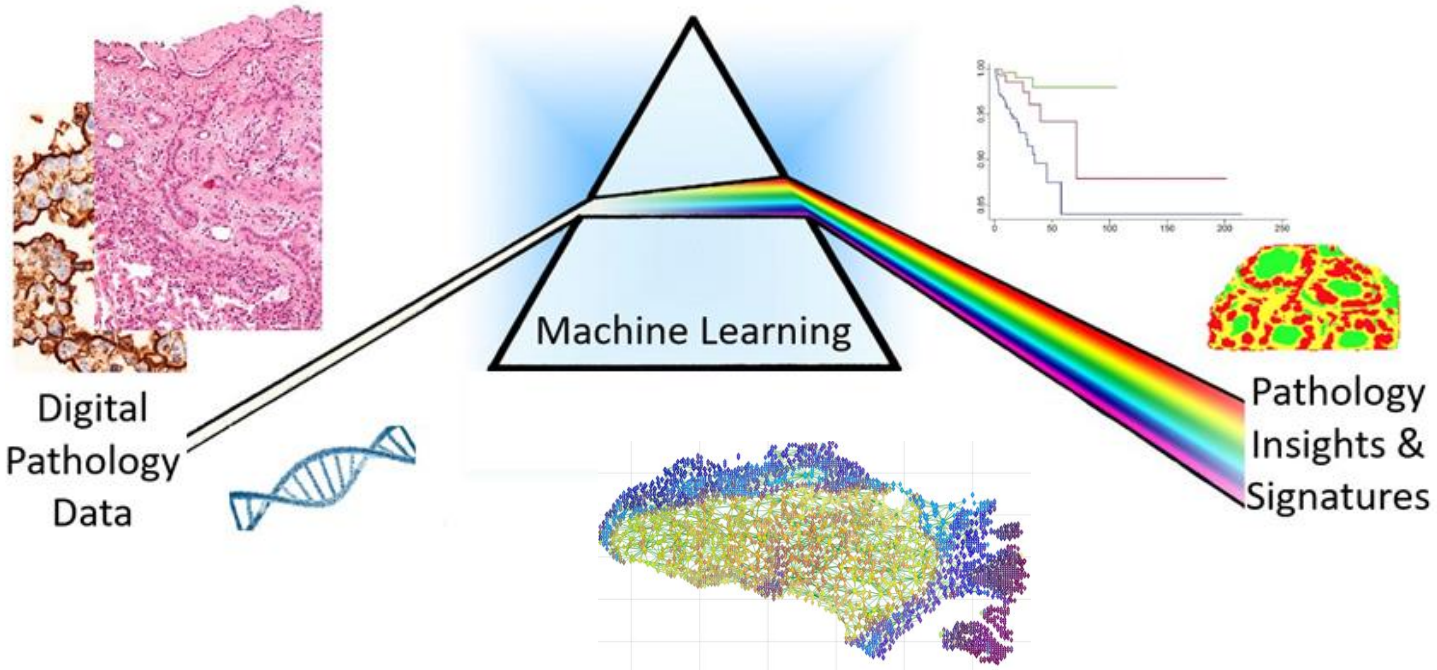

# Introduction

**Dr. Fayyaz Minhas**

Department of Computer Science
University of Warwick
https://warwick.ac.uk/fac/sci/dcs/teaching/material/cs909/

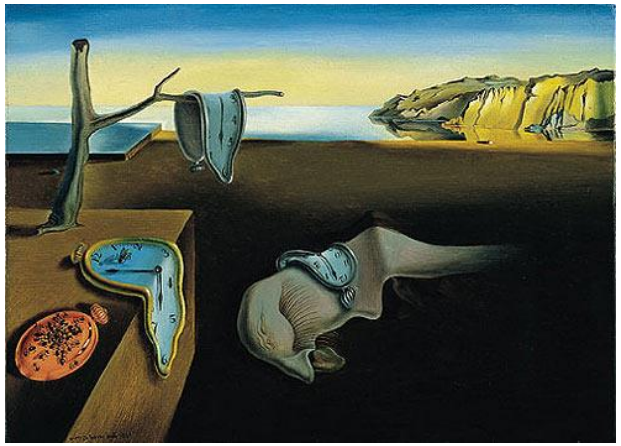
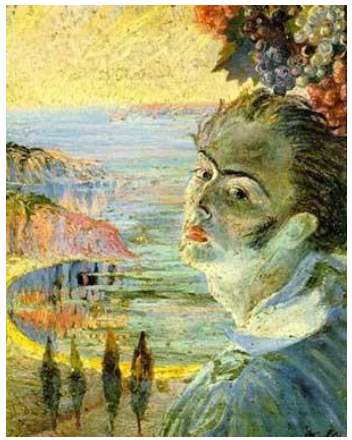
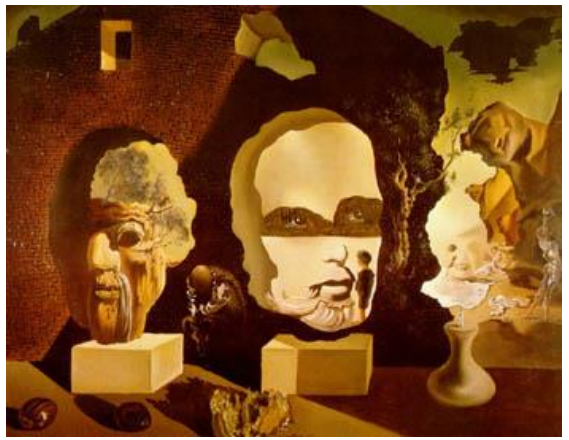
# Introduction

