Generative Approach Using Soft-Labels to Learn Uncertainties in Predicting Emotional Attributes
Kusha Sridhar, Carlos Busso

**Proposed Framework**
- Exploring annotator level uncertainties
- Generative modeling approach with soft-labels of emotional attributes
- Variational Autoencoder (VAE) with an Emotional Regressor (ER) attached to the bottleneck layer
- Multiple Monte Carlo from the latent space of the VAE to learn prediction uncertainties
- Soft-labels to train ER. Hard-labels to constrain the latent space of VAE

**Database and Features**
- The MSP-Podcast Corpus
  - Emotionally rich speaking turn from speaker appearing in various podcasts (2.75s to 11s in length)
  - Annotated for categorical and attribute-based emotion labels on Amazon Mechanical Turk
- Version 1.6: Train = 34,280 sentences
  - Test = 10,124 sentences from 50 speakers
  - Validation = 5,958 sentences from 40 speakers
- 42,567 sentences with speaker ID (1,078 speakers)

**Acoustic Features**
- Interspeech 2013 ComParE feature set extracted using Opensmile toolkit
- 65 LLDs and 6,373 HLDs

**Uncertainty Transfer Learning (UTL)**
- Exploring annotator level uncertainties

**Model Generalization**
- Cross-corpus experiments on the IEMOCAP corpus
  - Comparing proposed approach with other Monte Carlo Dropout (MCD) based approaches used for uncertainty modeling
  - Experiments with UTL in application to reject options

**Uncertainty Analysis and Reject Options**
- Novel generative modeling approach using soft-labels of emotional attributes in uncertainty predictions

**Conclusion**
- Can information from uncertainty prediction on one emotional attribute be transferred to another emotional attribute?
- Arousal and dominance uncertainties improve valence recognition performance but the vise versa is not true
- Transferred learned uncertainties lead to higher performance gains than self-learned uncertainties

**Gains up to 19.95% at 60% test coverage for valence**

<table>
<thead>
<tr>
<th>Attribute - Uncertainty</th>
<th>Approach</th>
<th>80%</th>
<th>60%</th>
<th>40%</th>
<th>20%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>VAE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proposed</td>
<td>0.5497</td>
<td>0.5717</td>
<td>0.6232</td>
<td>0.6433</td>
<td></td>
</tr>
<tr>
<td>MCD-NIL</td>
<td>0.5355</td>
<td>0.5661</td>
<td>0.6055</td>
<td>0.6459</td>
<td></td>
</tr>
<tr>
<td>MCD-nIL</td>
<td>0.5403</td>
<td>0.5700</td>
<td>0.6008</td>
<td>0.6452</td>
<td></td>
</tr>
<tr>
<td>MCD-sIL</td>
<td>0.5311</td>
<td>0.5407</td>
<td>0.5923</td>
<td>0.6304</td>
<td></td>
</tr>
<tr>
<td>MCD-nG</td>
<td>0.5349</td>
<td>0.5620</td>
<td>0.5725</td>
<td>0.6131</td>
<td></td>
</tr>
<tr>
<td><strong>VAE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proposed</td>
<td>0.5452</td>
<td>0.5700</td>
<td>0.6000</td>
<td>0.6300</td>
<td></td>
</tr>
<tr>
<td>MCD-NIL</td>
<td>0.5325</td>
<td>0.5500</td>
<td>0.5850</td>
<td>0.6090</td>
<td></td>
</tr>
<tr>
<td>MCD-nIL</td>
<td>0.5403</td>
<td>0.5700</td>
<td>0.6008</td>
<td>0.6304</td>
<td></td>
</tr>
<tr>
<td>MCD-sIL</td>
<td>0.5311</td>
<td>0.5407</td>
<td>0.5923</td>
<td>0.6304</td>
<td></td>
</tr>
<tr>
<td>MCD-nG</td>
<td>0.5349</td>
<td>0.5620</td>
<td>0.5725</td>
<td>0.6131</td>
<td></td>
</tr>
</tbody>
</table>

**Predicted attribute scores in terms of CCC as a function of prediction uncertainties**
- Entropy to quantify uncertainty

**Application in Reject Options for SER**
- At 60% test coverage, relative gains in CCC up to 16.85% for valence, 7.12% for arousal and 8.01% for dominance

**Computational complexity at inference**
- 74.90% faster than MCD based approaches

Uncertainty Analysis and Reject Options
- Novel generative modeling approach using soft-labels of emotional attributes in uncertainty predictions

**Experiments**
- Uncertainty Predictions
  - Using soft-labels
  - Reject Options using self-learned uncertainties
  - At 60% coverage, gains in CCC up to:
    - 16.85% for valence
    - 7.12% for arousal
    - 8.01% for dominance

**Uncertainty Transfer Learning**
- Works best with valence
- 19.95% gains in CCC

**Model Generalization**
- Cross-corpus results on IEMOCAP: Generalizes better than MCD based approaches

This work was funded by NSF career award IIS-1453781