

Lecture 1

What is Machine Learning?

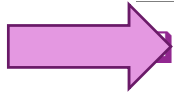
EE4563/ EL9123: INTRODUCTION TO MACHINE LEARNING

PROF. SUNDEEP RANGAN

Learning Objectives

- ❑ Provide examples of machine learning used today
- ❑ Given a new problem, qualitatively describe how machine learning can be used
 - Formulate a potential machine learning task
 - Identify the data needed for the task
 - Identify objectives
- ❑ Classify a machine learning task:
 - Supervised vs. unsupervised, regression vs. classification
- ❑ For supervised learning, identify the predictors and target variables
- ❑ Determine the role of expert knowledge in the task vs. data-driven learning

Outline



What is Machine Learning?

- Types of machine learning algorithms

- Classification
- Regression
- Unsupervised learning

- Why the hype today?

- Some slides from:

- A. Zisserman, “Machine Learning Introduction”
- Alpaydin, “Introduction to Machine Learning”

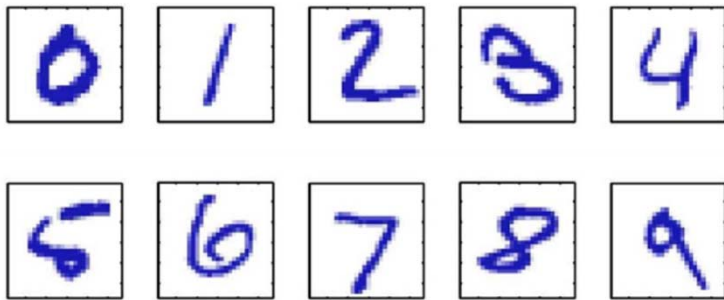
What is Machine Learning?

□ Learn to improve algorithms from data.

□ Why?

- Human expertise does not exist (navigating on Mars),
- Humans are unable to explain their expertise (speech recognition)
- Solution changes in time (routing on a computer network)
- Solution needs to be adapted to particular cases (user biometrics)

Example 1: Digit Recognition

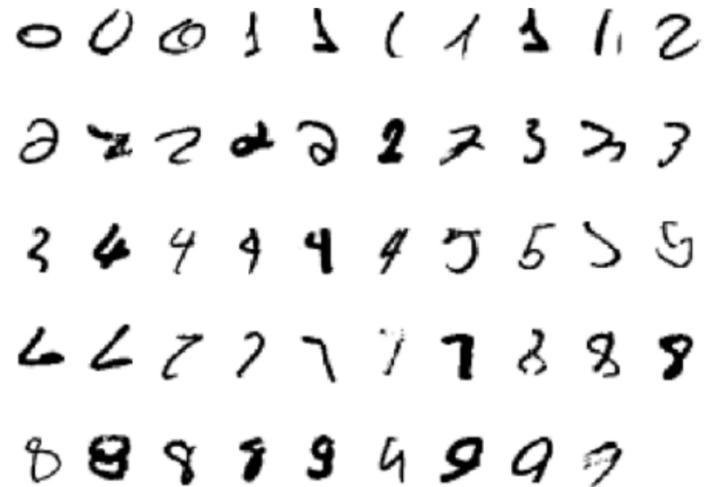


Images are 28 x 28 pixels

- ❑ Recognize a digit from the image
- ❑ Learn a function $f(x) \in \{0, 1, \dots, 9\}$, x is a 28 x 28 matrix
- ❑ Expert systems do not work well:
 - You can recognize the digits, but difficult to program a function $f(x)$ that works well
 - Try it!

Supervised Learning

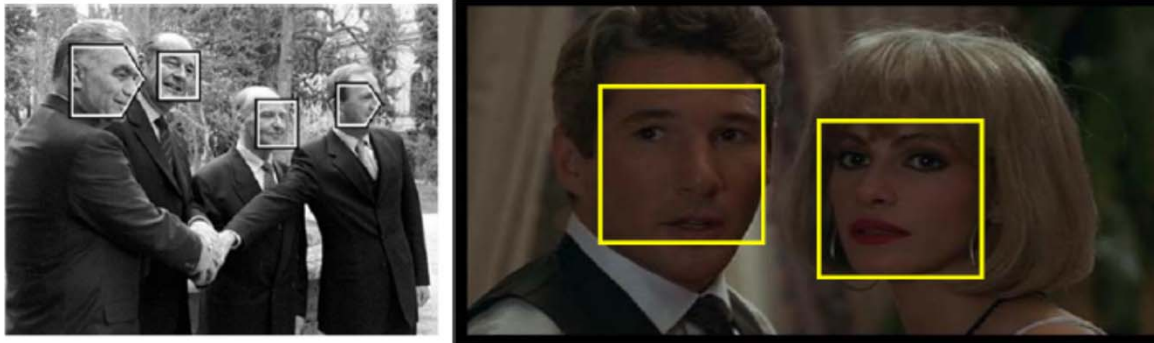
- Start with training data
- Ex: 6000 examples of each digit
- Learn a classifier $f(x)$ that matches label well on training data
- Given new data x use function to guess digit
- Current systems get <0.4% errors
- First commercial application:
 - Used by USPS for recognized zip codes on letters



