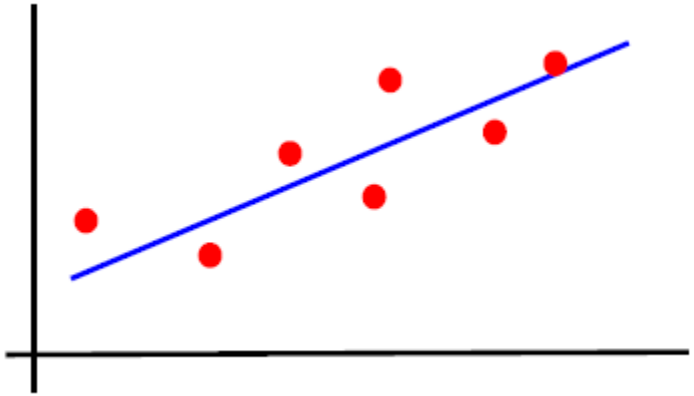
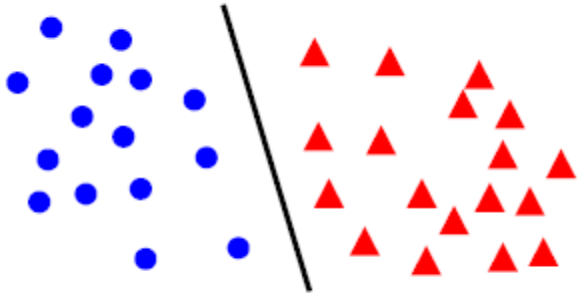
Machine Learning

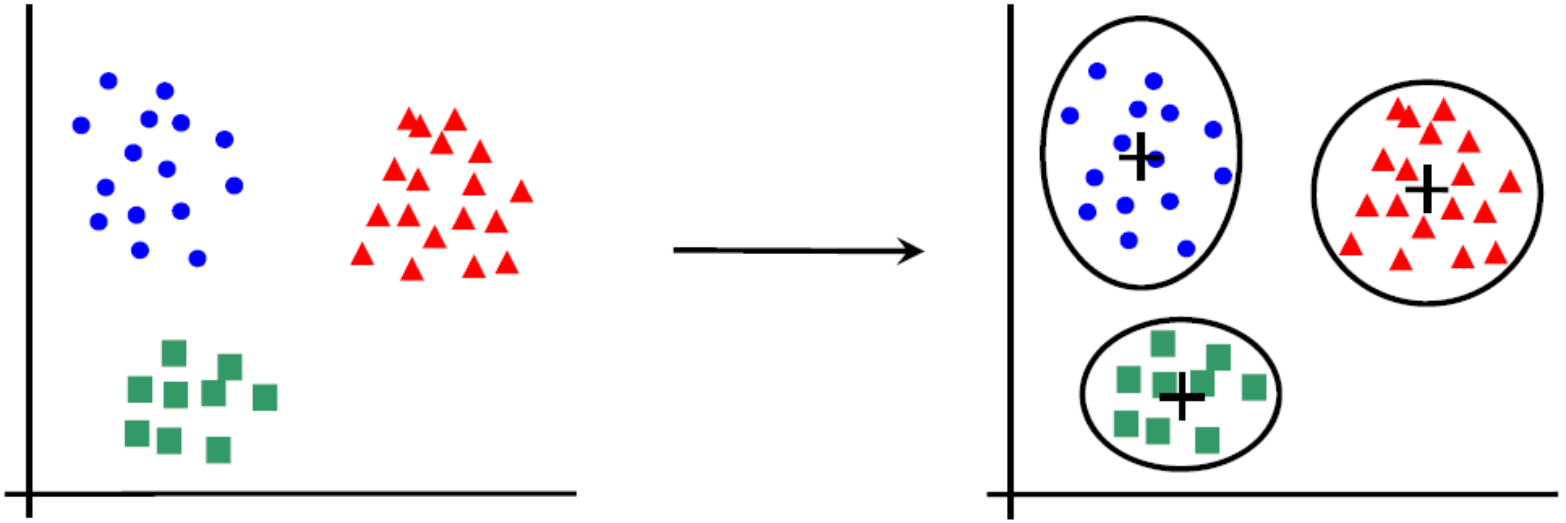
Introduction

Indian Institute of Information Technology
Sri City, Chittoor

Welcome to Machine Learning Class







Today's Agenda

- Course plan
  - Pre-requisite
  - Topics
  - Textbooks/References
  - Evaluation components
  - Honor code
- Introduction to machine learning
  - What is ML?
  - When do we use ML?
  - Applications
  - Relation with AI and DL
  - Relation with other fields
  - Different machine learning paradigms

Pre-requisite

- Probability
  - Distribution, random variable, expectation, conditional probability, variance, density
- Linear algebra
  - Matrix multiplication
  - Eigen vector
- Basic programming
  - Python (First Priority)
  - Matlab/C/C++ (Second Priority)

Topics

- Supervised Learning
  - Classification
  - Regression
- Unsupervised Learning
  - Clustering
- Reinforcement Learning
- Semi-supervised Learning

Textbooks/References

1. "Pattern Classification" by R. O. Duda, P. E. Hart and D. G. Stork.
2. "An Introduction to Statistical Learning" by Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani.
3. "Pattern Recognition and Machine Learning" by Christopher M. Bishop.
4. "Introduction to Machine Learning" by Ethem Alpaydin.
5. "Pattern Recognition: An Algorithmic Approach" by M. Narasimha Murty, V. Susheela Devi.
6. "Machine learning" by Tom Mitchell.

