Introduction to Machine Learning

Overview and learning goals

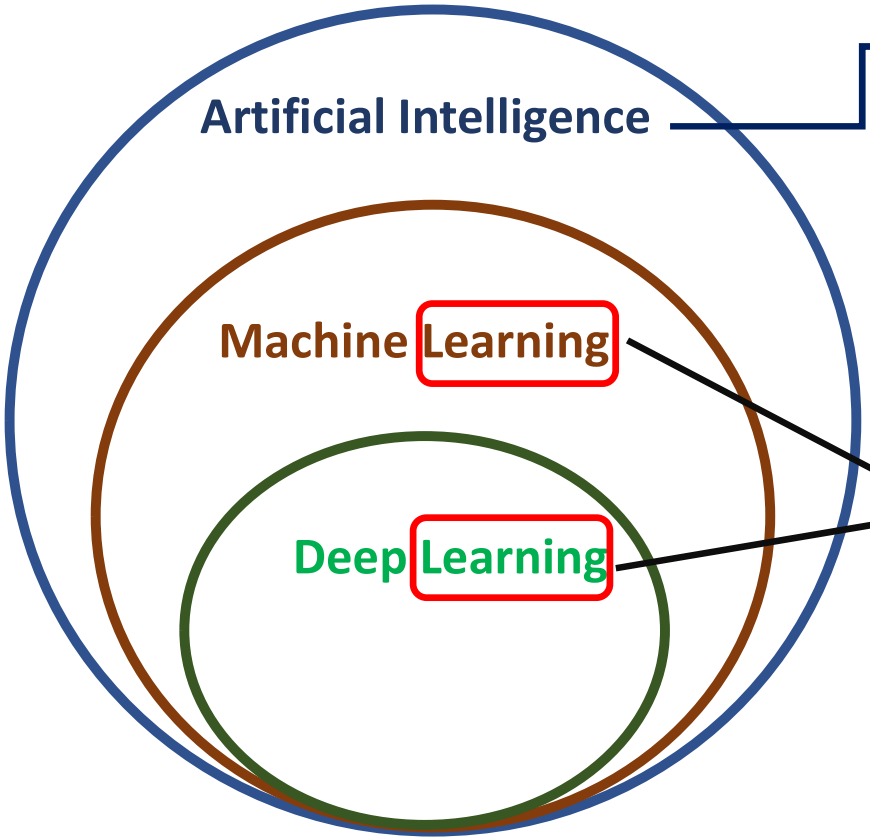
- Understand the basic terminology of the field
- Be able to apply the most fundamental machine learning algorithms and understand the basic principles behind the algorithms
- Gain knowledge on the various architectures of artificial neural networks and be able to select the ones suitable for your problems

What to expect from the course?

- What is Machine Learning ?
- Data Visualization/Analysis, Pandas, NumPy,……..
- Dimensionality Reduction
- Different Machine Learning Algorithms (Supervised, Unsupervised, metrics)
- Deep Learning (Neural Networks, back propagation, loss functions)
- CNN, RNN, LSTM
- ….. and more

Hands-On session will be conducted in parallel

Introduction

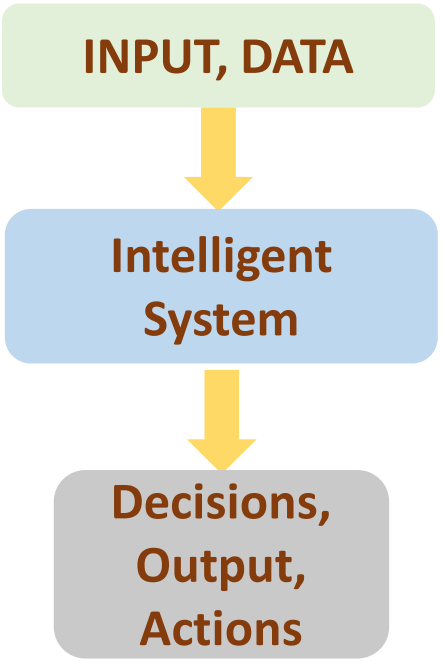


Any technique which enables computers to mimic human behaviour

- Learning is much deeper than memorization and information recall
- Learning is “a process that leads to change, which occurs as a result of experience and increases the potential for improved performance and future learning” (Ambrose et al, 2010, p.3)

Machine Learning

- Machine learning is a "*Field of study that gives computers the ability to learn without being explicitly programmed.*" : Arthur Samuel



- The function of a machine learning system can be:

✓ **descriptive**, meaning that the system uses the data to explain what happened
✓ **predictive**, meaning the system uses the data to predict what will happen
✓ **prescriptive**, meaning the system will use the data to make suggestions about what action to take

Data Driven Problem Solving

Area (sq.ft)	Price
250	250000
120	120000
310	310000
290	290000

Simple well-known solution.
(Price = Area *1000)
The above relation obtained in a trivial way, with one example.

Area (sq.ft)	Price
250	145500
120	212800
310	194390
290	

Not a trivial solution. There should be more parameters, (e.g., Age, Location)
Lot more data is needed to solve the above.

Remarks

General Strategy: Given many examples of (X,Y), learn an automated solution to predict Y given a new X, Y = F(X)

- Main Challenge: The data is becoming complex
- What is ***X*** is not a simple number?
  - A N-dim vector?
  - Entities other than numbers?
  - A picture?
  - A sound bite?

3.1
-2.6
0.41
1.89
15.2
⋮
9.23

3.9 m
₹ 8.2 L
Blue
Sedan

How do we get the machine to do this?

