
# Machine Learning Applications and Framework

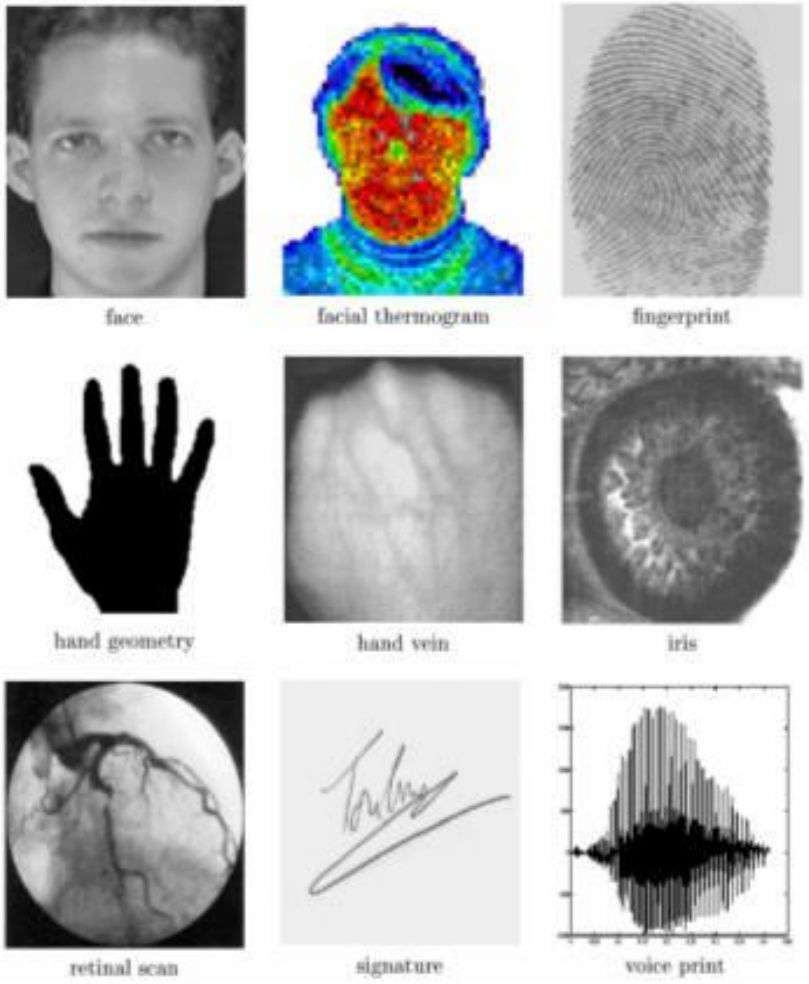
**Dr. Fayyaz ul Amir Afsar Minhas**
PIEAS Biomedical Informatics Research Lab
Department of Computer and Information Sciences
Pakistan Institute of Engineering & Applied Sciences
PO Nilore, Islamabad, Pakistan
http://faculty.pieas.edu.pk/fayyaz/

# Applications of Machine Learning

- An ability that I would like you to learn is to identify how to use machine learning in different domains.

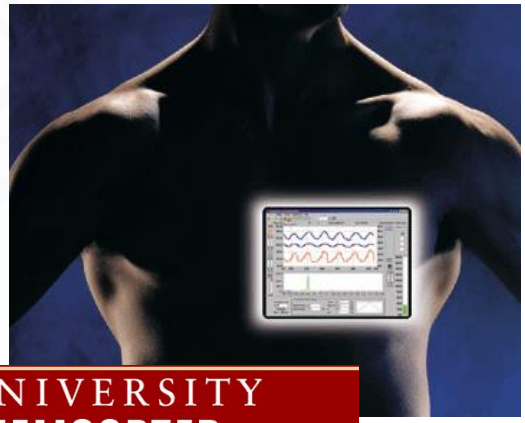
- Machine learning can be applied in a wide array of real-world applications

# Applications: Biometrics



# Applications






STANFORD UNIVERSITY
**AUTONOMOUS HELICOPTER**

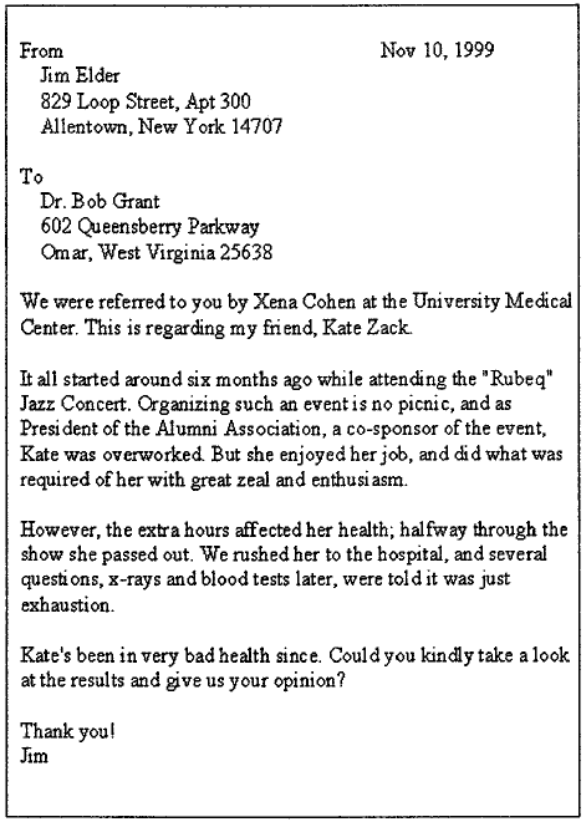
**Overview**

The goal of this project is to push the state-of-the-art in autonomous helicopter flight: extreme aerobatics under computer control.

**Featured Videos**

http://heli.stanford.edu/

# Handwriting Recognition / OCR



From
Jim Elder
829 Loop Street, Apt 300
Allentown, New York 14707

Nov 10, 1999

To
Dr. Bob Grant
602 Queensberry Parkway
Omar, West Virginia 25638

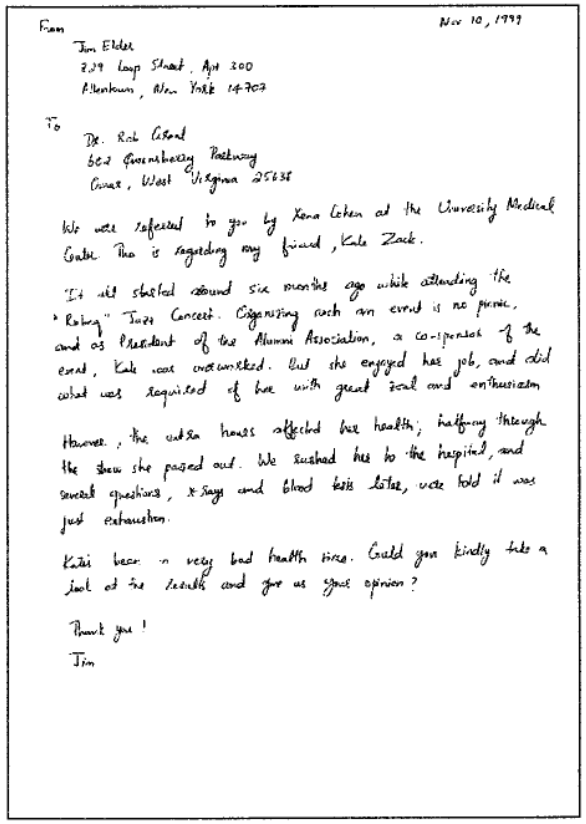
We were referred to you by Xena Cohen at the University Medical Center. This is regarding my friend, Kate Zack.

It all started around six months ago while attending the "Rubeq" Jazz Concert. Organizing such an event is no picnic, and as President of the Alumni Association, a co-sponsor of the event, Kate was overworked. But she enjoyed her job, and did what was required of her with great zeal and enthusiasm.

However, the extra hours affected her health; halfway through the show she passed out. We rushed her to the hospital, and several questions, x-rays and blood tests later, were told it was just exhaustion.

Kate's been in very bad health since. Could you kindly take a look at the results and give us your opinion?

Thank you!
Jim



From
Jim Elder
829 Loop Street, Apt 300
Allentown, New York 14707

Nov 10, 1999

To
Dr. Rob Grant
602 Queensberry Parkway
Omar, West Virginia 25638

We were referred to you by Xena Cohen at the University Medical Center. This is regarding my friend, Kate Zack.

It all started around six months ago while attending the "Rubeq" Jazz Concert. Organizing such an event is no picnic, and as President of the Alumni Association, a co-sponsor of the event, Kate was overworked. But she enjoyed her job, and did what was required of her with great zeal and enthusiasm.

However, the extra hours affected her health; halfway through the show she passed out. We rushed her to the hospital, and several questions, x-rays and blood tests later, were told it was just exhaustion.

Kate's been in very bad health since. Could you kindly take a look at the results and give us your opinion?

