YouTube Personality Detection

19

Kaifeng Chen[1], Penghao He[1], Andrew Lin[1], Yize Zang[1],
[1]School of Information, University of Michigan, Ann Arbor, MI 48105, USA

Overview

The objective of our project is to automatically predict the personality of video bloggers, each of whom recorded a one-minute monologue in front of the camera. We propose that the personality of a person can be modeled by using both non-verbal cues (e.g. expression) and verbal cues (monologues). Similar to sentiment analysis, such assessment can be helpful in many ways, such as AI speech generation and communications. We wish to automate the process and achieve a promising accuracy comparing with using human agents.

By considering the task of personality recognition, firstly, we can help further investigate the different contribution of verbal and non-verbal content in understanding human traits. Secondly, with a benchmark of previous researches, we can improve the process accuracy and provide insight in feature selection. Thirdly, we explore the feasibility of a fully automated system in personality detection, which can be applied in many different fields, such as customer service, interview assessment, or crime detections.

Related Work

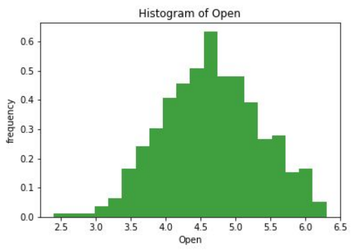
Recent research has investigated into associations between personality and individual differences in verbal usage. These research have identified the associations across different contexts, particularly text blogging, because of the abundant data available online.

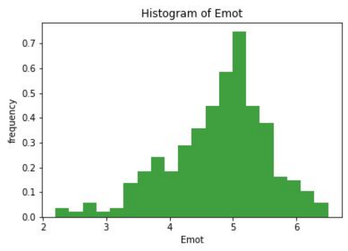
To see how personality traits correlate with verbal usage, there are two main approaches: 1) using word counts based on Linguistic Inquiry and Word Count (LIWC); 2) using n-grams to represent data. A systematic approach (2) shows that n-gram based method could outperform LIWC-based method on prediction of blogger personality. When studying the verbal differences caused by personality traits, previous studies mostly used the Big Five personality categories. In addition, verbal content (what people say) coupled with nonverbal cues (how people say it) can help improve performance of personality predictions (3).

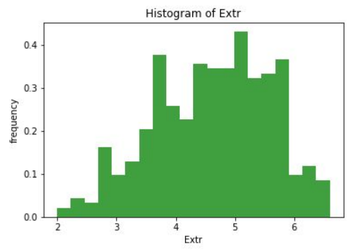
Data

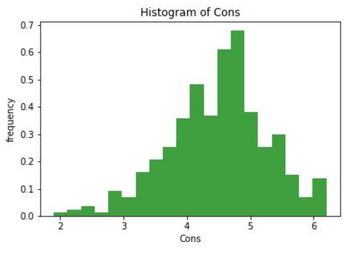
The YouTube Personality Dataset comes from Idiap Research Institute. The dataset includes video features from 404 vloggers. Each vlogger talks about a topic for one minute. Topics are from a broad range. These features are:

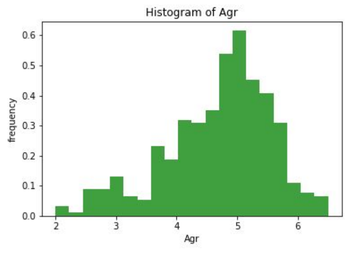
1. Nonverbal Behavioral Features:
   a. 21 features extracted from audios, including pitch, energy, peak, speaking time voice rate and number of turns;
   b. 4 features extracted from video images, including excitement and kinetic expressiveness.
2. Transcripts: Manually annotated for each video. There are around 10,000 unique words and 240,000 tokens.
3. Gender Identification: Female or male.
4. Big-Five Scores: The scores are estimated for each vlogger based on the crowdsourcing platform Amazon Mechanical Turk.[1] Scores are scaled from 1 to 7. Distributions of Big-Five Scores roughly follow Gaussian distribution.