General Strategy: Given many examples of (X,Y), learn an automated solution to predict Y given a new X, Y = F(X)

- There is too much information in raw data
- Relevant information is hidden probably?
- Leads to Feature Extraction: Extracting useful information (X) from raw data

3.1
-2.6
0.41
1.89
15.2
⋮
9.23

3.9 m
₹ 8.2 L
Blue
Sedan

Representation: From Raw data to Features

Area	Bedrooms	Bathrooms	Age	Parking	Basement	Price
240	3	2	10	No	Yes	250000

Raw Data

- Convert all data into a vector of real numbers: $\boldsymbol{X}$
  - ✓ Points in a feature space

$$\boldsymbol{X_i = (240, 3, 2, 10, No, Yes)} \longrightarrow \boldsymbol{X_i = (240, 3, 2, 10, 1, 0)}$$

- Convert all predictions into an integer/real number: $\boldsymbol{Y}$

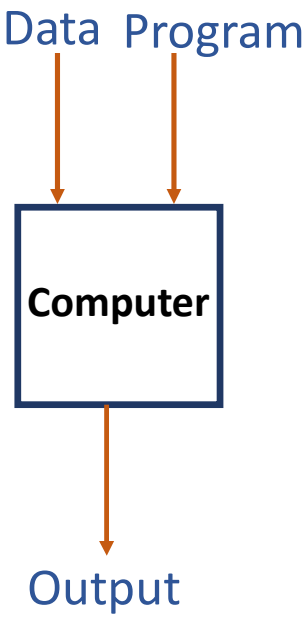
Representation: From Raw data to Features

Area	Bedrooms	Bathrooms	Age	Parking	Basement	Price
240	3	2	10	No	Yes	250000

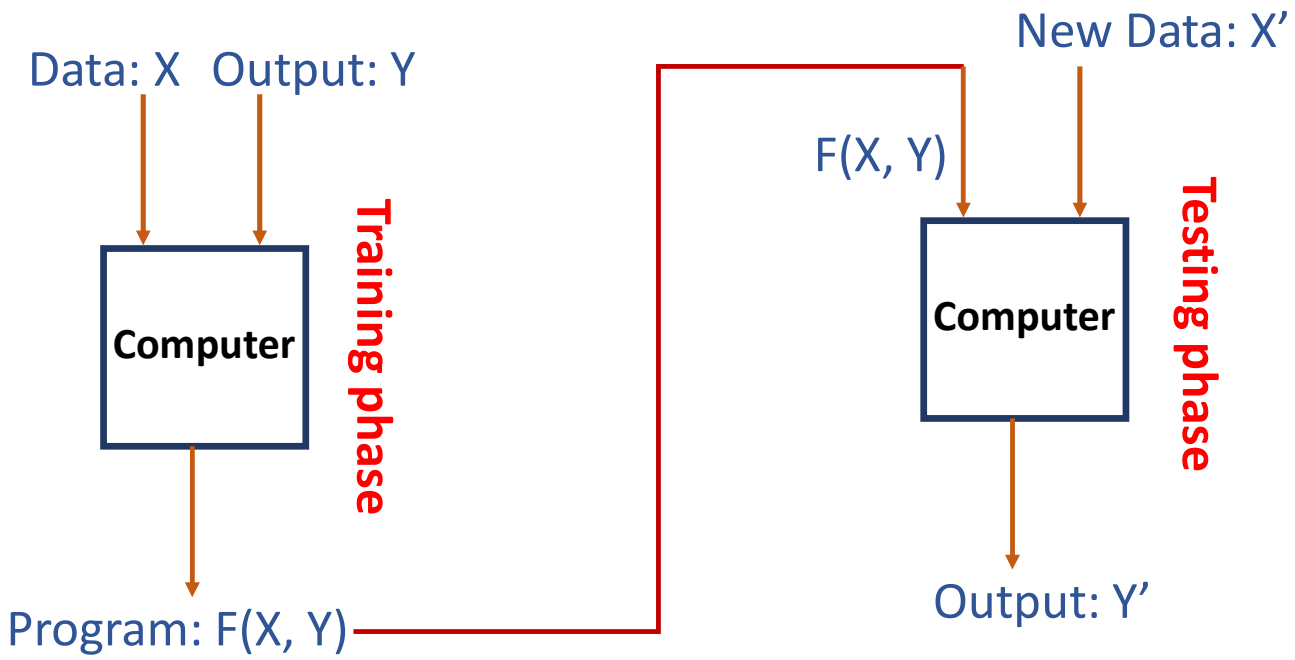
- We are given a set of *n* examples: (X_i, Y_i)
- Our goal is to learn a model: $Y = F(X)$, that captures the pattern of the training samples
- We can assume a model and learn its parameters
- Once we learn the model, we can predict the output, *y'* corresponding to any new input, **X'** : $Y' = F(X')$