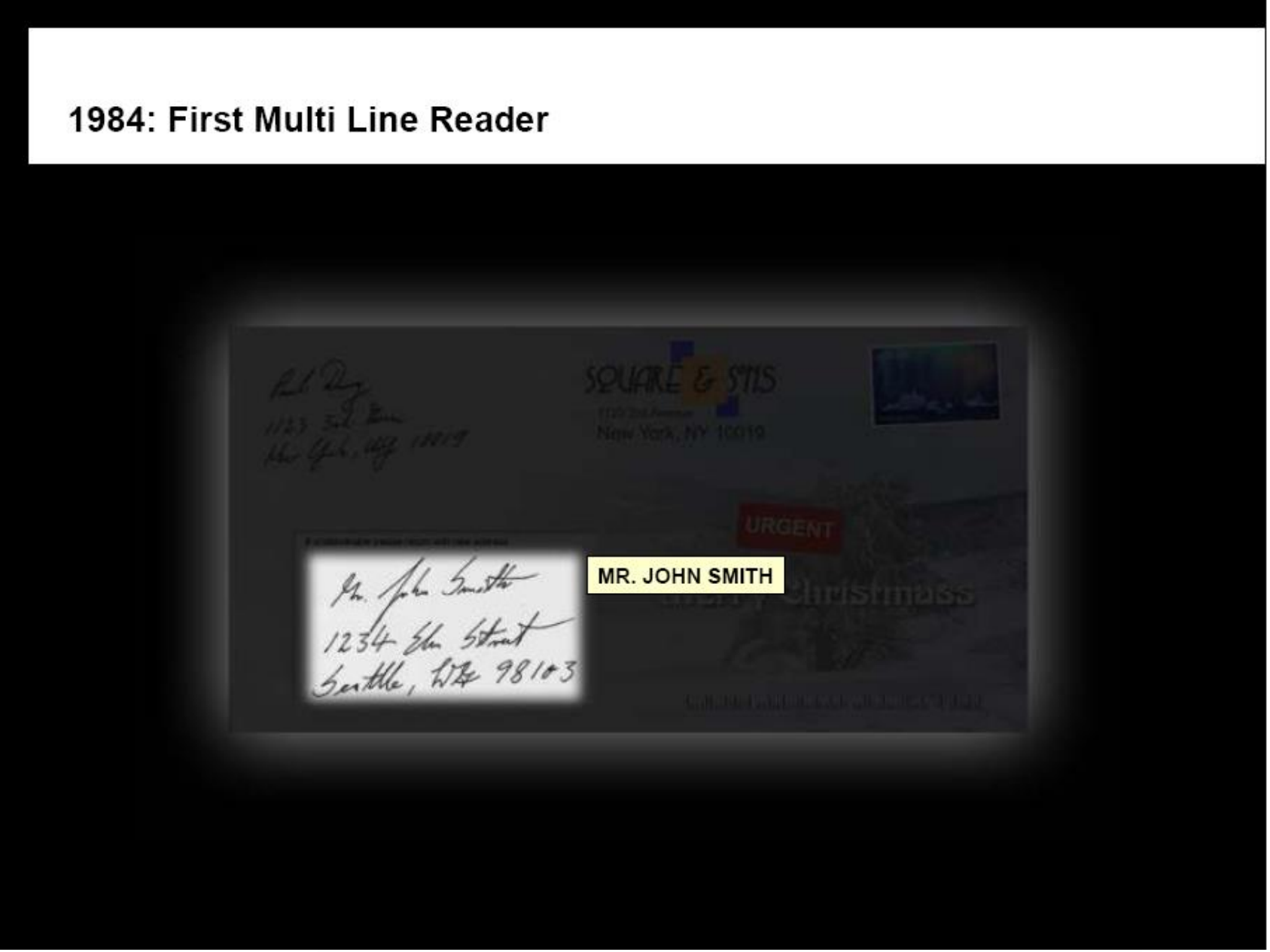
Thank you!
Jim

## The Letter

# 1978: First Postal Code Reader Worldwide

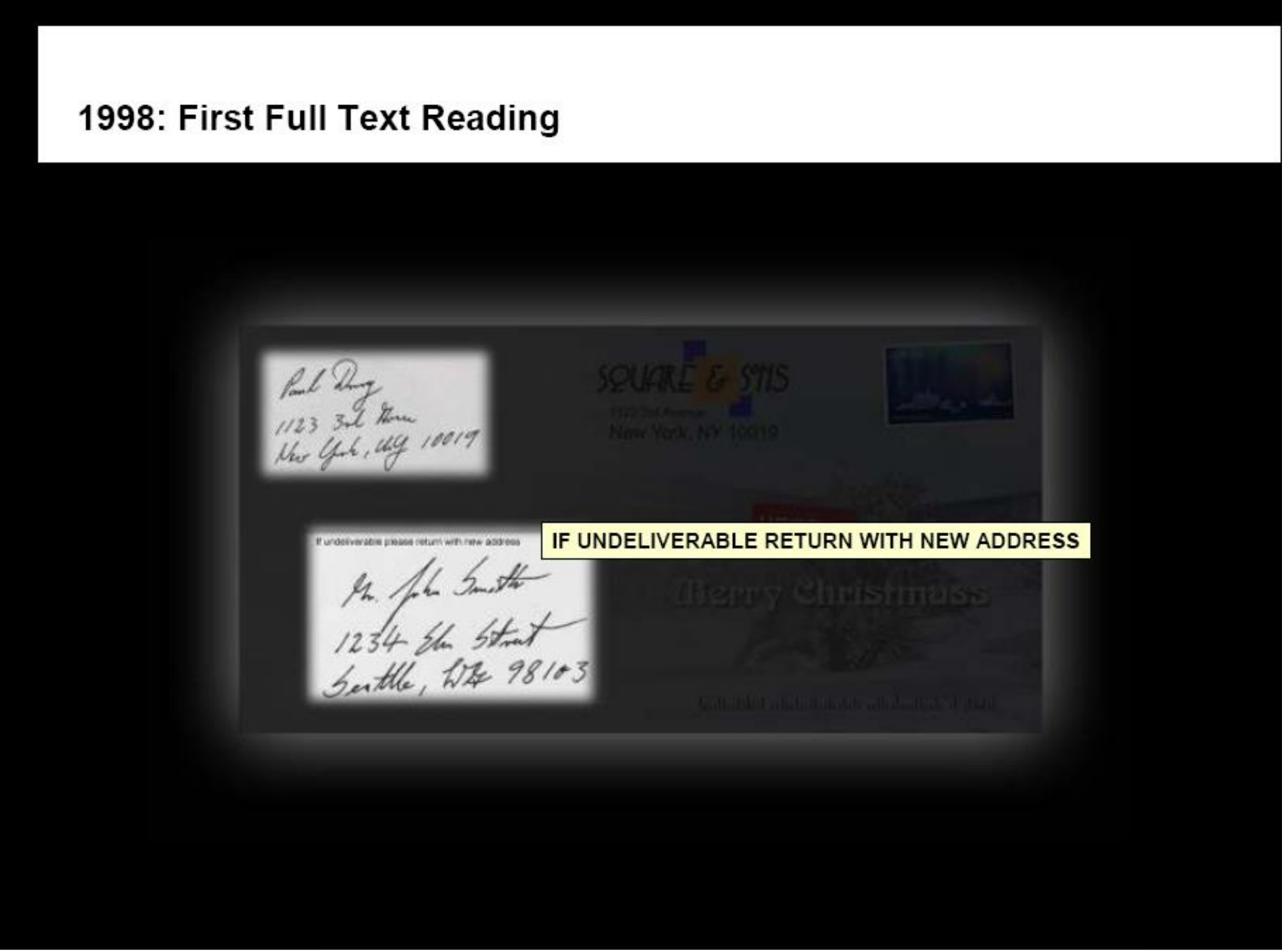
# 1982: First Address Reader Worldwide

# 1984: First Multi Line Reader



# 1996: First Sender's Address Reader

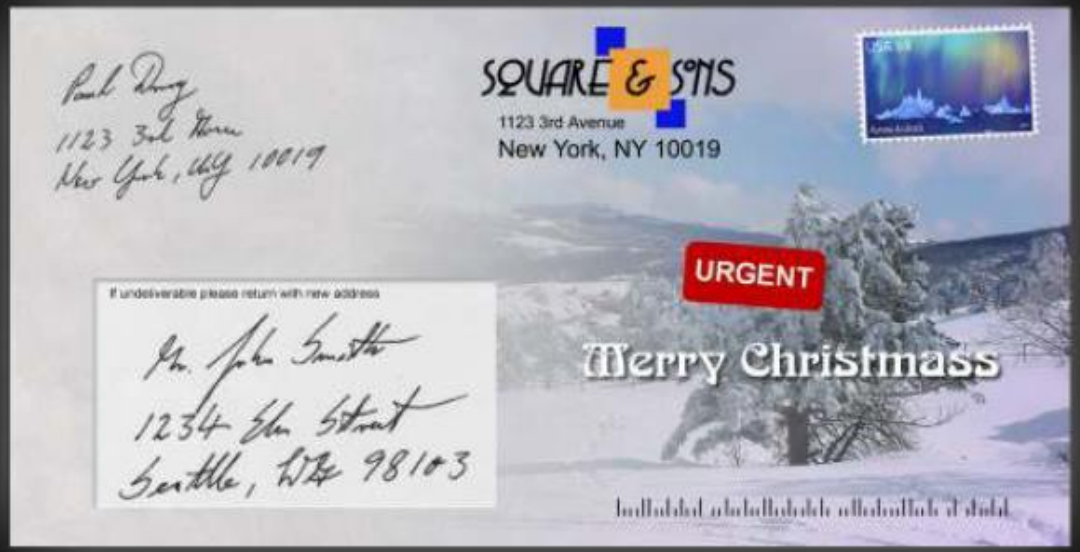
# 1998: First Full Text Reading



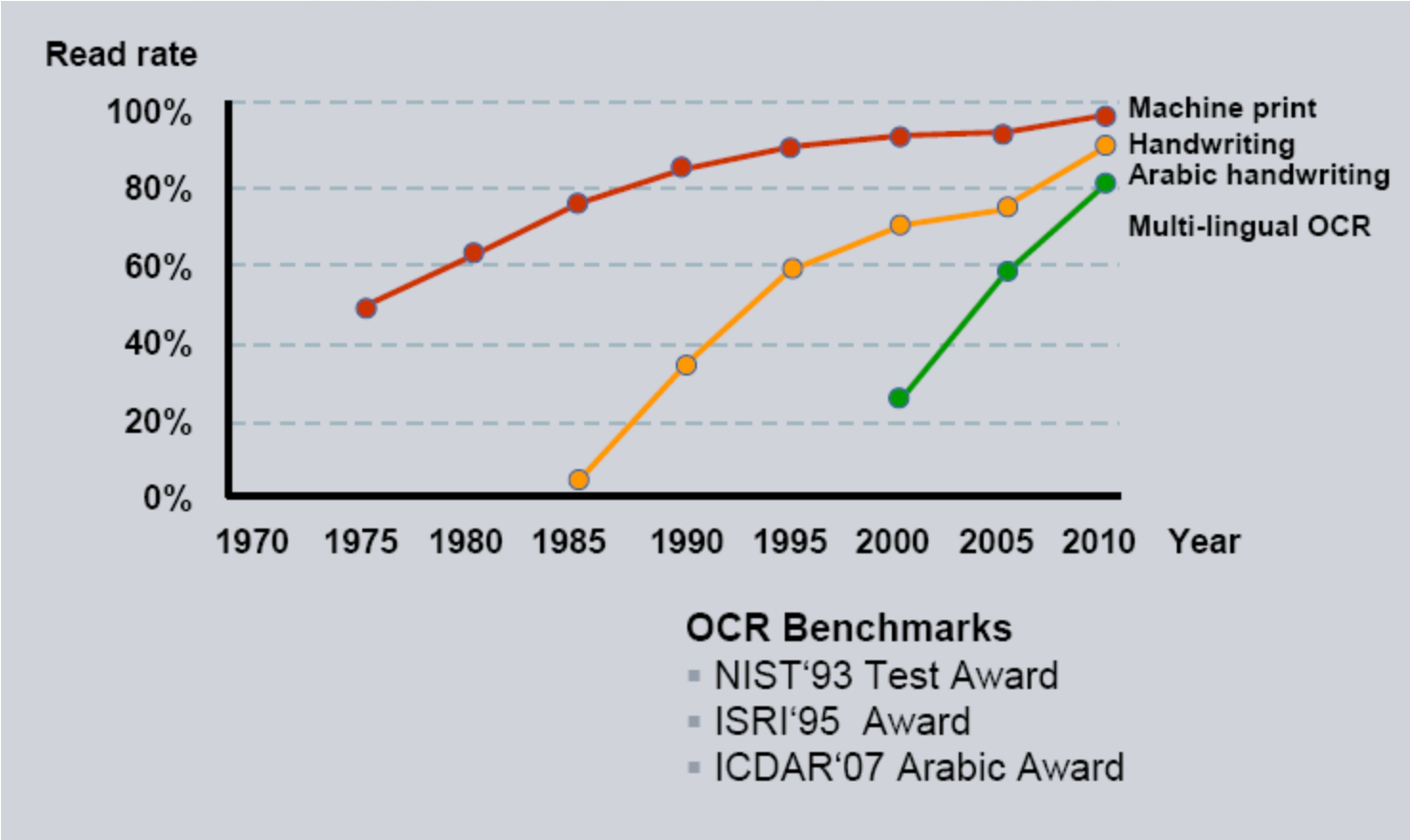
# 2000: First Graphics Recognition

## 2004: First Full Recognition

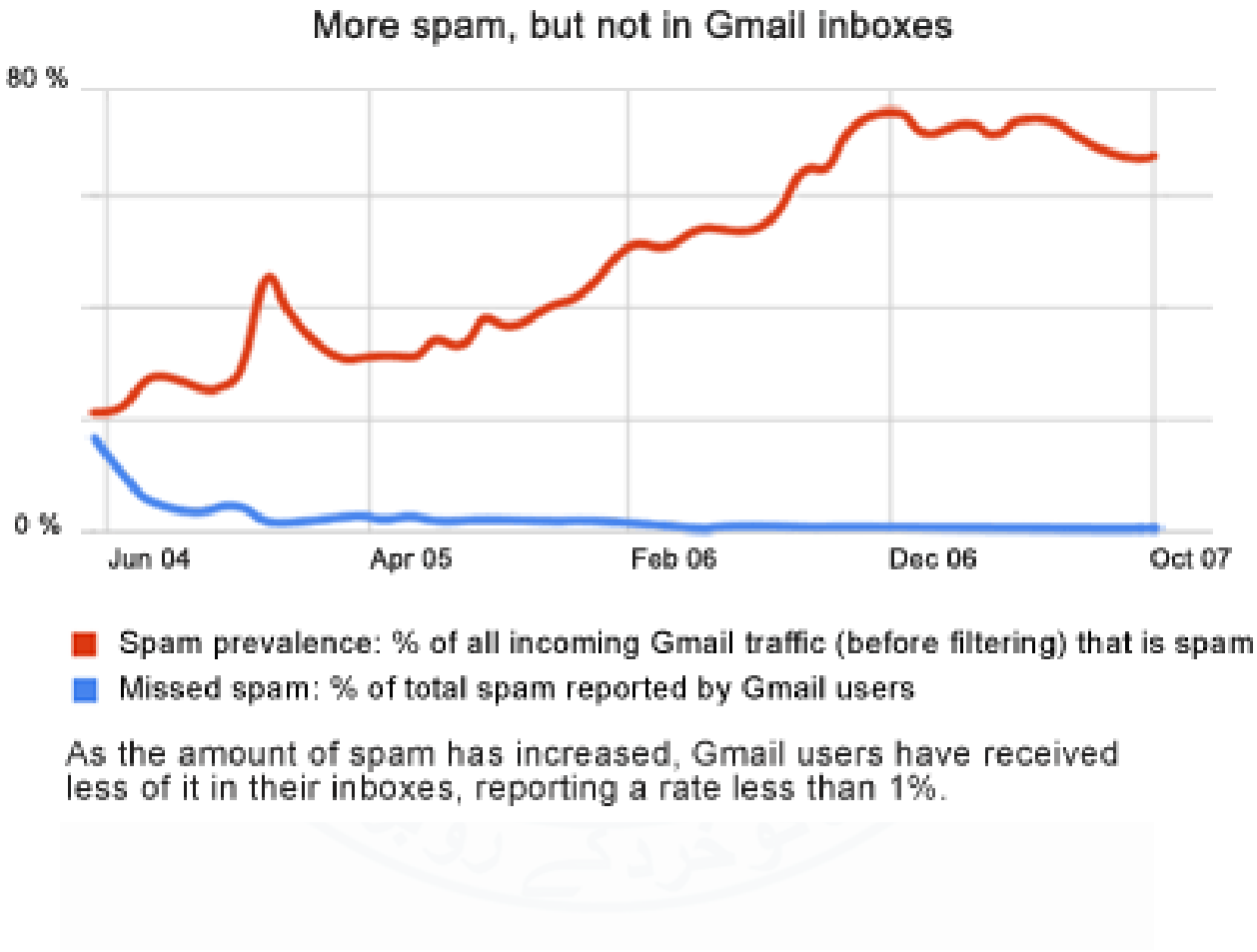
## 2008: Recognition on Both Sides of Envelope



