Personality & Psychology Detection in Language

Jim Pasko & Kristi Tsukida
January 18, 2011
EE517 Discussion

Psychology

9-11 Tribute

http://www.gothamgazette.com/article/fea/20050905/202/1558
Psychology of coping

- Effect of crisis on a population
- How long does emotional trauma last?
- Effect of preoccupation & other factors on coping
Psychology: LiveJournal Data

- Public LiveJournal blogs
  - 1084 high frequency users
  - Entries 2 months before & after 9/11
- Predefined word categories
  - "For example, the word **cried** falls into four categories: sadness, negative emotion, overall affect, and past-tense verb."

- Subset of relevant categories
<table>
<thead>
<tr>
<th>Linguistic Inquiry and Word Count</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Word count</strong></td>
<td></td>
</tr>
<tr>
<td>words/sentence</td>
<td></td>
</tr>
<tr>
<td>Dictionary words</td>
<td></td>
</tr>
<tr>
<td><strong>Words&gt;6 letters</strong></td>
<td></td>
</tr>
<tr>
<td>Total function words</td>
<td></td>
</tr>
<tr>
<td>Total pronouns</td>
<td></td>
</tr>
<tr>
<td><strong>Personal pronouns</strong></td>
<td></td>
</tr>
<tr>
<td>1st pers singular</td>
<td></td>
</tr>
<tr>
<td>1st pers plural</td>
<td></td>
</tr>
<tr>
<td>2nd person</td>
<td></td>
</tr>
<tr>
<td>3rd pers singular</td>
<td></td>
</tr>
<tr>
<td>3rd pers plural</td>
<td></td>
</tr>
<tr>
<td>Impersonal pronouns</td>
<td></td>
</tr>
<tr>
<td><strong>Impersonal pronouns</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Articles</strong></td>
<td></td>
</tr>
<tr>
<td>[Common verbs]</td>
<td></td>
</tr>
<tr>
<td>Auxiliary verbs</td>
<td></td>
</tr>
<tr>
<td>Past tense</td>
<td></td>
</tr>
<tr>
<td>Present tense</td>
<td></td>
</tr>
<tr>
<td>Future tense</td>
<td></td>
</tr>
<tr>
<td>Adverbs</td>
<td></td>
</tr>
<tr>
<td>Prepositions</td>
<td></td>
</tr>
<tr>
<td>Conjunctions</td>
<td></td>
</tr>
<tr>
<td>Negations</td>
<td></td>
</tr>
<tr>
<td>Quantifiers</td>
<td></td>
</tr>
<tr>
<td>Numbers</td>
<td></td>
</tr>
<tr>
<td>Swear words</td>
<td></td>
</tr>
<tr>
<td><strong>Social processes</strong></td>
<td></td>
</tr>
<tr>
<td>Family</td>
<td></td>
</tr>
<tr>
<td>Friends</td>
<td></td>
</tr>
<tr>
<td>Humans</td>
<td></td>
</tr>
<tr>
<td><strong>Affective processes</strong></td>
<td></td>
</tr>
<tr>
<td>Positive emotion</td>
<td></td>
</tr>
<tr>
<td>Negative emotion</td>
<td></td>
</tr>
<tr>
<td>Anxiety</td>
<td></td>
</tr>
<tr>
<td>Anger</td>
<td></td>
</tr>
<tr>
<td>Sadness</td>
<td></td>
</tr>
<tr>
<td><strong>Cognitive processes</strong></td>
<td></td>
</tr>
<tr>
<td>Insight</td>
<td></td>
</tr>
<tr>
<td>Causation</td>
<td></td>
</tr>
<tr>
<td>Discrepancy</td>
<td></td>
</tr>
<tr>
<td>Tentative</td>
<td></td>
</tr>
<tr>
<td>Certainty</td>
<td></td>
</tr>
<tr>
<td>Inhibition</td>
<td></td>
</tr>
<tr>
<td>Inclusive</td>
<td></td>
</tr>
<tr>
<td>Exclusive</td>
<td></td>
</tr>
<tr>
<td><strong>Perceptual processes</strong></td>
<td></td>
</tr>
<tr>
<td>See</td>
<td></td>
</tr>
<tr>
<td>Hear</td>
<td></td>
</tr>
<tr>
<td>Feel</td>
<td></td>
</tr>
<tr>
<td><strong>Biological processes</strong></td>
<td></td>
</tr>
<tr>
<td>Body</td>
<td></td>
</tr>
<tr>
<td>Health</td>
<td></td>
</tr>
<tr>
<td>Sexual</td>
<td></td>
</tr>
<tr>
<td>Ingestion</td>
<td></td>
</tr>
<tr>
<td><strong>Relativity</strong></td>
<td></td>
</tr>
<tr>
<td>Motion</td>
<td></td>
</tr>
<tr>
<td>Space</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td></td>
</tr>
<tr>
<td>Work</td>
<td></td>
</tr>
<tr>
<td>Achievement</td>
<td></td>
</tr>
<tr>
<td>Leisure</td>
<td></td>
</tr>
<tr>
<td>Home</td>
<td></td>
</tr>
<tr>
<td>Money</td>
<td></td>
</tr>
<tr>
<td>Religion</td>
<td></td>
</tr>
<tr>
<td>Death</td>
<td></td>
</tr>
<tr>
<td>Assent</td>
<td></td>
</tr>
<tr>
<td>Nonfluencies</td>
<td></td>
</tr>
<tr>
<td>Fillers</td>
<td></td>
</tr>
<tr>
<td>Punctuation</td>
<td></td>
</tr>
</tbody>
</table>

http://liwc.net/descriptiontable1.php
Psychology

- Emotional positivity
  - Positive: happy, good, nice
  - Negative: kill, ugly, guilty

- Cognitive processing
  - think, question, because

- Social orientation
  - talk, share, friends

- Psychological distancing
  - articles & long words (>6 letters)
  - low 1st person singular (no "I")
  - low discrepancy (would, could, should)
  - low present tense usage
Psychology

- Preoccupation
  - Custom "September 11 Dictionary"
  - 27 target words:
    - Osama, World Trade Center, hijack, Afghanistan
  - 3 Groups
    - High / Med / Low
Psychology: Results

"Urban stress" phenomenon

Temporary emotional and social response

Psychological shock
Psychology Results

- Male => more distancing
- Older => more positivity, distancing
- Preoccupation => less able to cope

- Pervasive change
- Long-term effects

- Internet is good for historical data mining

Criticism:
- No ground truth comparison
Recognizing Personality

F. Mairesse, M. Walker, M. Mehl, and R. Moore
Personality

1. Openness to Experience
2. Conscientiousness
3. Extraversion vs. Introversion
4. Agreeableness
5. Emotional Stability vs. Neuroticism
Personality

- Automatic recognition of personality features would be useful in many contexts

- Goal: accurately characterize personality according to the Big 5 traits by applying classification models to linguistic features
  - Written text
  - Spoken dialogue (transcribed or audio)

- Compare results to self and observer reports
### Personality - previous findings

<table>
<thead>
<tr>
<th>Level</th>
<th>Introvert</th>
<th>Extravert</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conversational</td>
<td>Listen</td>
<td>Initiate conversation</td>
</tr>
<tr>
<td>behaviour</td>
<td>Less back-channel behaviour</td>
<td>More back-channel behaviour</td>
</tr>
<tr>
<td>Topic selection</td>
<td>Self-focused</td>
<td>Not self-focused*</td>
</tr>
<tr>
<td></td>
<td>Problem talk, dissatisfaction</td>
<td>Pleasure talk, agreement, compliment</td>
</tr>
<tr>
<td></td>
<td>Strict selection</td>
<td>Think out loud*</td>
</tr>
<tr>
<td></td>
<td>Single topic</td>
<td>Many topics</td>
</tr>
<tr>
<td></td>
<td>Few semantic errors</td>
<td>Many semantic errors</td>
</tr>
<tr>
<td></td>
<td>Few self-references</td>
<td>Many self-references</td>
</tr>
<tr>
<td>Style</td>
<td>Formal</td>
<td>Informal</td>
</tr>
<tr>
<td></td>
<td>Many hedges (tentative words)</td>
<td>Few hedges (tentative words)</td>
</tr>
<tr>
<td>Syntax</td>
<td>Many nouns, adjectives, prepositions (explicit)</td>
<td>Many verbs, adverbs, pronouns (implicit)</td>
</tr>
<tr>
<td></td>
<td>Elaborated constructions</td>
<td>Simple constructions*</td>
</tr>
<tr>
<td></td>
<td>Many words per sentence</td>
<td>Few words per sentence</td>
</tr>
<tr>
<td></td>
<td>Many articles</td>
<td>Few articles</td>
</tr>
<tr>
<td></td>
<td>Many negations</td>
<td>Few negations</td>
</tr>
<tr>
<td>Lexicon</td>
<td>Correct</td>
<td>Loose*</td>
</tr>
<tr>
<td></td>
<td>Rich</td>
<td>Poor</td>
</tr>
<tr>
<td></td>
<td>High diversity</td>
<td>Low diversity</td>
</tr>
<tr>
<td></td>
<td>Many exclusive and inclusive words</td>
<td>Few exclusive and inclusive words</td>
</tr>
<tr>
<td></td>
<td>Few social words</td>
<td>Many social words</td>
</tr>
<tr>
<td></td>
<td>Few positive emotion words</td>
<td>Many positive emotion words</td>
</tr>
<tr>
<td></td>
<td>Many negative emotion words</td>
<td>Few negative emotion words</td>
</tr>
<tr>
<td>Speech</td>
<td>Received accent</td>
<td>Local accent*</td>
</tr>
<tr>
<td></td>
<td>Slow speech rate</td>
<td>High speech rate</td>
</tr>
<tr>
<td></td>
<td>Few disfluencies</td>
<td>Many disfluencies*</td>
</tr>
<tr>
<td></td>
<td>Many unfilled pauses</td>
<td>Few unfilled pauses</td>
</tr>
<tr>
<td></td>
<td>Long response latency</td>
<td>Short response latency</td>
</tr>
<tr>
<td></td>
<td>Quiet</td>
<td>Loud</td>
</tr>
<tr>
<td></td>
<td>Low voice quality</td>
<td>High voice quality</td>
</tr>
<tr>
<td></td>
<td>Non-nasal voice</td>
<td>Nasal voice</td>
</tr>
<tr>
<td></td>
<td>Low frequency variability</td>
<td>High frequency variability</td>
</tr>
</tbody>
</table>
Tool to count word categories in written essays
  - "small but significant correlations" between these categories and the various Big 5

Commercial package

Dictionary of 4500 words & stems

~80 output categories

Some traits may be more expressed through language, or the linguistic cues may be easier to analyze
  - Extraversion
### Linguistic Inquiry and Word Count (LIWC)

<table>
<thead>
<tr>
<th>Category</th>
<th>Subcategories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word count</td>
<td>words/sentence</td>
</tr>
<tr>
<td>Dictionary words</td>
<td>words</td>
</tr>
<tr>
<td>Words&gt;6 letters</td>
<td></td>
</tr>
<tr>
<td>Total function words</td>
<td></td>
</tr>
<tr>
<td>Total pronouns</td>
<td></td>
</tr>
<tr>
<td>Personal pronouns</td>
<td></td>
</tr>
<tr>
<td>1st pers singular</td>
<td></td>
</tr>
<tr>
<td>1st pers plural</td>
<td></td>
</tr>
<tr>
<td>2nd person</td>
<td></td>
</tr>
<tr>
<td>3rd pers singular</td>
<td></td>
</tr>
<tr>
<td>3rd pers plural</td>
<td></td>
</tr>
<tr>
<td>Impersonal pronouns</td>
<td></td>
</tr>
<tr>
<td>Articles</td>
<td></td>
</tr>
<tr>
<td>[Common verbs]</td>
<td></td>
</tr>
<tr>
<td>Auxiliary verbs</td>
<td></td>
</tr>
<tr>
<td>Past tense</td>
<td></td>
</tr>
<tr>
<td>Present tense</td>
<td></td>
</tr>
<tr>
<td>Future tense</td>
<td></td>
</tr>
<tr>
<td>Adverbs</td>
<td></td>
</tr>
<tr>
<td>Prepositions</td>
<td></td>
</tr>
<tr>
<td>Conjunctions</td>
<td></td>
</tr>
<tr>
<td>Negations</td>
<td></td>
</tr>
<tr>
<td>Quantifiers</td>
<td></td>
</tr>
<tr>
<td>Numbers</td>
<td></td>
</tr>
<tr>
<td>Swear words</td>
<td></td>
</tr>
<tr>
<td>Social processes</td>
<td></td>
</tr>
<tr>
<td>Family</td>
<td></td>
</tr>
<tr>
<td>Friends</td>
<td></td>
</tr>
<tr>
<td>Humans</td>
<td></td>
</tr>
<tr>
<td>Affective processes</td>
<td></td>
</tr>
<tr>
<td>Positive emotion</td>
<td></td>
</tr>
<tr>
<td>Negative emotion</td>
<td></td>
</tr>
<tr>
<td>Anxiety</td>
<td></td>
</tr>
<tr>
<td>Anger</td>
<td></td>
</tr>
<tr>
<td>Sadness</td>
<td></td>
</tr>
<tr>
<td>Cognitive processes</td>
<td></td>
</tr>
<tr>
<td>Insight</td>
<td></td>
</tr>
<tr>
<td>Causation</td>
<td></td>
</tr>
<tr>
<td>Discrepancy</td>
<td></td>
</tr>
<tr>
<td>Tentative</td>
<td></td>
</tr>
<tr>
<td>Certainty</td>
<td></td>
</tr>
<tr>
<td>Inhibition</td>
<td></td>
</tr>
<tr>
<td>Inclusive</td>
<td></td>
</tr>
<tr>
<td>Exclusive</td>
<td></td>
</tr>
<tr>
<td>Perceptual processes</td>
<td></td>
</tr>
<tr>
<td>See</td>
<td></td>
</tr>
<tr>
<td>Hear</td>
<td></td>
</tr>
<tr>
<td>Feel</td>
<td></td>
</tr>
<tr>
<td>Biological processes</td>
<td></td>
</tr>
<tr>
<td>Body</td>
<td></td>
</tr>
<tr>
<td>Health</td>
<td></td>
</tr>
<tr>
<td>Sexual</td>
<td></td>
</tr>
<tr>
<td>Ingestion</td>
<td></td>
</tr>
<tr>
<td>Relativity</td>
<td></td>
</tr>
<tr>
<td>Motion</td>
<td></td>
</tr>
<tr>
<td>Space</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td></td>
</tr>
<tr>
<td>Work</td>
<td></td>
</tr>
<tr>
<td>Achievement</td>
<td></td>
</tr>
<tr>
<td>Leisure</td>
<td></td>
</tr>
<tr>
<td>Home</td>
<td></td>
</tr>
<tr>
<td>Money</td>
<td></td>
</tr>
<tr>
<td>Religion</td>
<td></td>
</tr>
<tr>
<td>Death</td>
<td></td>
</tr>
<tr>
<td>Assent</td>
<td></td>
</tr>
<tr>
<td>Nonfluencies</td>
<td></td>
</tr>
<tr>
<td>Fillers</td>
<td></td>
</tr>
<tr>
<td>Punctuation</td>
<td></td>
</tr>
</tbody>
</table>

[http://liwc.net/descriptiontable1.php](http://liwc.net/descriptiontable1.php)
**Linguistic Features**

- Frequency counts of 88 LIWC categories
- MRC Psycholinguistic Database
- Only LIWC and MRC used in essay corpus models

**MRC FEATURES (Coltheart, 1981):**

Number of letters (Nlet), phonemes (Nphon), syllables (Nsyl), Kucera-Francis written frequency (K-F-freq), Kucera-Francis number of categories (K-F-ncats), Kucera-Francis number of samples (K-F-nsamp), Thorndike-Lorge written frequency (T-L-freq), Brown verbal frequency (Brown-freq), familiarity rating (Fam), concreteness rating (Conc), imageability rating (Imag), meaningfulness Colorado Norms (Meanc), meaningfulness Paivio Norms (Meanp), age of acquisition (AOA)
Method

- Collect individual corpora
- Collect associated personality ratings
- Extract relevant features from the texts
- Build statistical models based on extracted features
- Test the models on the linguistic output of unseen individuals
Corpora

- Essay Corpus - 2479 essays written by psychology students, who were given 20 minutes to write down whatever occurred to them
  - Self-reported personality

- EAR corpus - a set of recorded conversations and associated transcripts
  - smaller corpus, but each utterance contains additional information (prosody etc.)
  - Self- and observer-reported personality

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Essays</th>
<th>EAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source of language</td>
<td>Written</td>
<td>Spoken</td>
</tr>
<tr>
<td>Personality reports</td>
<td>Self reports</td>
<td>Self and observer</td>
</tr>
<tr>
<td>Number of words</td>
<td>1.9 million</td>
<td>97,468</td>
</tr>
<tr>
<td>Subjects</td>
<td>2,479</td>
<td>96</td>
</tr>
<tr>
<td>Words per subject</td>
<td>766.4</td>
<td>1,015.3</td>
</tr>
</tbody>
</table>
Additional Features for EAR models

- Utterance type (applied to EAR corpus only)
  - command, assertion, question, etc.
  - homegrown classifier

<table>
<thead>
<tr>
<th>Label</th>
<th>Fraction</th>
<th>Labelling accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assertion</td>
<td>73.0%</td>
<td>0.95</td>
</tr>
<tr>
<td>Command</td>
<td>4.3%</td>
<td>0.50</td>
</tr>
<tr>
<td>Prompt</td>
<td>7.0%</td>
<td>0.57</td>
</tr>
<tr>
<td>Question</td>
<td>15.7%</td>
<td>1.00</td>
</tr>
<tr>
<td>All</td>
<td>100%</td>
<td>0.88</td>
</tr>
</tbody>
</table>

- Prosodic features (EAR corpus)
  - Pitch, intensity, time, rate

- The only manual effort was in transcription
**Binary classifiers**

- Majority Class (baseline)
- Naive Bayes
- Nearest Neighbor
- J48 decision tree
- JRip rule set
- AdaboostM1
- Support Vector Machine
Binary Classification Results

- Binary classification models evaluated using self reports (essay corps) and both self and observer reports (EAR corpus)

- All results averaged over 10-fold cross-validation
Essay corpus results

<table>
<thead>
<tr>
<th>Trait</th>
<th>Base</th>
<th>J48</th>
<th>NN</th>
<th>NB</th>
<th>JRIP</th>
<th>ADA</th>
<th>SMO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extraversion</td>
<td>50.04</td>
<td>54.44*</td>
<td>53.27*</td>
<td>53.35*</td>
<td>52.70</td>
<td>55.00*</td>
<td>54.93*</td>
</tr>
<tr>
<td>Emotional stability</td>
<td>50.08</td>
<td>51.09</td>
<td>51.62</td>
<td>56.42*</td>
<td>55.90*</td>
<td>55.98*</td>
<td>57.35*</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>50.36</td>
<td>53.51*</td>
<td>50.16</td>
<td>53.88*</td>
<td>52.63</td>
<td>52.71</td>
<td>55.78*</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>50.57</td>
<td>51.37</td>
<td>52.10</td>
<td>53.80</td>
<td>52.71</td>
<td>54.45*</td>
<td>55.29*</td>
</tr>
<tr>
<td>Openness to experience</td>
<td>50.32</td>
<td>54.24*</td>
<td>53.07</td>
<td>59.57*</td>
<td>58.85*</td>
<td>59.09*</td>
<td>62.11*</td>
</tr>
</tbody>
</table>

- statistically significant improvement over the majority class baseline (two-tailed paired t-test, $p < .05$)

Table 12: Classification accuracy with two equal size bins on the essays corpus, using self-reports. Models are the majority class baseline (Base); J48 decision tree (J48); Nearest neighbour (NN); Naive Bayes (NB); JRip rule set (JRIP); AdaboostM1 (ADA); Support vector machines (SMO).

- Openness to experience easiest to classify
Figure 1: J48 decision tree for binary classification of extraversion, based on the essays corpus and self-reports.
LIWC vs MRC feature set

- Best-performing classifiers were trained on exclusively LIWC and MRC
- Solely LIWC-based classifiers performed better for the essay corpus
  - They also performed better than the full set

<table>
<thead>
<tr>
<th>Feature set</th>
<th>None</th>
<th>LIWC features</th>
<th>MRC features</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Base</td>
<td>NB</td>
<td>ADA</td>
</tr>
<tr>
<td>Classifier</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Set size</td>
<td>0</td>
<td>88</td>
<td>88</td>
</tr>
<tr>
<td>Extraversion</td>
<td>50.04</td>
<td>52.71</td>
<td><strong>56.34</strong></td>
</tr>
<tr>
<td>Emotional stability</td>
<td>50.08</td>
<td>56.02</td>
<td><strong>55.33</strong></td>
</tr>
<tr>
<td>Agreeableness</td>
<td>50.36</td>
<td>54.12</td>
<td><strong>52.71</strong></td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>50.57</td>
<td>53.92</td>
<td><strong>54.48</strong></td>
</tr>
<tr>
<td>Openness to experience</td>
<td>50.32</td>
<td>58.92</td>
<td><strong>58.64</strong></td>
</tr>
</tbody>
</table>

* statistically significant improvement over the majority class baseline (two-tailed paired t-test, p < .05)
Classification results: EAR corpus

- Extraversion is the easiest trait to classify

<table>
<thead>
<tr>
<th>Data</th>
<th>Trait</th>
<th>Base</th>
<th>J48</th>
<th>NN</th>
<th>NB</th>
<th>JRIP</th>
<th>ADA</th>
<th>SMO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obs</td>
<td>Extra</td>
<td>47.78</td>
<td>66.78</td>
<td>59.33</td>
<td>73.00•</td>
<td>60.44</td>
<td>73.00•</td>
<td>65.78</td>
</tr>
<tr>
<td>Obs</td>
<td>Emot</td>
<td>51.11</td>
<td>62.56</td>
<td>58.22</td>
<td>73.89•</td>
<td>56.22</td>
<td>48.78</td>
<td>60.33</td>
</tr>
<tr>
<td>Obs</td>
<td>Agree</td>
<td>47.78</td>
<td>48.78</td>
<td>51.89</td>
<td>61.33•</td>
<td>51.89</td>
<td>52.89</td>
<td>56.33</td>
</tr>
<tr>
<td>Obs</td>
<td>ConsC</td>
<td>47.78</td>
<td>57.67</td>
<td>61.56</td>
<td>67.67•</td>
<td>61.56</td>
<td>60.22•</td>
<td>57.11</td>
</tr>
<tr>
<td>Obs</td>
<td>Open</td>
<td>47.78</td>
<td>52.22</td>
<td>46.78</td>
<td>57.00</td>
<td>49.67</td>
<td>50.56</td>
<td>55.89</td>
</tr>
<tr>
<td>Self</td>
<td>Extra</td>
<td>47.78</td>
<td>48.78</td>
<td>49.67</td>
<td>57.33</td>
<td>50.56</td>
<td>54.44</td>
<td>49.89</td>
</tr>
<tr>
<td>Self</td>
<td>Emot</td>
<td>51.11</td>
<td>45.56</td>
<td>46.78</td>
<td>50.44</td>
<td>46.78</td>
<td>41.89</td>
<td>44.33</td>
</tr>
<tr>
<td>Self</td>
<td>Agree</td>
<td>52.22</td>
<td>47.89</td>
<td>50.89</td>
<td>58.33</td>
<td>56.89</td>
<td>55.22</td>
<td>52.33</td>
</tr>
<tr>
<td>Self</td>
<td>ConsC</td>
<td>51.11</td>
<td>33.44</td>
<td>45.56</td>
<td>39.33</td>
<td>43.11</td>
<td>46.11</td>
<td>53.22</td>
</tr>
<tr>
<td>Self</td>
<td>Open</td>
<td>51.11</td>
<td>52.00</td>
<td>42.22</td>
<td>61.44</td>
<td>45.00</td>
<td>56.00</td>
<td>47.78</td>
</tr>
</tbody>
</table>

- statistically significant improvement over the majority class baseline (two-tailed paired t-test, p < .05)

Table 14: Classification accuracy with two equal size bins on the EAR corpus, for observer ratings (Obs) and self-reports (Self). Models are majority class baseline (Base); J48 decision tree (J48); Nearest neighbour (NN); Naive Bayes (NB); JRip rules set (JRIP); AdaboostM1 (ADA); Support vector machines (SMO).
Feature set comparison (EAR)

- Openness to experience was better classified using prosodic features only
  - Results are from a Naive Bayes classifier

<table>
<thead>
<tr>
<th>Feature set</th>
<th>None</th>
<th>Type</th>
<th>LIWC</th>
<th>MRC</th>
<th>Prosody</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set size</td>
<td>0</td>
<td>4</td>
<td>88</td>
<td>14</td>
<td>11</td>
</tr>
<tr>
<td>Extraversion</td>
<td>47.78</td>
<td>45.67</td>
<td>68.89</td>
<td>68.78</td>
<td>67.56</td>
</tr>
<tr>
<td>Emotional stability</td>
<td>51.11</td>
<td>60.22</td>
<td>69.89</td>
<td>60.78</td>
<td>61.78</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>47.78</td>
<td>57.56</td>
<td>54.00</td>
<td>58.67</td>
<td>50.44</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>47.78</td>
<td>59.67</td>
<td>60.22</td>
<td>66.78</td>
<td>52.11</td>
</tr>
<tr>
<td>Openness to experience</td>
<td>47.78</td>
<td>53.11</td>
<td>61.11</td>
<td>54.00</td>
<td>64.56</td>
</tr>
</tbody>
</table>

- • statistically significant improvement over the majority class baseline (two-tailed paired t-test, p < .05)

Table 15: Classification accuracies for the EAR corpus with observer reports using the Naive Bayes classifier, for different feature sets (None=baseline, Type=utterance type). Models performing better than with the full feature set are in bold.
Another interesting decision tree

Word count

≤ 1284  > 1284

Metaphysical issues  Extravert

≤ 0.25  > 0.25

Commas  Articles

≤ 8.72  > 8.72  ≤ 3.51  > 3.51

Eating  Extravert  Extravert  Space

≤ 0.51  > 0.51  ≤ 3.22  > 3.22

Introvert  Sad  Extravert  Frequency of use

≤ 0.15  > 0.15  ≤ 6072  > 6072

Introvert  Extravert  Extravert  Introvert
Regression Models used

- Mean value baseline
- Linear regression
- M5 regression tree
- M5 regression tree with linear models
- REPTree
- Support Vector Machine for regression
Regression Model Results

- Baseline: mean of all personality scores in training set

- Evaluated using relative error
  - low relative error means the model performed better than the baseline
  - 100 % relative error means the model performed equivalently to the baseline

- 10-fold cross validation
Regression Results - Essay Corpus

- Associate a scalar personality value instead of a binary category with an individual
  - much more difficult, so results don't seem as impressive
- Openness to experience produces the best results
- M5 performs best with emotional stability and openness to experience

<table>
<thead>
<tr>
<th>Trait</th>
<th>Base</th>
<th>LR</th>
<th>M5R</th>
<th>M5</th>
<th>REP</th>
<th>SMO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extraversion</td>
<td>100.00</td>
<td>99.17</td>
<td>99.31</td>
<td>99.22</td>
<td>99.98</td>
<td>100.65</td>
</tr>
<tr>
<td>Emotional stability</td>
<td>100.00</td>
<td>96.87●</td>
<td>99.75</td>
<td>96.43●</td>
<td>99.35</td>
<td>98.35</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>100.00</td>
<td>98.92</td>
<td>99.86</td>
<td>99.22</td>
<td>99.78</td>
<td>100.28</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>100.00</td>
<td>98.68</td>
<td>100.62</td>
<td>98.56</td>
<td>100.47</td>
<td>99.30</td>
</tr>
<tr>
<td>Openness to experience</td>
<td>100.00</td>
<td>93.58●</td>
<td>97.68●</td>
<td>93.27●</td>
<td>99.82</td>
<td>94.19●</td>
</tr>
</tbody>
</table>

● statistically significant improvement over the mean value baseline (two-tailed paired t-test, p < .05)
LIWC vs MRC feature set

- Again, using just LIWC features usually produces more accurate results than MRC features
  - Also generally outperforms the full feature set (both LIWC and MRC)

- Exception: Extraversion is more accurately described by only MRC features using the Linear Regression model, which outperforms baseline

