# Unsupervised Personality Recognition for Social Network Sites

Fabio Celli
*CLIC-CIMeC*
*University of Trento*
*Italy*
*fabio.celli@unitn.it*

***Abstract*—In this paper, we present a system for personality recognition that exploits linguistic cues and does not require supervision for evaluation. We run the system on a dataset sampled from a popular Social Network: FriendFeed. We adopted the five classes from the standard model known in psychology as the "Big Five": extraversion, emotional stability, agreeableness, conscientiousness and openness to experience. Making use of the linguistic features associated with those classes the system generates one personality model for each user. The system then evaluates the models by comparing all the posts of one single user (users that have only one post are discarded). As evaluation measures the system provides accuracy (measure of the reliability of the personality model) and validity (measure of the variability of writing style of a user). The analysis of a sample of 748 Italian users of FriendFeed showed that that the most frequent personality type is represented by the model of an extravert, insecure, agreeable, organized and unimaginative person.**

***Keywords-Social Network Sites; Personality Recognition; Information Extraction; Natural Language Processing*.**

## I. Introduction and Related Work

Personality is a crucial aspect of social interaction. Under the computational perspective it can be very useful for marketing and for interesting tasks such as stylometry and sentiment analysis. Recent studies showed that there is a connection between the personality of individual users and their behavior online (see Amichai-Hamburger and Vinitzky [1]). Social Network Sites (SNSs henceforth, see Boyd and Ellison [2] for definitions and history) are huge, virtually infinite, corpora where authors (users) and sentences (posts) are found together. Many scholars used data from social networks for personality classification. In 2006 a pioneering work by Oberlander et al. classified four traits of blog authors' personality using n-grams as features. Some very recent works such as Quercia et al. [10] and Golbeck et al. [4] predicted personality of users from social network data. In particular Golbeck et al. predicted personality from some users' profiles on Facebook using machine learning techniques. Golbeck's work is supervised because it required that subjects completed a personality test for evaluation. Here we introduce a novel technique for personality recognition that does not require subjects.

In the following section, we will present a system that builds on the fly one personality model for each user in a corpus in an unsupervised way and performs automatic evaluation of the models comparing all of his/her posts. Then, in Section 3, we will present the results of the analysis of personality on FriendFeed. In Section 4, we will conclude introducing possible directions for future works.

## II. Unsupervised Personality Recognition

The large amount of data available from Social Network Sites allows us to predict users' personality from text in a computational way, but there are at least four nontrivial problems:

1) The definition of personality, which is a very fuzzy and subjective notion;
2) The annotation of personality in the data from SNSs, that would require personality judgements by the author themselves or by other native speakers.
3) The construction of one model for each user in the dataset.
4) The evaluation of personality models.

In the next paragraphs we are going to discuss the solutions for those problems we adopted for the unsupervised personality recognition system.

### A. Definition of Personality

Psychologists describe personality along five dimensions known as the "Big Five" (see Goldberg [5]), a model introduced by Norman in 1963 [8], obtained from factor analysis of personality description questionnaires that has become a standard over the years. The five dimentions are the following:

- Extraversion (E) (sociable vs shy)
- Emotional stability (S) (calm vs insecure)
- Agreeableness (A) (friendly vs uncooperative)
- Conscientiousness (C) (organized vs careless)
- Openness (O) (insightful vs unimaginative)

Those dimentions can be represented computationally as continuous numerical variables with 2 poles: one positive

 ISBN: 978-1-61208-176-2 

(1) and one negative (0). Once we have the numerical values for each attribute (one attribute is one dimention in the "Big Five"), we can easily calculate whether a user has one trait of personality (y) or not (n) or we have no information about that trait (o). From this representation, we can formalize a personality model for each user simply taking the majority class for each attribute/dimension from all posts the user made. In the end personality models are formalized as string of five charachters: one for each attribute, which one can take three possible vaues: positive (y), negative (n) or balanced (o). For example a the string `ynooy` is the model of an extravert, nervous and open-minded user.

### *B. Dataset*

The dataset is a sample of 748 Italian FriendFeed users (1065 posts). It is a subset of the dataset sampled by Celli et al. [3]. The dataset has been collected from FriendFeed public URL, where new posts are publicly available. The dataset was already processed with a language identifier, whose performance is correct at 88%. This made easier the extraction of the Italian subset.

Our unsupervised system does not require direct annotation of the dataset, but just a set of correlations between linguistic factors and personality traits to build models. Either Mairesse et al., Golbek et al. and Quercia et al. report sets of correlations between some cues and the dimentions of personality in the "Big Five". In our system we used a set taken from Mairesse et al. because it is the largest one and it is more focused on linguistics.