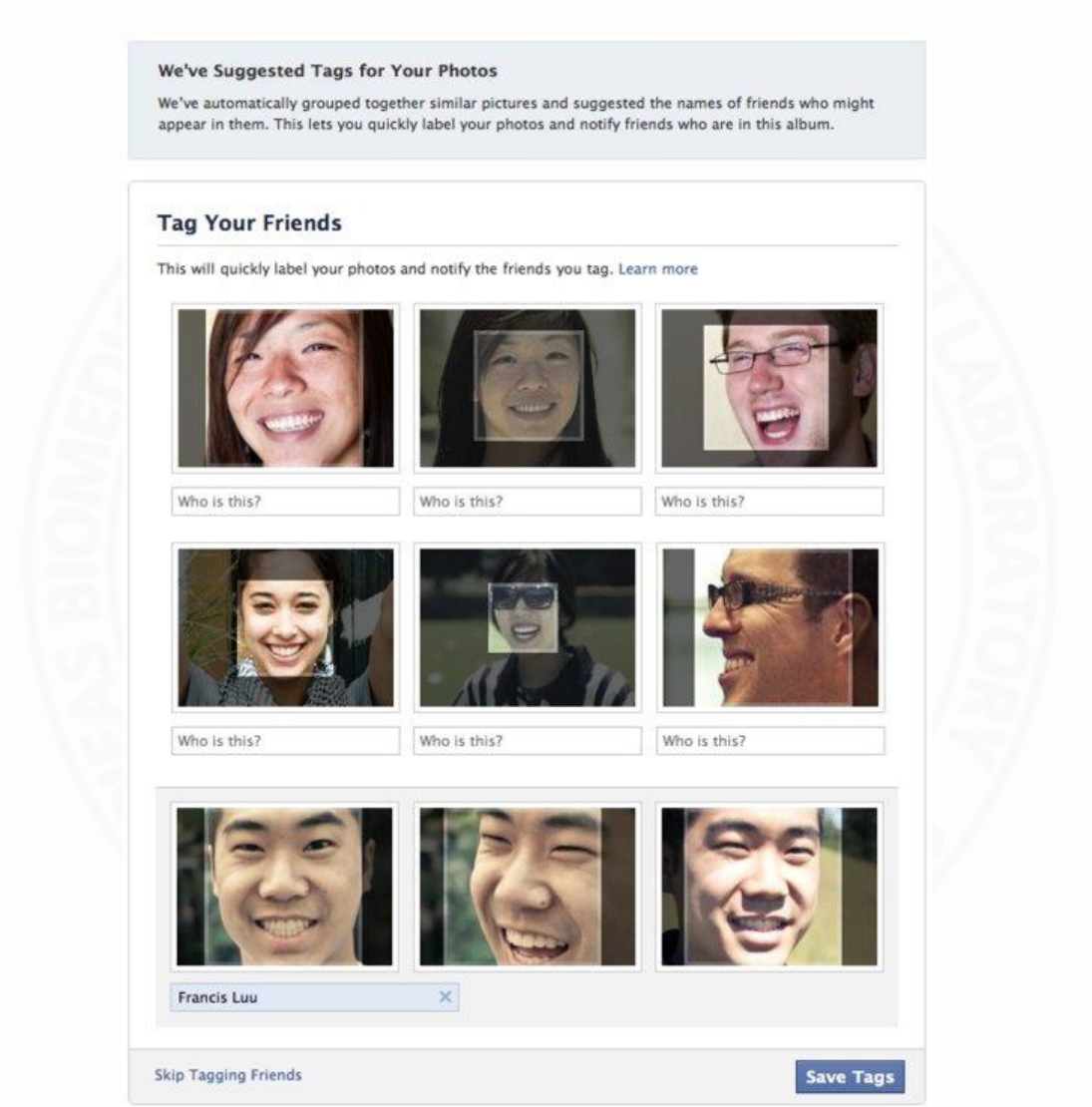
# OCR Accuracy



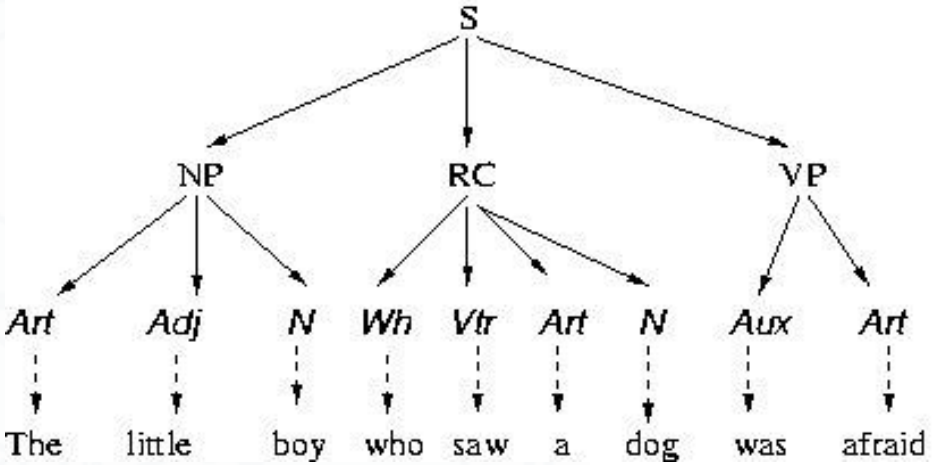
# Gmail: ML in NLP

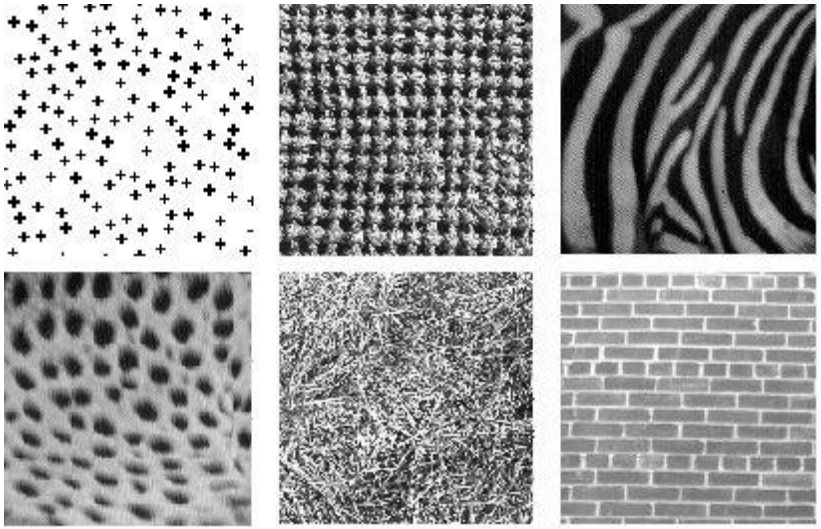


# Facebook Friends Tagging



# Applications of PR



# Recommender Systems


- Recommend movies based on user preferences, interests and likes
- Similar ideas for facebook…
  - Find friends that share your interests

# Computer / Network Security

- Prediction of threats
- Prediction of bugs / vulnerabilities in software
- Identification of malicious activity
- Identification of malicious software / viruses
- Attacking through side channels
  - Keyboard acoustics

# Keyboard acoustics

Keyboard Acoustic Emanations Revisited

LI ZHUANG, FENG ZHOU, and J. D. TYGAR
University of California, Berkeley

Text recognized by the HMM classifier, with cepstrum features (underlined words are wrong),

> the big money fight has drawn the shoporo od dosens
> of companies in the entertainment industry as well
> as attorneys gnnerals on states, who fear the
> fild shading softwate will encourage illegal acyivitt,
> srem the grosth of small arrists and lead to lost cobs and
> dimished sales tas revenue.

Text after spelling correction using trigram decoding,

> the big money fight has drawn the support of dozens
> of companies in the entertainment industry as well as
> attorneys generals in states, who fear the film sharing
> software will encourage illegal activity, stem the growth
> of small artists and lead to lost jobs and finished sales
> tax revenue.

Original text. Notice that it actually contains two typographical errors, one of which is fixed by our spelling corrector.

> the big money fight has drawn the support of dozens
> of companies in the entertainment industry as well as
> attorneys gnnerals in states, who fear the file sharing
> software will encourage illegal activity, stem the growth
> of small artists and lead to lost jobs and dimished sales
> tax revenue.

# PCR in HCI/CV

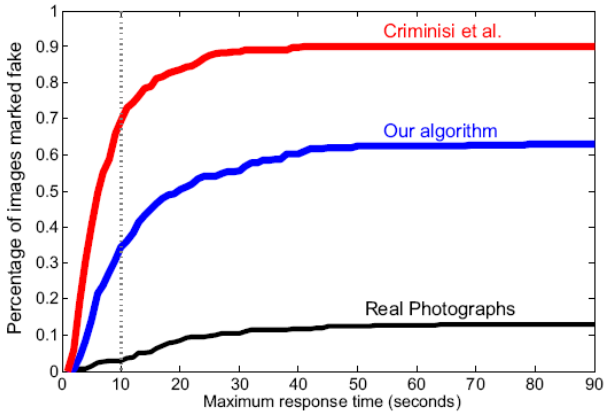
- Gesture Recognition

# Scene Completion Using Millions of Photographs

James Hays Alexei A. Efros
Carnegie Mellon University

Original Image

Output



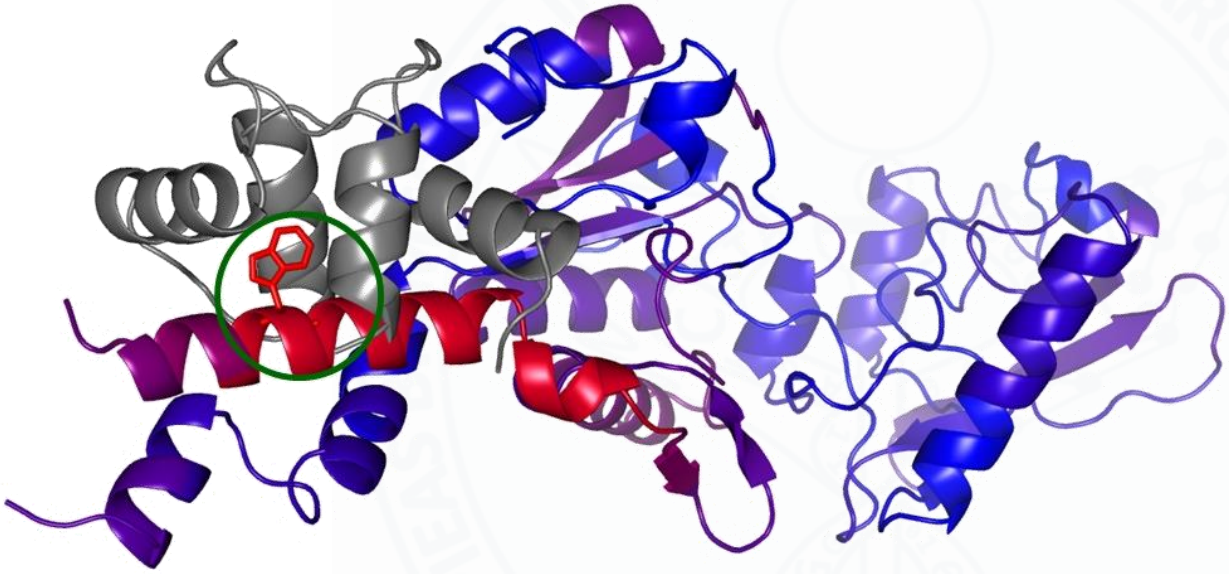
# Deep Learning

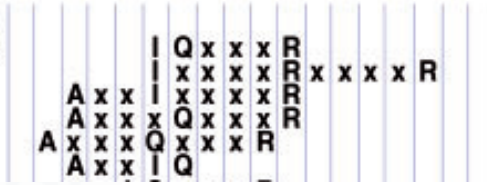
- Google Brain
  - https://en.wikipedia.org/wiki/Google_Brain

# Applications at PIEAS BIR Lab

- Development of Intelligent Computational Solutions to problems in Biology and Medicine