Usual Programming vs Machine Learning

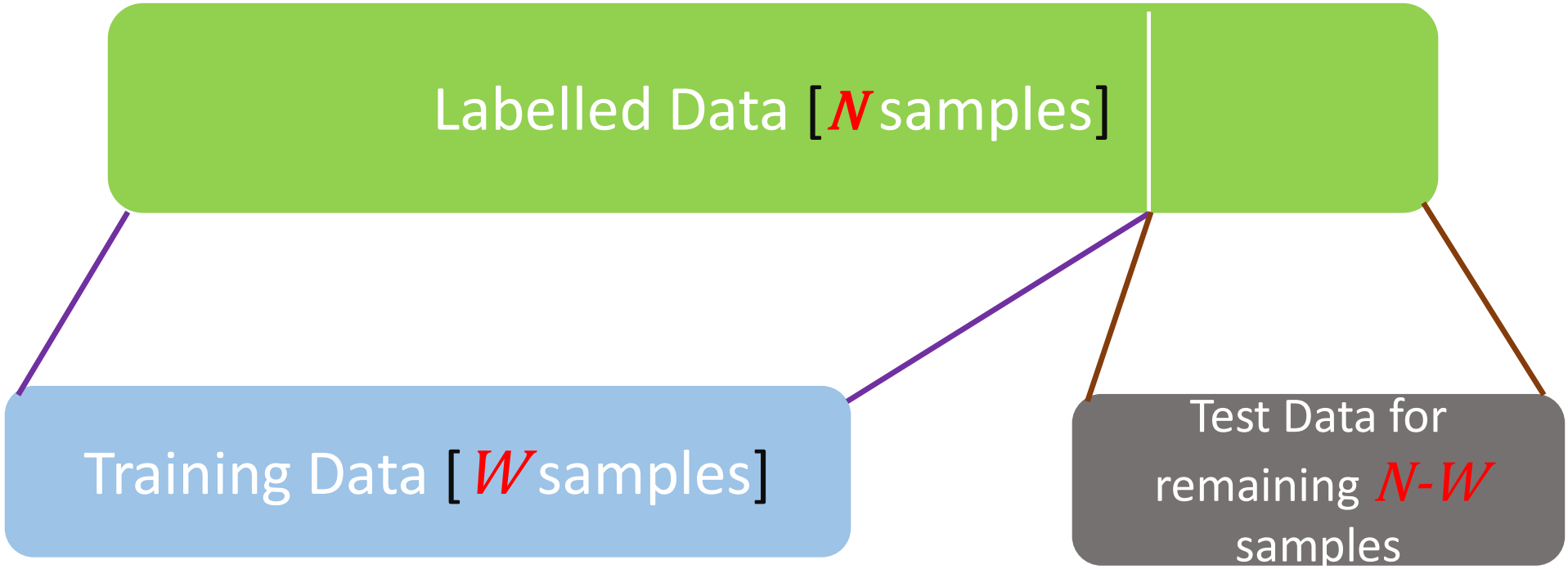
Programming:



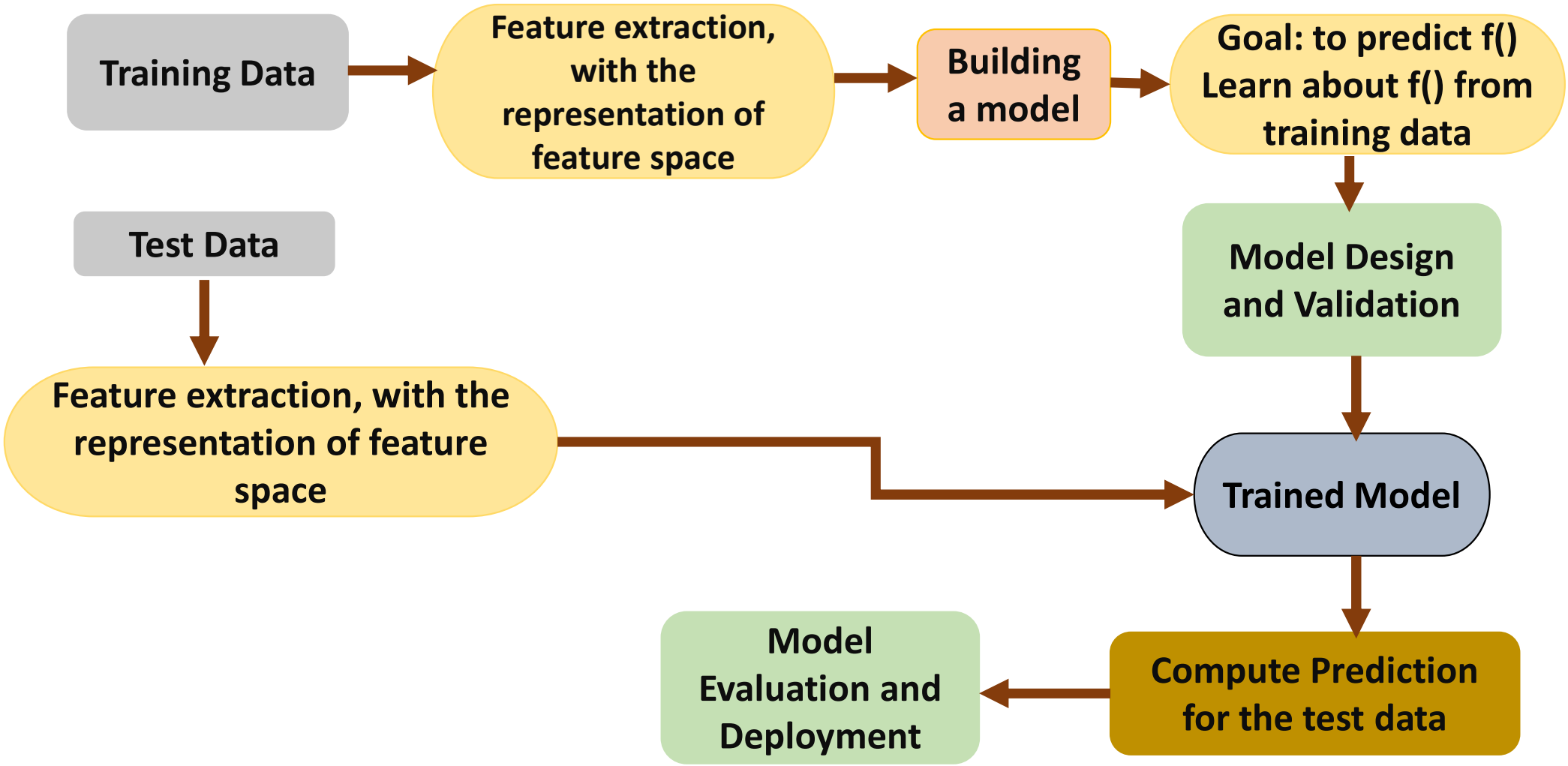
Machine Learning:



ML Based on Training-Testing Data

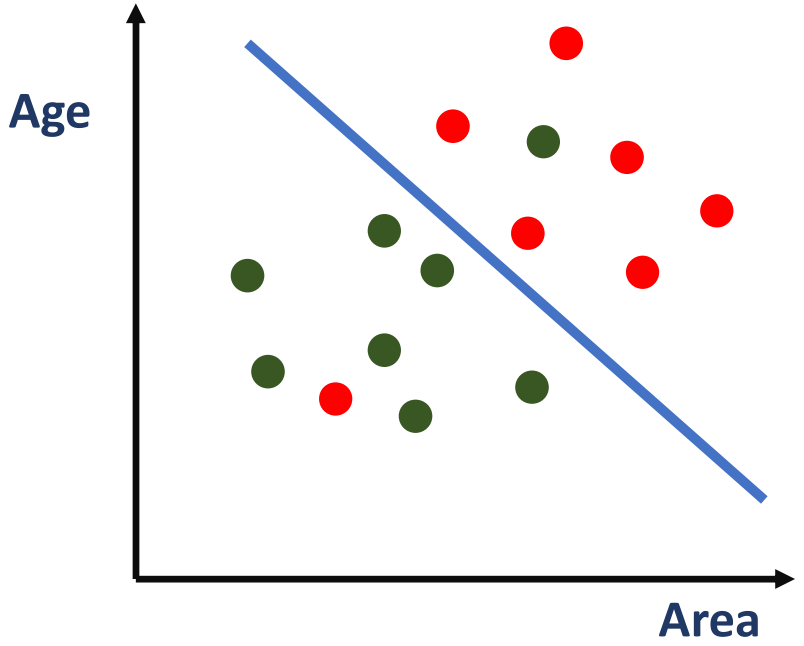


- Take care to not leak information from Test Data into the Model

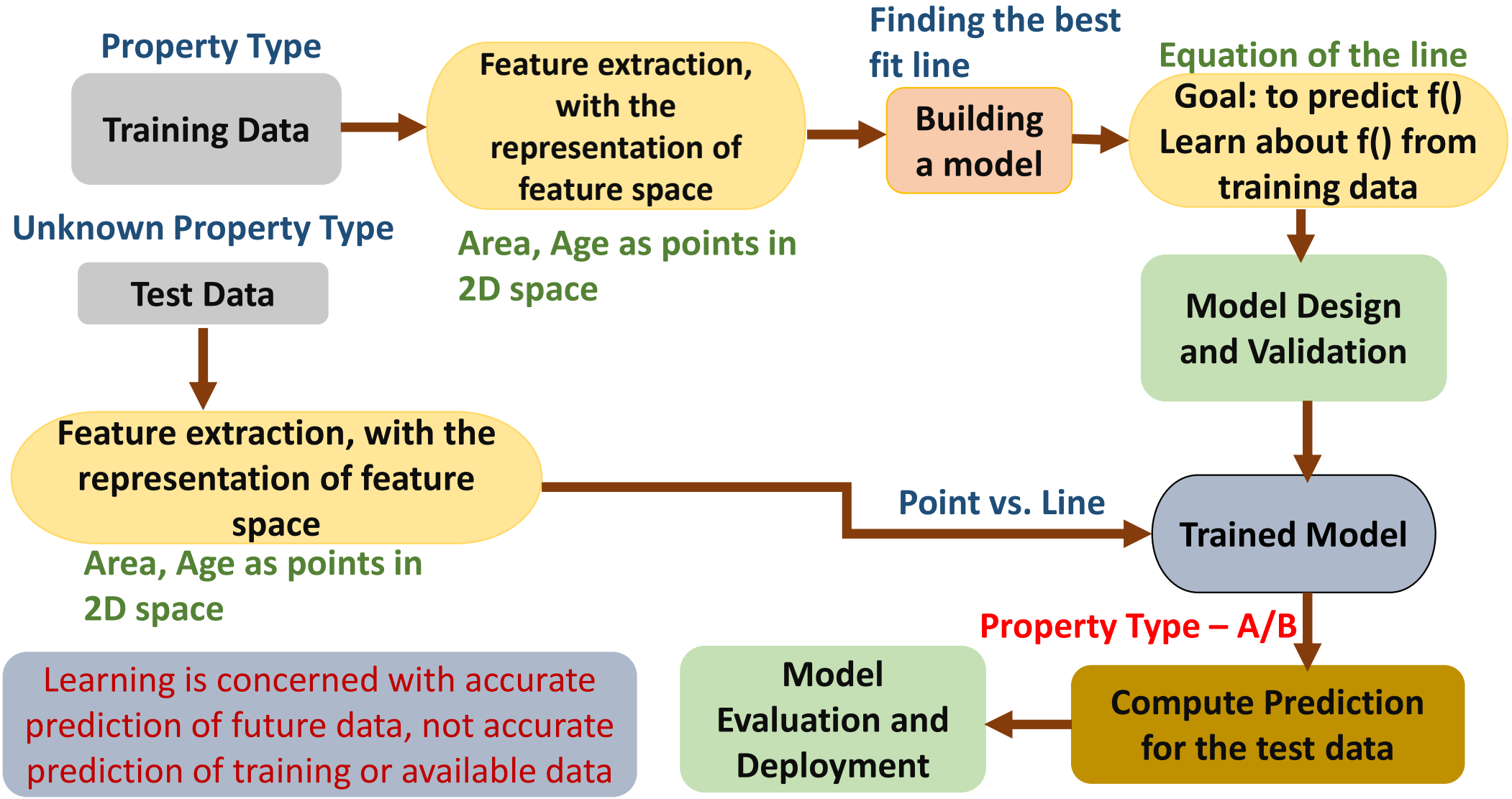

Training Data
Feature extraction, with the representation of feature space
Building a model
Goal: to predict f()
Learn about f() from training data
Model Design and Validation
Test Data
Feature extraction, with the representation of feature space
Trained Model
Compute Prediction for the test data
Model Evaluation and Deployment

