of statuses posted between 12-18, (4)number of statuses posted between 18-24, (5) number of statuses posted between 00-06, and (6)

number of statuses posted on weekdays (7) number of statuses posted on weekends

- **Status statistics features**

4 features related to the statuses of the user . (1)total number status updates of a user , (2) average number of words in a status , (3) average number of characters in a status , (4) number of words with more than 6 letters

- **Language based features**

17 features related to language
(1) First-person singular words , (2) First-person plural (3)second person ,(4) Third person singular , (5) Third person plurall ,(6)Indefinite pronouns , (7)Articles ,(8)Common verbs ,(9) Auxillary Verbs , (10) Future tense ,(11) Adverbs ,(12)prepositions ,(13) conjunctions ,(14) negations ,(15) quantifiers , (16) swear words used 17) Number of names of the persons used .

- **Psychology features**

13 features from the below categories which include the emotional sent of the user .

i)Affective features ;

(1) Anxiety words ,(2) Positive emotion words , (3) negative emotion words , (4) Anger words , (5) sadness (6) Happiness words

ii) Social features :

(1) Words related to family and friends

iii) Biological processes :

(1)Body (2) health

iv)Personal concerns :

(1)work (2)achievement (3) money (4) religion (5)death

- **Unigrams & Bigrams based:** A set of unigrams specific to each of the traits can be predetermined based on some metric and these unigrams can be included in the feature set for the trait. These unigrams can be extracted from the facebook data along with essays dataset. To do this, first we calculate the word count for each word in the entire corpus and for each trait separately. Now we have 6 different word counts for each word.A word is related more to a trait if, it occurs more in the statuses related to the persons with that particular trait as well as the word occurs less in the statuses of the person not related to that same particular trait. For a word W if x is the word count for the trait t, and y is the totlal word count. Then, the word W is more related to the trait t if x/(y-x) is huge and x is more than a threshold value (both x and y could be less. So x > th=reshold number). Synonyms for each of the unigrams are also handled .

- **Smileys based:** 30-40 smileys are classified into 3 categories (positive, negative and neutral feelings). Number of smileys in each of the category

- **Complexity based:** A feature which stores the complexity of the words that were used by the user. A complexity of a word is pre-calculated based of some factors and each word has a complexity value.

- **Meta features :** 4 features based on the writing style of the user are calculated (1) number of punctuations used , (2) Number of words in the statuses , (3) average number of words in statuses (4) numbers used

- **Other features :**

6 features which are not included in the above categories (1) total number of statuses per user, (2)number of capitalized words, (3) number of capital letters, (4) number of words that are used more than once,(5) number of urls, and (6) number of occurrences of the string PROPNAME — a string used in the data to replace proper names of persons for anonymisation purposes

Classification of the traits:

We built a different classifier model for each of the 5 personality traits from the training data which distinctively classifies a user into 2 subclasses yes/no assigning a value to that trait in the personality vector of the user using SVM.

Training: Each user has one or more posts. All these posts are combined and a feature vector for each user is produced. Now there are feature vectors and the category (for each trait, y/n for each trait) for each user is labelled. Using this data, input is sent to SVM and a classifier is built. This is done five times for each trait. So, now there are five classifiers.We use facebook status updates of abbout 180 users to generate the training data set

Testing: In the similar manner the users in the test data are mapped to a feature vector. Now there are the feature vectors, their category(for each trait, y/n for each trait) and classifier for each of the five traits. Now the data is sent to SVM again for testing and the results based on the classifier is obtained. Now we calculate the Precision and Recall from the results produced by the SVM and the original classes the users belong to which we have. Testing is done on facebook statuses of 70 users .

We compare the results of the svm calssifier taking various possible combination of the features .we give the score as a weighted average of precission and recall .

		Predicted Class	
		YES	NO
Actual Class	YES	a	b
	NO	c	d

Table 1: Binary classification confusing matrix

The weights are $W_{yes} = (a+b) / (a+b+c+d)$, $W_{no} = (c+d)/(a+b+c+d)$. Precission for yes class is $P_{yes}= a/(a+c)$ and $P_{no}=a/(a+b)$. Now overall precission

$P= W_{yes}*P_{yes} + W_{no}*P_{no}$ and similarly recall R is calculated . F measure = $2* ((P*R)/(P+R))$

We use th above defind precission and recall as the distribution of YES and NO in the dataset is not equal .

Classification Results :

Value	EXT	NEU	AGR	CON	OPN
yes	96	99	134	130	176
no	154	151	116	120	74

Table 2: Distribution of personality types in Facebook data

We have a total of about 165 features belonging to 10 categories of feature sets . Of these features we take combinations to see which feature set gives resutls for a trait .

For each set of features a classifier is built for each of the traits .