# Applications in Bioinformatics





High predicted binding propensity

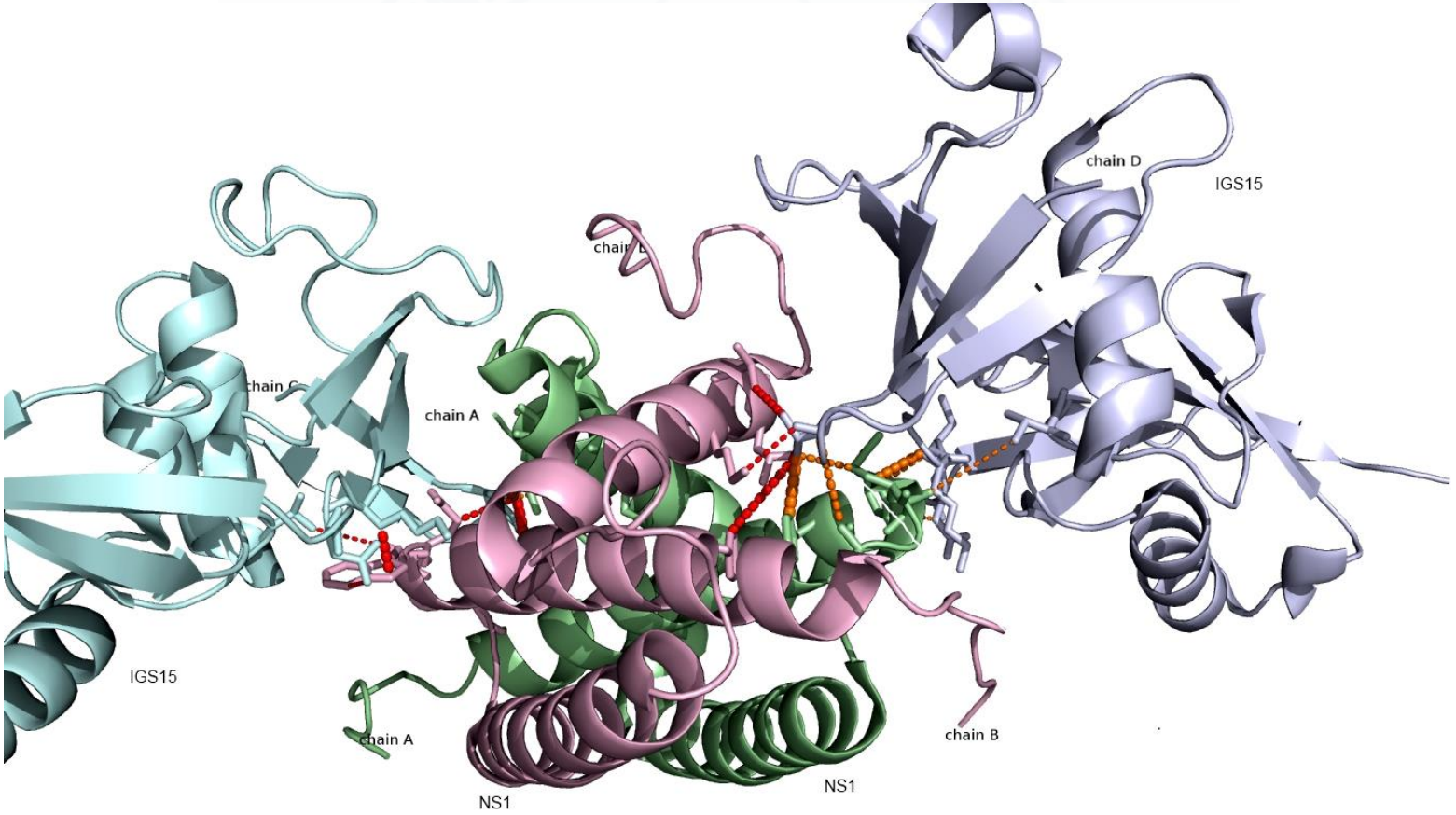
Low predicted binding propensity

Calmodulin

PDB: 1YRT

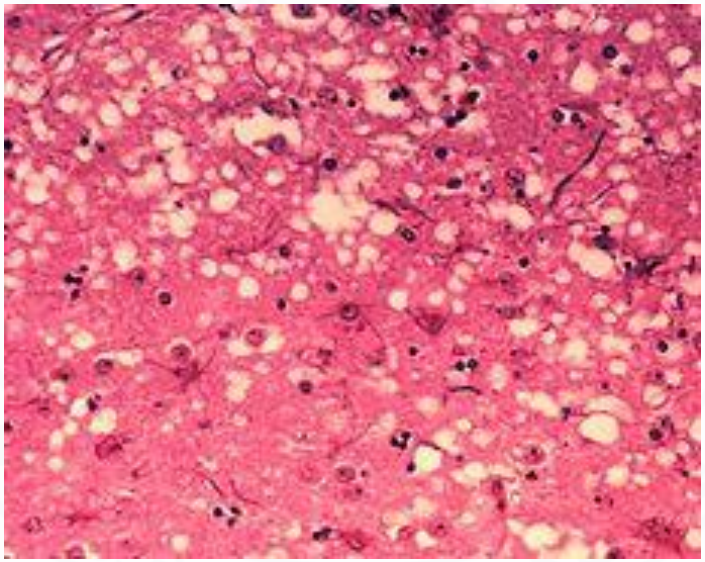
# Applications in Bioinformatics

- Predict protein interfaces

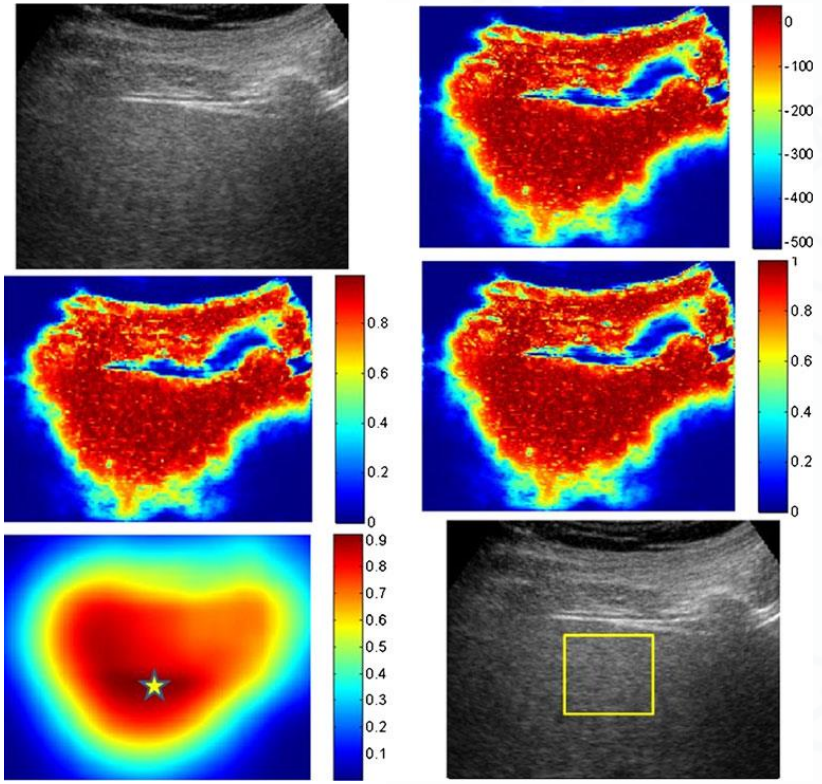


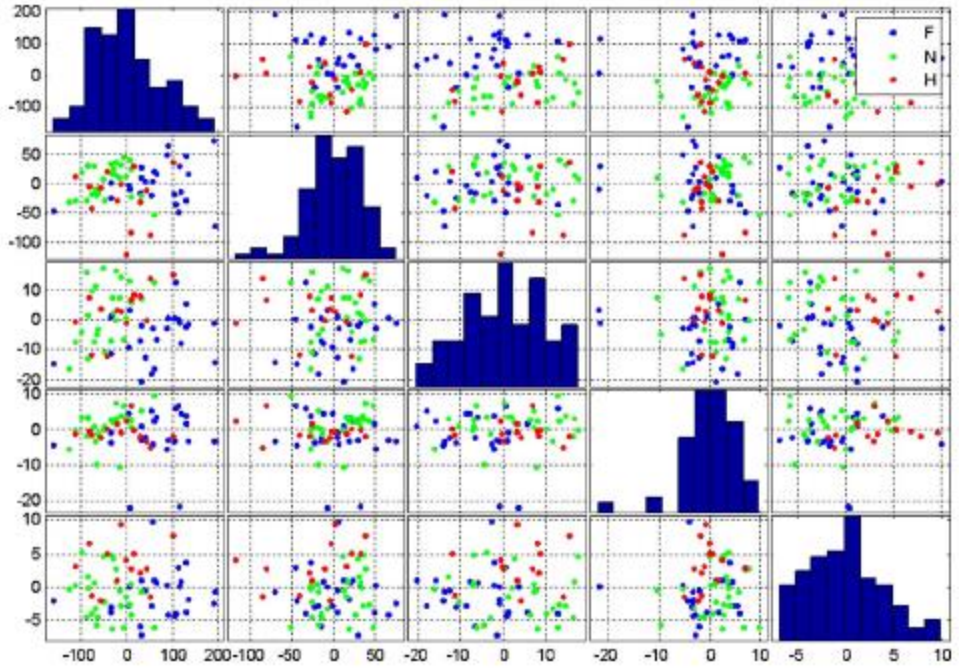
# Applications in Bioinformatics

- Prion prediction using ML



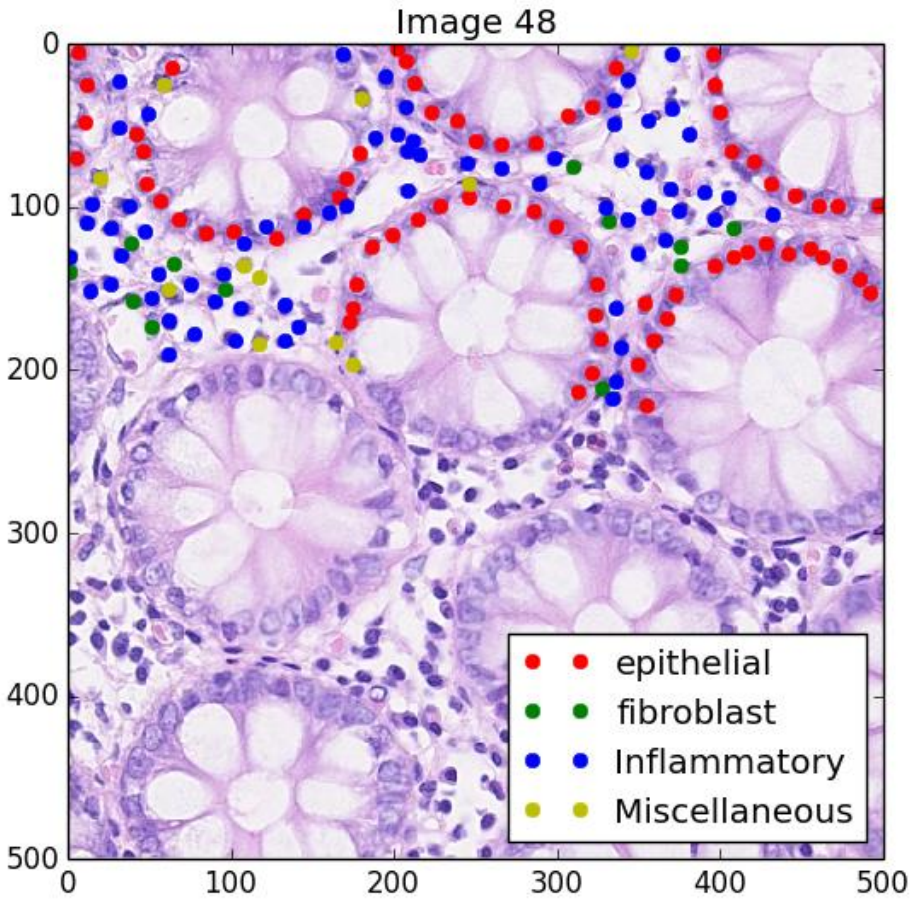
# Medical Image Processing Applications



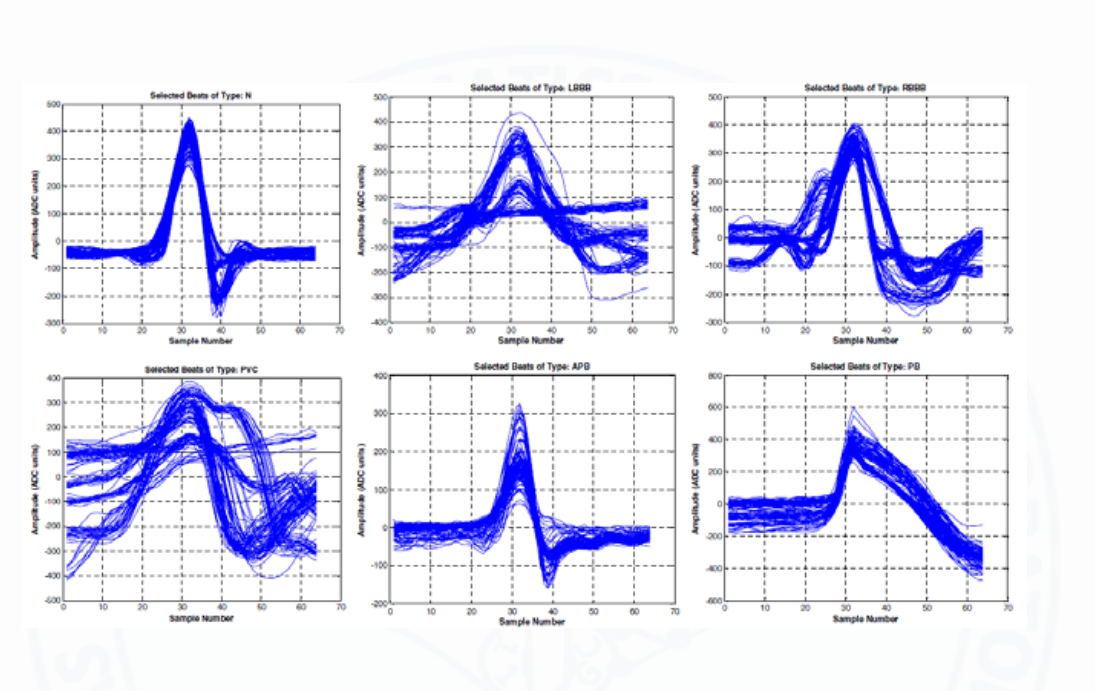


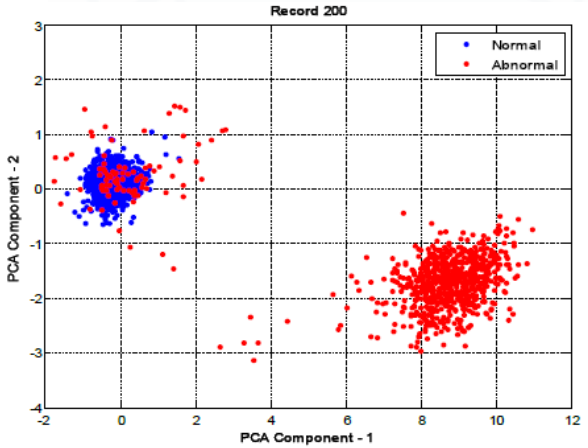
# Medical Image Processing Applications

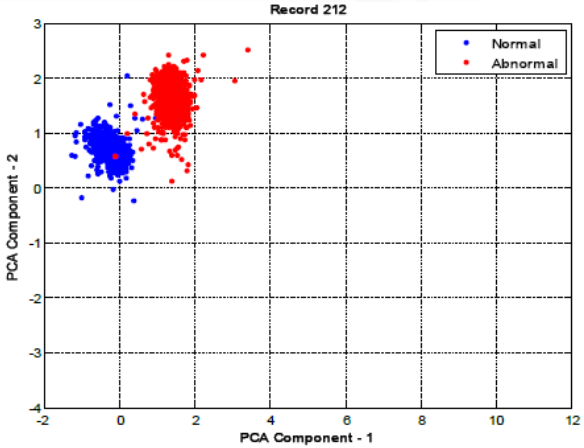
- Detection of cancerous cells



# Applications in signal analysis

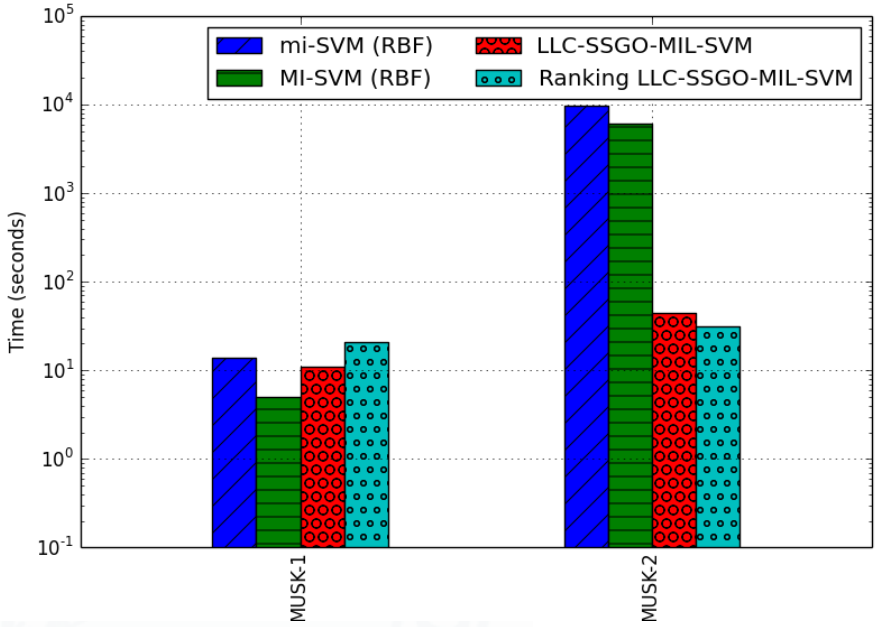


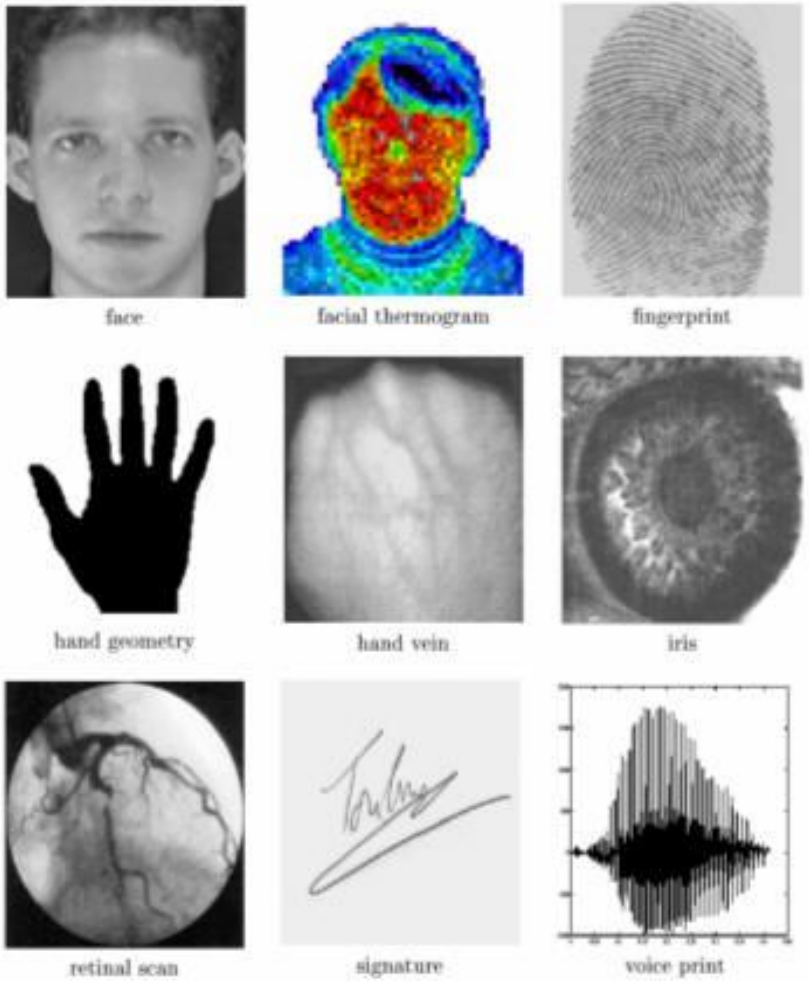



# BMI Lab Projects

- Development of open-source machine learning tools and packages
  - PyLemmings: **Py**thon Based **L**arg**e** **M**argin **M**ultiple **I**nstance **L**earning **S**ystem
  - CAFÉ-Map: **C**ontext **A**ware **Fe**ature **Map**ping