Data Representation

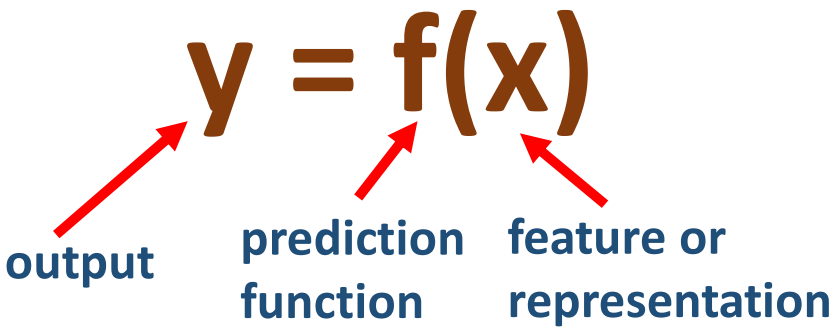
Area	Age	Property
230	15	A
120	6	B
202	2	B
398	11	A
274	8	?



Feature Space Representation


Property Type
Training Data
Feature extraction, with the representation of feature space
Area, Age as points in 2D space
Finding the best fit line
Building a model
Equation of the line
Goal: to predict f()
Learn about f() from training data
Model Design and Validation
Unknown Property Type
Test Data
Feature extraction, with the representation of feature space
Area, Age as points in 2D space
Point vs. Line
Trained Model
Property Type – A/B
Compute Prediction for the test data
Model Evaluation and Deployment
Learning is concerned with accurate prediction of future data, not accurate prediction of training or available data

Summary – Machine Learning Framework

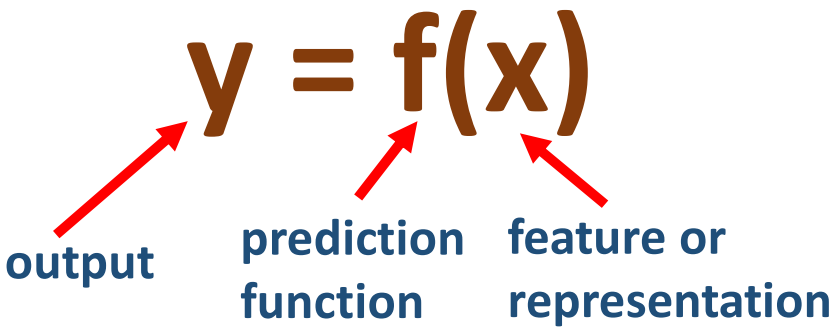


- Note: Training set and testing set come from the same distribution

- **Training**: given a training set of labeled examples $\{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\}$, estimate the prediction function **f** by minimizing the prediction error

- **Testing**: apply **f** to the test example **x'** and output the predicted value **y = f(x')**

Summary – Machine Learning Framework



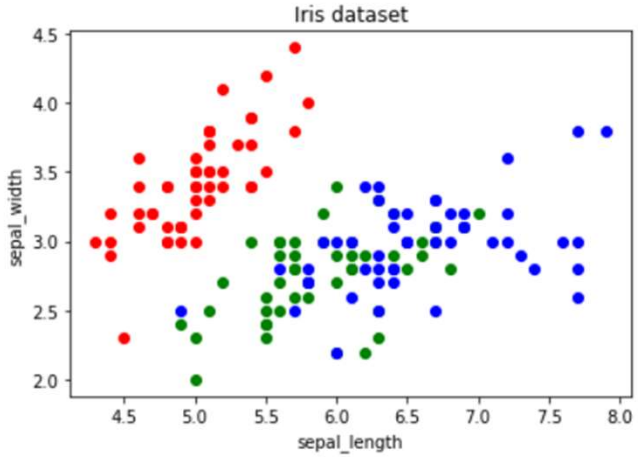
- The input is converted to a vector **x**
- The output is a value indicated by **y**
- Depending on the nature of **x** and **y**, we define
  1) Regression
  2) Classification
  3) ……