Features	**cEXT**	**cNEU**	**cAGR**	**cCON**	**cOPN**
Time	0.53 0.58 61.42 0.559	0.425 0.264 60 0.322	0.61 0.53 51.5 0.57	0.55 0.786 54.28 0.652	0.763 0.763 78.57 0.763
Network	0.329 0.51 35.71 0.39	0.372 0.40 44.2 0.388	0.605 0.484 48.57 0.50	0.559 0.786 47 0.65	0.779 0.54 45.714 0.69
Unigrams and Bigrams	0.53 0.58 61.42 0.559	0.438 0.30 54.28 0.354	0.61 0.48 50 0.54	0.61 0.681 51.42 0.635	0.862 0.818 65.714 0.889
Language	0.55 0.406 62.857 0.459	0.35 0.157 54.28 0.207	0.619 0.796 54.28 0.683	0.572 0.707 47.14 0.624	0.877 0.89 71.428 0.830
Social +Affective	0.55 0.08 62.85 0.124	0.216 0.085 54.87 0.108	0.615 0.716 52.85 0.67	0.54 0.77 44.28 0.63	0.85 1 78.57 0.83
Biology + personal	0.478 0.86 40 0.643	0.487 0.85 40 0.633	0.766 0.205 47 0.277	0.35 0.152 40 0.189	0.829 0.963 75.714 0.8217
Complexity + meta + smileys	0.685 0.243 65.7 0.34	58.57	0.518 0.436 41.42 0.472	0.628 0.836 58.57 0.74	0.885 1 78.57 0.88
others	0.423 0.114 61.8 0.148	68	0.62 0.819 55.714 0.719	0.64 0.813 58.57 0.71	0.855 1 78.57 0.83

Table3 precission , recall , accuracy , fmeasure . for each of the traits in the order .

To improve performance, we iteratively combine different sets of features as presented in Table 3

. We start with the pair that has a value greater than 0.5 for
all 3 measures. If there is more than one pair, we choose the one with better results and fewer features. At each round, we add all the unused feature sets one by one to create new combinations. Among these, we choose the combination that has the highest performance and we compare it to the previous round. This process stops when performance of the current round does not improve w.r.t the previous one.

Trait	Features	Measures
cEXT	unigrams + network	0.671 0.54 70 0.581
cNEU	Affective + Lanuguage + network + other + Time	0.617 0.97 60 0.753
cAGR	Language + affective + social +smileys	0.619 0.82 55.714 0.703
cCON	Language + Biology + Complexity + smileys	0.617 0.80 52 0.705
cOPN	Social+ affective + biology + personal + Complexity + smileys	0.739 0.91 62.857
		0.834

Table 4 : Final combinations of the features that gave best performance for each of the traits .

To assess which features are important to predict the personality traits, we use Pearson's correlation of a feature and the trait score .

Features	EXT	NEU	AGR	CON	OPN
Network size	0.14	0.001	-0.06	-0.044	0.067
Betweenness	0.139	-0.055	-0.016	-0.103	0.123
N Betweenness	0.109	-0.077	0.033	-0.099	0.106
Brokerage	-0.11	-0.077	0.031	-0.101	0.108
N Brokerage	-0.109	-0.076	-0.03	-0.09	0.031
Transitivity	-0.12	-0.07	0.015	-0.101	0.113

Table 5 : Correlation results between network features and traits .

Time Features :

We also observed correlation of time features with the personlaity traits . Conscientiousness users showed a negative correlation(-0.12) and with the time of posts from 0-6 AM and extroverts showed a positive correlation(0.17) with the time of statuses from 6-12 AM .

Social Features :

Conscientiousness users showed a negative correlation (-0.11) with the statuses related to family and friends and openness users showed a positive correlation (0.08) with the same feature .

Biology and Personal Features :

Extroverts showed a postive correlation (0.11) with work features and conscientiousness users showed a negative correlation(-0.13) with health features

Affective Features:

Extroverts and agreaable users showed positive correlations with the positive-words . Neuoritic users howed a positive correlation (0.12)with swear words .

Language Features :

Agreeable users showed a positive correlation (0.17) with the third person nouns used . Neurotics have a positive correlation (0.11)with anger words and *PROPNAME* count(0.09) a negative correlation (-0.9)with positive-words and greeting words . Conscientiousness users showed a negative correlation(-0.9) and third person nouns (0.15)with negative – words and also with third person pronouns (0.1) and second person pronouns (0.12) . Agreeable users showed a positive correlation with third person nouns (0.13)

Conclusions

We had some interesting findings on a set of 250 users and 9917 status updates. Even with a fairly small set of training examples we can perform well, hence Facebook status updates do contain important cues of their author's personality types. There is no single kind of features that gives the best results for all personality traits. Advantages of this are that training examples from different social media platforms can be used in combination to train more accurate models and that such models are also applicable on social network sites for which no training data is available. Aside from the work we have presented in this report, there is clear potential in more fine grained feature selection to improve the classification results.