**Zero-shot Learning with Verb Attribute Induction**

Rowan Zellers, Yejin Choi

University of Washington

---

**Verb attribute induction**

**Motivation:** to predict semantic and physical attributes of verbs (Fried and Palmer, 2014; Siegel and McKeown, 2000)

- Use GloVe vectors to model the dictionary definition
- Use a Bi-directional GRU to predict semantic and physical attributes of verbs

**New dataset: Verbs with Attributes**

We build a dataset of verb attributes containing 1711 verb-attribute pairs and 24 distinct attributes, motivated by verb semantics (Baker et al., 1998, Levin, 1993, Croft, 2012)

- **Attributes**
  - Verbal aspect
  - Temporal duration
  - Motion dynamics
  - Social dynamics
  - Transitivity
  - Post-conditions
  - Body parts
  - Emotional connotation

**Our model**

- Use a Bi-directional GRU to model the dictionary definition
- Use GloVe vectors to model verbs
- Pretraining and rebalancing critical

**Conclusion**

- Attributes as inductive bias towards zero-shot learning
- First results on two challenging tasks
- Dictionaries and embeddings have complementary information
- Attributes help improve zero-shot performance
- More attributes could be categorized and used
- End-to-end attribute learning desirable

**Download at:**

[github.com/uwnlp/verb-attributes](https://github.com/uwnlp/verb-attributes)

---

**Zero-shot activity recognition**

**Motivation:** to recognize novel verb-based activities at test time via attribute prediction. This is commonly done for object recognition (Krause et al., 2009, Arifin et al., 2016), however activity attributes are more difficult (Wang et al., 2014).

**Our model**

- Use ResNet152 as the underlying CNN (He et al., 2015)
- Predict attribute and embedding representations and add the resulting label distributions

**MFC**

- Top-k attribute prediction
- Attribute lookup

**DeViSE**

- Top-1 activity prediction

**Ours**

- Top-1 activity prediction

**Table:**

<table>
<thead>
<tr>
<th>Model</th>
<th>Acc</th>
<th>Using Att WV</th>
<th>DeViSE</th>
<th>Ours</th>
<th>DeViSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most Freq Class</td>
<td>61.33</td>
<td>×</td>
<td>10.50</td>
<td>17.60</td>
<td>18.10</td>
</tr>
<tr>
<td>CAAP</td>
<td>68.15</td>
<td>×</td>
<td>37.56</td>
<td>39.29</td>
<td>41.46</td>
</tr>
<tr>
<td>GloVe</td>
<td>66.60</td>
<td>×</td>
<td>4.79</td>
<td>19.98</td>
<td>18.10</td>
</tr>
<tr>
<td>BiGRU</td>
<td>66.05</td>
<td>×</td>
<td>4.79</td>
<td>19.98</td>
<td>18.10</td>
</tr>
<tr>
<td>NBoW</td>
<td>65.41</td>
<td>×</td>
<td>4.79</td>
<td>19.98</td>
<td>18.10</td>
</tr>
<tr>
<td>NBoW + GloVe</td>
<td>67.32</td>
<td>×</td>
<td>4.79</td>
<td>19.98</td>
<td>18.10</td>
</tr>
<tr>
<td>BiGRU + GloVe</td>
<td>68.43</td>
<td>×</td>
<td>4.79</td>
<td>19.98</td>
<td>18.10</td>
</tr>
</tbody>
</table>