Training examples

Each sample must be labeled by hand who knows truth

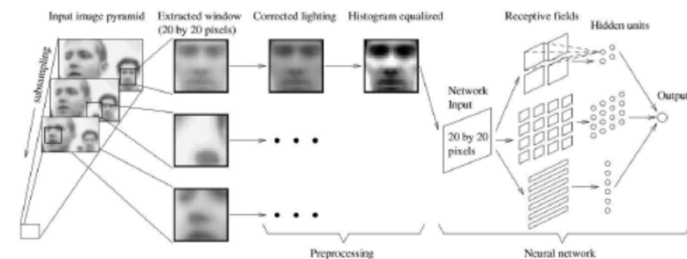
Example 2: Face Detection



- ❑ Also a supervised learning problem
- ❑ For each image region, determine if
 - Face or non-face

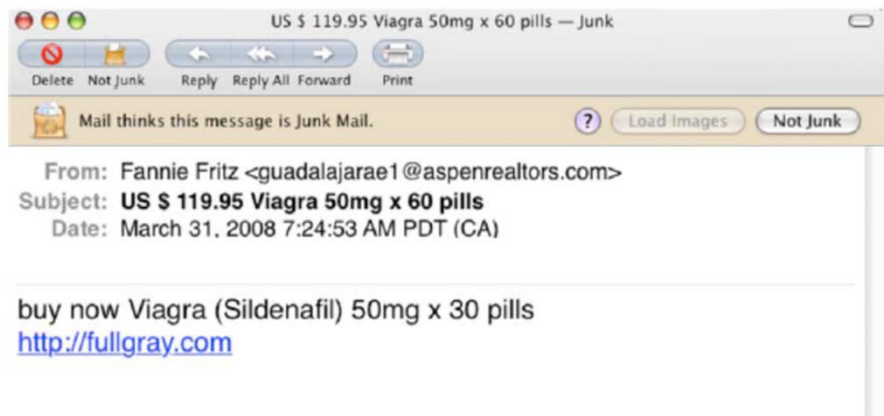
Training Data

- ❑ Typical early face recognition datasets:
 - ❑ 5000 faces
 - All near frontal
 - Vary age, race, gender, lighting
 - ❑ 10^8 non faces
 - ❑ Faces are normalized (scale, translation)
 - ❑ “functions” that work well may be very complex
- ❑ Many more datasets are available now:
 - See <http://www.face-rec.org/databases/>
 - You can use this for your project!



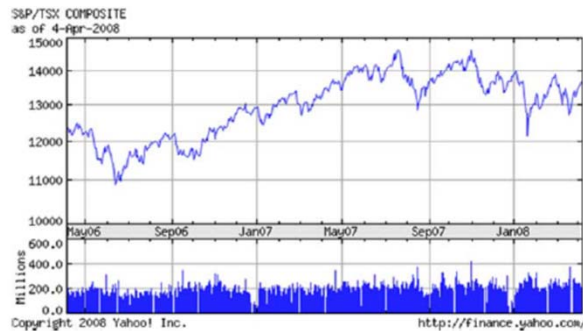
Rowley, Baluja and Kanade, 1998

Example 3: Spam Detection



- ❑ Classification problem:
 - Is email junk or not junk?
- ❑ For ML, must represent email numerically
 - Common model: bag of words
 - Enumerate all words, $i = 1, \dots, N$
 - Represent email via word count
 $x_i = \text{num instances of word } i$
- ❑ Challenge:
 - Very high-dimensional vector
 - System must continue to adapt (keep up with spammers)

Example 4: Stock Price Prediction



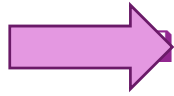
- ☐ Can you predict the price of a stock?
- ☐ What variables would you use?
- ☐ What is a non-machine learning approach?