Evaluation Components

- Mid-Exam: 20%
- End-Exam: 30%
- Assignments: 30%
- Scheduled Quiz: 10%
- Class Participation: 10%

Honor Code

Do's

- Write down the code independently
- Submit the assignment within the deadline
- Read the books/references for detail description of the topics

Don'ts

- copy, refer to, or look at any official or unofficial previous years' solutions in preparing the answers

Introduction to ML

- What is ML?
- Terminologies used in ML
- When do we use ML?
- Applications
- Relation with AI and DL
- Relation with other fields
- Different machine learning paradigms

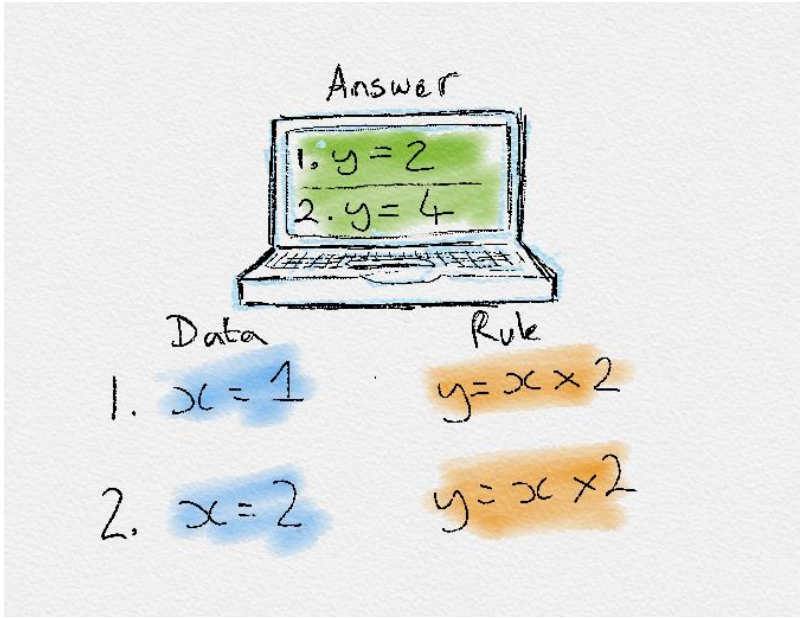
What is ML?

- Machine learning (ML) is the study of computer algorithms that improve automatically through experience.

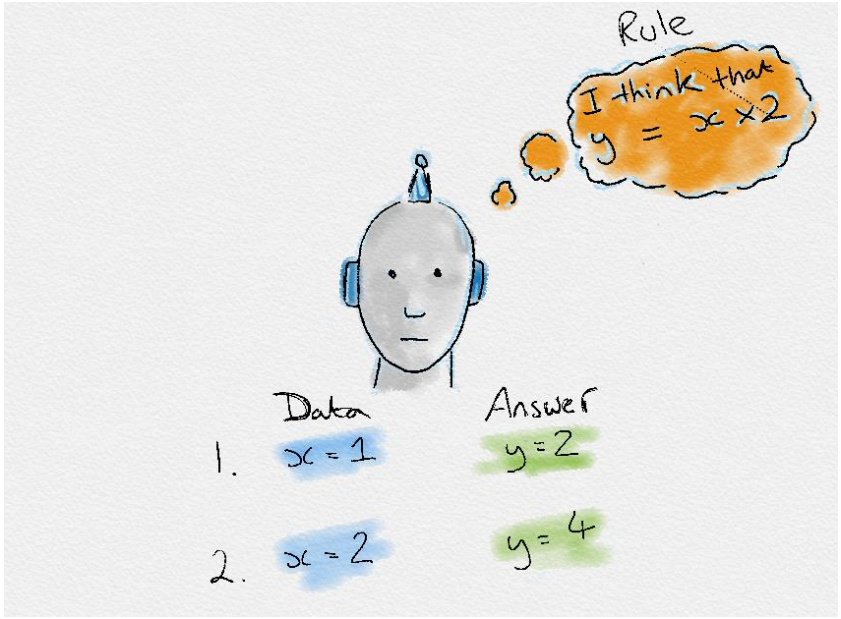
- Machine-learning algorithms use statistics to find patterns in massive amounts of data.

- Traditionally, software engineering combined human created rules with data to create answers to a problem. Instead, machine learning uses data and answers to discover the rules behind a problem – **F. Chollet, Deep Learning with Python**

What is ML?



Traditional Programming



Machine Learning

Terminologies used in ML

- ML systems learn how to make inference from the input data samples to produce useful predictions on un-seen (test) data.
- Input data:
  - labelled examples: A labelled example includes feature(s) and the label. {features, label}: (x, y)For e.g.:

Features:	Label
Normal RBC, Normal HgB	Healthy
Low RBC, Low HgB	Anaemic

  - unlabelled examples: An unlabelled example contains features but not the label. {features, ?}: (x, ?)
    - For e.g.:

Features:
Housing type: 4BHK, Price: 40,000
Housing type: 4BHK, Price: 15,000
Housing type: 2BHK, Price: 25,000
Housing type: 2BHK, Price: 8,000

Terminologies used in ML

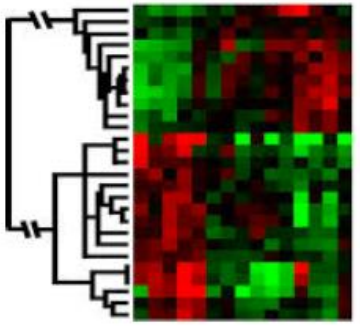
- Machine Learning Model:
  - A ML model defines the relationship between the features and label.
    - For e.g.: An anaemia diagnostic model might associate certain features strongly with “anaemic“ or “healthy”, and predict the labels based on the association rules it inferred.
  - Two Phases of ML model development
    - **Training** means creating or **learning** the model.
    - **Testing/Inference** means applying the trained model to unlabelled examples.

When do we use ML?