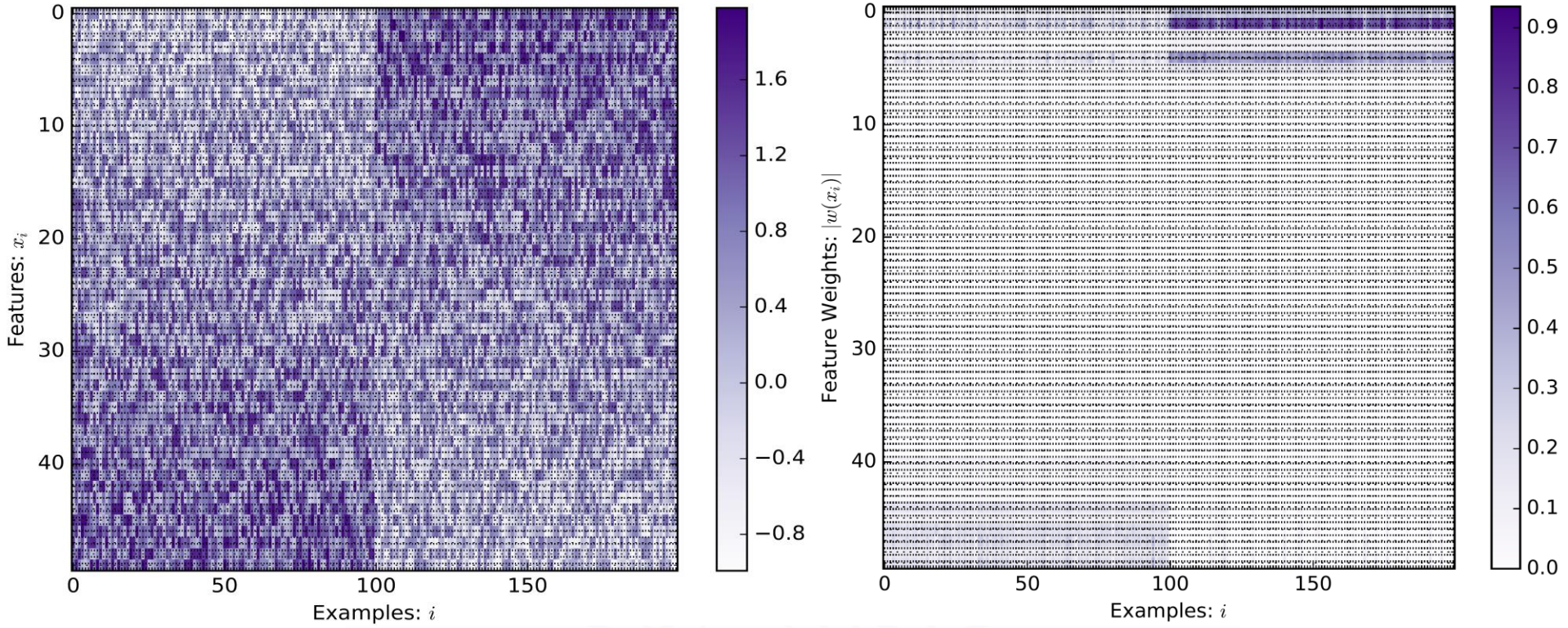



# BMI Lab Projects: Biometrics



# Example

- Tall and Thin Data

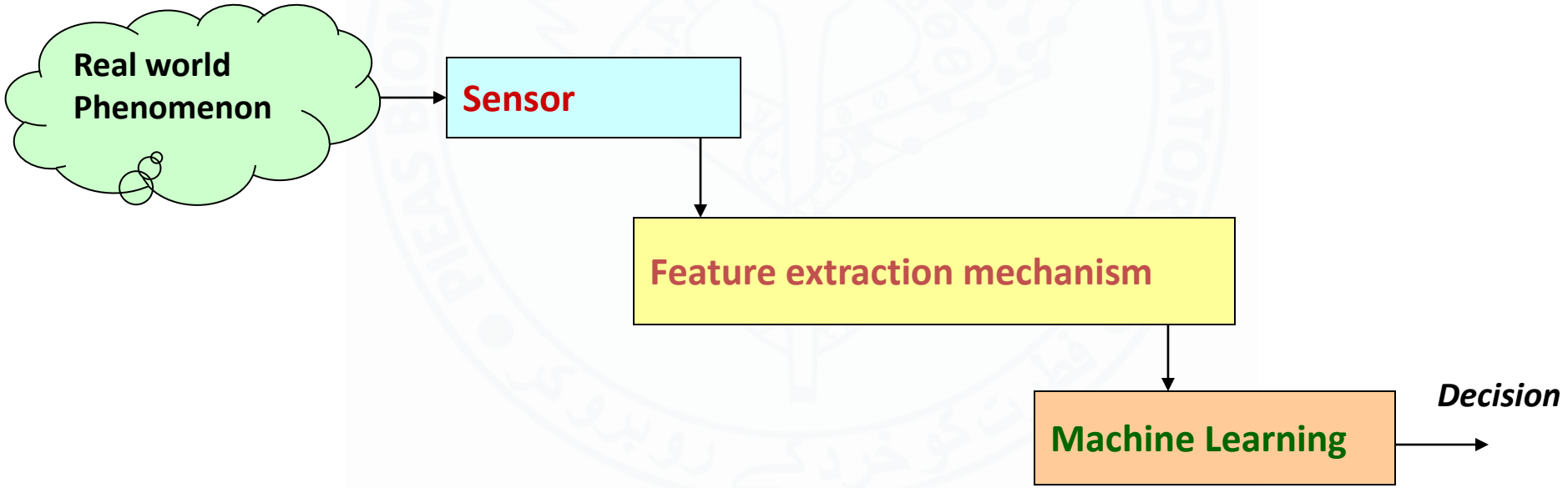


# Applications in Data Science

- How to design a recommender systems for websites?

# Constructs of a PR System

- Identify the objective
  - Identify the unit of classification (example)
    - Image block, protein sequence, ….

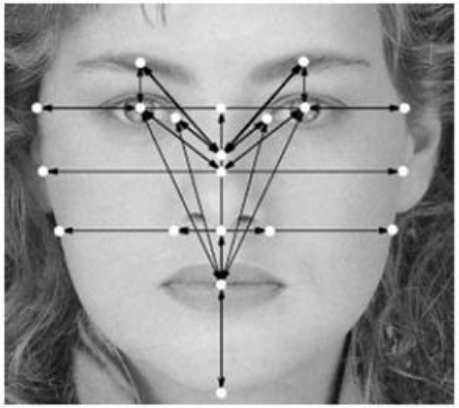


# Components of a PR System

- Sensor
  - Responsible for getting raw data from an object
    - Examples
      - Camera for face recognition system
      - ECG for cardiac disease diagnosis
  - Multiple sensors can be combined to provide a better picture
  - Sensors can introduce noise into the PR system

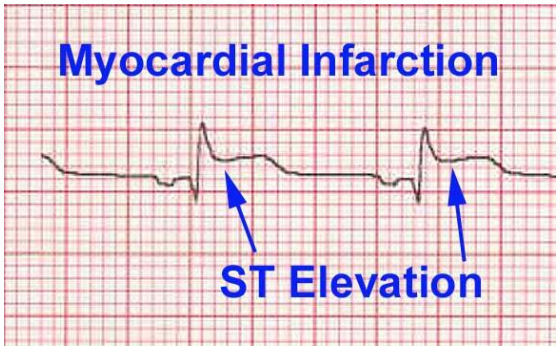
# Components of a PR System…

- Feature Extraction
  - Usually (almost all the time!) the amount of raw data obtained from sensor(s) is too large and redundant
    - Example
    - 256x256 image acquired by a camera for face recognition contains 256x256x24 bits of data but what is the information we are looking for?
    - We are looking for features that purely characterize a face and as a result enable us to distinguish amongst faces



# Components of a PR System…

- Feature Extraction…
  - Consider the design of a system for automated diagnosis of Myocardial Infarction
  - The usual sensor is an ECG machine which samples the electrical activity of the heart at some sample rate (e.g. 1000Hz)
  - We need some feature which would enable us to recognize particular ailments



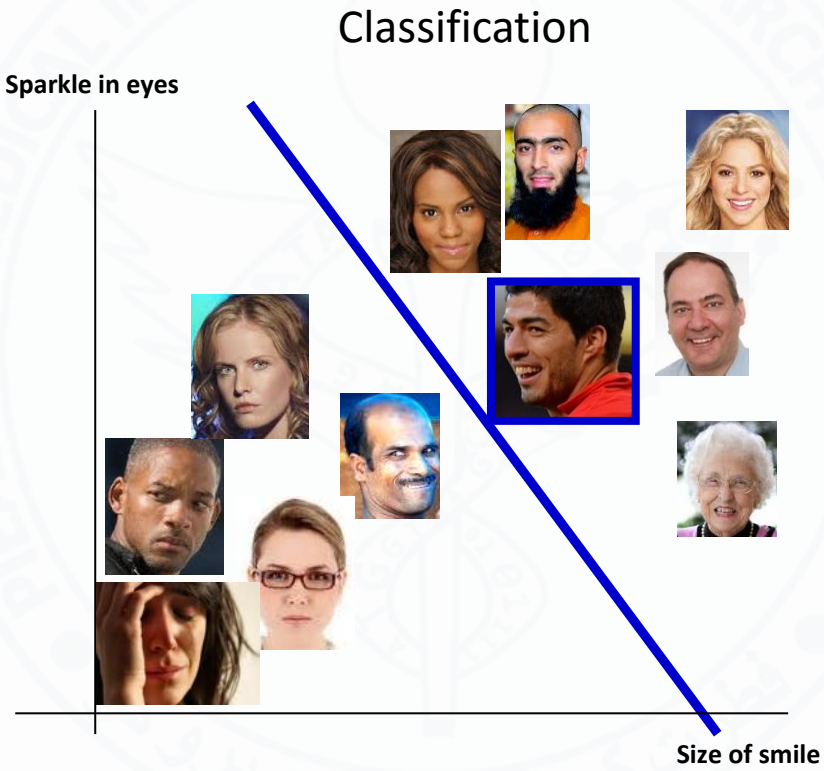
# Components of a PR System…

- Feature Extraction…
  - Thus a feature extraction mechanism works as an information processor which takes in raw data and outputs information in the form of a feature vector which describes an object
  - We need descriptors called features which tend to remain somewhat constant over objects belonging to the same class but are different for objects belonging to other classes so we can discriminate easily between classes

# Components of a PR System…

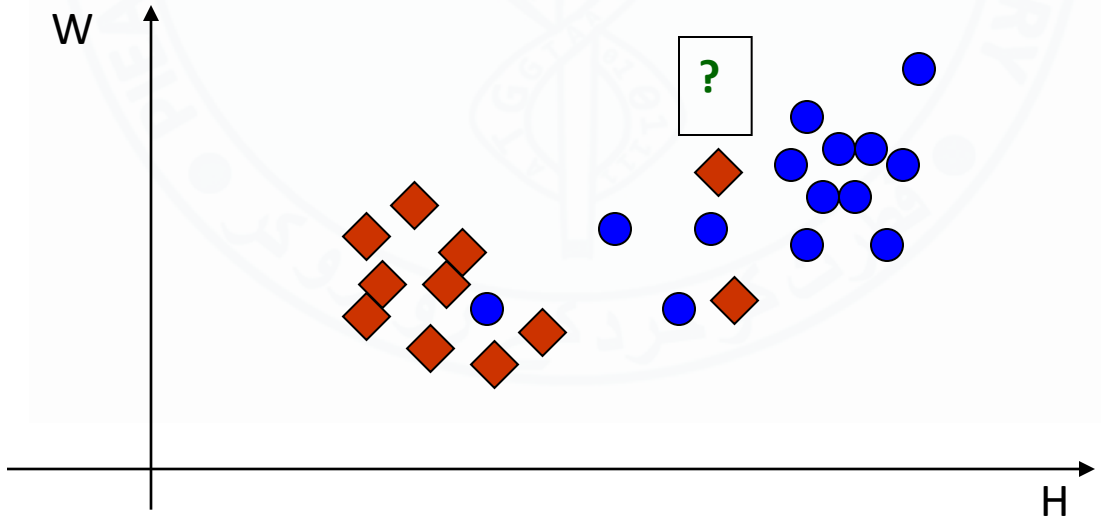
- Feature Extraction…
  - Points to note
    - A feature extraction mechanism
      - Extracts information in the form of a feature vector about an object which would enable our system to recognize it
      - Provides a form of dimensionality reduction aiming at removal of redundancy in data while maintaining discrimination between objects to be recognized
      - Computes numeric or symbolic information from the observations collected by sensor(s)
      - Helps the PR system to ignore noise effects of sensor(s)
      - Is problem specific

# Example



# Components of a PR System…

- Machine Learning
  - The job of the last stage of the PR system is to classify/describe objects on the basis of their features
  - Assigning labels to objects

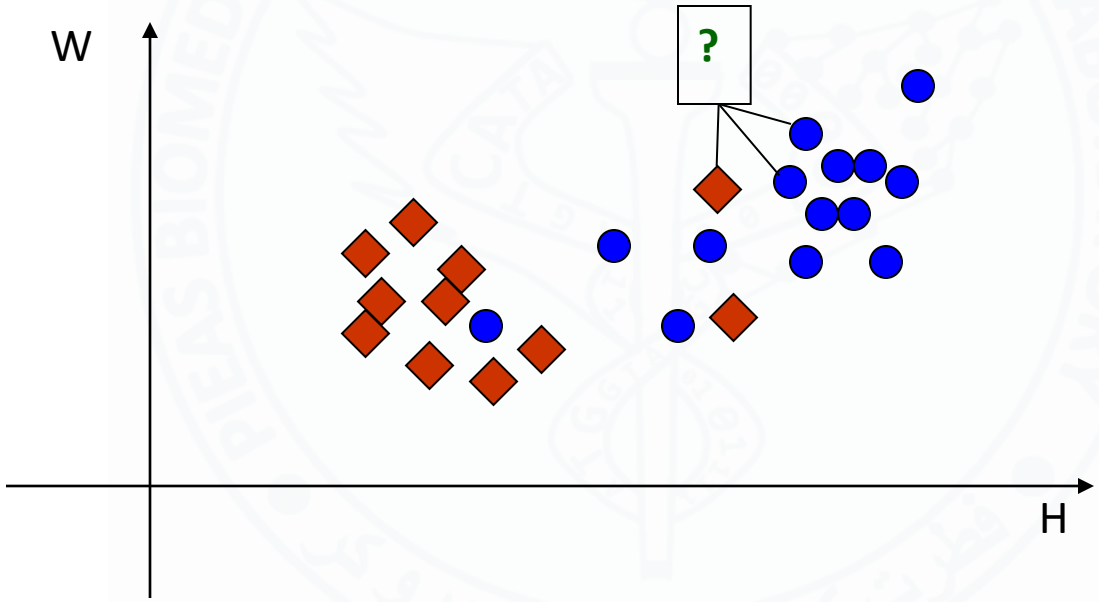


# Components of a PR System…

- Classification/Description
  - Approaches
    - Using Apriori Knowledge
      - Use already know rules to make a decision
      - Example:
        - » ST deviation of more than 0.1mV in the ecg is indicative of Ischemia
    - Supervised Learning
      - Assumes that a set of already classified patterns (called the training set or training examples) is available and a learning strategy can be used to assign labels to unknown patterns (on the basis of the learning data)
        - » Input: A set of labeled examples (training feature vectors & their labels)
        - » Task: Find the boundary (discriminant) between classes
        - » Output: Given a unlabeled object, it uses the discriminant to assign a class label to it

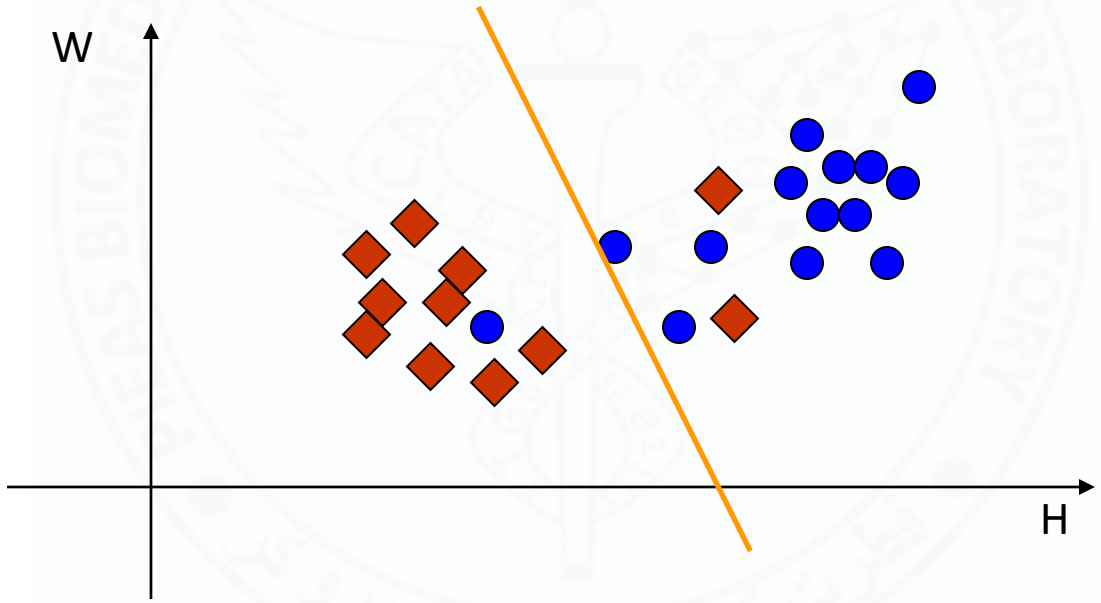
# Classification Approaches: Supervised

- Example (k=3)-Nearest Neighbor Classification



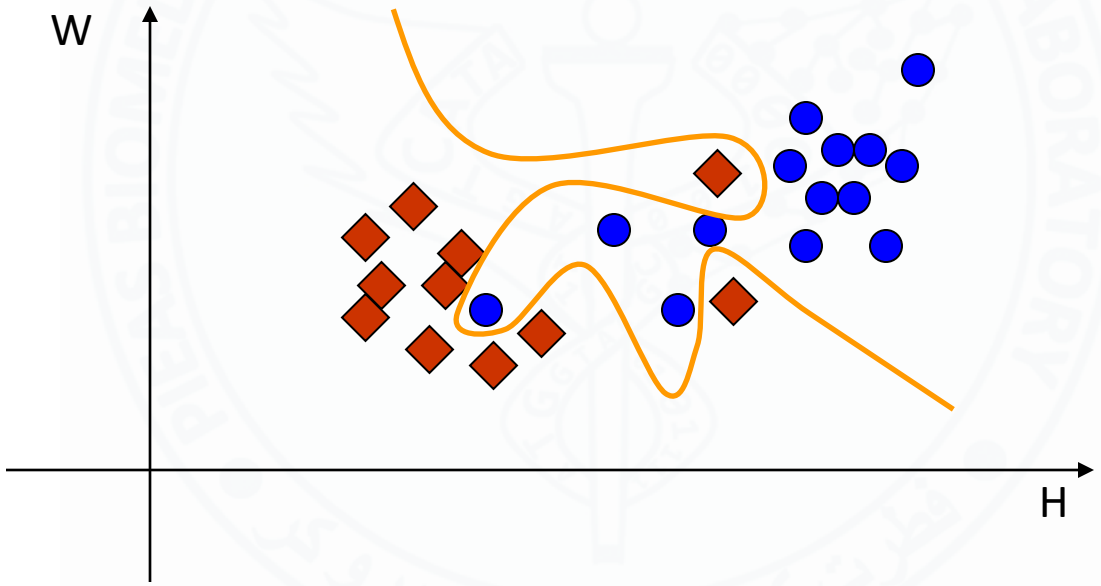
# Classification Approaches: Supervised…

- Linear Classifier



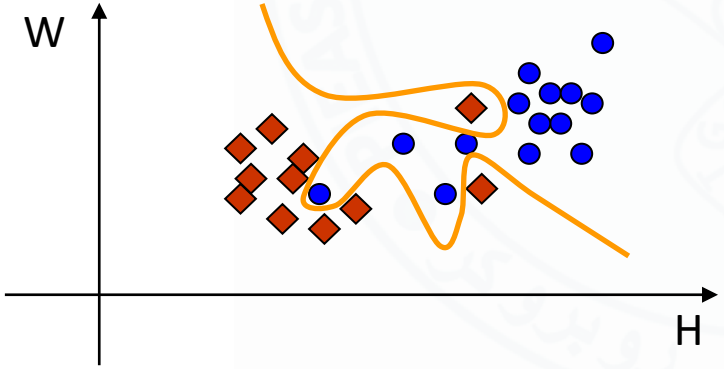
# Classification Approaches: Supervised…

- Nonlinear Classification boundary

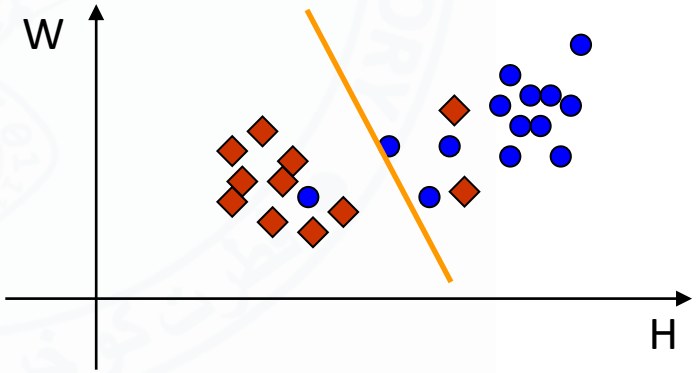


# Classification Approaches: Supervised…

- Generalization vs. Memorization
  - A particular issue in classification is the tradeoff between memorization vs. generalization




Has great memorization but may generalize poorly




Has lesser memorization but may generalize better

# Life Cycle


Collect data
Select features
Select model
Train classifier
Evaluate classifier

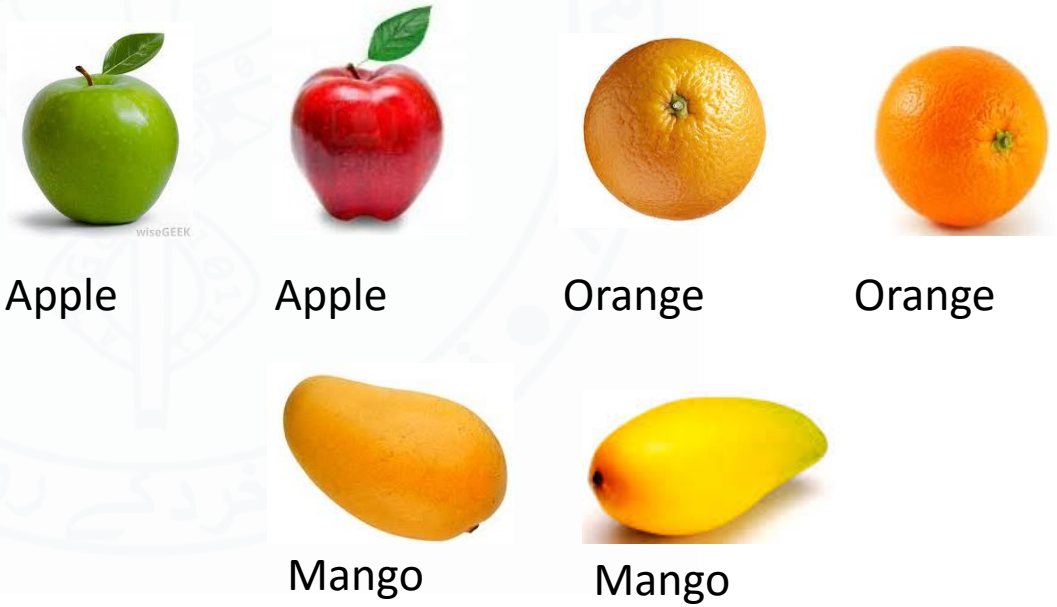
# Large Margin Machines for different problems

- Classification: SVM
- Multiple Class Classification
- One-Class Classification: One Class SVM
- Feature Selection: 1-norm SVM
- Ranking: Ranking SVM
- Structured-output prediction: Structured SVM
- Multiple instance learning: miSVM, MISVM
- Multi-task learning: Multi-task SVM
- Multi-label learning
- Semi-Supervised learning
- Transductive learning
- Self-taught learning
- Online learning
- Active learning
- Transfer learning
- Regression: Support Vector Regression
- Clustering: Support Vector Clustering
- …

# Supervised Classification

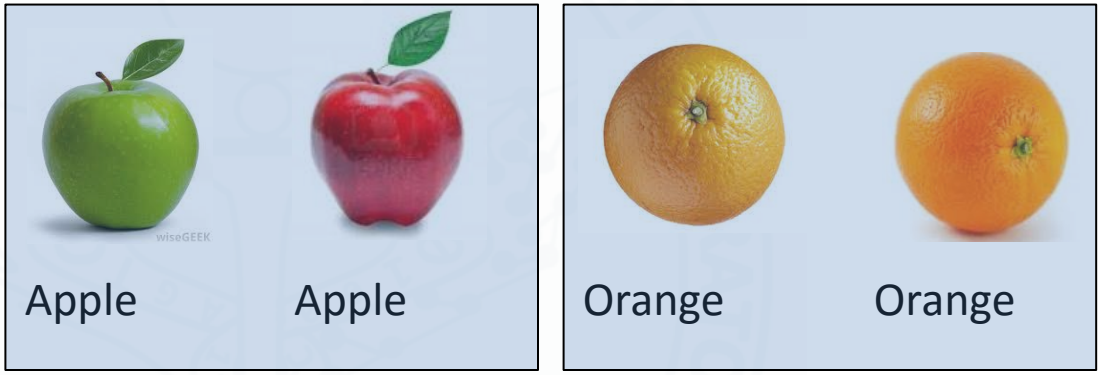
- Supervised Classification
  - Apple or orange
  - Inductive: Infer a rule for classification and use it to label unknown examples


- Multi-class Classification
  - Apple or orange or mango
  - For a binary classifier we can use
  - One vs. All
    - Apple vs. (Orange,Mango)
    - Orange vs. (Apple, Mango)
    - Mango vs, (Apple, Orange)
  - One against One
    - Apple vs. Orange
    - Apple vs. Mango
    - Orange vs. Mango



# One Class Classification & Feature Selection

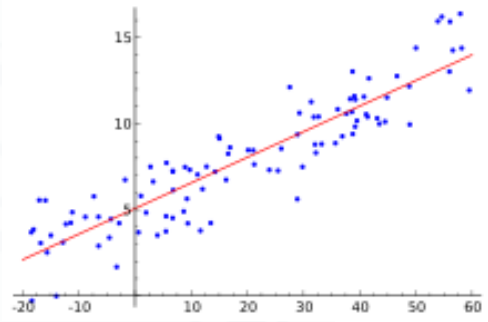
- **One Class Classification**
  - Apple or not
  - Orange or not
  - One-Class SVM




- **Feature Selection**
  - Select only the required features for classification
  - 1-norm SVM

# Regression & Clustering

- Regression
  - Price of the apple vs. prices of the orange
  - Can be multi-variable in both input and output
  - Support Vector Regression




- **Clustering**
  - **Unsupervised learning**
  - **Support Vector Clustering**
  - **Examples in one clusters should be similar (based on some criteria) to each other and different from other examples**
  - **Example: Apple sorting**

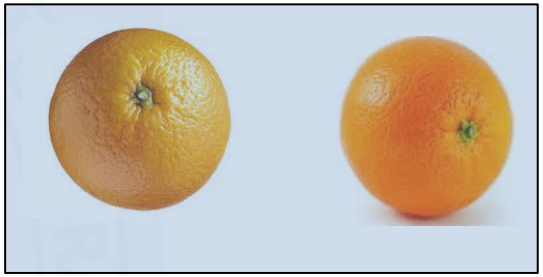
# Exploratory Data Analysis & Visualization

- Analysis of data sets to summarize their main characteristics
  - PCA
  - Clustering
    - SOM
  - Visualization of classifiers
    - Classification trees
    - Hierarchial clustering
    - Nomograms

