

Administrivia - Background

- Calculus:

$$E = mc^2 \Rightarrow \frac{\partial E}{\partial c} = 2mc$$

- Linear algebra:

$$\mathbf{A}\mathbf{u}_i = \lambda_i\mathbf{u}_i; \quad \frac{\partial}{\partial \mathbf{x}}(\mathbf{x}^T \mathbf{a}) = \mathbf{a}$$

- See PRML Appendix C

- Probability:

$$p(X) = \sum_Y p(X, Y); \quad p(x) = \int p(x, y) dy; \quad \mathbb{E}_x[f] = \int p(x) f(x) dx$$

What is Machine Learning (ML)?

- Algorithms that automatically improve performance through experience
- Often this means define a model by hand, and use data to fit its parameters

Why ML?

- The real world is complex – difficult to hand-craft solutions.
- ML is the preferred framework for applications in many fields:
 - Computer Vision
 - Natural Language Processing, Speech Recognition
 - Robotics
 - ...

Hand-written Digit Recognition



Belongie et al. PAMI 2002

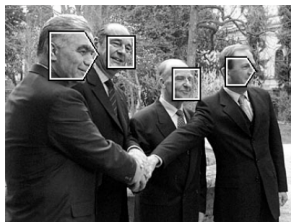
- Difficult to hand-craft rules about digits

Hand-written Digit Recognition

$$\mathbf{x}_i = \begin{array}{|c|} \hline \text{4} \\ \hline \end{array} \quad \mathbf{t}_i = (0, 0, 0, 1, 0, 0, 0, 0, 0, 0)$$

- Represent input image as a vector $\mathbf{x}_i \in \mathbb{R}^{784}$.
- Suppose we have a target vector \mathbf{t}_i
 - This is **supervised learning**
 - Discrete, finite label set: perhaps $\mathbf{t}_i \in \{0, 1\}^{10}$, a **classification** problem
- Given a **training set** $\{(\mathbf{x}_1, \mathbf{t}_1), \dots, (\mathbf{x}_N, \mathbf{t}_N)\}$, learning problem is to construct a “good” function $\mathbf{y}(\mathbf{x})$ from these.
 - $\mathbf{y} : \mathbb{R}^{784} \rightarrow \mathbb{R}^{10}$

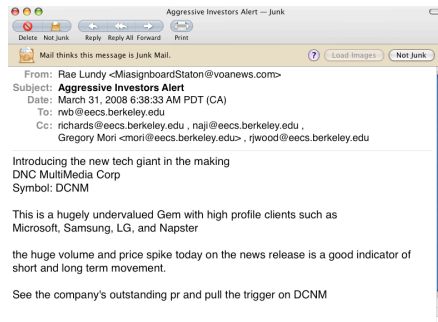
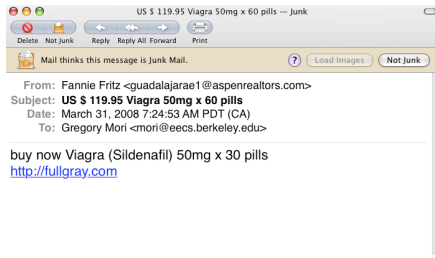
Face Detection



Schneiderman and Kanade, IJCV 2002

- **Classification** problem
- $t_i \in \{0, 1, 2\}$, non-face, frontal face, profile face.

Spam Detection



- **Classification** problem
- $t_i \in \{0, 1\}$, non-spam, spam
- x_i counts of words, e.g. Viagra, stock, outperform, multi-bagger