**Machine Learning**

Machine Learning: algorithms that use "experience" to improve their performance

We use machine learnning situations where it is very challenging (or impossible) to define the rules by hand: e.g.

• face detection
• speech recognition
• stock prediction
• driving a car
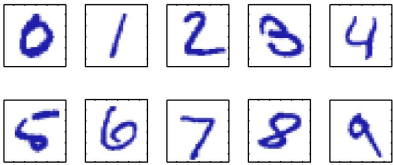• medical diagnosis

1

---

**Machine Learning**

Machine Learning: write programs with thousands/millions of undefined constants.

Learn through experience how to set those constants.

Machine learning algorithms are getting better and better and better…..

2

---

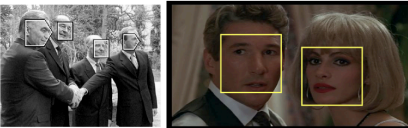Example 1: hand-written digit recognition



Images are 28 x 28 pixels

Represent input image as a vector $\mathbf{x} \in \mathbb{R}^{784}$
Learn a classifier $f(\mathbf{x})$ such that,
$$f : \mathbf{x} \to \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$$
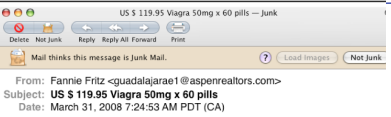
3

---

Example 2: Face detection



• Again, a supervised classification problem
• Need to classify an image window into three classes:
  • non-face
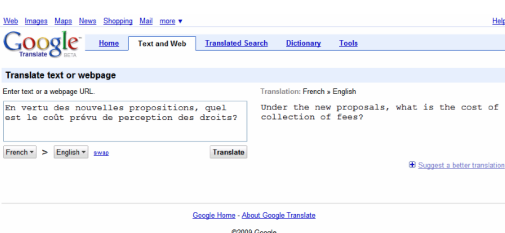  • frontal-face
  • profile-face

21

---

Example 3: Spam detection



• This is a classification problem
• Task is to classify email into spam/non-spam
• Data $x_i$ is word count, e.g. of viagra, outperform, "you may be surprised to be contacted" …
• Requires a learning system as "enemy" keeps innovating

5

---

**Example 4: Machine translation**



**What is the anticipated cost of collecting fees under the new proposal?**

6

---
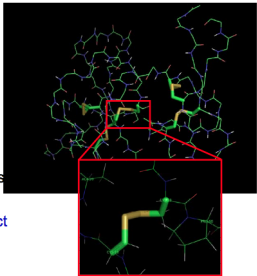
## Example 5: Computational biology

x                                    y

AVITGA**C**ERDLQ**C**G
KGT**CC**AVSLWIKSV
RV**C**TPVGTSGED**C**H
PASHKIPFSGQRMH
HT**C**P**C**APNLA**C**VQT
SPKKFK**C**LSK

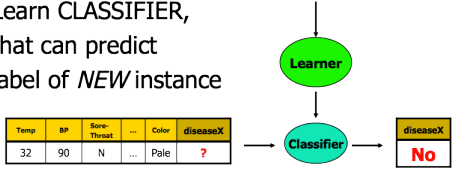Protein Structure and Disulfide Bridges

Regression task: given sequence predict
3D structure

Protein: 1IMT

---

- Given "labeled data"

| Temp. | BP. | Sore Throat | ... | Colour | diseaseX |
|-------|-----|-------------|-----|--------|----------|
| 35 | 95 | Y | ... | Pale | No |
| 22 | 110 | N | ... | Clear | Yes |
| : | : | | | : | : |
| 10 | 87 | N | ... | Pale | No |

- Learn CLASSIFIER,
  that can predict
  label of *NEW* instance

**Learner**

**Classifier**

| Temp | BP | Sore-Throat | ... | Color | diseaseX |
|------|-----|-------------|-----|-------|----------|
| 32 | 90 | N | ... | Pale | ? |

diseaseX
**No**

8

---

Use random variables to represent everything about the world

Space of possible random variables and classifiers indexed by
parameters  which are knobs we turn to create different
classifiers.

**Learning:  the problem of estimating joint
probability density functions, tuning the knobs,
given samples from the function.**

9

---

**This is huge!**

growing flood of online data

recent progress in algorithms and theoretical foundations

computational power

never-ending industrial applications.

10