ML is used when:

- Human expertise does not exist (navigating on Mars)
- Humans can't explain their expertise (speech recognition)
- Models must be customized (personalized medicine)
- Models are based on huge amounts of data (genomics)



Learning isn't always useful:

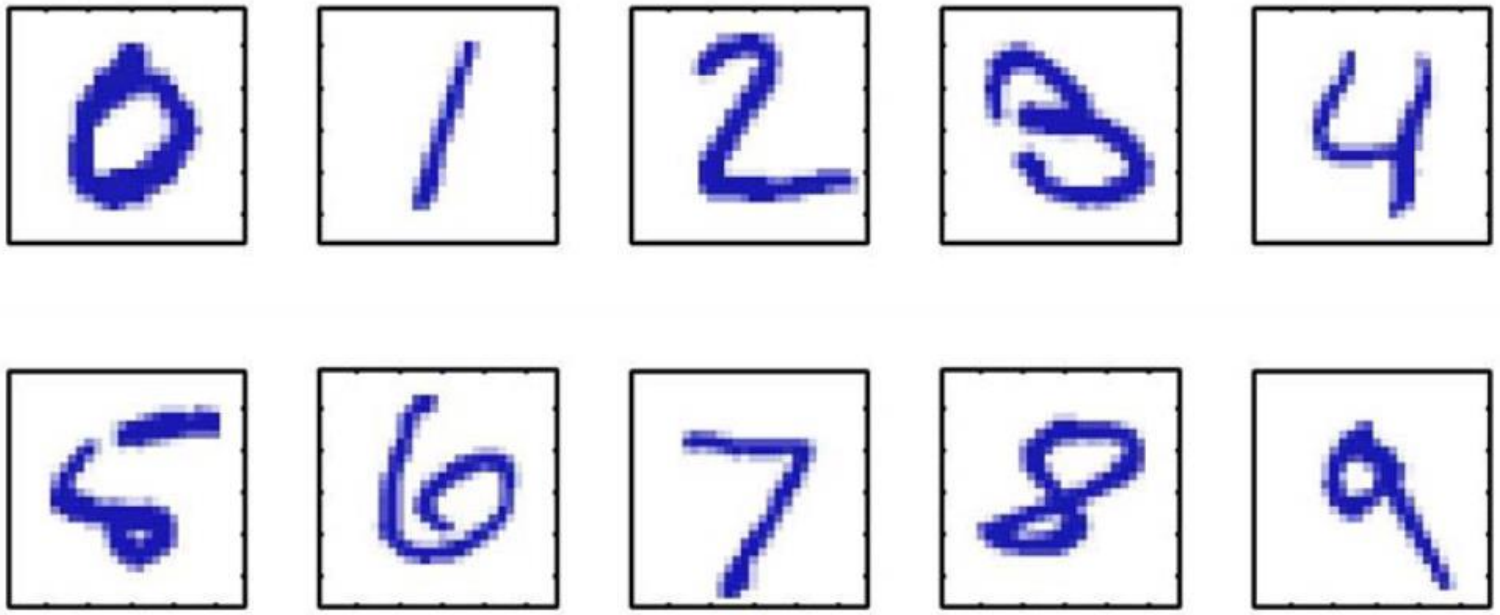
- There is no need to "learn" to calculate payroll

Slide adapted from **E Alpaydin**

Applications

- Hand-written digit recognition
- Speech recognition
- Face detection
- Object classification
- Email spam detection
- Computational biology
- Autonomous cars
- Computer-aided diagnosis

