

# **REPORT**

## **Group 52: Personality Recognition**

### **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 gold standard labeled datasets: Essays and MyPersonality. Essays is a large dataset of stream-of-consciousness 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 with 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

### *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.

**Unigrams 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 total 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  $x > \text{threshold number}$ ).

**Network based:** Network features like the number for friends, etc are also used which is provided in the MyPersonality data set.

**Timings based:** Features are extracted based on the time in which he posted the post. The day is divided into 4 parts (each part 6hrs). If the user posted a post in

6pm-10pm twice, then the feature corresponding to this time is of value 2.

**Smileys based:** 30-40 smileys are classified into 3 categories (positive, negative and neutral feelings). If the user uses 5 smileys belonging to positive category, then the value of the feature related to positive smiley is 5.

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

**Others:** Various other things like number of words in caps, letters in caps, superlative words used, punctuations.

### *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.

**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.

## **Experiments**

## **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.