Machine Learning in Many Fields

- ❑ **Retail:** Market basket analysis, Customer relationship management (CRM)
- ❑ **Finance:** Credit scoring, fraud detection
- ❑ **Manufacturing:** Control, robotics, troubleshooting
- ❑ **Medicine:** Medical diagnosis
- ❑ **Telecommunications:** Spam filters, intrusion detection
- ❑ **Bioinformatics:** Motifs, alignment
- ❑ **Web mining:** Search engines
- ❑ ...

Outline

□ What is Machine Learning?



□ Types of machine learning algorithms

- Classification
- Regression
- Unsupervised learning

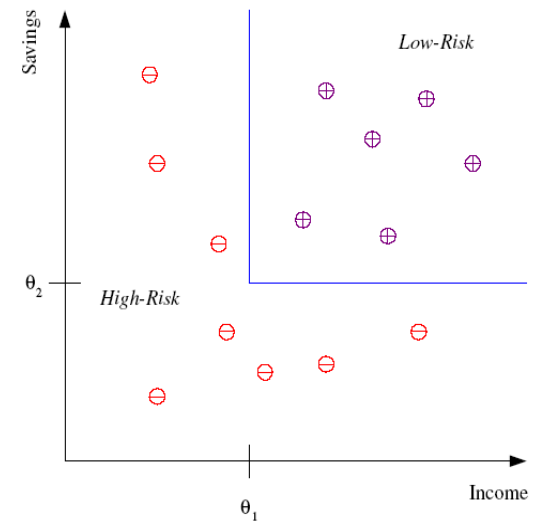
□ Why the hype today?

□ Some slides from:

- A. Zisserman, “Machine Learning Introduction”
- Alpaydin, “Introduction to Machine Learning”

Classification

- ❑ Example: Credit score
- ❑ Determine if customer is high-risk or low-risk
- ❑ Select some **features**:
 - Example: income & savings
 - Represent as a vector $x = (x_1, x_2)$
- ❑ Learn a function from **features** to **target**
 - Use past training data
 - Need to get this data
- ❑ This function is an example of a **decision tree**.



Regression

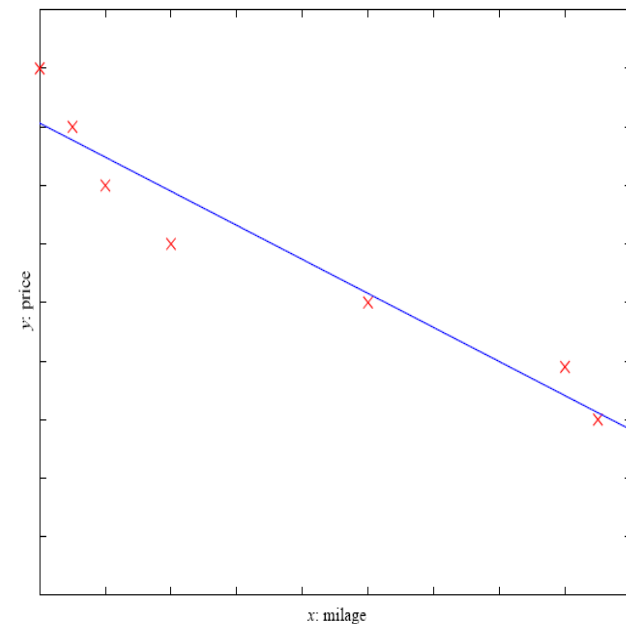
□ Target variable y is continuous-valued

□ Example:

- Predict y = price of car
- From x = mileage, size, horsepower, ..
- Can use multiple predictor

□ Assume some form of the mapping

- Ex. Linear: $y = \beta_0 + \beta_1 x$
- Find parameters β_0, β_1 from data



Regression Example

Machine Learning Repository

Center for Machine Learning and Intelligent Systems

Diabetes Data Set

Download: [Data Folder](#), [Data Set Description](#)

File Names and format:

- (1) Date in MM-DD-YYYY format
- (2) Time in XX:YY format
- (3) Code
- (4) Value

The Code field is deciphered as follows:

33 = Regular insulin dose
34 = NPH insulin dose
35 = UltraLente insulin dose
48 = Unspecified blood glucose measurement
57 = Unspecified blood glucose measurement
58 = Pre-breakfast blood glucose measurement
59 = Post-breakfast blood glucose measurement
60 = Pre-lunch blood glucose measurement
61 = Post-lunch blood glucose measurement
62 = Pre-supper blood glucose measurement
63 = Post-supper blood glucose measurement
64 = Pre-snack blood glucose measurement
65 = Hypoglycemic symptoms
66 = Typical meal ingestion
67 = More-than-usual meal ingestion
68 = Less-than-usual meal ingestion
69 = Typical exercise activity
70 = More-than-usual exercise activity
71 = Less-than-usual exercise activity
72 = Unspecified exercise activity

❑ Predict blood glucose level

❑ Many possible predictors:

- Recent past levels
- Insulin dose
- Time of last meal
- ...