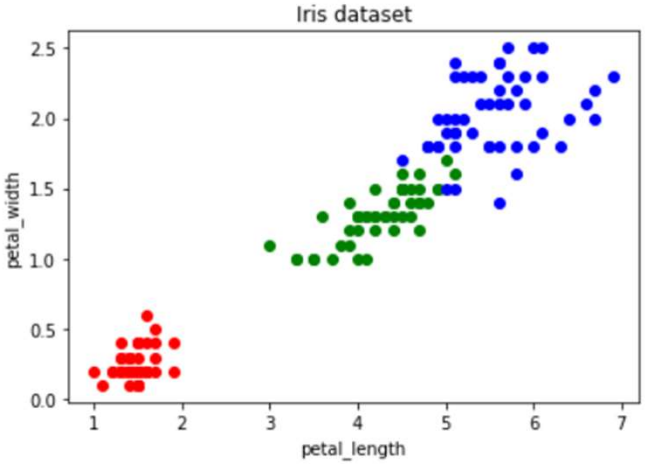
Representations

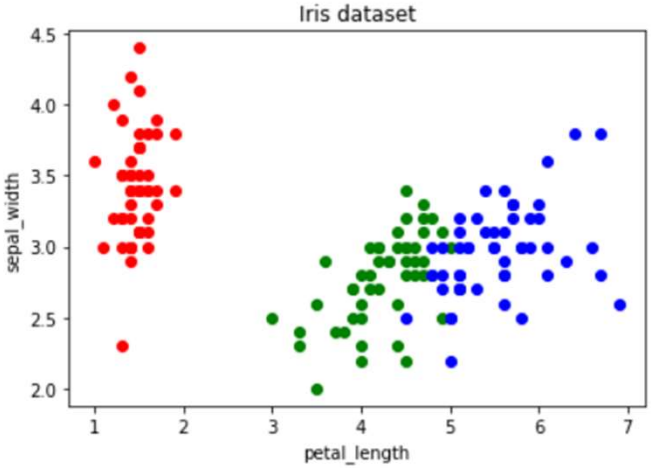
- Representations in machine learning refer to the way data is transformed or encoded into a format that is suitable for a learning algorithm to process

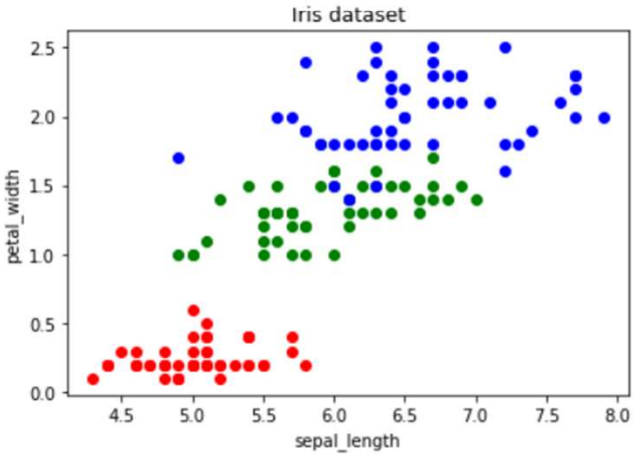
Sepal Length	Sepal Width	Petal Length	Petal Width	Species
5.1	3.5	1.4	0.2	A
5.4	3.7	1.1	0.1	A
5.2	2.7	3.9	1.0	B
6.6	2.9	3.5	1.2	B
5.8	2.8	5.1	2.4	C
7.7	3.7	6.7	2.2	C
--------	------	-----	-----	-----

