Lecture 10 – Multi-Class Classification

Dan Sheldon

October 15, 2012

Today

- Decision-trees odds and ends
  - Pruning
  - Why entropy works
- Multi-class classification

A Real Classification Problem

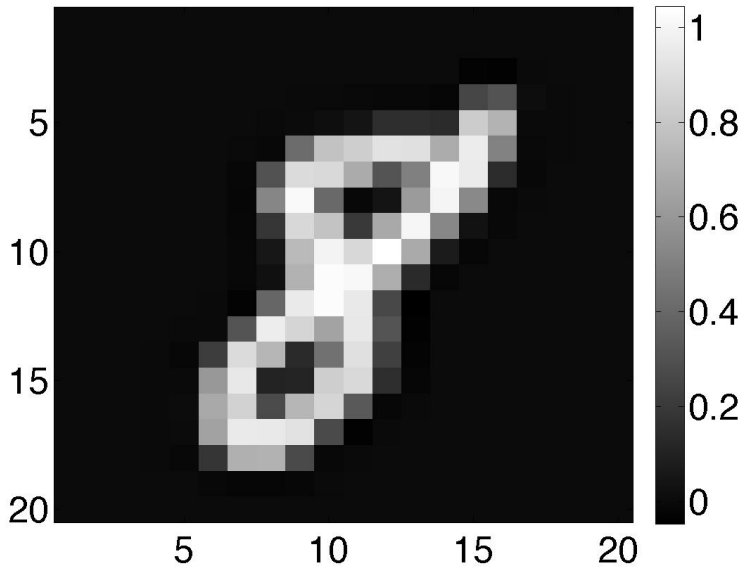
Classify handwritten digits.

$$y \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$$

We don't know how to solve this yet

Hand-written digit classification

Input: 20×20 grayscale image



Unroll the image into a vector

$$\begin{bmatrix} x_1 & x_{21} & \dots & x_{381} \\ x_2 & x_{22} & \dots & x_{382} \\ & & \vdots & \\ x_{20} & x_{40} & \dots & x_{400} \end{bmatrix}$$

Feature vector $\mathbf{x} \in \mathbb{R}^{400}$

$$\mathbf{x} = (x_1, \dots, x_{400})^T$$

Multi-class Classification

Input: $\mathbf{x} \in \mathbb{R}^d$ (continuous or discrete)
Labels: $y \in \{1, \ldots, K\}$

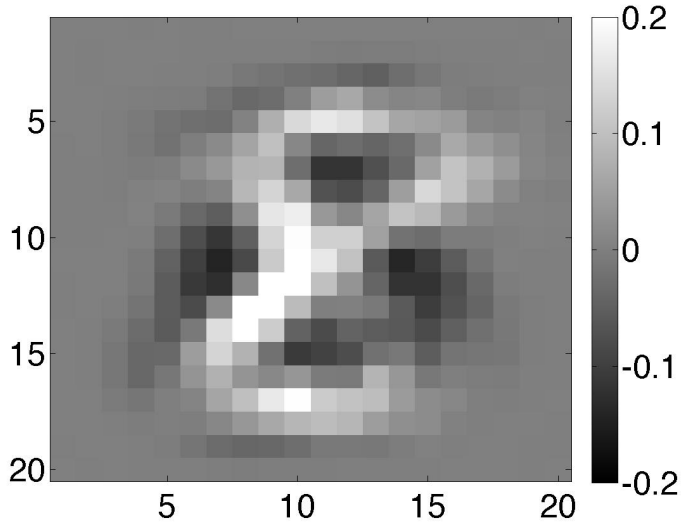
Group 1: Decision Trees

Group 2: Logistic regression
(Use one or more binary $(y \in \{0, 1\})$ classifiers)

