RBM Classifier for MNIST

ISPR - Midterm 2
Assignment 3

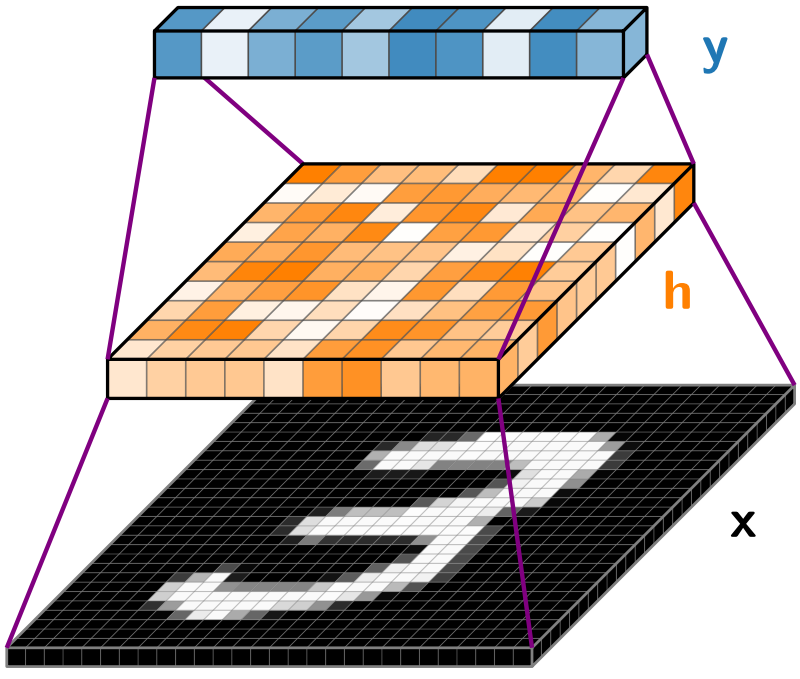
Filippo Baroni

May 3, 2021

RBM Classifier

Model

- Visible units $\mathbf{x}$ (input).
  Size: $28 \times 28 \implies 784$.
- Half-visible units $\mathbf{y}$ (output).
  Size: 10.
- Hidden units $\mathbf{h}$.
  Size: 100.



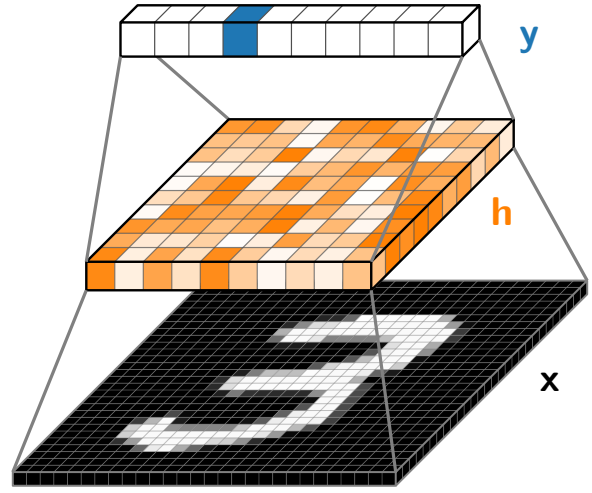
Energy

$$E(\mathbf{x}, \mathbf{y}, \mathbf{h}) = -\mathbf{h}^t \mathbf{W} \mathbf{x} - \mathbf{h}^t \mathbf{U} \mathbf{y} - \\ - \mathbf{b}^t \mathbf{x} - \mathbf{c}^t \mathbf{h} - \mathbf{d}^t \mathbf{y}$$

Training and Prediction

Training

- Treat **x** and **y** as visible units.
- Use one-hot encoding for **y**.
- Contrastive divergence, 1 step.

