Automated Classification of Classroom Climate using Audio Analysis

Anusha James  
Research Associate,  
Dauwels Lab, School of EEE, NTU  
jamesanusha@ntu.edu.sg
Research Team

PI: Assoc Prof Justin Dauwels, Nanyang Technological University, Singapore

Dr. Tomasz Maszczyk, Research Fellow, Nanyang Technological University, Singapore

Dr. Rebacca Bull, Principal Research Scientist at the National Institute of Education, Singapore

Dr. Ana Moreno Nunez, Research Fellow, National Institute of Education, Singapore

Project Officer, Chua Yi Han Victoria
Overview

1. Introduction
2. Design/Methods
3. Data/Results
4. Conclusion
5. Next Steps…
“Classroom Climate” What does that mean?

Prevailing mood, attitudes, standards, and tone that students feel when they are in a classroom

Negative classroom climate can feel hostile, chaotic, and out of control.

Positive classroom climate feels safe, respectful, welcoming, and supportive of student learning.

Establishing a positive climate in class is important for effective teaching.
Examples of Positive climate classroom
Examples of Negative climate classroom
Classroom’s climate doesn’t just happen—it’s created!

This needs improvement of teacher instruction

- **Analysis of teacher conversations** and ways to improve it.
- Provide teachers with **qualitative and formative feedback** on their instruction to improve teaching strategies.

Teacher instruction techniques to be assessed using CLASS scores.
Classroom Assessment Scoring System (CLASS)

- Assess the interactional quality of classroom between students and teachers in preschool classrooms.
- 2 of its 10 subscales related to the climate in the classrooms.
- Current efforts to assess teacher training rely on manual coding by trained observers – tedious and time-consuming.
Can we predict **emotional climate** of kindergarten classrooms?

**Can we?**

Yes. Provide actionable feedbacks to teachers, students and for policy makers, thus contributing smart classroom in the future

**That’s great! How?**

Design technological platforms that analyze real-life data in learning environments, and generate automatic objective assessments in real-time

### Problem of Speaker Diarization

- Who spoke when?

**Is anyone speaking?**

(Speech Activity Detection)

**Which speaker in the audio is speaking?**

(Speaker Diarization)
What’s new in this research?

• All the studies rely on speech features to automatically predict types of classroom activities, while in our study we aim to predict classroom assessment scores; specifically, as a first step, we attempt inferring the classroom climate.

• Unlike studies which investigate structured class settings, an ecological data set is considered here. Learning environments are more dynamic and less controlled than structured dialogs.
System design

CLASSROOM VIDEO RECORDINGS

SPEAKER DIARIZATION: WHO SPOKE WHEN?

SPEECH FEATURES

KIDS

TEACHER

OVERLAP

SPEECH FEATURES

Mean and std. deviation of Opensmile features + Conversational features = 5940 features
92 classrooms in multiple preschools in Singapore by researchers of the National Institute of Education (NIE).

**Duration of video and Team size**

15-20 minutes 10-15 students

**Class activities**

Small team activities (children sitting at tables with teacher walking around), and Teacher-Student discussions (students sitting around teacher).

**Climate coding**

2 independent annotators scored the overall classroom climate of each video according to the rubrics outlined in the CLASS manual.

**Dimensions for coding**

Positive climate - dimensions of positive affect, relationships, positive communication, and respect during teacher-student and student-student interactions.

Negative climate - dimensions of negativity, punitive control and disrespect.
Challenges

• Varying classroom settings

• Video captured by one stationary camera

• Audio recorded by a microphone worn by the teacher.
  • Captures teacher and children speech, background noise (crying and feet stamping)

• The speech of the children was not always captured with adequate fidelity and intelligibility, since
  • the teacher is often walking around and
  • the children might be far away from the microphone and might be speaking softly;
  • So the children speech is occasionally misclassified as noise.
Challenges

• State-of-the-art diarization algorithms were designed for single non-moving microphone.