Hand-written Digit Recognition



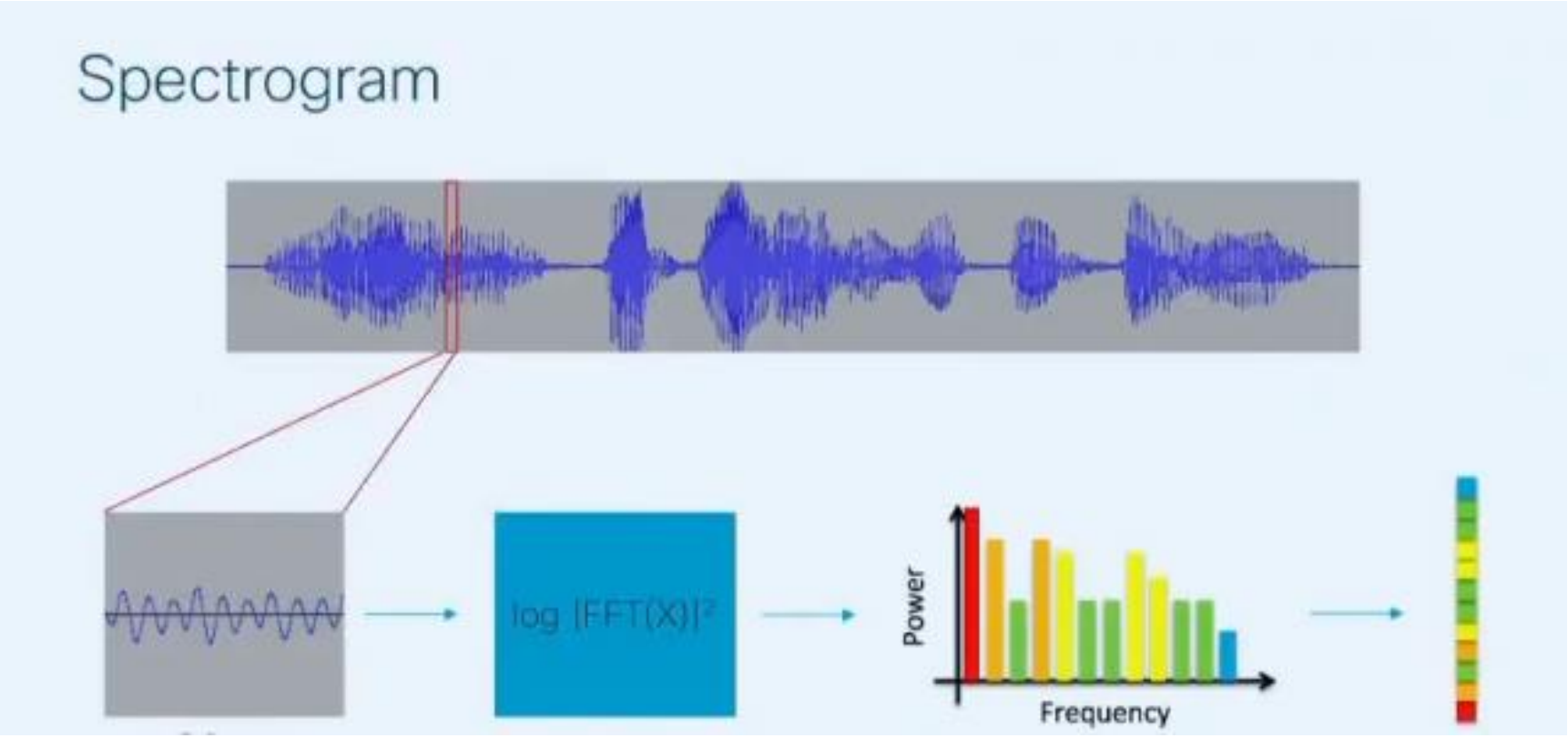
Images are 28 x 28 pixels

Represent input image as a vector $\mathbf{x} \in \mathbb{R}^{784}$

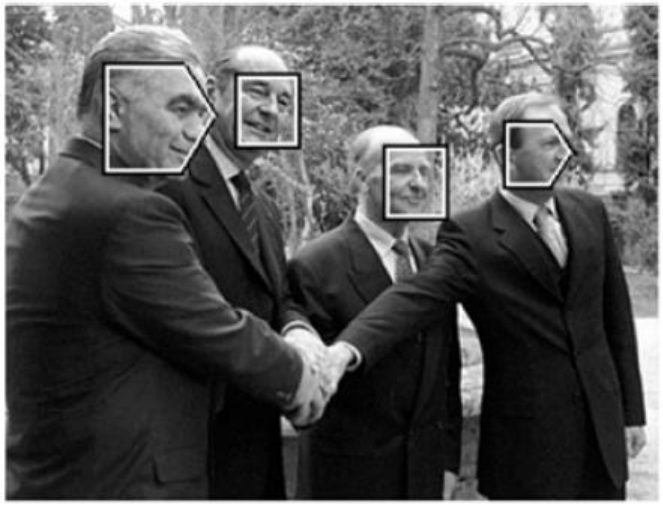
Learn a classifier $f(\mathbf{x})$ such that,

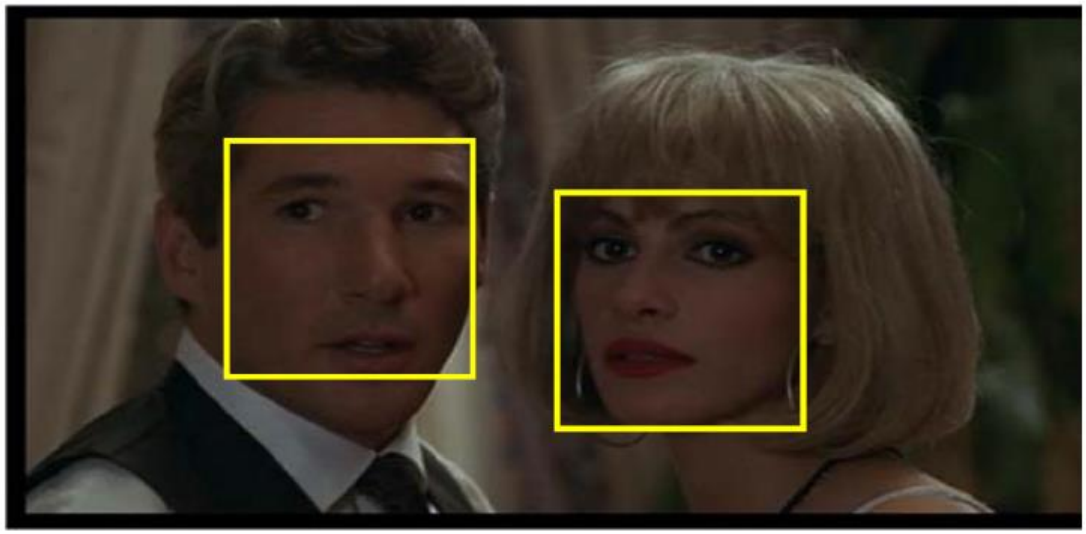
$$f : \mathbf{x} \rightarrow \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$$

