An Exploration of Joint Learning for Personality Prediction

Ben Verhoeven & Walter Daelemans

CLiPS, University of Antwerp

ATILA 2015
Antwerp, 16 October 2015
Personality types

Big Five (OCEAN)

- Openness to experience
  - Inventive/curious vs. consistent/cautious
- Conscientiousness
  - Efficient/organized vs. easy-going/careless
- Extraversion
  - Outgoing/energetic vs. solitary/reserved
- Agreeableness
  - Friendly/compassionate vs. analytical/detached
- Neuroticism (emotional stability)
  - Sensitive/nervous vs. secure/confident

Do the test at http://www.outofservice.com/bigfive/
CLiPS Stylometry Investigation Corpus

General details
- Verhoeven & Daelemans, 2014
- Freely available at www.clips.uantwerpen.be
- Fifth year of data collection starting

Corpus details
- Reviews and essays (short and long)
- Document metadata
- Author metadata
CSI Corpus

Document metadata
- Date
- Genre
- Review: truthfulness, category and topic
- Essay: grade

Author metadata
- Birth date
- Gender & Sexual Preference
- Region of origin
- Personality (Big Five & MBTI)
CSI Corpus

Statistics - version 2015-10

> 1000 reviews
> 500 essays

Average personality

- Openness 49.2
- Conscientiousness 46.9
- Extraversion 53.9
- Agreeableness 43.0
- Neuroticity 53.4
Simple Features

- Word n-grams [1,2]
- Character n-grams [3,4]
- LIWC features
- POS n-grams [1,2]
Our hypotheses were:

▶ People with different personalities may lie/deceive in a different way, which will give us more clues for predicting their personality

▶ Some personality types may be better at deceiving others, which may cause the accuracy of deception detection to drop
Personality interacts with deception?

**Exploration**

- ✓ Personality recognition within deceptive and truthful reviews separately
- ✓ Add deception as a feature when learning personality
- × Joint learning of deception and personality
  - ▶ Probably only on OPN & EXT
Personality types

Structure

- Each trait has a value between 0 and 100
- Usual order: OCEAN = OPN, CON, EXT, AGR, NEU
- E.g. 24-50-89-76-04
- Making classes with binarization
- \( N = [0:50]; Y = [50:] \)
Personality interacts with deception?

<table>
<thead>
<tr>
<th></th>
<th>OPN</th>
<th>CON</th>
<th>EXT</th>
<th>AGR</th>
<th>NEU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Majority</td>
<td>50.1</td>
<td>53.5</td>
<td>59.8</td>
<td>56.1</td>
<td>55.1</td>
</tr>
<tr>
<td>Weighted Random</td>
<td>50.0</td>
<td>50.2</td>
<td>51.9</td>
<td>50.8</td>
<td>50.5</td>
</tr>
<tr>
<td>Best simple</td>
<td><strong>62.0</strong></td>
<td>53.5</td>
<td>54.3</td>
<td>54.9</td>
<td>53.0</td>
</tr>
<tr>
<td>Truthful</td>
<td>60.3</td>
<td>51.1</td>
<td>56.7</td>
<td>50.8</td>
<td>52.3</td>
</tr>
<tr>
<td>Deceptive</td>
<td>56.9</td>
<td>55.3</td>
<td><strong>58.3</strong></td>
<td>54.6</td>
<td>51.3</td>
</tr>
<tr>
<td>Deception feat</td>
<td>61.5</td>
<td>53.1</td>
<td>54.8</td>
<td>54.7</td>
<td>53.7</td>
</tr>
</tbody>
</table>

All data = ± 800 reviews
Split data = ± 400 reviews each
Personality interacts with deception

Tentative conclusions Debunked?

**TRUE** Neuroticism cannot be detected in Dutch reviews

**FALSE** Other personality traits are a lot harder (if not impossible) to detect in deceptive reviews, except for OPN

- Style obfuscation?
Joint learning of traits

Classes
All traits together, e.g. YNNNY, makes 32 classes
Evaluation on traits separately

Best results on reviews (train-dev)

<table>
<thead>
<tr>
<th></th>
<th>OPN</th>
<th>CON</th>
<th>EXT</th>
<th>AGR</th>
<th>NEU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Majority</td>
<td>50.1</td>
<td>53.5</td>
<td>59.8</td>
<td>56.1</td>
<td>55.1</td>
</tr>
<tr>
<td>Weighted Random</td>
<td>50.0</td>
<td>50.2</td>
<td>51.9</td>
<td>50.8</td>
<td>50.5</td>
</tr>
<tr>
<td>Best simple</td>
<td>62.0</td>
<td>53.5</td>
<td>54.3</td>
<td>54.9</td>
<td>53.0</td>
</tr>
<tr>
<td>Joint</td>
<td>57.4</td>
<td>55.2</td>
<td>56.4</td>
<td>53.9</td>
<td>56.8</td>
</tr>
</tbody>
</table>

- LIWC features are often important
- POS-tags also contribute
- Improved over best simple system for CON, EXT, NEU
Joint learning of traits

Other corpora

- PAN-2015 Dutch Twitter dataset (cancelled)
  - Very small, few authors (can't keep them apart)
  - Enormously skewed
- CSI Essays (version 2015-10)
Trends in Personality Prediction for Dutch

Big trends

- Openness is consistently the best recognizable trait
- Not a lot of structure in what else works

Small trends

- Personality is better detected when truthfulness is steady
- Adding deception as feature does not help
- LIWC features work well for joint learning of personality
- Joint learning helps the performance of previously lesser recognizable traits
Questions?
ben.verhoeven@uantwerpen.be

@verhoevenben
@clipsua