Building NLP Classifiers Cheaply With Transfer Learning + Weak Supervision

**Problem**

- Text classification models are very useful, but there’s a high initial and recurring labeling cost.
- Both Weak Supervision and Transfer Learning have proved to be very successful.
- The challenge is how to combine them to enable training of text models efficiently without AI engineers in the loop. We study this question on an important problem: hate speech and anti-semitism.

**Data / Task**

- Train: 24,738 unlabeled
- Dev: 733
- Test: 438
- Definitely not an easy task!

**Approach**

- We have evidence that weak supervision combined with transfer learning helps build text classifiers cheaply.
- For future work we would try other transfer learning approaches like fine-tuning Bert and the OpenAI Transformer.

**Analysis**

- Weak labels increase ULMFiT ROC-AUC of by +0.15
- We were able to train a good classifier with just 1,000 labels
- Error Analysis: 65% of errors are because tweet ends in “jew”

**Results**

<table>
<thead>
<tr>
<th>Method</th>
<th>Precision</th>
<th>Recall</th>
<th>ROC-AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label Model</td>
<td>80%</td>
<td>23%</td>
<td>0.73</td>
</tr>
<tr>
<td>Logistic Regression</td>
<td>94%</td>
<td>12%</td>
<td>0.77</td>
</tr>
<tr>
<td>XGBoost</td>
<td>94%</td>
<td>12%</td>
<td>0.76</td>
</tr>
<tr>
<td>Feed Forward NN</td>
<td>94%</td>
<td>11%</td>
<td>0.70</td>
</tr>
</tbody>
</table>

**Classifier Baselines**

- ULMFiT with Weak Supervision: 94% Precision / 39% Recall
- ULMFiT Only: ROC-AUC: 0.654170615532118
- ULMFiT with Weak Supervision: ROC-AUC: 0.8066116670278043

**Conclusions**

- We have evidence that weak supervision combined with transfer learning helps build text classifiers cheaply.
- For future work we would try other transfer learning approaches like fine-tuning Bert and the OpenAI Transformer.

**References**


Abraham Starosta
starosta@stanford.edu