- Designing bespoke data science / machine learning models
  - Computational Biology & Pathology



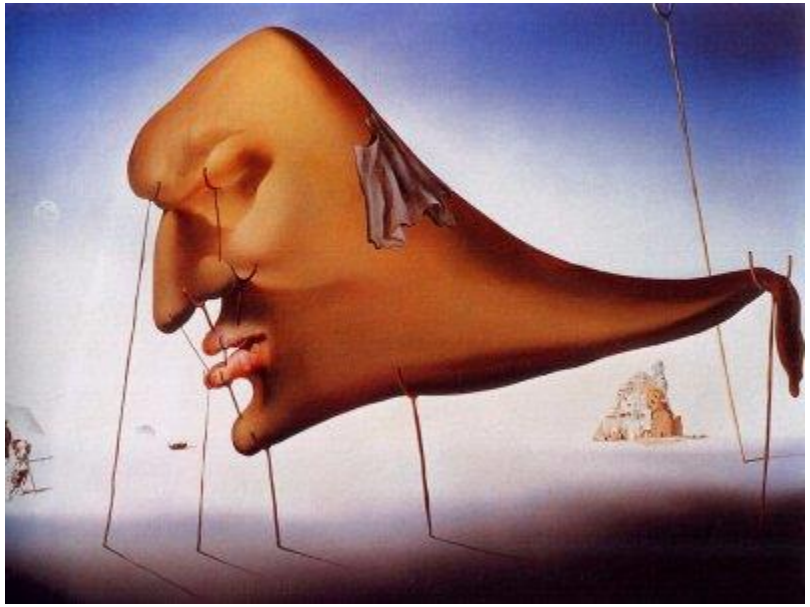

# Paintings by two different painters



# Who’s painting is this?

# And this?



*learning from data for generalization to unseen cases*

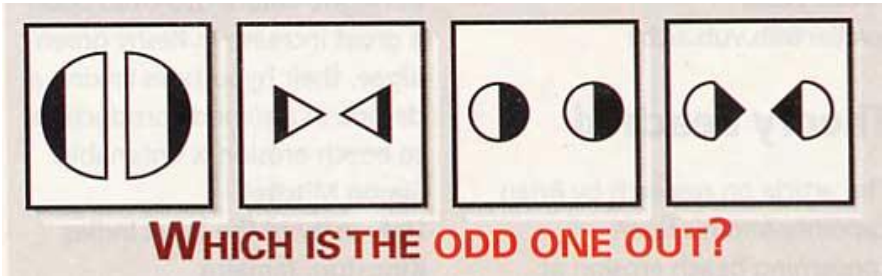
Apples


Oranges

# What is this?

# How many categories (clusters) are there?

# Find the odd one out!



# Predict the series

- 1,1,2,3,5,8,13,…

# Question?

- Consider the vectors
  - $X_1=[1\ 2\ 1\ 4]^T$
  - $X_2=[2\ 4\ 2\ 4]^T$
  - $X_3=[0\ 0\ 0\ 4]^T$
  - $X_4=[3\ 6\ 3\ 4]^T$
  - $X_5=[4\ 8\ 4\ 4]^T$
- To store each vector, how many dimensions (or variables) do we need?

