Neural networks

Sparse coding - feature extraction

FEATURE EXTRACTION

Topics: feature learning

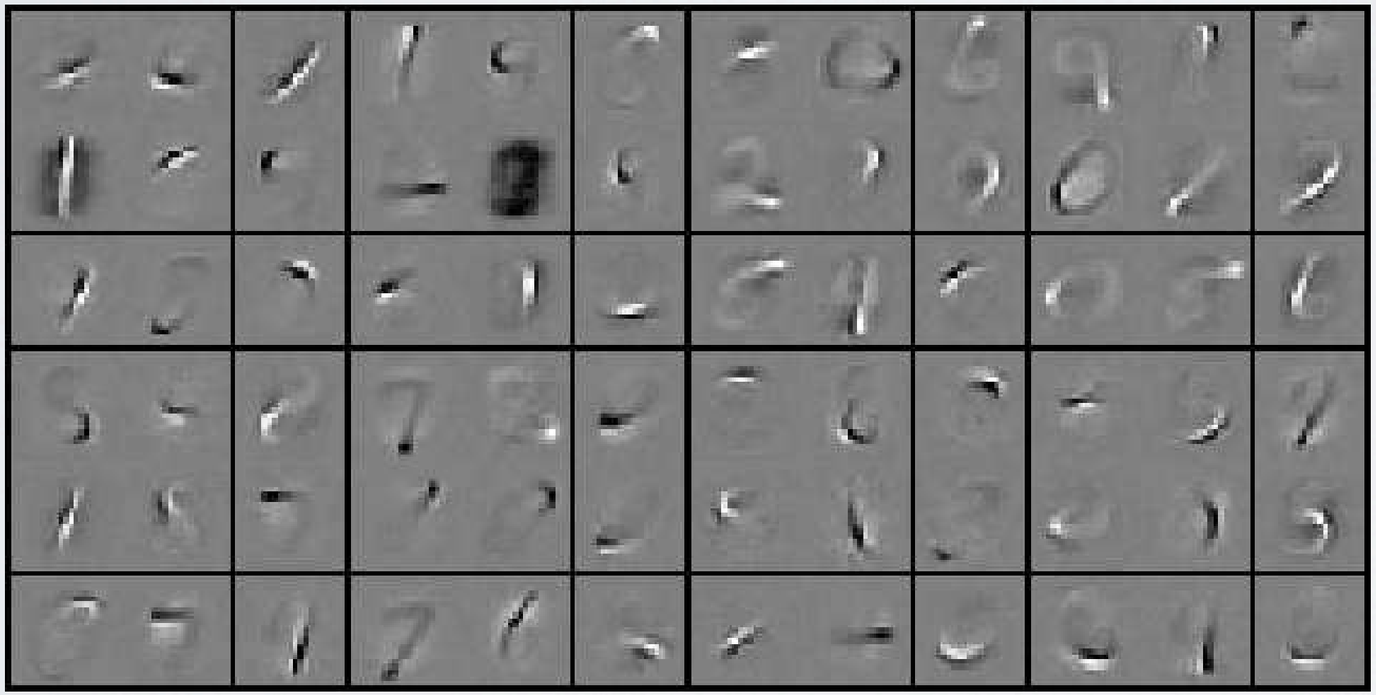
- A sparse coding model can be used to extract features
  - given a labeled training set $\{(\mathbf{x}^{(t)}, y^{(t)})\}$
  - train sparse coding dictionary only on training inputs $\{\mathbf{x}^{(t)}\}$
    - this yields a dictionary $\mathbf{D}$ from which to infer sparse codes $\mathbf{h}(\mathbf{x}^{(t)})$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

- When classifying test input $\mathbf{x}$, must infer its sparse representation $\mathbf{h}(\mathbf{x})$ 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.

FEATURE EXTRACTION

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

Digits → English handwritten characters			
Training set size	Raw	PCA	Sparse coding
100	**39.8%**	25.3%	39.7%
500	54.8%	54.8%	**58.5%**
1000	61.9%	64.5%	**65.3%**

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