Feature Space



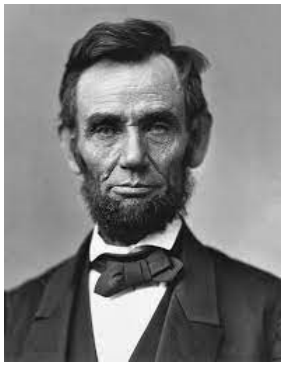


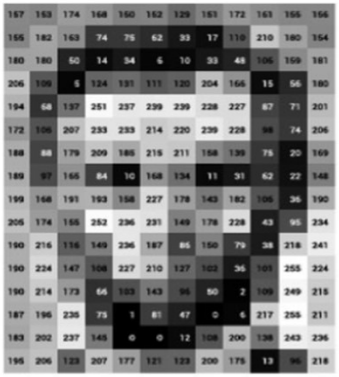


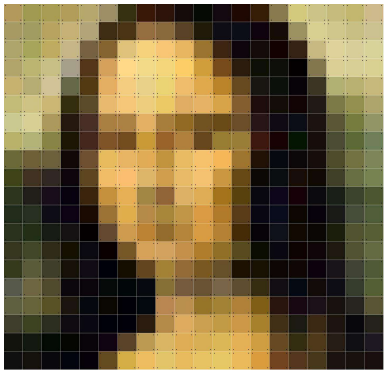


Representations

- Images: Raw Pixel Representation, Deep Learning Based Features



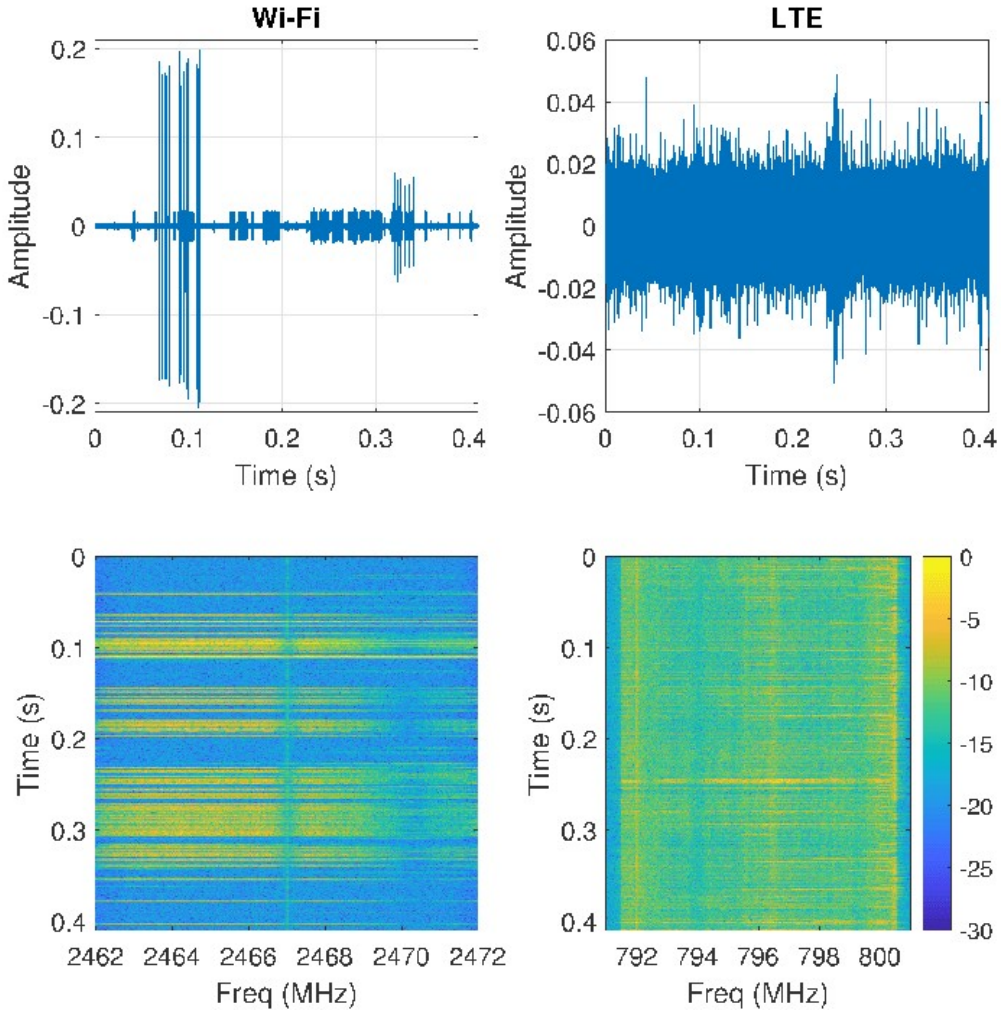




- The sum of all the pixels
- The number of boundary pixels
- Edge detection

Representations

- Sound: Waveform Representation, Spectogram Representation, MFCC



Reference: Towards Low-Complexity Wireless Technology Classification Across Multiple Environments, DOI:10.1016/j.adhoc.2019.101881

Representations – Textual Data

- Text Data: N-grams, Bag of Words, Term Fequency-Inverse Document Frequency, Word Embeddings

  Sentence: The weather is sunny today

N-gram	N-gram Generated Sentence	Number of N-gram Features
Unigram (1-Gram)	“The”, “weather”, “is”, “sunny”, “today”	5
Bigram (2-Gram)	“The weather”, “The is”, “The sunny”, ……	10
Trigram (3-Gram)	“The weather is”, “weather is sunny”, ….	3

Representations – Textual Data

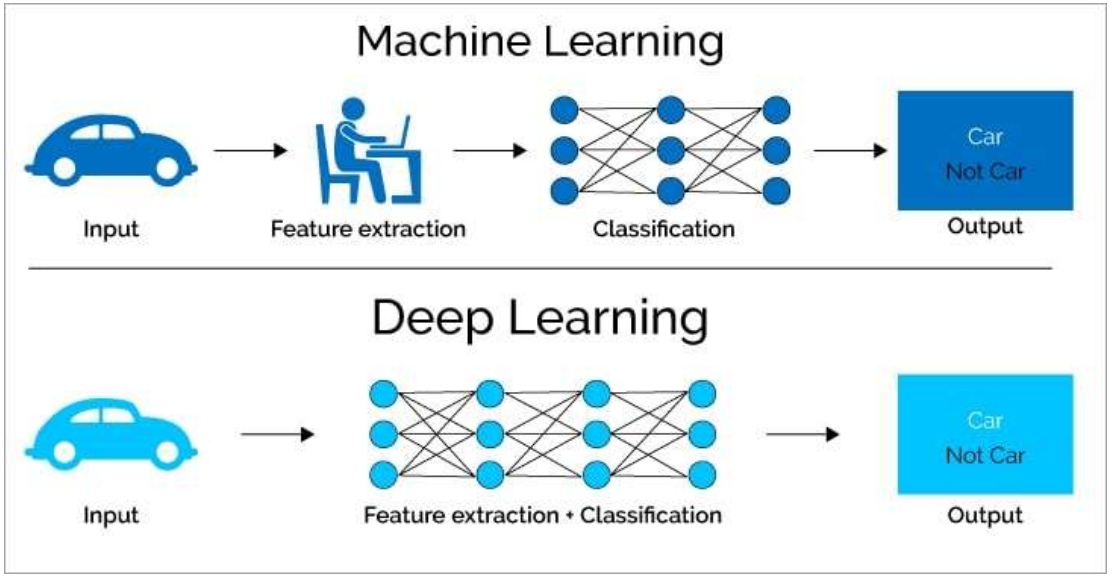
- Text Data: N-grams, Bag of Words, Term Fequency-Inverse Document Frequency, Word Embeddings
  - Sentence 1: The weather is sunny today
  - Sentence 2: The weather was rainy yesterday

	1 The	2 weather	3 is	4 sunny	5 today	6 was	7 rainy	8 yesterday	Length
1	1	1	1	1	1	0	0	0	5
2	1	1	0	0	0	1	1	1	5

- Vector of Sentence 1: [1 1 1 1 1 0 0 0]
- Vector of Sentence 2: [1 1 0 0 0 1 1 1]