# Multiple Instance Learning (MIL)

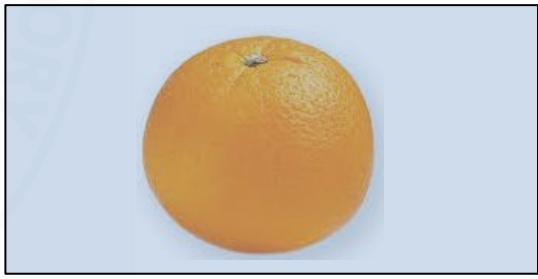
- Examples come in bags
- Positive bags
  - Each Has at least one positive example
- Negative bags
  - All examples are negative
- Tasks
  - Classify Bags
  - Classify instances
- Regular binary classification is a special case of MIL
  - Each bag is of size 1
    - Positive bags reduce to positive examples
    - Negative bags become negative examples

Positive Bag-1 of apples or oranges



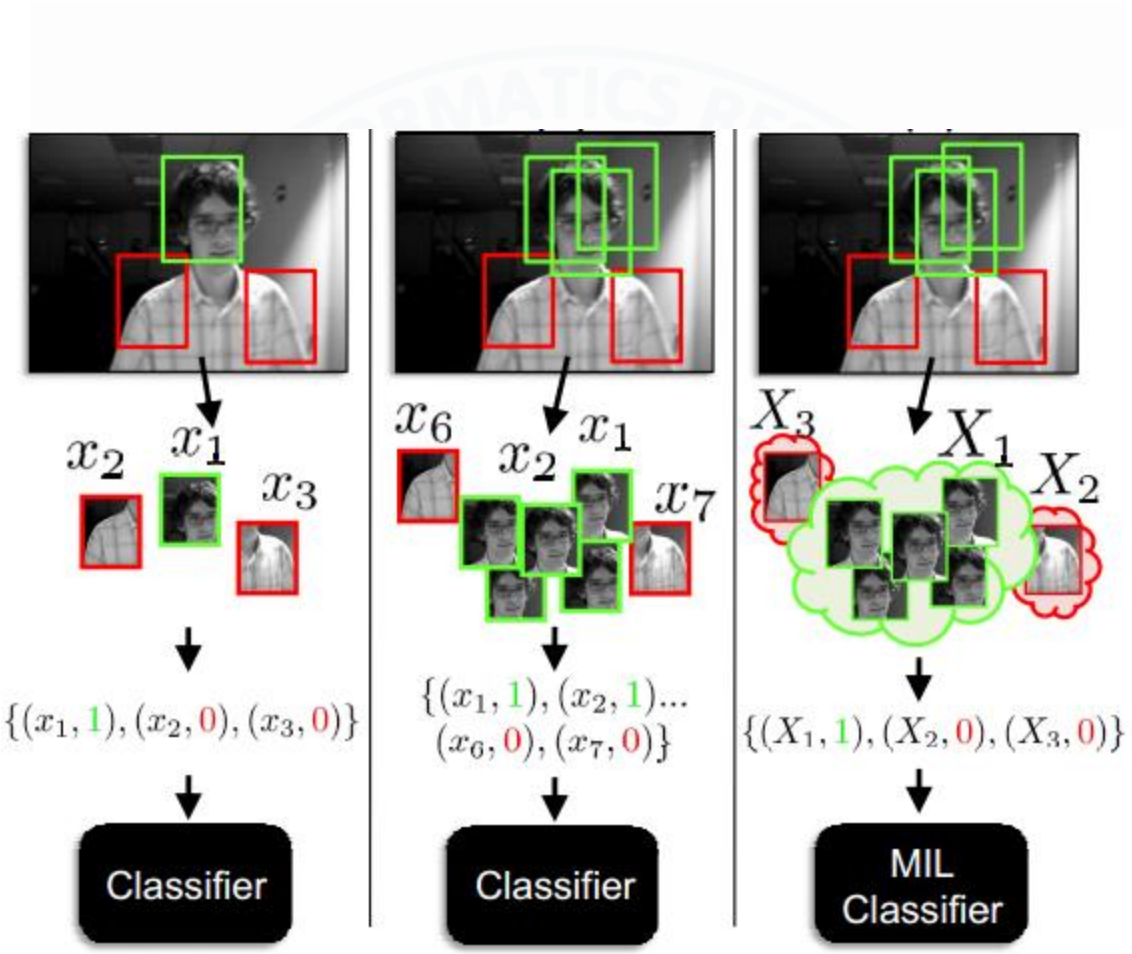
Negative Bag-1 of oranges

Positive Bag-2 of apples or oranges



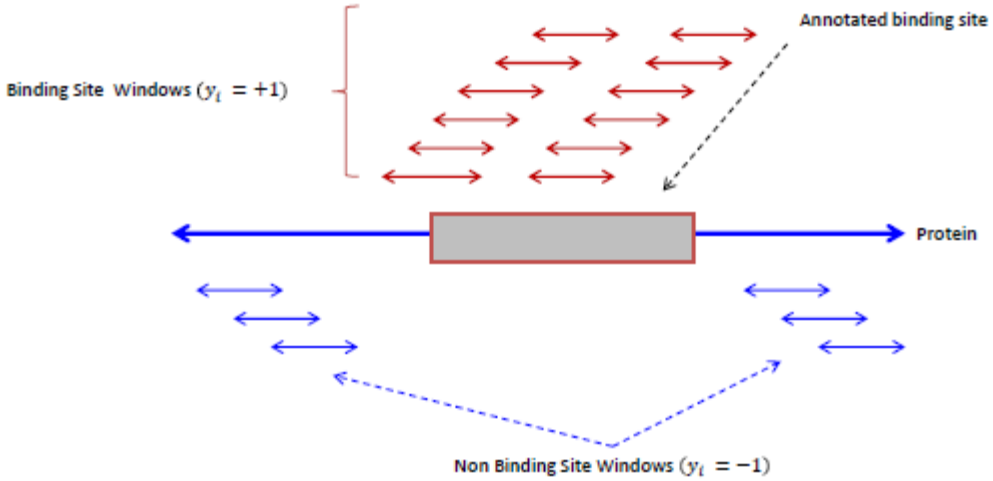
Negative Bag-2 of oranges

# Applications of MIL




Babenko, Boris, Ming-Hsuan Yang, and Serge Belongie. "Robust Object Tracking with Online Multiple Instance Learning." *IEEE Trans. Pattern Anal. Mach. Intell.* 33, no. 8 (August 2011): 1619–32. doi:10.1109/TPAMI.2010.226.

# Applications of MIL

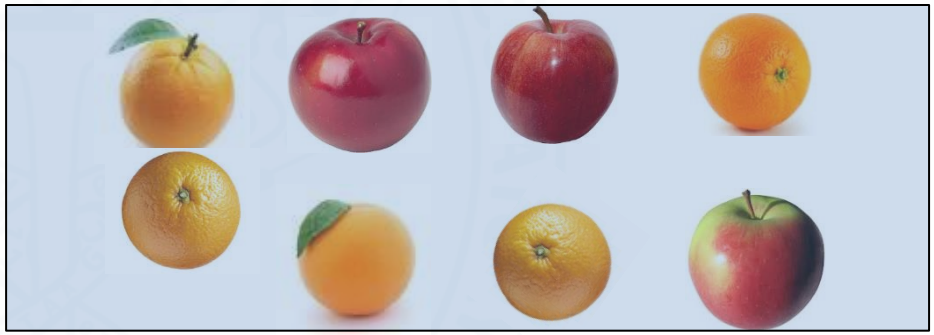


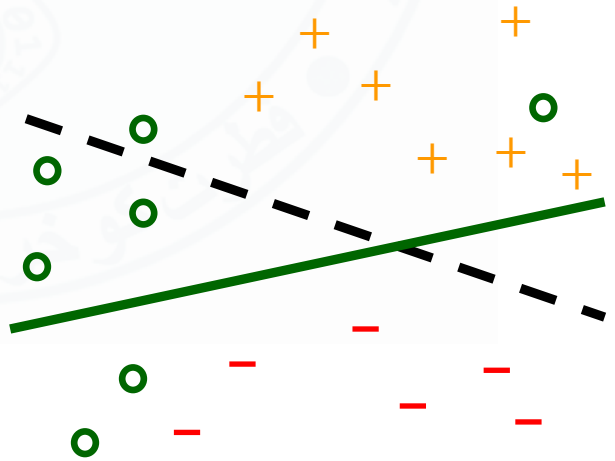

- Minhas, Fayyaz ul Amir Afsar, and Asa Ben-Hur. 2012. "Multiple Instance Learning of Calmodulin Binding Sites." *Bioinformatics* 28 (18): i416–22. doi:10.1093/bioinformatics/bts416.

# Semi-Supervised Classification

- Use unlabeled data set in conjunction with labeled data
- Useful in cases when the number of labeled examples is small
  - Difficult to obtain labeled examples
- Semi-Supervised SVM
- Can use indirect labeling constraints over unlabeled data as well when available
  - Let's say we know that a pair of example has the same (but unknown) or different labels

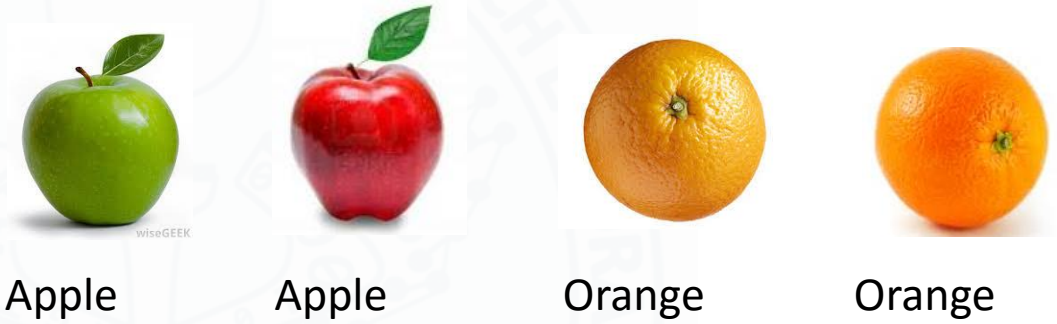
Apple Apple Orange Orange

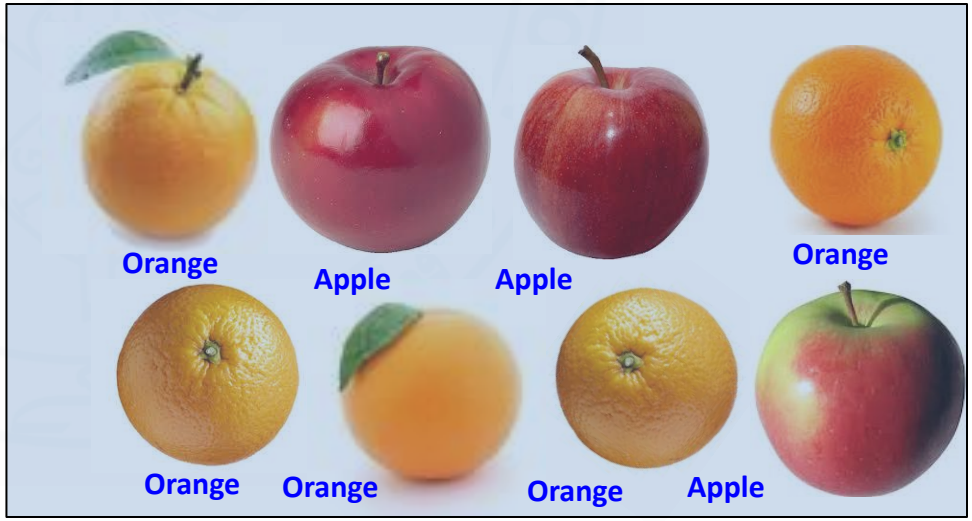




# Transductive SVM

- **Use unlabeled data set in conjunction with labeled data to specifically label the unlabeled data only**
- **Transductive SVM**



# Ranking SVM

- **Assign a rank to an input example**
- **Apple grading**
  - **U.S. Extra Fancy**
  - **U.S. Fancy**
  - **U.S. No. 1**
  - **U.S. No. 1 Hail**
  - **U.S. Utility**

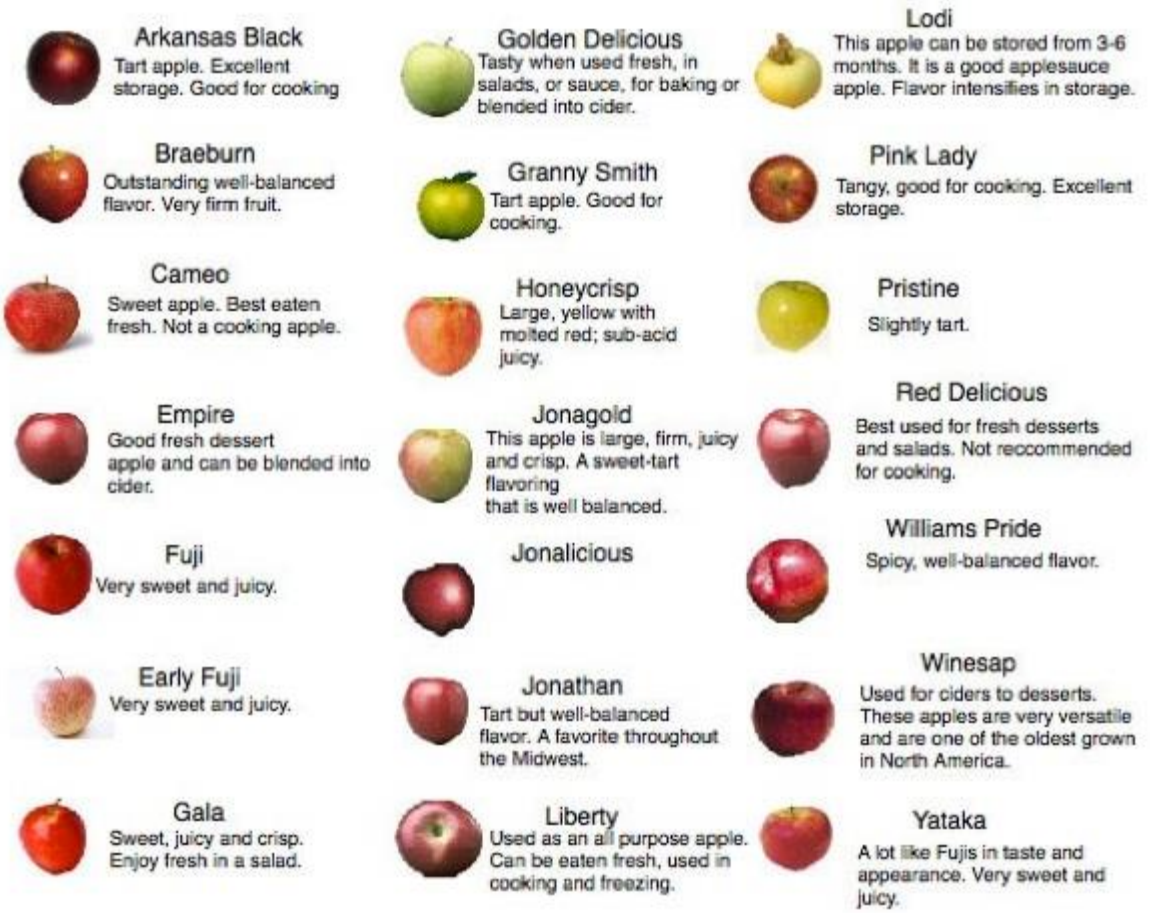
If $i$ is 'US Extra fancy' and $j$ is any other category then we want the score of $i$ to be higher than that of $j$, i.e.,:

$$f(x_i) \geq f(x_i) + 1 - \xi_{ij}$$

These are the constraints in a ranking SVM. The objective function minimizes the sum of slacks and maximizes the margin.