Speech Recognition

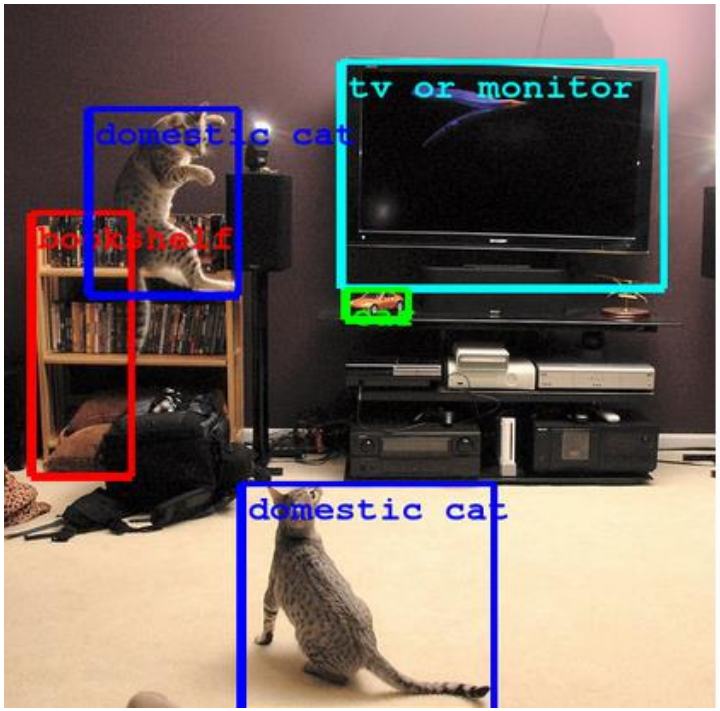


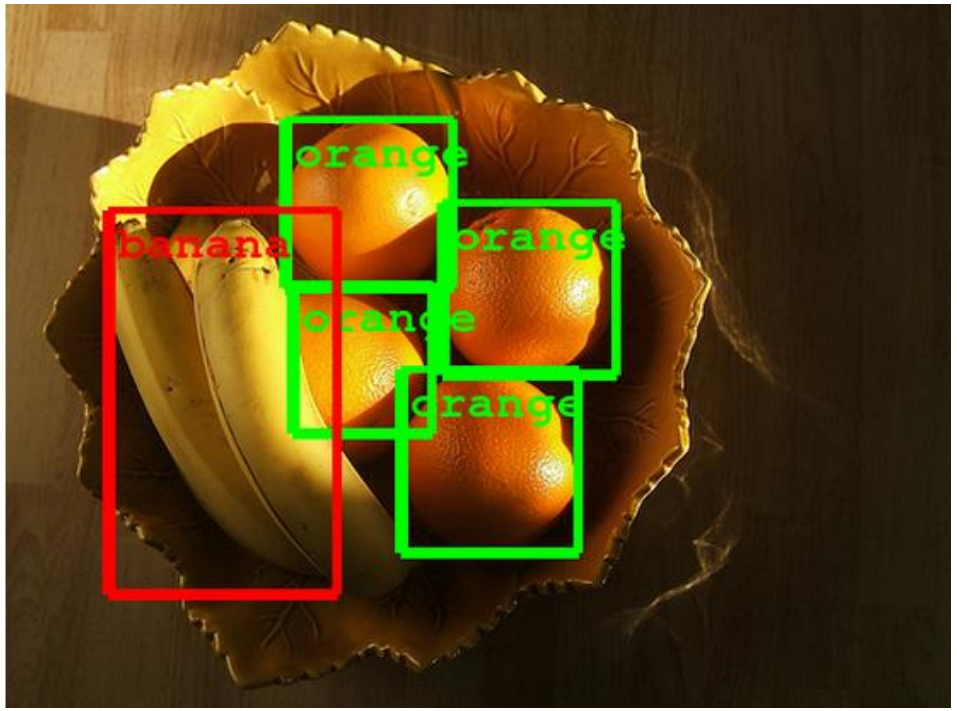
Face Detection



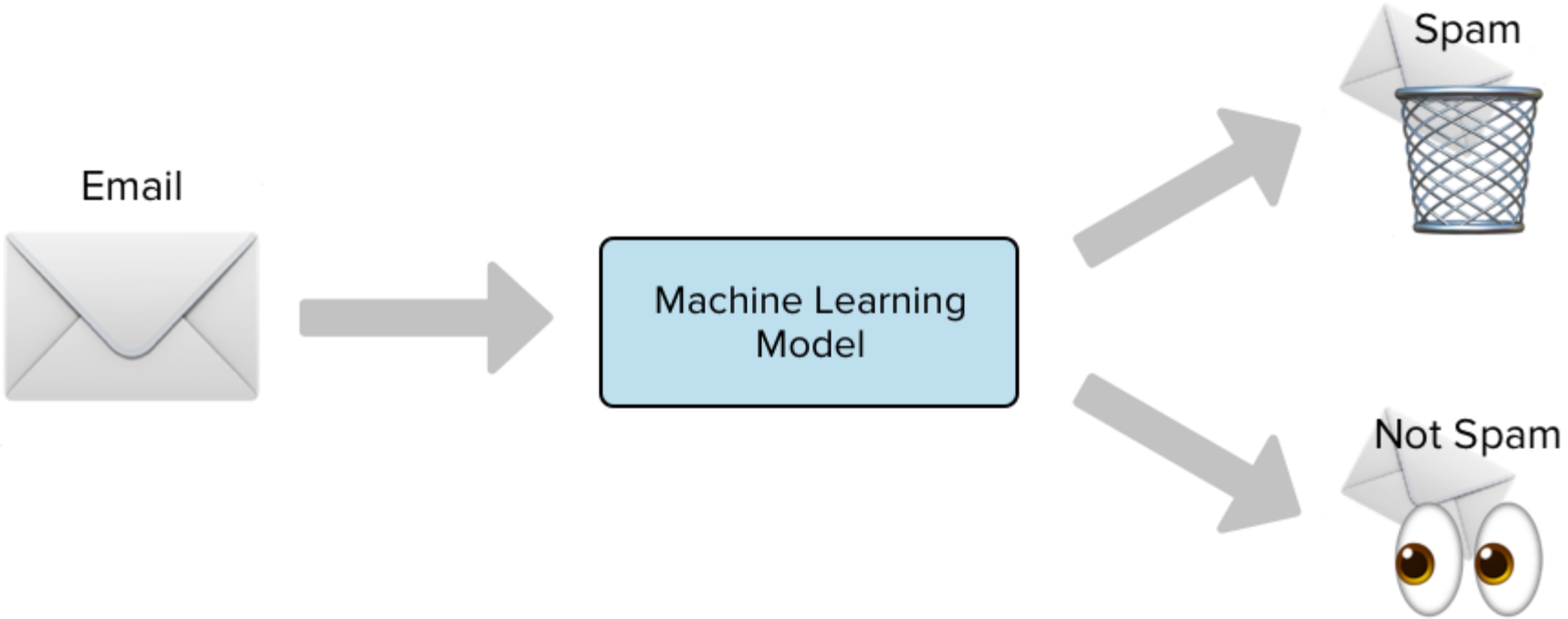


Object Classification





Email Spam Detection

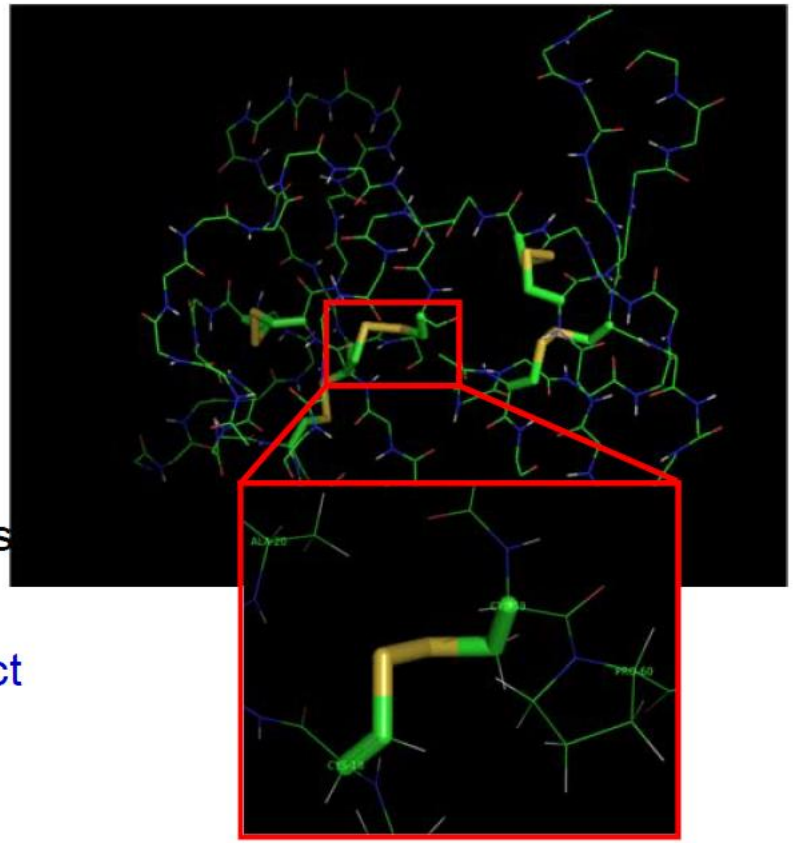


Computational Biology

x

y

