



Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 116 (2017) 604–611

Procedia
Computer Science

www.elsevier.com/locate/procedia

2nd International Conference on Computer Science and Computational Intelligence 2017, ICCSCI
2017, 13-14 October 2017, Bali, Indonesia

Personality Prediction System from Facebook Users

Tommy Tandera, Hendro, Derwin Suhartono*, Rini Wongso, and Yen Lina Prasetyo

*Computer Science Department, School of Computer Science, Bina Nusantara University,
Jl. K. H. Syahdan No. 9 Kemanggis, Jakarta 11480, Indonesia*

Abstract

The use of social networks is increasing rapidly. Various informations are shared widely through social media, i.e. Facebook. Information about users and what they expressed through status updates are such important assets for research in the field of behavioral learning and human personality. Similar researches have been conducted in this field and it grows continually till now. This study attempts to build a system that can predict a person's personality based on Facebook user information. Personality model used in this research is Big Five Model Personality. While other previous researches used older machine learning algorithm in building their models, this research tries to implement some deep learning architectures to see the comparison by doing comprehensive analysis method through the accuracy result. The results succeeded to outperform the accuracy of previous similar research with the average accuracy of 74.17%.

© 2017 The Authors. Published by Elsevier B.V.

Peer-review under responsibility of the scientific committee of the 2nd International Conference on Computer Science and Computational Intelligence 2017.

Keywords: personality prediction; big five model personality; facebook; machine learning; deep learning

1. Introduction

Social media has become the most widely used communication and interaction tool between people over the past few years. Direct interaction between people is decreasing as people tend to communicate indirectly through smartphones. Thus, it is quite difficult to recognize person's personality. However, what's written in social media might help us to get the information needed as people spend much time checking social media and expressing their feelings and thoughts through statuses, comments, and updates. Facebook has the largest users reaching 1.8 billion

* Corresponding author. Tel.: +62-21-5345830; fax: +62-21-5300244.

E-mail address: dsuhartono@binus.edu

users with around 800 million users spending about 40 minutes a day using it ¹. Facebook users generally express their feelings and opinions through status updates or comments. Although Facebook is currently more widely used to share photos and videos, this research focuses on users' linguistic aspect which is their status updates. Studies in the field of psychology showed that there is a correlation between personality and the linguistic behavior of a person ^{2,3}. This correlation can be effectively analyzed and illustrated using natural language processing approach. Therefore, the goal of this research is to build a prediction system that can automatically predict user personality based on their activities in Facebook.

There are several personality models used in predicting personality, such as Big Five Personality, MBTI (Myers-Briggs Type Indicator) or DISC (Dominance Influence Steadiness Conscientiousness). However, after some considerations and literature review process, Big Five Personality is used in this study as it is the most popular and precise in telling someone's personality traits. Traits in this model consist of Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism.

The corpus used in this study consists of 2 (two) datasets. The first dataset consists of 250 users with around 10,000 statuses obtained from *myPersonality* project sample data, and the second dataset consists of 150 users which are collected manually. Prediction system is built using some linguistic features with different approach. The first used closed vocabulary that includes some features such as LIWC (Linguistic Inquiry and Word Count) and SPLICE (Structured Programming for Linguistic Cue Extraction). SNA (Social Network Analysis) is also included in the process because all the features' scores are provided by *myPersonality* dataset. All features in the first approach are specifically used in the older machine learning algorithm implementation. The second approach used open vocabulary approach. It is word embedding which is specifically used in deep learning technique implementation. This study uses some machine learning algorithms which are widely used in previous researches. To the best of our knowledge, deep learning implementation in this field is still hard to find. Hence, we implement some deep learning algorithms so that improvement to the prediction system can be made.

2. Related Work

Previous study on personality prediction has been done by using social media Facebook and some features such as LIWC features, SNA features, time-related features, and others ⁴. Their research is very similar with ours especially for the dataset (250 dataset from *myPersonality*) and the features (LIWC and SNA features). Another research in personality prediction based on Facebook status were done by using two approaches such as open-vocabulary DLA (Differential Language Analysis) and LIWC features ⁵. By using Facebook, a research defining features with bag-of-words and token (unigrams) approaches were conducted as well. Other study was done to make a personality prediction system by using Twitter with LIWC and MRC as features ⁶.

All mentioned above researches did personality prediction by using social media in English based on Big Five Personality models. Recent research was conducted to make a personality prediction system using Twitter in Bahasa based on Big Five Personality models ^{7,8}. Other research on personality prediction was done using deep learning technique to classify Big Five Personality models from social media Facebook ⁹.

3. Methodology

3.1. Dataset

The dataset used in this study is divided into two parts. The first dataset obtained from *myPersonality* ¹⁰ consists of 250 data of Facebook users with approximately 10,000 statuses with given personality label based on the Big Five Personality Traits model. The distribution of the *myPersonality* dataset based on the personality type is presented in Table 1 below.

متن کامل مقاله

دریافت فوری ←

ISIArticles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات