Objective Language Feature Analysis in Children with Neurodevelopmental Disorders during Autism Assessment

Manoj Kumar¹, Rahul Gupta², Daniel Bone³, Nikolaos Malandrakis⁴, Somer Bishop⁵, Shrikant Narayanan¹

¹Signal Analysis and Interpretation Lab, University of Southern California ²Department of Psychiatry, University of California San Francisco

Motivation

- Growing prevalence of ASD. Among American children: 1 in 68
- Marked by delayed and impaired language production and use: Echolalia (meaningless repetition), neologism, etc.
- To come up with objective linguistic measures that quantify and describe behavioral characteristics
- Aid language-specific assessment and improve overall clinical diagnosis

Background

- Linguistic norms: Continuous affect measures extracted from transcriptions (Sentiment analysis, Document polarity)
- Recent works - explore linguistic norms beyond affect & scalability to large corpus
- Autism Diagnostic Observation Schedule: Semi-structured, module-specific ASD assessment tool
- Different categorical codes combined into Calibrated Severity Score (CSS).

Classification Summary

- To test the discriminative power of psycho-linguistic norms over word usage distributions (Maximum Entropy Classifier)
- Train-validation-test split : 8-1-1; 10-fold CV

<table>
<thead>
<tr>
<th>Classifier Setup</th>
<th>Accuracy(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maxent</td>
<td>66.7</td>
</tr>
<tr>
<td>Maxent + LIWC</td>
<td>69.1</td>
</tr>
<tr>
<td>Maxent + Psycho-Linguistic Norm</td>
<td>69.7</td>
</tr>
</tbody>
</table>

Table 1. Performance of classification systems. Chance = 52.3%

- Feature selection returned Gender Ladenness (F1); and Affect (F2) from negative valence conversations

Classification System

- Train Data
- POS Tagging + Feature Extraction
- N-grams
- SVM
- Frequency Filter
- Entropy Filter
- MaxEnt Classifier
- SVM
- Fusion
- Parameter Tuning
- Test Data
- Labels

Fig 1. Overview of the classifier system. Best-estimate clinical diagnosis used as ground truth.

Norm-Severity Correlates

- To analyze how closely lexical indicators are indicative of the underlying behavioral characteristic
- Correlation analysis with Calibrated Severity Scores (CSS) driven by existing hypotheses from ASD literature

<table>
<thead>
<tr>
<th>Norm</th>
<th>Child</th>
<th>Psychologist</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concrete (F1)</td>
<td>0.09</td>
<td>-0.10</td>
</tr>
<tr>
<td>Valence (F1)</td>
<td>-0.15</td>
<td>-0.20</td>
</tr>
<tr>
<td>Gender Ladenness (F1)</td>
<td>-0.07</td>
<td>0.32</td>
</tr>
<tr>
<td>Affect (F2)</td>
<td>0.08</td>
<td>0.30</td>
</tr>
</tbody>
</table>

Table 2. Partial correlations for selected psycho-linguistic descriptors. Controlled for age, gender, verbal IQ (p<0.05)

- Findings in parallel with Bone et al. 2014., where psychologist’s prosodic patterns are indicative of child’s autism severity

Discussion

- Significant classification accuracy with MaxEnt. No significant increase with lexical norms
- Existence of variation in conduct of Sadness, Anger and Fear questions - child’s response and psychologist’s follow-up
- Psychologist’s affect influenced by child’s diagnosis
- Selected frequent N-grams of different diagnostic groups:

<table>
<thead>
<tr>
<th>Child</th>
<th>Psychologist</th>
</tr>
</thead>
<tbody>
<tr>
<td>I_DONT_DON_T_KNOW, AND_I, UM_U, BUT_I</td>
<td>FEEL, WHEN, IT_FEEL, OTHER_PEOPLE, MAKES_YOU, DO_YOU</td>
</tr>
<tr>
<td>MY_BROTHER, IN_THE, I_GET, LIKE_I, IJUST</td>
<td>YOU_FEEL, WHEN_YOURE, HOW_DOES, CAN_YOU, FEEL_INSIDE</td>
</tr>
</tbody>
</table>

Table 3. Selected significant N-grams returned by MaxEnt classifier

Future Work

- Automate lexical analysis using ASR decoded hypothesis/lattices
- Integrate audio/video modality in the classification setup

References