Neural networks
Sparse coding - feature extraction
FEATURE EXTRACTION

**Topics:** feature learning

- A sparse coding model can be used to extract features
  - given a labeled training set \( \{(x^{(t)}, y^{(t)})\} \)
  - train sparse coding dictionary only on training inputs \( \{x^{(t)}\} \)
    - this yields a dictionary \( D \) from which to infer sparse codes \( h(x^{(t)}) \)
  - train favorite classifier on transformed training set \( \{(h(x^{(t)}), y^{(t)})\} \)

- When classifying test input \( x \), must infer its sparse representation \( h(x) \) first, then feed it to the classifier
FEATURE EXTRACTION

**Topics:** feature learning

- When trained on handwritten digits:

Self-taught Learning: Transfer Learning from Unlabeled Data
Raina, Battle, Lee, Packer and Ng.
**Topics:** self-taught learning

- **Self-taught learning:**
  - when features trained on different input distribution
- **Example:**
  - train sparse coding dictionary on handwritten digits
  - use codes (features) to classify handwritten characters

<table>
<thead>
<tr>
<th>Digits → English handwritten characters</th>
<th>Training set size</th>
<th>Raw</th>
<th>PCA</th>
<th>Sparse coding</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100</td>
<td>39.8%</td>
<td>25.3%</td>
<td>39.7%</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>54.8%</td>
<td>54.8%</td>
<td>58.5%</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>61.9%</td>
<td>64.5%</td>
<td>65.3%</td>
</tr>
</tbody>
</table>

Self-taught Learning: Transfer Learning from Unlabeled Data
Raina, Battle, Lee, Packer and Ng.