AVITGACERDLQCG
KGTCCAVSLWIKSV
RVCTPVGTSGEDCH
PASHKIPFSGQRMH
HTCPCAPNLACVQT
SPKKFKCLSK

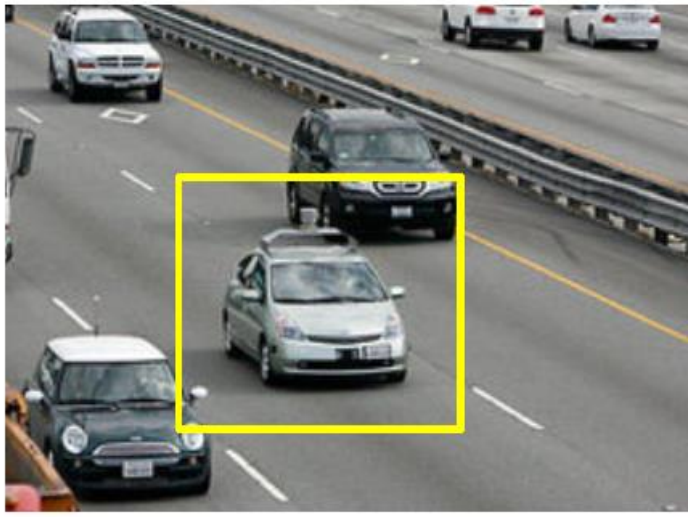


Protein Structure and Disulfide Bridges

Regression task: given sequence predict 3D structure

Protein: 1IMT

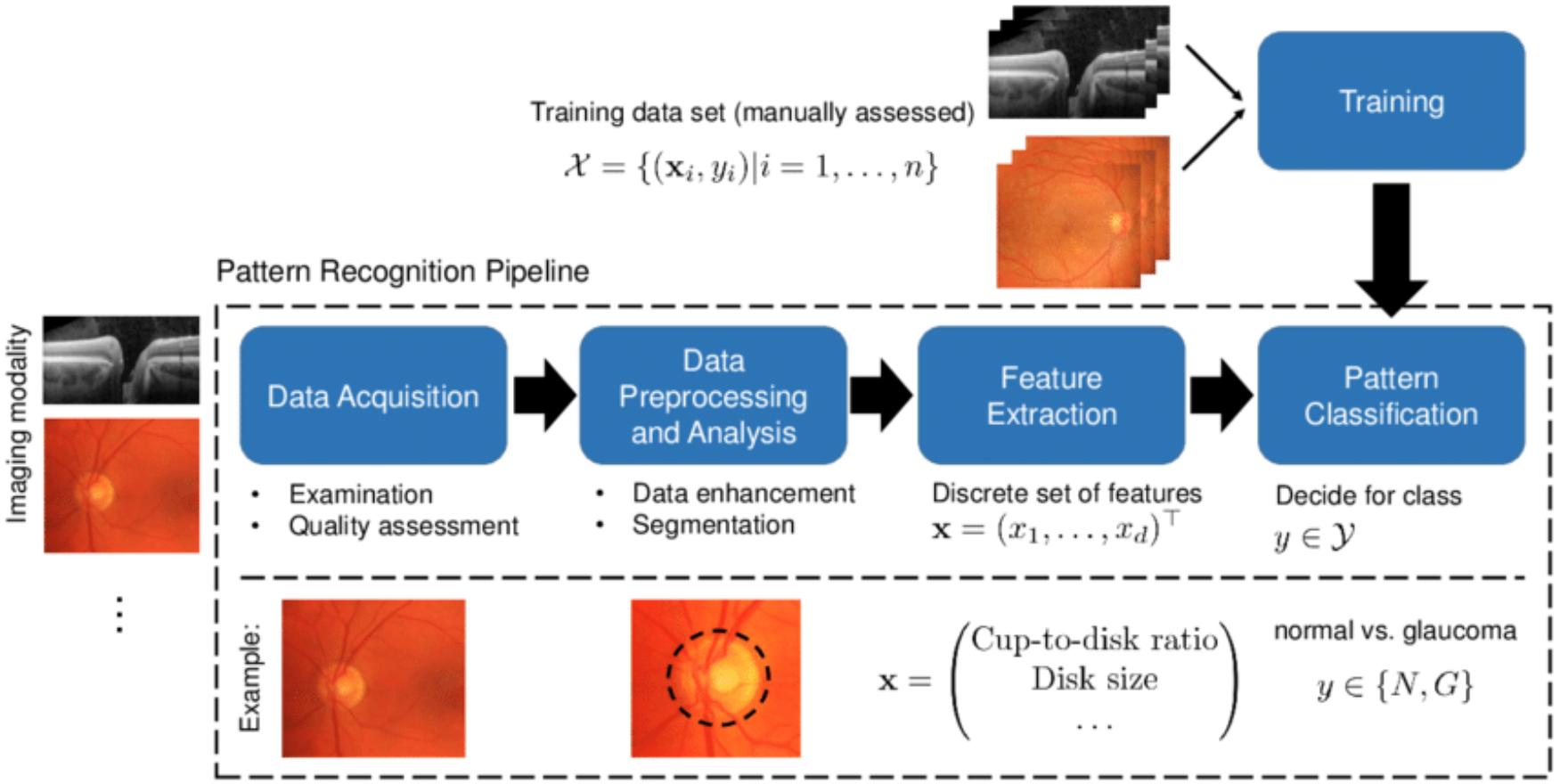
