Applications of Machine Learning applications in support of improving the everyday pedagogical practices of teachers

Data and project overview

APT draws on a corpus of video of K-12 teachers of all grade and subjects. Teachers who are seeking National Board certification (the highest recognition of professional practice) submit video and a number of associated artifacts for reviewers to consider. Over the years this corpus has been looked at by education researchers for evidence of particular practices. In some cases, reviewers have looked across the corpus for examples of science teachers meeting aspects of the Next Generation Science Standards. Other education researchers, who are seeking to advance the profession, have identified particular practices as being highly associated with effective or ambitious teaching, e.g., TeachingWorks High Leverage Practices or Danielson’s Framework. In every case domain experts reviewed the video and identified the relevant portion with a tag. The APT project was conceived as a way to automate and scale this identification process, and offer it in near realtime to educators.

For the purposes of our work over the last year, we have focused on a subset of the corpus that has been tagged with TeachingWorks High Leverage Practices. Rather than using speech to text processors to generate transcripts, we relied on videos in the collection that were already transcribed. To date, all of our analysis has focused on a few tags.

Challenges

- Limited training set size
- Noisy, discourse is messy, inconsistent lengths
- No "bad" practice examples
- Does sufficient information exist in the audio transcripts to ascertain best practice behavior?
- Can such information be extracted as features useful to a machine learning model?
- What are the unknowns?
- Do we sufficiently understand the problem?
- What are the adjacent fields of inquiry?
- What information and methods are useful?

Research

Experiments

- Term Frequency – Inverse Document Frequency Matrix (TF-IDF)

<table>
<thead>
<tr>
<th>Term</th>
<th>TF-IDF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Math</td>
<td>0.576</td>
</tr>
<tr>
<td>Sci</td>
<td>0.424</td>
</tr>
<tr>
<td>Learn</td>
<td>0.223</td>
</tr>
<tr>
<td>Teach</td>
<td>0.139</td>
</tr>
<tr>
<td>Lead</td>
<td>0.076</td>
</tr>
<tr>
<td>Team</td>
<td>0.048</td>
</tr>
</tbody>
</table>

- Inquiry 5 - Tone Analysis

- Decision Tree Based On Question Distribution Across TW08 Transcripts

Boundary

- Full Corpus Transcript to Cosine Similarity View

- Figure 5: Similarity Check Example using Cosine Metric and Jaccard Index

- Jaccard Index: Intersection of Tag Sets / Union of Tag Sets

- K-Nearest Neighbor Results

- Cosine Similarity: Inner product of document vectors from the TF-IDF matrix

- Jaccard Index: Intersection of Tag Sets / Union of Tag Sets

Combinatorial

- Combinatorial Challenges

- What is the current state of art?

- Students studying to become teachers spend time in classrooms in apprenticeship roles of a variety of modes, with time under tutelage being the biggest variance. Some new K-12 teachers engage in induction programs where more experienced educators observe their teaching, but these observations do not happen during every lesson that the new teachers teach. In higher education, the selection of graduate student instructors (GSI) frequently is not based on pedagogical expertise, but rather might be based on financial need of the GSI, or in limited instances demonstrated expertise with the subject matter. While GSI might receive post class student evaluations, it is rare for a GSI to get direct teaching support by a more experienced educator.

As the success of using digital video in sports improvement has grown, educators are beginning to look to video as a method for improving practice of teachers.