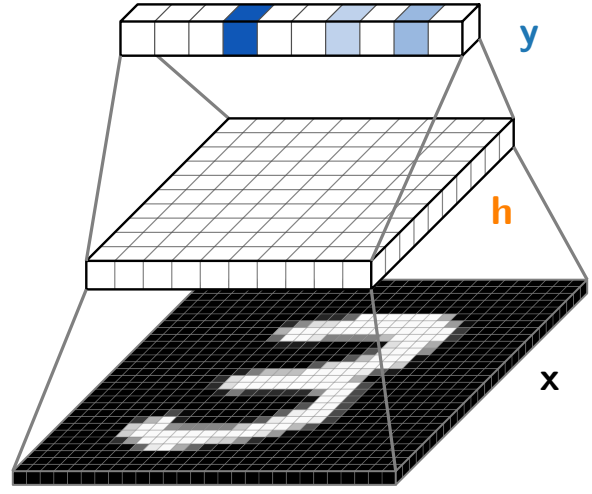


Prediction

- Treat **y** as hidden units.
- Compute the exact conditional distribution

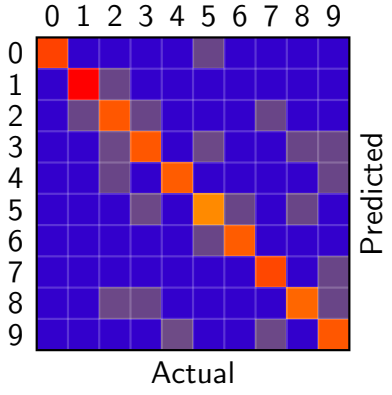
$$p(y|\mathbf{x}) \propto e^{d_y} \prod_{j=0}^{9} \left(1 + e^{c_j + U_{jy} + \sum_i W_{ji} x_i}\right).$$

- Pick the digit y^* that maximizes $p(y^*|\mathbf{x})$.



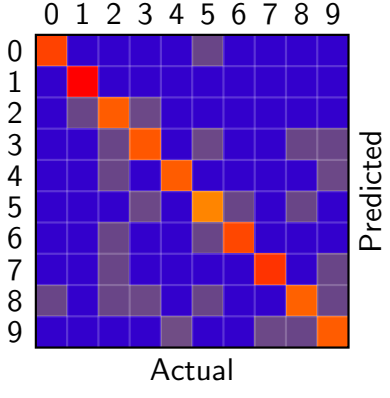
Results

Test



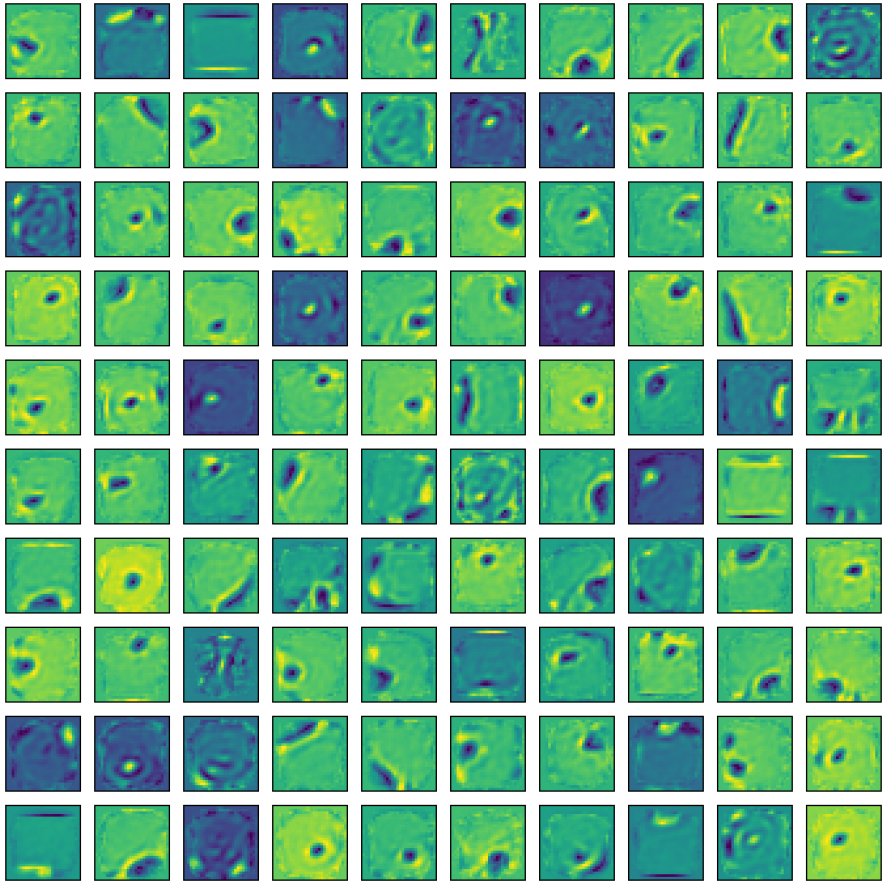
Accuracy:
93.1%

Train



Accuracy:
92.8%

Neuron activations



Experiments

j