# Learning to drive


# Questions

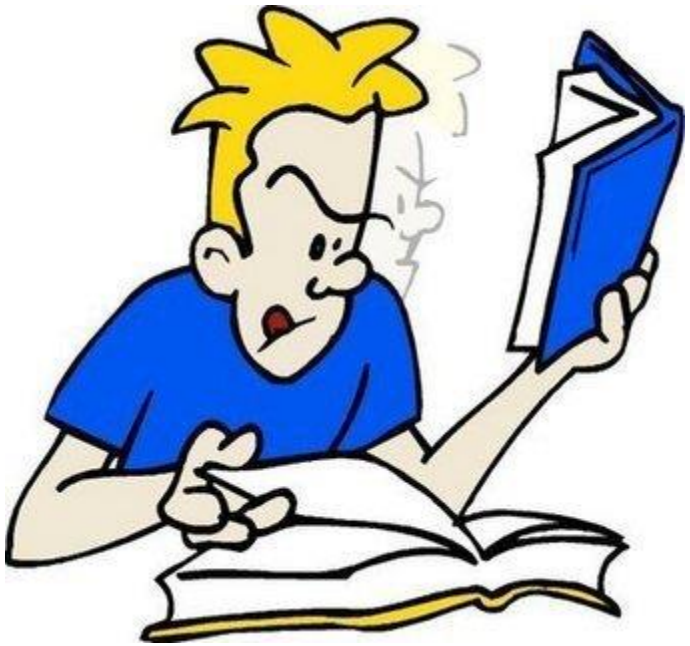
| | |
|---|---|
| • How were you able to recognize that the object shown was indeed an apple? | Classification |
| • How were you able to discriminate between the paintings from two different painters? | Classification |
| • How were you able to find out the different types of apples in the picture? | Clustering |
| • How did you manage to find the next number in the series? | Regression |
| • How were you able to find which dimension was redundant? | Dimensionality Reduction |
| • How were you able to find the odd one out? | Anomaly Detection |
| • Learning to drive / write? | Reinforcement learning |

# Definitions

- Computers are _____.
  - **Dumb**
- Making a machine (computer) perform the same tasks which you have just done is called _____
  - **Artificial Intelligence**
- If you make a computer learn to do these tasks using existing data, then this is called _____
  - **Machine Learning (learning from data)**
- Mining for patterns in data (for understanding it better and making predictions) is
  - **Data Mining**
- Solving a machine learning task using a deep layered network of artificial models of biological neurons (artificial neural network), is called _____
  - **Deep Learning (Learning from data with neural modeling)**

# Well Humans learn, don't they?

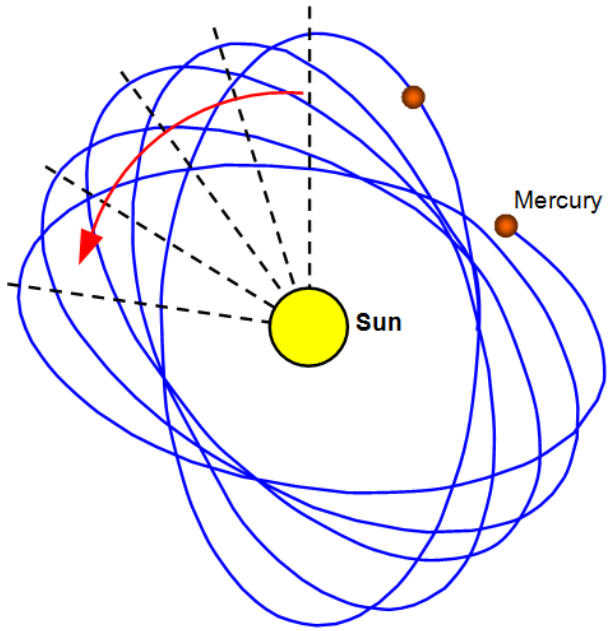
- Of course!
- In many many ways!!
  - Instruction based learning
  - Rote Learning
  - Informal Learning
  - Active Learning
  - Enculturation
  - Experience based learning



- And we have achieved great things with this!!

# Example: Human Learning

- Science is based on developing and testing hypothesis that "explain" our universe
- For example:
  - Newton's Formula F = ma explains the motion of an object of mass m when a force F is applied to it
  - Scientists observed that Newtonian mechanics does not "explain" the motion of mercury properly
  - This led to the development of theory of relativity by Einstein which explains it!!
- We constantly try to develop and refine models of the world and the universe
- However sometimes it gets hard!