### *C. Building the Personality Model*

Mairesse et al. provides a long list of correlation coefficients between linguistic factors and the personality traits. These coefficients are obtained from an essay corpus where authors and external observers provided personality judgements following the "Big Five" model. In order to develop an unsupervised personality recognition system we need to turn those coefficients into features that can be automatically extracted from text. Among those linguistic factors that correlates with certain aspects of personality there are some regarding topic (for example if a person writes about job, leisure, music, other people), some regarding word usage (for example the frequency of words used, the use of negative particles, first person pronouns, fillers, swares) and some regarding psychological aspects (for example age of acquisition of the word used, length of the words used, expression of positive and negative feelings). Factors are supposed to be valid for the western culture. We picked up and adapted 22 features from Mairesse et al. They are:

1) **all punctuation** (ap): the count of `. , ; :` in the post,
2) **commas** (cm): the count of `,` in the post,
3) **reference to other users** (du): the count of the pattern `@` in the post,
4) **exclamation marks** (em): the count of `!` in the post,
5) **external links** (el): the count of external links in the post,
6) **first person singular pronouns** (im): the number of first person singular pronouns in the post,
7) **negative particles** (np): the count of negative particles in the post,
8) **negative emotions** (ne): the count of emoticons expressing negative feelings in the post,
9) **numbers** (nb): the count of numbers in the post,
10) **parenthesis** (pa): the count of parenthetical phrases in the post,
11) **positive emotions** (pe): the count of emoticons expressing positive feelings in the post,
12) **prepositions** (pp): the count of prepositions in the post,
13) **pronouns** (pr): the count of pronouns in the post,
14) **question marks** (qm): the count of `?` in the post,
15) **long words** (sl): the count of words longer than 6 letters in the post,
16) **self reference** (sr): the count of first person (singular and plural) pronouns in the post,
17) **swears** (sw): total count of vulgar expressions in the post,
18) **type/token ratio** (tt): defined in the formula below,
19) **word count** (wc): words in the post,
20) **first person plural pronouns** (we): count of first person plural pronouns in the post,
21) **second person singular pronouns** (yu): count of second person singular pronouns in the post,
22) **mean word frequency** (mf): simple mean of the frequency of words in the post, defined in the formula below.

$$tt = \frac{w - T}{T} \quad mf = \frac{\Sigma wf}{T}$$

where $w$ is the count of words already used in the sentence, $T$ is the total word count in the sentence and $wf$ is the frequency count of the word in the dataset. Table I (from Mairesse et al.) shows how the linguistic features used correlate with personality traits. First the system extracts a random sample of the dataset for statistical purposes. The size of the sample can be decided a-priori, in this case we sampled 500 posts. From this sample the system extracts mean and standard deviation for each feature. The mean word frequency (feature mf) in this case is calculated using an external corpus of Italian (CORISsmall, see [11]) but in principle it can be calculated also from the dataset itself as relative frequency. Results are summarized in Table II. In the second step the system processes the entire dataset building a personality model for each post applying the following rules: if a sentence shows a feature correlating positively with one personality trait and the frequency of that feature is higher than mean plus standard deviation for that feature,

 ISBN: 978-1-61208-176-2 

| F. | E | S | A | C | O |
|---|---|---|---|---|---|
| ap | -.08** | -.04 | -.01 | -.04 | -10** |
| cm | -.02 | .01 | -.02 | -.01 | .10** |
| du | -.07** | .02 | .01 | .01 | .06** |
| el | -.05* | -.02 | -.01 | -.03 | .09** |
| em | -.00 | -.05* | .06** | .00 | -.03 |
| in | -.04* | .01 | -.01 | -.03 | -.01 |
| im | .05* | -.15** | .05* | .04 | -.14** |
| np | -.08** | .12** | .11** | -.07** | .01 |
| ne | -.03 | -.18** | -.11** | -.11** | .04 |
| nb | -.03 | .05* | -.03 | -.02 | -.06** |
| pa | -.06** | .03 | -.04* | -.01 | .10** |
| pe | .07** | .07** | .05* | .02 | .02 |
| pp | .00 | .06** | .04 | .08** | -.04 |
| pr | .07** | .12** | .04* | .02 | -.06** |
| qm | -.06** | -.05* | -.04 | -.06** | .08** |
| sr | .07** | -.14** | -.06** | -.04 | -.14** |
| sl | -.06** | .06** | -.05* | .02 | .10** |
| sw | -.01 | .00 | -.14** | -.11** | .08** |
| tt | -.05** | .10** | -.04* | -.05* | .09** |
| wc | -.01 | .02 | .02 | -.02 | .06** |
| we | .06** | .07** | .04* | .01 | .04 |
| yu | -.01 | .03 | -.06** | -.04* | .11** |
| mf | .05* | -.06** | .03 | .06** | -.07** |

