**Motivation**

Declarative supervision from NL can enable learning with limited or no labeled examples

- Show my important emails.
- What are my important emails?
- If the subject says 'urgent', it is almost certainly important.
- Most emails from John are important.
- Emails that I reply to are usually important.
- Unimportant emails are often sent to a list.

- Quantifier adjectives and adverbs are explicit denoters of generality.
- Use semantics of quantifiers to drive classifier training.

**Idea**

- NL encodes key properties that aid statistical learning.
  - "Emails that I reply to are usually important"

  1. Features important for a learning problem
     - $x$: repliedTo=true
   2. Class labels
      - $y$: Important
   3. Type of Relationship b/w features and labels
      - $P(y|x)$
   4. Strength of Relationship
      - Specified by quantifier?

- Convert statements to quantitative assertions
  - $P(label|important | replied=true) \approx \hat{p}_{\text{weekly}}$

- Simplest model: approximate quantifiers as point probabilities
  - Pre-register estimates (purely subjective)

<table>
<thead>
<tr>
<th>Frequency quantifier</th>
<th>Probability value</th>
</tr>
</thead>
<tbody>
<tr>
<td>always, certainly, definitely, all</td>
<td>0.95</td>
</tr>
<tr>
<td>usually, normally, generally, likely</td>
<td>0.70</td>
</tr>
<tr>
<td>mostly, majority</td>
<td>0.60</td>
</tr>
<tr>
<td>often, half</td>
<td>0.50</td>
</tr>
<tr>
<td>sometimes, frequently, some, many</td>
<td>0.30</td>
</tr>
<tr>
<td>few, occasionally</td>
<td>0.20</td>
</tr>
<tr>
<td>never, rarely</td>
<td>0.05</td>
</tr>
</tbody>
</table>

**Approach**

- Mapping Language to Constraints
  - Identify key properties to convert each NL explanation of a concept to a quantitative assertion
    - $P(f|s) = P(l_{xy}|s) P(l_{type}|l_{xy},s) P(l_{quant}|s)$
  - Classifier for constraint types: $P(y|x), P(x|y) \& P(y)$
    - Based on syntactic and dependency parse features

- Each assertion acts as a constraint during model training.

- Classifier Training via Posterior Regularization
  - PR imbeds human-provided advice in learned models

  **Example:** have predictions from the classifier agree with NL advice

**Evaluation**

- Datasets: Three domains, include synthetic and real concepts

- **Classification Performance**

<table>
<thead>
<tr>
<th>Approach</th>
<th>Avg Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bayes Optimal</td>
<td>0.831</td>
</tr>
<tr>
<td>LNQ</td>
<td>0.754</td>
</tr>
<tr>
<td>LR (n=10)</td>
<td>0.757</td>
</tr>
<tr>
<td>LNQ (no-quantification)</td>
<td>0.679</td>
</tr>
<tr>
<td>LNQ (no-quantification)</td>
<td>0.545</td>
</tr>
<tr>
<td>Human learner</td>
<td>0.754</td>
</tr>
</tbody>
</table>

- Learning from Natural Quantification (LNQ) consistently achieves performance comparable with learning from a small number of labeled examples (LR with n=10).
- Learning is due to differential associative strength of quantifiers. For Shape tasks, LNQ is competitive with humans learning from NL advice.

- Empirical semantics of quantifiers

  - Empirical values may differ significantly from pre-registered beliefs, or show large spreads (not meaningfully modeled as point probability values).
  - LNQ is robust to changes in values of probability estimates.