Autonomous Cars



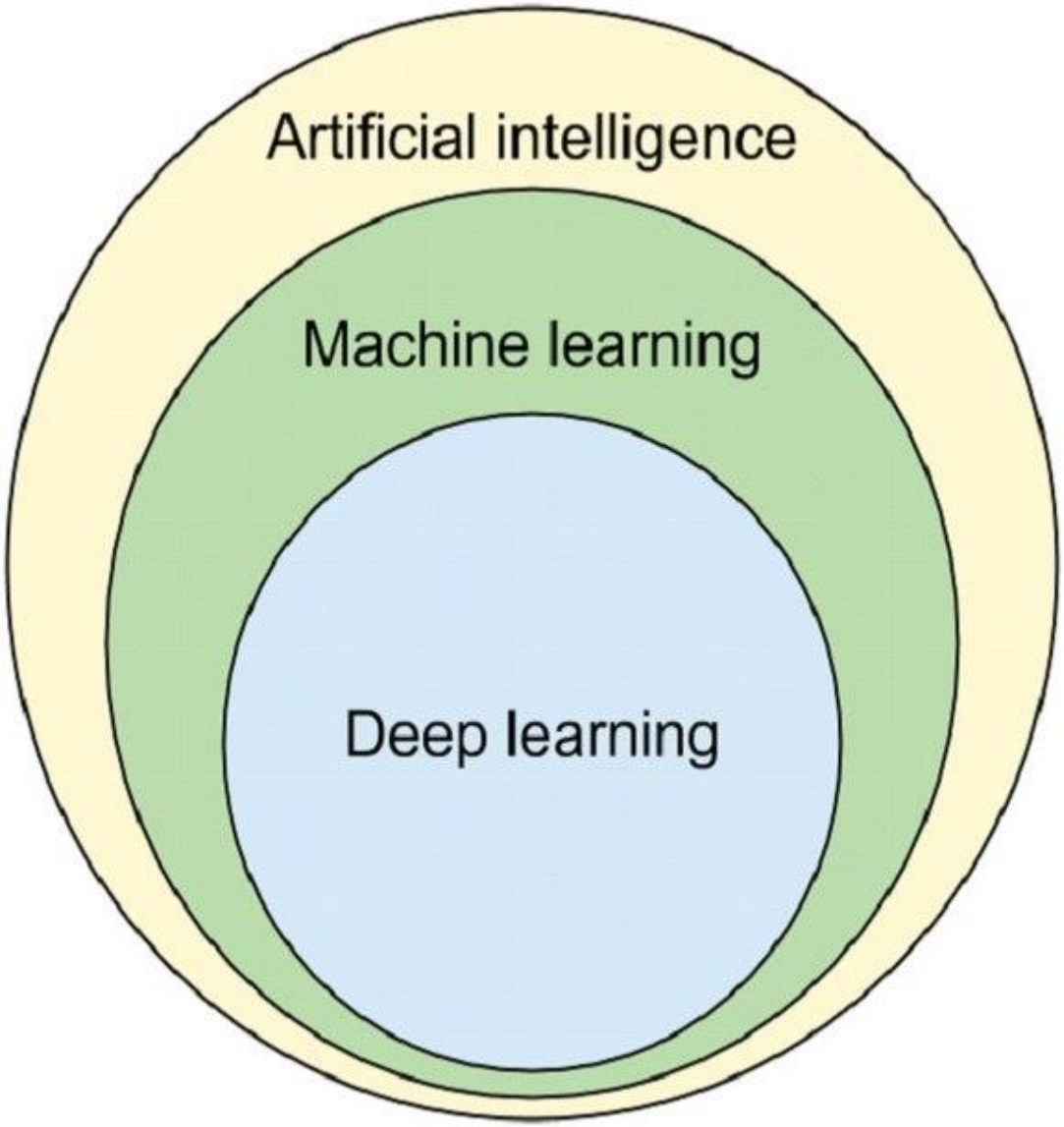
- Nevada made it legal for autonomous cars to drive on roads in June 2011
- As of 2013, four states (Nevada, Florida, California, and Michigan) have legalized autonomous cars

Penn's Autonomous Car →
(Ben Franklin Racing Team)