Table I
FEATURES USED IN THE SYSTEM AND THEIR PEARSON'S CORRELATION COEFFICIENTS WITH PERSONALITY TRAITS AS REPORTED IN MAIRESSE ET AL. 2007. * = $p$ SMALLER THAN .05 (WEAK CORRELATION), ** = $p$ SMALLER THAN .01 (STRONG CORRELATION)

| feature | mean | sd | min | max |
|---|---|---|---|---|
| ap | 1 | 2 | 0 | 28 |
| cm | 0 | 1 | 0 | 19 |
| du | 0 | 0 | 0 | 3 |
| el | 0 | 0 | 0 | 3 |
| em | 0 | 0 | 0 | 7 |
| im | 0 | 0 | 0 | 3 |
| np | 0 | 0 | 0 | 4 |
| ne | 0 | 0 | 0 | 1 |
| nb | 1 | 4 | 0 | 64 |
| pa | 0 | 0 | 0 | 3 |
| pe | 0 | 0 | 0 | 2 |
| pp | 1 | 2 | 0 | 32 |
| pr | 0 | 0 | 0 | 8 |
| qm | 0 | 0 | 0 | 3 |
| sr | 0 | 0 | 0 | 4 |
| sl | 6 | 6 | 0 | 71 |
| sw | 0 | 0 | 0 | 1 |
| tt | 0.971 | 0.048 | 0.706 | 1 |
| wc | 7 | 7 | 1 | 79 |
| we | 0 | 0 | 0 | 2 |
| yu | 0 | 0 | 0 | 2 |
| mf | 101264 | 87192 | 68 | 567704 |

Table II
SUMMARY OF THE BEHAVIOR OF FEATURES ASSOCIATED TO PERSONALITY TRAITS IN THE DATASET.

then the system increases the score of that personality trait. If a sentence shows a feature whose frequency is higher than mean plus standard deviation and it correlates negatively with one personality trait, the system decrease the score of that personality trait. Then numerical values are turned into nominal ones ("y","n" and "o") simply checking if a value is positive, negative or it is zero. In the end the majority class of each personality trait is calculated for each user and the resulting string is taken as the user's personality model.

### D. Evaluation of Personality Models

The evaluation method is based on the assumption that one user has one and only one personality and that this personality emerges at different degrees from user's posts. Hence the system evaluates the personality model comparing many posts of the same user. The drawback of this method is that we can only evaluate models for users that have more than one post in the dataset, and we have to discard all the other users.

The unsupervised system takes all the models built from the posts of a user and compares each value of the string. This evaluation method provides two measures, accuracy ($a$) and validity ($v$), defined in the formulas below:

$$a = \frac{tp + tn}{tp + tn + fp + fn} \qquad v = 1 - \frac{a}{P}$$

where $P$ is the number of posts of one user; $tp$ is the sum of each personality attribute matching within the same user (for example "y" and"y", "n" and "n", "o" and "o"); $tn$ is the sum of opposite attributes within the same user ("y" and "n", "n" and "y"); $fp$ is the sum of possible attributes turned to the balance value within the same user ("y" to "o" and "n" and "o") and $fn$ is the sum of the zero attributes turned to positive or negative ("o" to "y" and "o" and "n"). Accuracy gives a measure of the reliability of the personality model and validity gives information about how much the model is valid for all the user's posts, in other words how much the user writes expressing the same personality traits. A low validity score means that the user shows variability in his/her writing style.

## III. ANALYSIS AND DISCUSSION

We filtered out group posts (because many users with different personalities can post in a group) and kept only single users from the dataset. Most users (592) have just one post and the models obtained from those users were not considered reliable (accuracy is set to 0). Excluding the users with only one post the average accuracy is 0.631 and the average validity is 0.729. Accuracy is in line with the classification accuracies reported by Mairesse et al. 2007 for observer ratings evaluation. This fact is very encouraging because it is a clue that we developed a system that implements Mairesse's model completely automatically. The results of the frequency of personality models in the sample is reported in Table III. Below rank 7 models become more and more sparse, with a long tail of models appearing only once. Thay do not appear in Table III.