❑ Check out data in:

<https://archive.ics.uci.edu/ml/datasets/Diabetes>

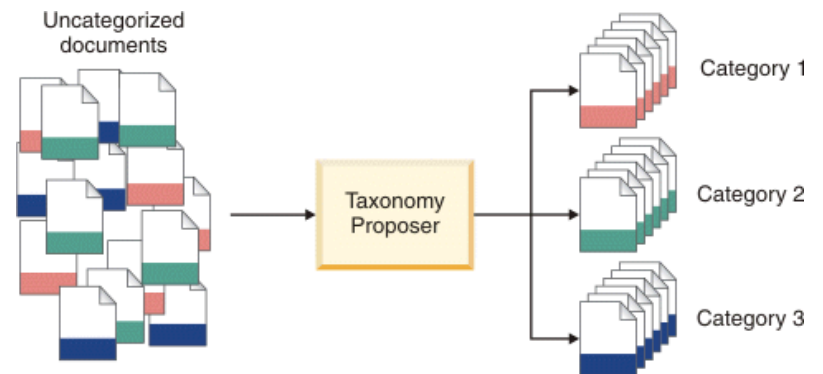


NYU

TANDON SCHOOL
OF ENGINEERING

Unsupervised Learning

- ❑ Learning “what normally happens”
- ❑ No output
- ❑ Clustering: Grouping similar instances
- ❑ Example applications
 - Customer segmentation
 - Image compression: Color quantization
 - Bioinformatics: Learning motifs



Example: Document classification

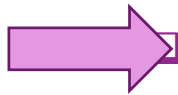
http://www.ibm.com/support/knowledgecenter/SSBRAM_8.7.0/com.ibm.classify.ccenter.doc/c_WBG_Taxonomy_Proposer.htm

Outline

- What is Machine Learning?

- Types of machine learning algorithms

- Classification
- Regression
- Unsupervised learning



- Why the hype today?

- Some slides from:

- A. Zisserman, “Machine Learning Introduction”
- Alpaydin, “Introduction to Machine Learning”

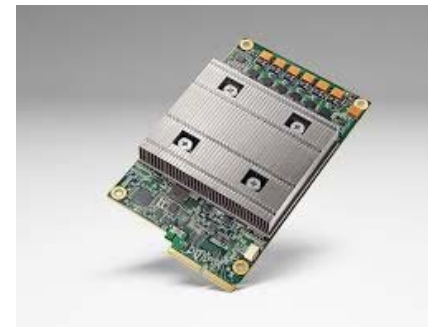
What ML is Doing Today?

- ❑ Autonomous driving
- ❑ Jeopardy
- ❑ Very difficult games: Alpha Go
- ❑ Machine translation
- ❑ Many, many others...



Why Now?

- ❑ Machine learning is an old field
 - Much of the pioneering statistical work dates to the 1950s
- ❑ So what is new now?
- ❑ Big Data:
 - Massive storage. Large data centers
 - Massive connectivity
 - Sources of data from Internet and elsewhere
- ❑ Computational advances
 - Distributed machines, clusters
 - GPUs and hardware



Google Tensor Processing Unit (TPU)

Journals

- ☐ Journal of Machine Learning Research www.jmlr.org
- ☐ Machine Learning
- ☐ Neural Computation
- ☐ Neural Networks
- ☐ IEEE Trans on Neural Networks and Learning Systems
- ☐ IEEE Trans on Pattern Analysis and Machine Intelligence
- ☐ Journals on Statistics/Data Mining/Signal Processing/Natural Language Processing/Bioinformatics/...

Conferences

- ❑ International Conference on Machine Learning (ICML)
- ❑ European Conference on Machine Learning (ECML)
- ❑ Neural Information Processing Systems (NIPS)
- ❑ Uncertainty in Artificial Intelligence (UAI)
- ❑ Computational Learning Theory (COLT)
- ❑ International Conference on Artificial Neural Networks (ICANN)
- ❑ International Conference on AI & Statistics (AISTATS)
- ❑ International Conference on Pattern Recognition (ICPR)

Exercise

- ❑ Break into small groups
- ❑ Take a field that interests you:
 - Ex. Driving a car, understanding social networks, finding a good date, recommend a movie to watch, ...
- ❑ Identify a specific task that can be done with machine learning
 - What is the objective of the task?
 - What is the data you need?
 - What type of ML problem is this? Classification, regression, ...
 - How would your approach compare to an expert-driven method?