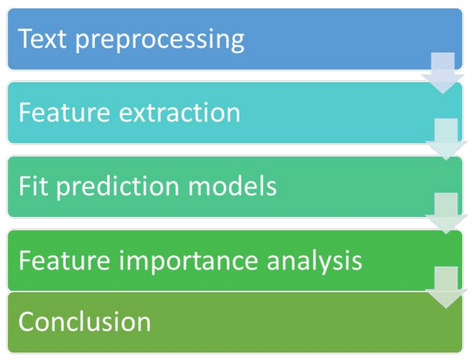






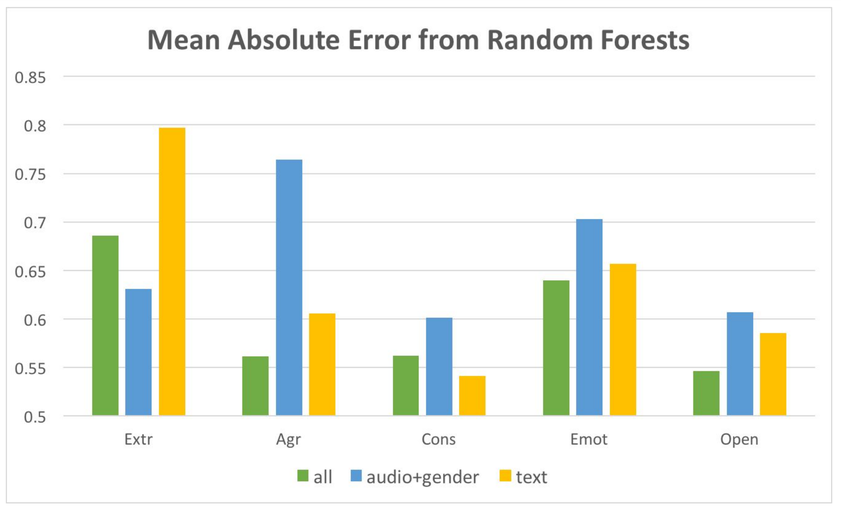


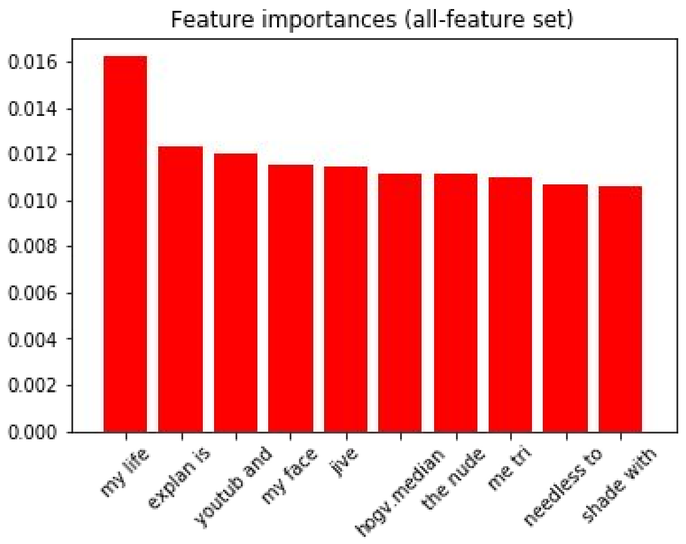
Method



1. Text preprocessing: We applied tokenization, stemming, and punctuation removal on all transcripts. We didn't remove stop words because we anticipate they are important for personalities.
2. Feature extraction: From preprocessed text, we built a n-gram language model (n=1,2).
   We constructed 3 feature sets: (1) Nonverbal features and gender identification, (2) n-grams only (3) nonverbal, n-grams features and gender.
3. Fit prediction models: We feed each feature set into three supervised learning models: Random Forest, SVM and Lasso Regression. 5-fold cross validation is applied to estimate accuracy.

Evaluation





According to our experiment results, which is shown in the table below, Random Forests performs the best compared with the other two algorithms. And if we apply all the features (including text features and non-text features), the performance is relatively the best, which demonstrates the fact that feature set that includes both text features and non-text features are the best.

	Lasso	SVM	Random Forest
All Features	0.607452	0.609007	0.595190
Non-Verbal Features + Gender	0.634724	0.632659	0.643385
Unigrams + Bigrams	0.612471	0.612015	0.606658

References

[1] Biel, Joan-Isaac, and Daniel Gatica-Perez. "The youtube lens: Crowdsourced personality impressions and audiovisual analysis of vlogs." *IEEE Transactions on Multimedia* 15.1 (2013): 41-55.
[2] Ligia Maria Batrinca, Nadia Mana, Bruno Lepri, Fabio Pianesi, and Nicu Sebe. 2011. Please, tell me about yourself: automatic personality assessment using short self-presentations. In *Proceedings of the 13th international conference on multimodal interfaces* (ICMI '11). ACM, New York, NY, USA, 255-262. DOI=http://dx.doi.org/10.1145/2070481.2070528
[3] Joan-Isaac Biel, Vagia Tsiminaki, John Dines, and Daniel Gatica-Perez. 2013. Hi YouTube!: personality impressions and verbal content in social video. In *Proceedings of the 15th ACM on International conference on multimodal interaction* (ICMI '13). ACM, New York, NY, USA, 119-126. DOI: http://dx.doi.org/10.1145/2522848.2522877