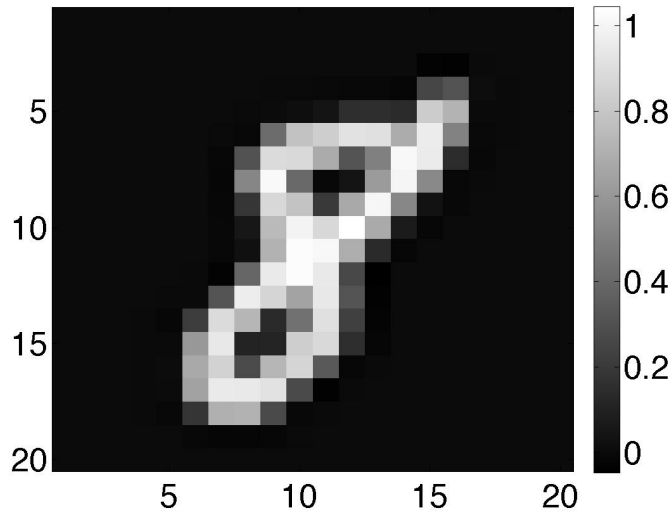
Hint: think about prediction first, then training.

Visualization

Format weight vector as an image:



$\mathbf{w}_8$



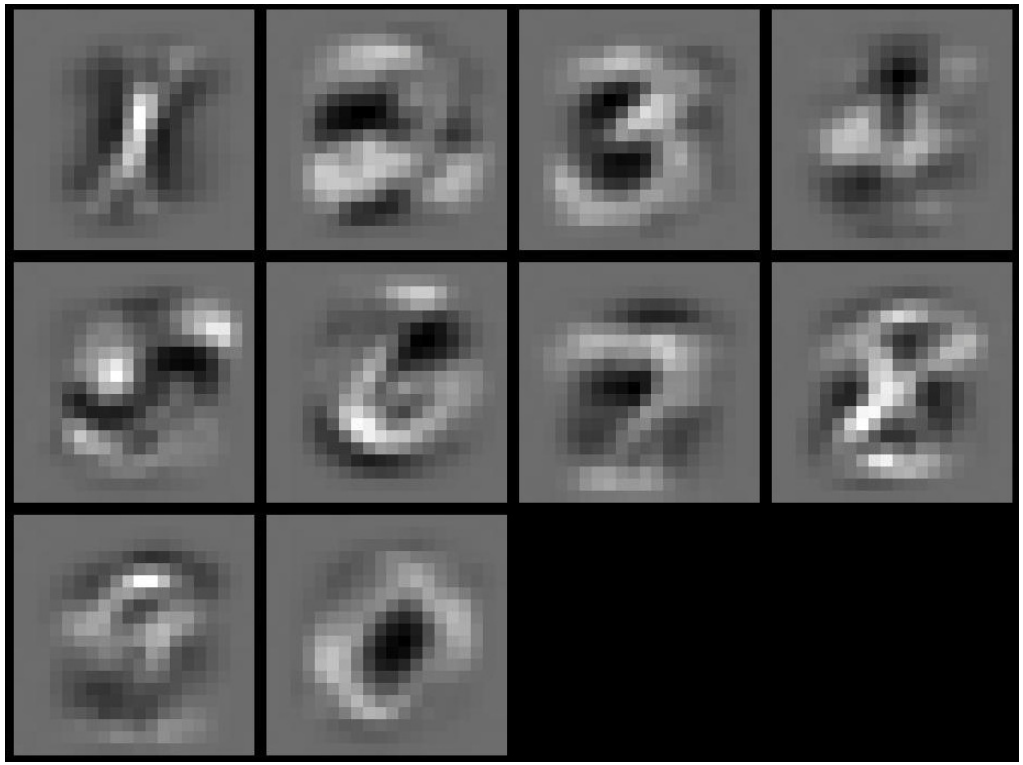
$\mathbf{x}$

Recall that

$$\text{Prediction} = \begin{cases} 1 & \mathbf{w}^T\mathbf{x} \geq 0 \\ 0 & \mathbf{w}^T\mathbf{x} < 0 \end{cases}$$

Dot product = multiply together corresponding pixels and add

Visualization: One vs. All



Visualization: One vs. One