The most frequent personality type in the Italian subset of FriendFeed is represented by the model of an extravert, insecure, agreeable, organized and unimaginative person. It is interesting to note that the features "insecure" and

 ISBN: 978-1-61208-176-2

| Rank | Model | Rel.Freq. |
|---|---|---|
| 1 | ynyyn | 16.6% |
| 2 | ynyon | 12.1% |
| 3 | onoyn | 7.6 % |
| 4 | ooooo | 7.6% |
| 5 | ynoyn | 4.5% |
| 6 | yoooo | 4.5% |
| 7 | ynooo | 3.8% |
| 8 | ynoyo | 3.8% |
| 9 | ynoon | 3.2% |
| 10 | onyoo | 3.2% |
| 11+ | *others* | 33.1% |

Table III
FREQUENCY OF PERSONALITY MODELS.

"unimaginative" is present in the first four positions of the ranking and that no shy people is found in the first six positions. Pearson's correlation test reveal that there is a strong (+0.79) and highly significant correlation ($p$-value = .0003) between the accuracy and personality model types, meaning that there are certain personality types that express strongly and reliably their personality in written language, and others that do not. Althought there is no correlation ($p$-value = .413) between personality and posting activity, once filtered out the long tail of users with sparse personality models, emerges that there is one personality type that produces more posts than others, that is the extravert, insecure, friendly, not particularly precise and unimaginative person (`ynyon`).

A manual look to the data reveals that there are some users (the ones with higher validity) that are focused on a topic, and sometimes this topic is clear from their username: for example "styleandthecity", or such users as "ultimora" or "cronaca24", which appear to be journalists and have a very recognizable and normalized style, but not the same personality model.

## IV. CONCLUSIONS AND FUTURE WORK

In this work, we described and developed an unsupervised system for personality recognition that does not require subjects for evaluation. It exploits existing correlations between language cues and personality traits providing accuracy and validity as evaluation measures. We showed that it is possible to extract personality information from SNSs in an unsupervised way with acceptable accuracy with a process that is completely automatic. The results reported here show that the distribution of personality models in SNSs has a high peak of people sharing the same personality traits and a long tail of people with a unique personality model. Results also show that validity is a good measure of the recognizability of the style of a user.

In the future, we would like to improve the system exploiting different correlation sets. We would also like to sample and automatically annotate large corpora of Social Network data in order to facilitate the research in this field.

## ACKNOWLEDGEMENT

This work has been realized also thanks to the support from the Provincia autonoma di Trento and the Fondazione Cassa di Risparmio di Trento e Rovereto.

## REFERENCES

[1] Amichai-Hamburger, Y. and Vinitzky, G. Social network use and personality. In *Computers in Human Behavior*. 26(6). pp. 1289–1295. 2010.

[2] Boyd, D. and Ellison, N. Social Network Sites: Definition, history, and scholarship. In *Journal of Computer-Mediated Communication* 13(1). pp. 210–230. 2007.

[3] Celli, F. and Di Lascio and F.M.L. and Magnani, M. and Pacelli, and B., Rossi, L. *Social Network Data and Practices: the case of Friendfeed*. Advances in Social Computing, pp. 346–353. Series: Lecture Notes in Computer Science, Springer, Berlin. 2010.

[4] Golbeck, J. and Robles, C., and Turner, K. Predicting Personality with Social Media. In *Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems*, pp. 253–262. 2011.

[5] Goldberg, L., R. The Development of Markers for the Big Five factor Structure. In *Psychological Assessment*, 4(1). pp. 26–42. 1992.

[6] Mairesse, F., and Walker, M. Words mark the nerds: computational models of personality recognition through language. In: *Proceedings of the 28th Annual Conference of the Cognitive Science Society*. pp. 543-548. 2006.

[7] Mairesse, F. and Walker, M. A. and Mehl, M. R., and Moore, R, K. Using Linguistic Cues for the Automatic Recognition of Personality in Conversation and Text. In *Journal of Artificial intelligence Research*, 30. pp. 457–500. 2007.

[8] Norman, W., T. Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality rating. In *Journal of Abnormal and Social Psychology*, 66. pp. 574–583. 1963.

[9] Oberlander, J., and Nowson, S. Whose thumb is it anyway? classifying author personality from weblog text. In *Proceedings of the 44th Annual Meeting of the Association for Computational Linguistics ACL*. pp. 627–634. 2006.

[10] Quercia, D. and Kosinski, M. and Stillwell, D., and Crowcroft, J. Our Twitter Profiles, Our Selves: Predicting Personality with Twitter. In *Proceedings of SocialCom2011*. pp. 180–185. 2011

[11] Rossini Favretti R. and Tamburini F., and De Santis C. CORIS/CODIS: A corpus of written Italian based on a defined and a dynamic model. In A Rainbow of Corpora: Corpus Linguistics and the Languages of the World, Wilson, A., Rayson, P. and McEnery, T. (eds.), Lincom-Europa, Munich. pp: 27–38. 2002.

 ISBN: 978-1-61208-176-2