MOTIVATION

Background:
- Classifier performance degrades when training and testing conditions are different.
- Supervised domain adaptation is normally used to improve the base classifier’s performance.
- The performance increase depends on the data used for adaptation.

Proposed Solution:
- Active learning can be used to annotate the most useful samples to the classifier.
- Adjust hyperplane while maintaining learned information.
- Conservative approach that incrementally modifies the hyperplane with consistent samples.

Framework

The Problem

Domain Adaptation
Adapt SVM:
\[ f(x) = f^s(x) + \Delta f(x) \]
\[ = f^s(x) + \Delta w^T \phi(x) \]
\[ \min \frac{1}{2} ||\Delta w|| + C \sum_{i=1}^{N} \xi_i \]
\[ s.t. \xi_i \geq 0, y_i (f^s(x) + \Delta w^T \phi(x)) \geq 1 - \xi_i \]

Databases

Source: USC-IEMOCAP
- 12 hours of recordings
- Scripts and improvised scenarios
- Turns are annotated with emotions
- Angry, Happy, Sad and Neutral

Target: MSP-IMPROV
- Over 9 hours of recordings
- Improvised scenarios
- Turns are labeled with four emotions
- Angry, Happy, Sad and Neutral

Proposed Approach
Active learning
- Identify samples with low confidence
- Annotate samples

while stopping criteria is not met do
- Select subset \( N_a \) that the classifier predicted correctly
- Adapt classifier using subset \( N_a \)

Stopping Criteria
- Criterion 1: \( N_a \) doesn't contain labels of all classes
- Criterion 2: \( N_a \) contains labels of only one class
- Criterion 3: All samples are used

Experimental Settings and Results

Settings
- Interspeech 2013 feature set
- Feature Selection
- Correlation Feature selection 6373 \( \rightarrow \) 3000
- Forward Feature Selection 3000 \( \rightarrow \) 300
- SVM Classifier with a linear kernel
- Four class balanced classification problem
  - Angry, Happy, Sad, Neutral
  - Random under-sampling

Results
- Active learning select 200 samples
- Results are the average of 20 trials
- Baseline approach is adapting with all samples

When to stop
No adaptation: 45.5 %
Baseline approach: 46.7 %

<table>
<thead>
<tr>
<th>Criteria</th>
<th># Samples</th>
<th>F1 Score</th>
<th>After Adaptation</th>
<th># Iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st iteration</td>
<td>64.4</td>
<td>47.78 %</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Criterion 1</td>
<td>117.8</td>
<td>48.28 %</td>
<td>3.71</td>
<td></td>
</tr>
<tr>
<td>Criterion 2</td>
<td>123.6</td>
<td>48.13 %</td>
<td>4.71</td>
<td></td>
</tr>
<tr>
<td>Criterion 3</td>
<td>200</td>
<td>45.47 %</td>
<td>5.71</td>
<td></td>
</tr>
</tbody>
</table>

Key Point
Carefully selecting the samples used in adaptation yields better performance

DISCUSSION

Conclusions:
- Proposed an algorithm for incremental supervised SVM domain adaptation.
- We showed the importance of selecting the data used for adaptation.
- We used a portion of the labeled dataset, converging to a stable performance after 3 to 5 iterations.

Future Work:
- Modify the optimization function so that we can make use of all of the available data.
  - Introducing a variable regularization parameter for each instance.
- Extend the proposed algorithm to other classifiers.

Acknowledgment: This study was funded by a NSF CAREER award IIS-1453781