<table>
<thead>
<tr>
<th>Feature set</th>
<th>None</th>
<th>LIWC features (LR, M5, SMO)</th>
<th>MRC features (LR, M5, SMO)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression model</td>
<td>Base</td>
<td>LR 99.39 M5 99.25 SMO 100.8</td>
<td>LR 98.79 M5 98.79 SMO 99.13</td>
</tr>
<tr>
<td>Extraversion</td>
<td>100.00</td>
<td>96.71 (96.42) 98.03</td>
<td>99.49 99.54 99.89</td>
</tr>
<tr>
<td>Emotional stability</td>
<td>100.00</td>
<td>98.50 (98.52) 99.52</td>
<td>99.75 99.81 99.31 (99.31)</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>100.00</td>
<td>98.23 (98.14) 99.46</td>
<td>99.23 99.23 99.16</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>100.00</td>
<td>93.50 (93.70) 94.14</td>
<td>97.44 (97.44) 97.26</td>
</tr>
<tr>
<td>Openness to experience</td>
<td>100.00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Statistically significant improvement over the mean value baseline (two-tailed paired t-test, p < .05)
Regression Results - EAR corpus

- Self reports are difficult to model
- Observed reports are easier, particularly for extraversion, emotional stability, and conscientiousness
- Linear Regression and Support Vector Machines perform poorly
  - probably require a bigger training set

<table>
<thead>
<tr>
<th>Data</th>
<th>Trait</th>
<th>Base</th>
<th>LR</th>
<th>M5R</th>
<th>M5</th>
<th>REP</th>
<th>SMO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obs</td>
<td>Extraversion</td>
<td>100.00</td>
<td>179.16</td>
<td>82.16</td>
<td>80.15</td>
<td>79.94</td>
<td>140.05</td>
</tr>
<tr>
<td>Obs</td>
<td>Emotional stability</td>
<td>100.00</td>
<td>302.74</td>
<td>92.03</td>
<td>86.75</td>
<td>100.51</td>
<td>162.05</td>
</tr>
<tr>
<td>Obs</td>
<td>Agreeableness</td>
<td>100.00</td>
<td>242.68</td>
<td>96.73</td>
<td>111.16</td>
<td>99.37</td>
<td>173.97</td>
</tr>
<tr>
<td>Obs</td>
<td>Conscientiousness</td>
<td>100.00</td>
<td>188.18</td>
<td>82.68</td>
<td>90.85</td>
<td>98.08</td>
<td>131.75</td>
</tr>
<tr>
<td>Obs</td>
<td>Openness to experience</td>
<td>100.00</td>
<td>333.65</td>
<td>101.64</td>
<td>119.53</td>
<td>102.76</td>
<td>213.20</td>
</tr>
<tr>
<td>Self</td>
<td>Extraversion</td>
<td>100.00</td>
<td>204.96</td>
<td>104.50</td>
<td>118.44</td>
<td>99.94</td>
<td>176.51</td>
</tr>
<tr>
<td>Self</td>
<td>Emotional stability</td>
<td>100.00</td>
<td>321.97</td>
<td>104.10</td>
<td>108.39</td>
<td>99.91</td>
<td>233.19</td>
</tr>
<tr>
<td>Self</td>
<td>Agreeableness</td>
<td>100.00</td>
<td>349.87</td>
<td>106.90</td>
<td>110.84</td>
<td>101.64</td>
<td>201.80</td>
</tr>
<tr>
<td>Self</td>
<td>Conscientiousness</td>
<td>100.00</td>
<td>177.12</td>
<td>103.39</td>
<td>120.29</td>
<td>107.33</td>
<td>124.91</td>
</tr>
<tr>
<td>Self</td>
<td>Openness to experience</td>
<td>100.00</td>
<td>413.70</td>
<td>107.12</td>
<td>122.68</td>
<td>126.31</td>
<td>233.01</td>
</tr>
</tbody>
</table>
Feature set comparison - EAR corpus

Only focus on the 3 regression tree models

LIWC predict observed extraversion and conscientiousness well

Prosodic features only provide the best model for emotional stability
  speech cues are useful to detect neuroticism
Regression tree - Extraversion

Word count

≤ 675

Mean pitch

≤ 231

Intensity variation

≤ 6.39

2.86

> 231

3.23

> 675

Word count

≤ 1299

3.83

> 1299

4.24
Questions?