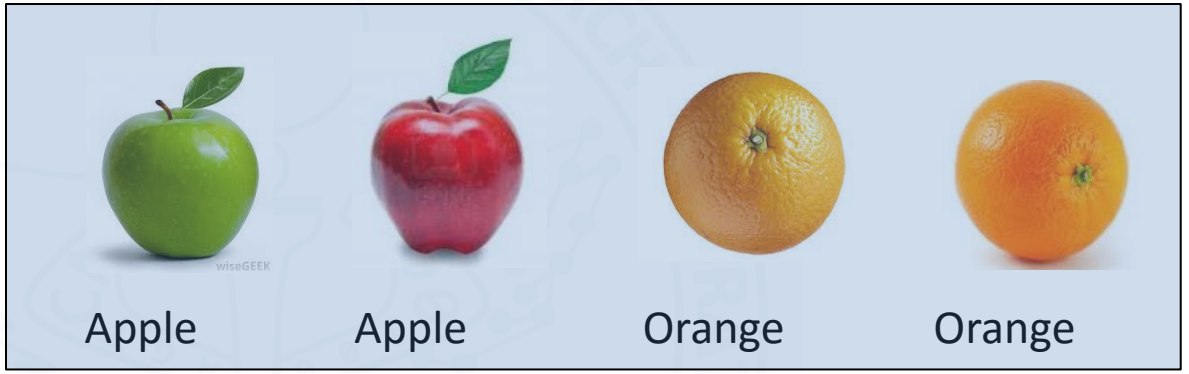
# Multi-task learning

- **Simultaneously predicting related tasks**
  - **Predict Apple Grade and Apple Variety Simultaneously**
  - **Tasks are related**
- **If the tasks are related to each other then it might be useful to do multi-task learning instead of training independent classifiers**

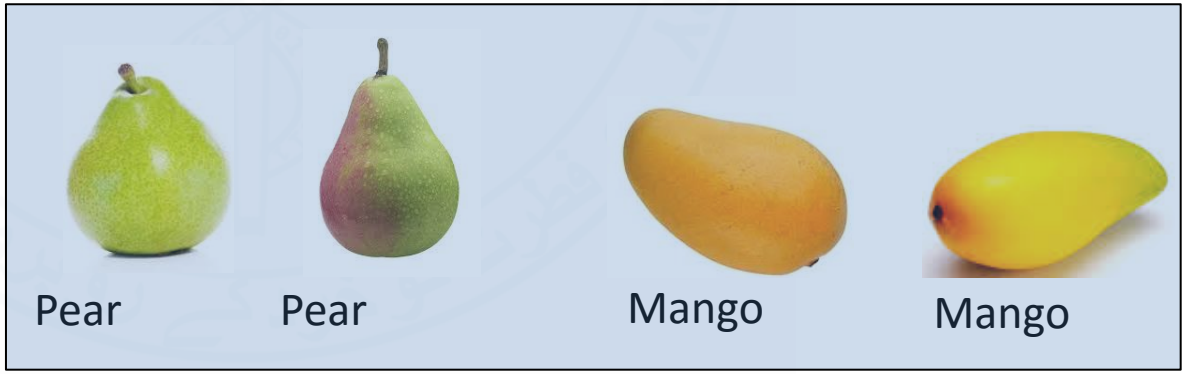


# Transfer learning

- **Use information from one task to learn to classify another**




**Task-1**




**Task-2**

# Multi-label learning

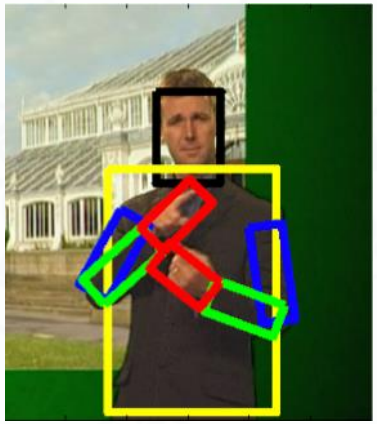
- In certain problems, an example can belong to more than one class
  - Example
    - Does a picture have:
      - Apples
      - Oranges
      - Mangoes

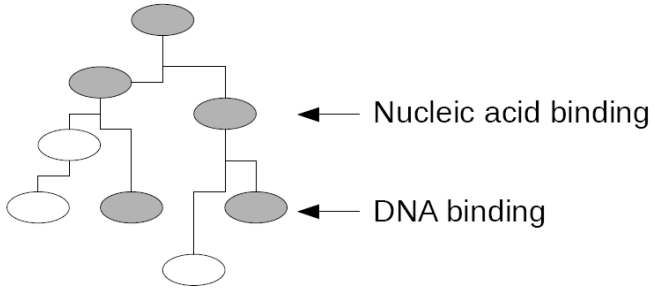
Apple, Orange, Mangoes

Apple

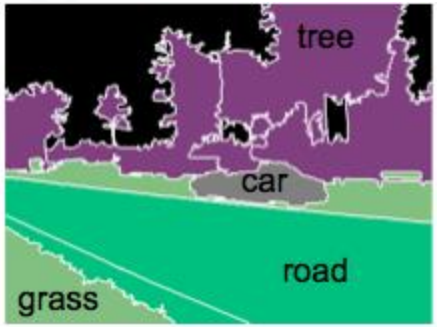
# Structured output learning

- When the output is a structured object
  - Single label or real value
  - Vector
  - Directed Acyclic Graph
  - Relationships between output variables
- Structured SVM
- Most generic
  - Can be used for
    - Multi-class
    - Multi-label
    - …



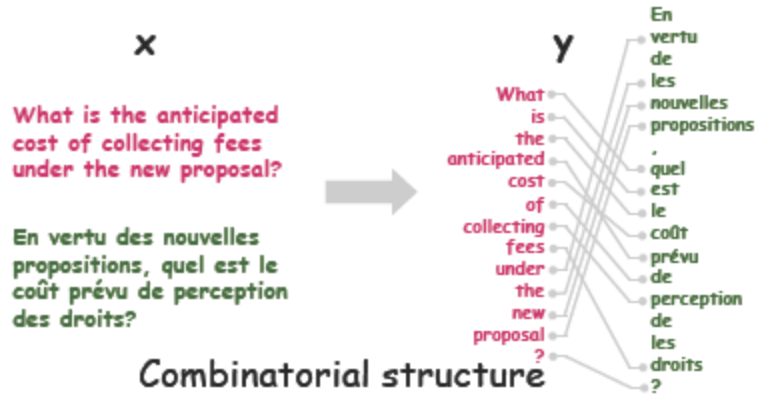



(0, 0, 1, 1, 1, 0, 1, 1)

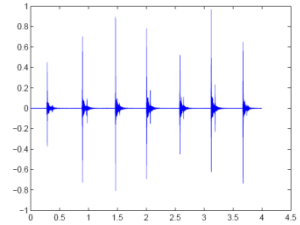


# Structured output learning

- Machine translation



# Structured output learning



- Sequence labeling as structured output learning
  - Given a sequence, predict the labels
  - Example:
    - Finding what keys were pressed using audio recording of keyboard emnations
      - Uses a hidden markov model
      - Can use a structured SVM here

Text recognized by the HMM classifier, with cepstrum features (underlined words are wrong),

> the big money fight has drawn the shoporo od dosens of companies in the entertainment industry as well as attorneys gnnerals on states, who fear the fild shading softwate will encourage illegal acyivitt, srem the grosth of small arrists and lead to lost cobs and dimished sales tas revenue.

Text after spelling correction using trigram decoding,

> the big money fight has drawn the support of dozens of companies in the entertainment industry as well as attorneys generals in states, who fear the film sharing software will encourage illegal activity, stem the growth of small artists and lead to lost jobs and finished sales tax revenue.

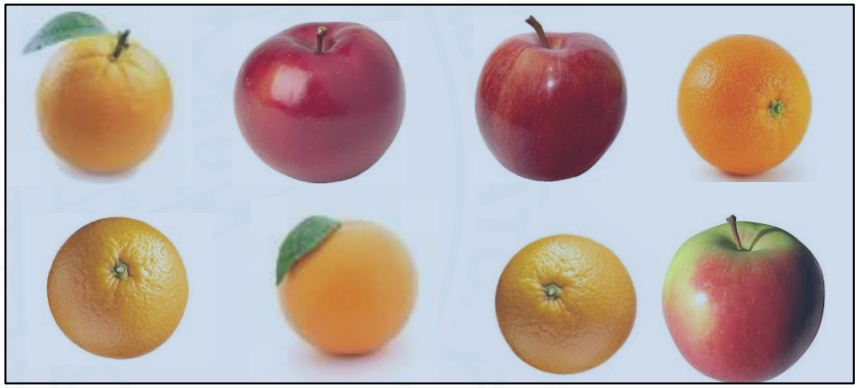
Original text. Notice that it actually contains two typographical errors, one of which is fixed by our spelling corrector.

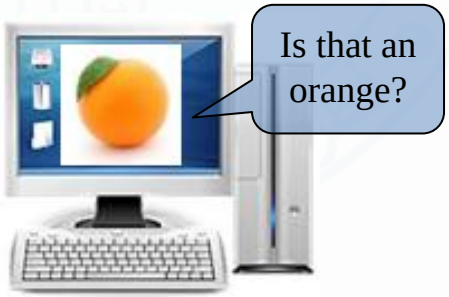
> the big money fight has drawn the support of dozens of companies in the entertainment industry as well as attorneys gnnerals in states, who fear the file sharing software will encourage illegal activity, stem the growth of small artists and lead to lost jobs and dimished sales tax revenue.

# Active Learning

- **Active Learning**
  - Special case of semi-supervised learning in which a learning algorithm is able to interactively query the user (or some other information source) to obtain the desired outputs at new data points
  - Useful when
    - Manual labeling is expensive
      - Small set of labeled examples
      - Large set of unlabeled examples
    - Can ask about the labels of some examples in the unlabeled set from an oracle
  - Don't ask too much and the oracle can also make errors: Proactive learning

Apple Apple Orange Orange







The Oracle (The matrix, 1999)

# Online Learning

- **Online Learning**
  - **Learn and unlearn concepts**
  - **Concepts can change over time**
  - **Example: Price of apples vs. their grade**
  - **Online learning allows incremental learning over time so we don't have to retrain the classifier every time**

*Online learning is a method of machine learning in which data becomes available in a sequential order and is used to update our best predictor for future data at each step, as opposed to batch learning techniques which generate the best predictor by learning on the entire training data set at once.*

The illiterate of the 21st century will not be those who cannot read and write, but those who cannot learn, unlearn, and relearn.

- Alvin Toffler

# Representation Learning

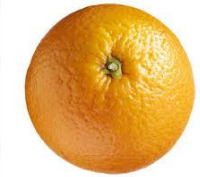
- Learn how to represent data for different machine learning tasks
  - Replacement of feature engineering or hand crafted features
- Deep Learning
- Sparse representations

**Google's Artificial Brain Learns to Find Cat Videos**

# Self-Taught Learning

- **Self-taught learning**
  - **Learn features automatically using completely unrelated images**

Apple



Orange

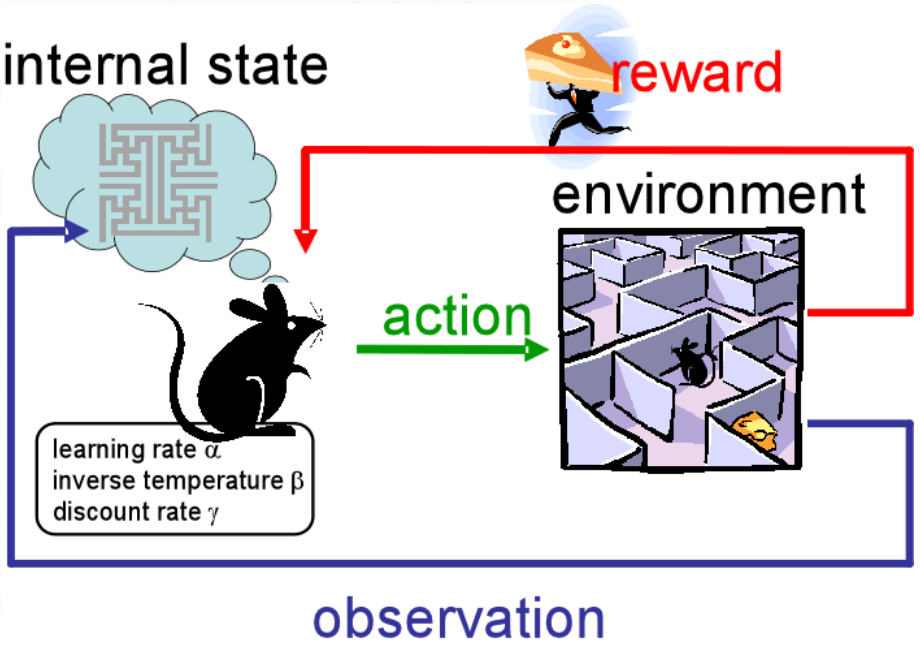
# Recommender Systems

- Predict the rating or preferences
  - Collaborative Filtering
    - Recommendations are based on a user's past behavior (items previously purchased or selected or rating of products, etc.)
  - Content Based Filtering
    - Use characteristic of an item to recommend additional items with similar properties

Gomez-Uribe, Carlos A., and Neil Hunt. 2015. "The Netflix Recommender System: Algorithms, Business Value, and Innovation." *ACM Trans. Manage. Inf. Syst.* 6 (4): 13:1–13:19. doi:10.1145/2843948.

# Reinforcement Learning

- Reinforcement learning
  - Learn to fly a plane or play Tic-Tac-Toe by yourself based only on rewards or penalties
  - If a function that gives rewards or penalties based on how the machine behaves, you can develop a reinforcement algorithm for such a problem
    - Maximize rewards!



# End of Lecture

We want to make a machine that will be proud of us.

- Danny Hillis