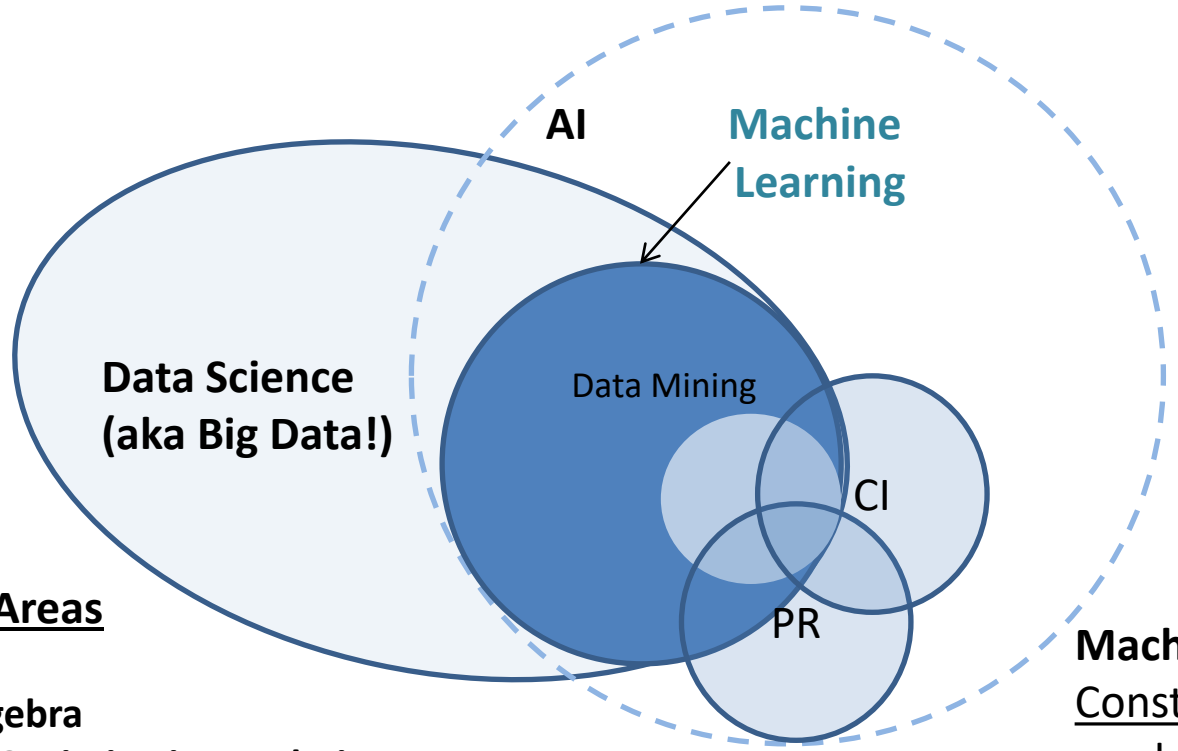


# Why do we need computers?

- …ATTCGAGGATTACACCGTAAGAAATTT…
- …ATCGCCTGATTACATATATACCGTTGG…
- ….AGATTAAATCGTTCGATTCACATTGAC
- **Deduction vs. Induction Reasoning**
- **High dimensions**
- **Required Reading**
  - Halevy, Alon, Peter Norvig, and Fernando Pereira. “The Unreasonable Effectiveness of Data.” *IEEE Intelligent Systems*, 2009.

# Machine Learning

- **A unique junction of computer science, applied mathematics statistics and the world!**




**The Machine Learning Centric View of AI**
(not to scale)

**Related Areas**
**Statistics**
**Linear Algebra**
**Calculus, Optimization Techniques**
**High Performance Computing**
**Algorithms, Data structures and Programming**
**Information Retrieval, NLP, Computer Vision, Signal Analysis**

**Machine Learning** is the Construction of algorithms that can learn from data to "explain" the data and make predictions

# Machine Learning Overview

- **Types**
  - Supervised
    - Classification, Regression, Reinforcement learning
  - Unsupervised
    - Classification, Visualization, Representation
  - Semi-Supervised
- **Approaches**
  - Discriminative
  - Generative
- **Classes of Algorithms**
  - Distance Based
  - Neural Networks
  - Deep Learning
  - Large Margin Methods (Kernels)
  - Ensemble Techniques
  - Logic Based
  - Probabilistic, Bayesian Networks
- **Philosophies**
  - Applied
    - Pattern Recognition
  - Theoretical
  - Developmental
    - Making new learning algorithms
  - **Hybrid**

# This course

- **Aims**
  - Enable students to develop data mining solutions for real world problems
  - Provide a strong base for development of novel data mining and machine learning algorithms
- **Contents**
  - **Weeks 1-6 (Classical ML)**
    - **Introduction**
    - **Classification**
    - **Experiment Design & Theory**
    - **ML Problems**
    - **Dimensionality Reduction**
    - **Ensemble Learning**
  - **Week 7-10 (Deep Learning)**
    - **Multi-layer Perceptron**
    - **Convolutional Neural Networks**
    - **Residual Neural Networks**
    - **Autoencoders**
    - **Advanced Topics (Sequence Learning, Attention, Generative, Geometric)**

# Evaluation

- Two Hour Examination:
  - Meng: 50%
  - MSc: 40%
- Assignments:
  - Assignment-1: 25% of final grade
  - Assignment-2:
    - Meng: 25% of final grade
    - MSc: 35% of final grade

# End of Lecture-1

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

- Danny Hillis