• Ground truth CLASS scores are coded w.r.t. visual and audio information, while we only used audio information to predict CLASS climate scores.
### Speaker Diarization LOOCV results

**Average of Normalized Confusion matrices (%)**

<table>
<thead>
<tr>
<th>True \ Estimate</th>
<th>Teacher</th>
<th>Children</th>
<th>Overlap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher</td>
<td>77.3 (2773.4 sec)</td>
<td>21 (741.2 sec)</td>
<td>1.7 (61.4 sec)</td>
</tr>
<tr>
<td>Children</td>
<td>23.7 (618.4 sec)</td>
<td>71.6 (1885.6 sec)</td>
<td>4.7 (142.6 sec)</td>
</tr>
<tr>
<td>Overlap</td>
<td>75.4 (483sec)</td>
<td>21.6 (169.2sec)</td>
<td>3.1 (29.9sec)</td>
</tr>
</tbody>
</table>

**Confusion matrix normalized by total time (%)**

<table>
<thead>
<tr>
<th>True \ Estimate</th>
<th>Teacher</th>
<th>Children</th>
<th>Overlap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher</td>
<td>40.2</td>
<td>10.7</td>
<td>0.9</td>
</tr>
<tr>
<td>Children</td>
<td>9.0</td>
<td>27.3</td>
<td>2.1</td>
</tr>
<tr>
<td>Overlap</td>
<td>7.0</td>
<td>2.5</td>
<td>0.4</td>
</tr>
</tbody>
</table>

Confusion matrices showing the actual / hypothesized speakers association for **8 classroom conversations**.
Speaker Diarization LOOCV results

Boxplot of accuracy obtained for speakers in 8 classroom conversations
Feature Selection

5940 features → Kruskal-wallis test → 116 features → Correlation based Feature Selection → 20 features
Statistics of salient features

Distribution of p-values

Number of salient features associated with teacher and children speech, and overlap segments.
How well our classroom climate prediction is doing?

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Crossvalidation accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear SVM (Support Vector Machine)</td>
<td>0.71</td>
</tr>
<tr>
<td>Radial SVM (SVMG)</td>
<td>0.73</td>
</tr>
<tr>
<td>Logistic Regression(LR)</td>
<td>0.78</td>
</tr>
<tr>
<td>K Nearest Neighbors (kNN)</td>
<td>0.73</td>
</tr>
<tr>
<td>Decision Tree (DT)</td>
<td>0.60</td>
</tr>
<tr>
<td>AdaBoost Classifier (AB)</td>
<td>0.70</td>
</tr>
<tr>
<td>Random Forest Classifier(RF)</td>
<td>0.74</td>
</tr>
<tr>
<td>Naïve Bayes(NB)</td>
<td>0.79</td>
</tr>
<tr>
<td>Multilayer Perceptron (MLP)</td>
<td>0.80</td>
</tr>
</tbody>
</table>

Training set: 12 hours of audio
Test set: 80 classroom = 27 hours of audio

70 – 80 % accuracy in predicting climate of a real noisy KG classroom dataset!
How well our classroom climate prediction is doing?
EduBrowser: Cloud platform for TBL

EduBrowser

Report

Today, 10 November 2017

AUDIO RECORDING

Viewing Mode

Comments: 広い門の下には、この男のほかに誰もいない。

RESULTS

Speaking Time: 15mins 2sec (Teachers) 5mins 8sec (Students)
Overlap: 4
Turn taking: 26
Climate: POSITIVE

Filename: 354.Obs.4_EL.5G (354).mp3
Duration: 20:10:36 (20mins 10 secs)
Description: Childcare study of Teacher and 5 Groups

Climate: POSITIVE

65%
Conclusion

• Prediction of classroom climate might be possible from non-speech verbal and prosodic cues, even from recordings captured by a single microphone worn by the teacher in noisy classrooms.

• Could be useful in the development of automated feedback tools for aiding professional teacher development.

• More testing is warranted to investigate how such a system can be adapted to suit the needs of professional development and classroom assessment.
What will I do next?

Investigate video features and further validate and evaluate our system on a larger and more diverse dataset across more teachers, classroom sessions and class activities.