Introduction

Goal: Identify the speaker from an unknown audio signal.
Find the highest probability speaker matching an audio signal from a repository of known speakers.

Results

Model Performance

<table>
<thead>
<tr>
<th></th>
<th>Training</th>
<th>Dev</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>98%</td>
<td>97%</td>
<td>97%</td>
</tr>
<tr>
<td>Loss</td>
<td>0.76</td>
<td>0.81</td>
<td>0.81</td>
</tr>
</tbody>
</table>

Model & Hyperparameters

<table>
<thead>
<tr>
<th>Architecture</th>
<th>Mini Batch</th>
<th>Optimizer</th>
<th>Activations</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNN + FC</td>
<td>32</td>
<td>Adam</td>
<td>ReLu + SoftMax</td>
</tr>
</tbody>
</table>

Discussions

- The model predicts the speaker with remarkable accuracy.
- The prediction is quick and done in constant time. O(1).
- The prediction errors have high correlation with high noise levels.
- L2 Regularization helped reduce variance.

Future Work and Citations

- Study the impact of additional recording sources, i.e., video, open space, conferences.
- Test model against additional datasets and analyze the impact on performance.
- Algorithm applied to streaming audio signals.