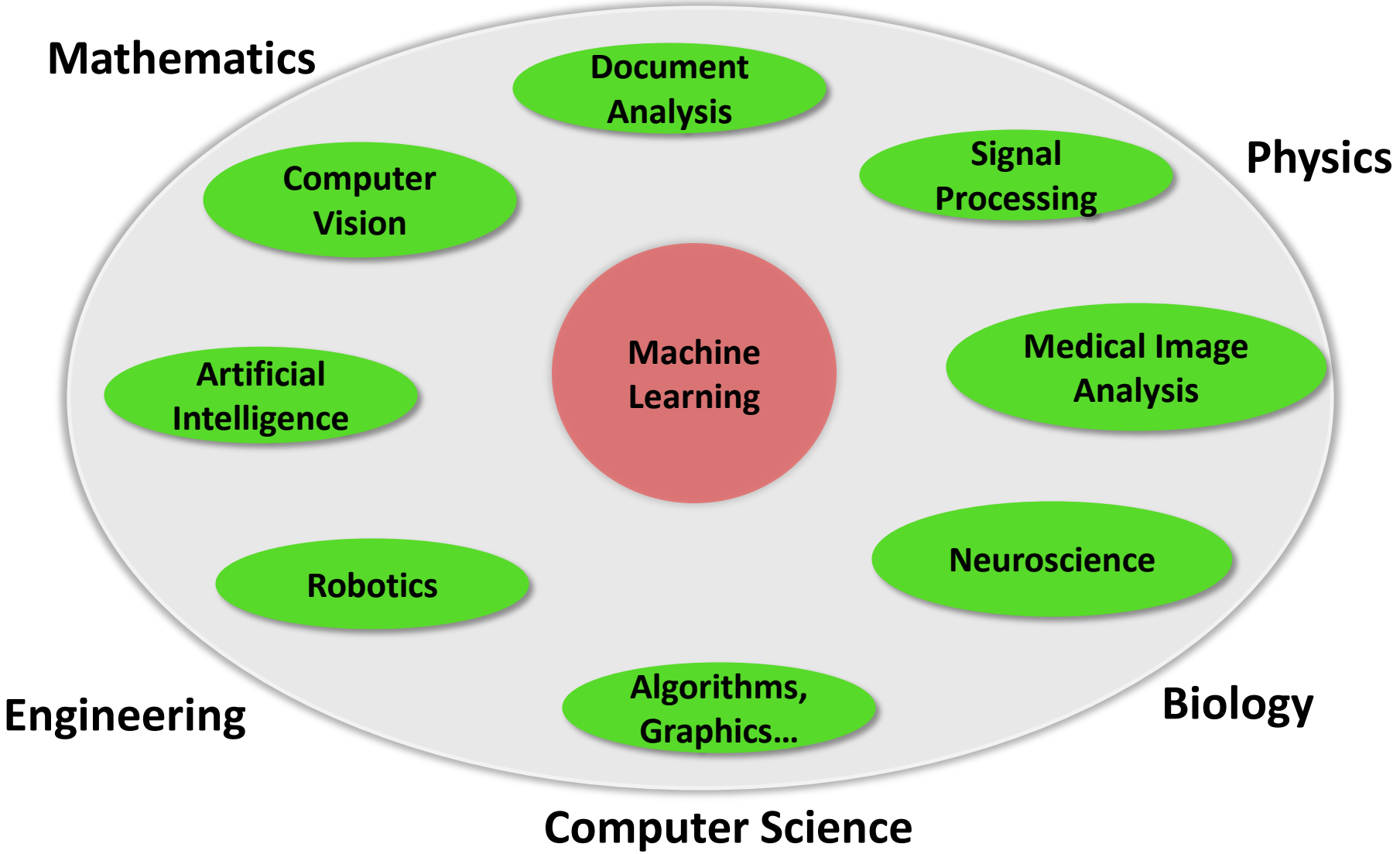
Computer-aided Diagnosis



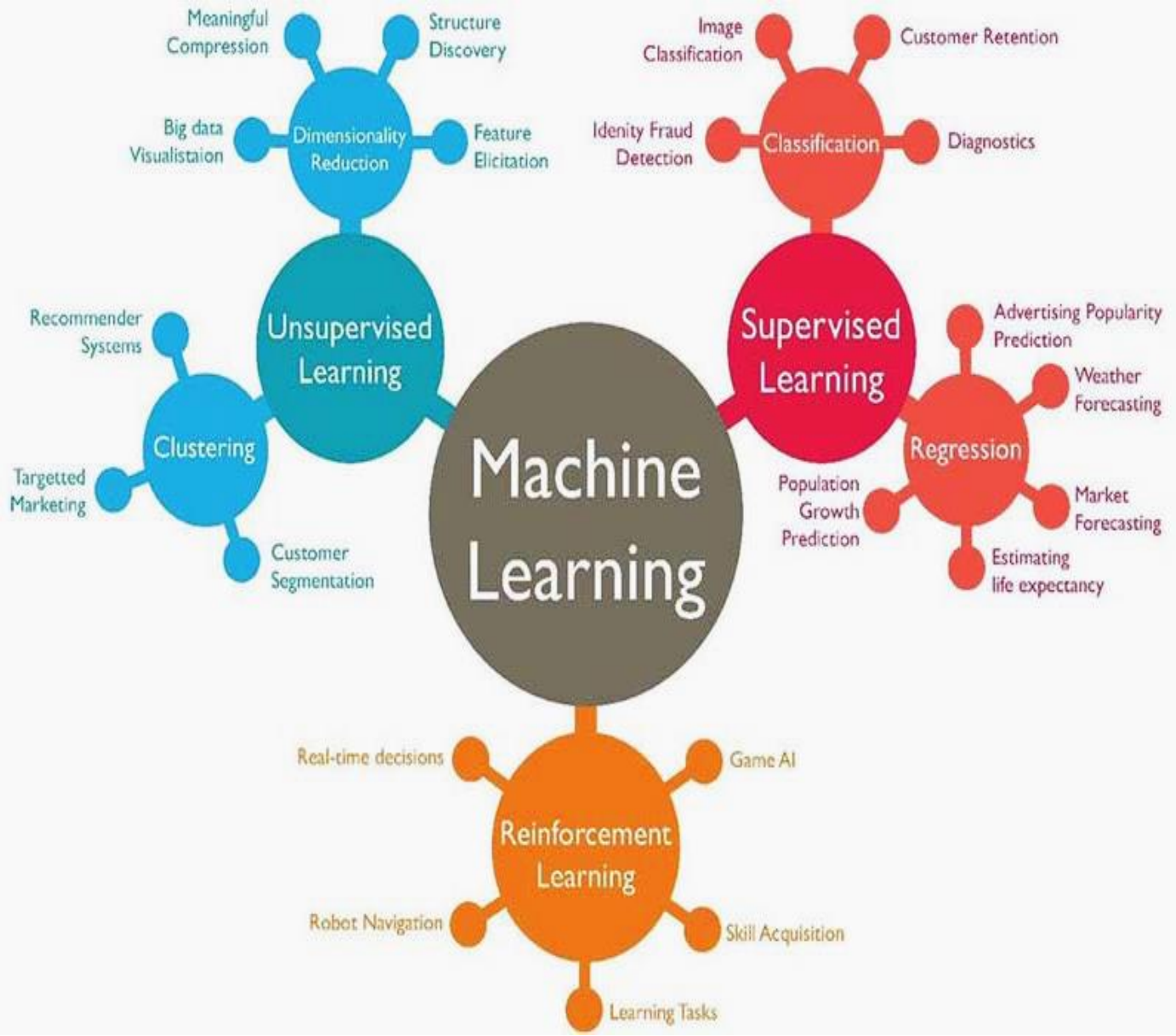
Relation with AI and DL



Relation with Other Fields



Different Machine Learning Paradigms



Thank You: Question?