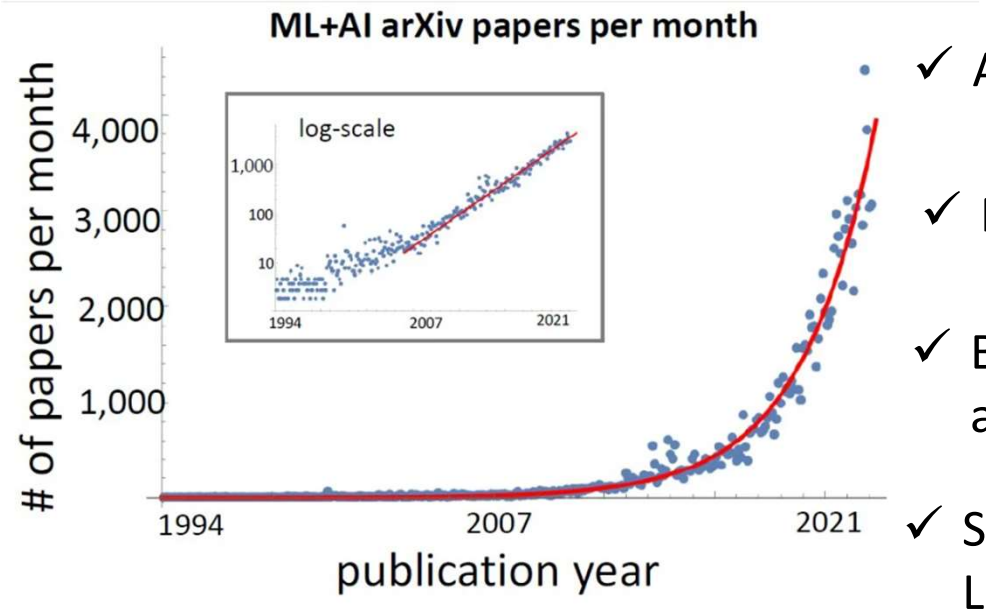
Deep Learning

- Deep Learning is a specialized subset of Machine learning
- Imitates the data processing function of the human brain
- Relies on a layered structure of algorithms called as Artificial Neural Network
- Deep learning models require a large amount of data to train, but requires little human intervention to function properly



The Deep Learning algorithm doesn't need a software engineer to identify features but is capable of automatic feature engineering through its neural network.
(Source: softwaretestinghelp.com)

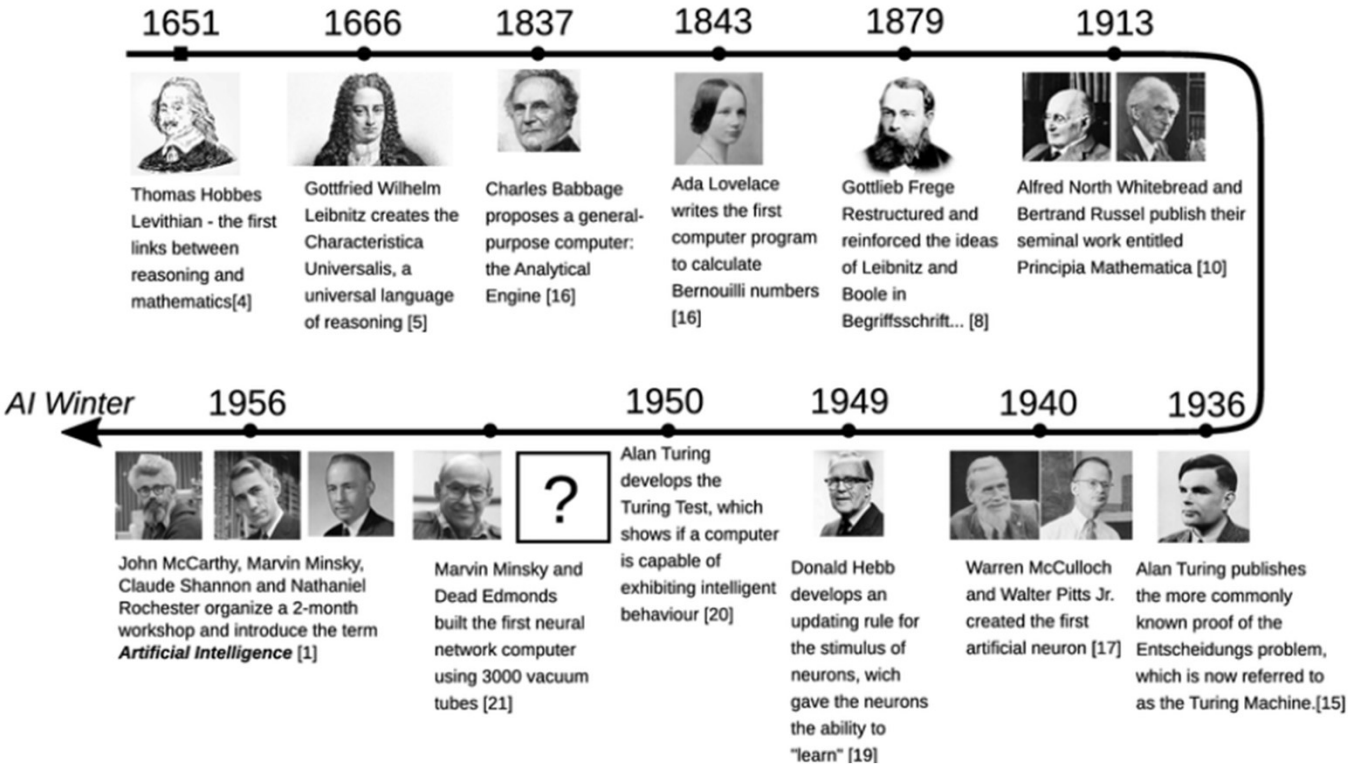
Why sudden interest in AI?



More People, Papers, Results, Funding, Positive Feedback.

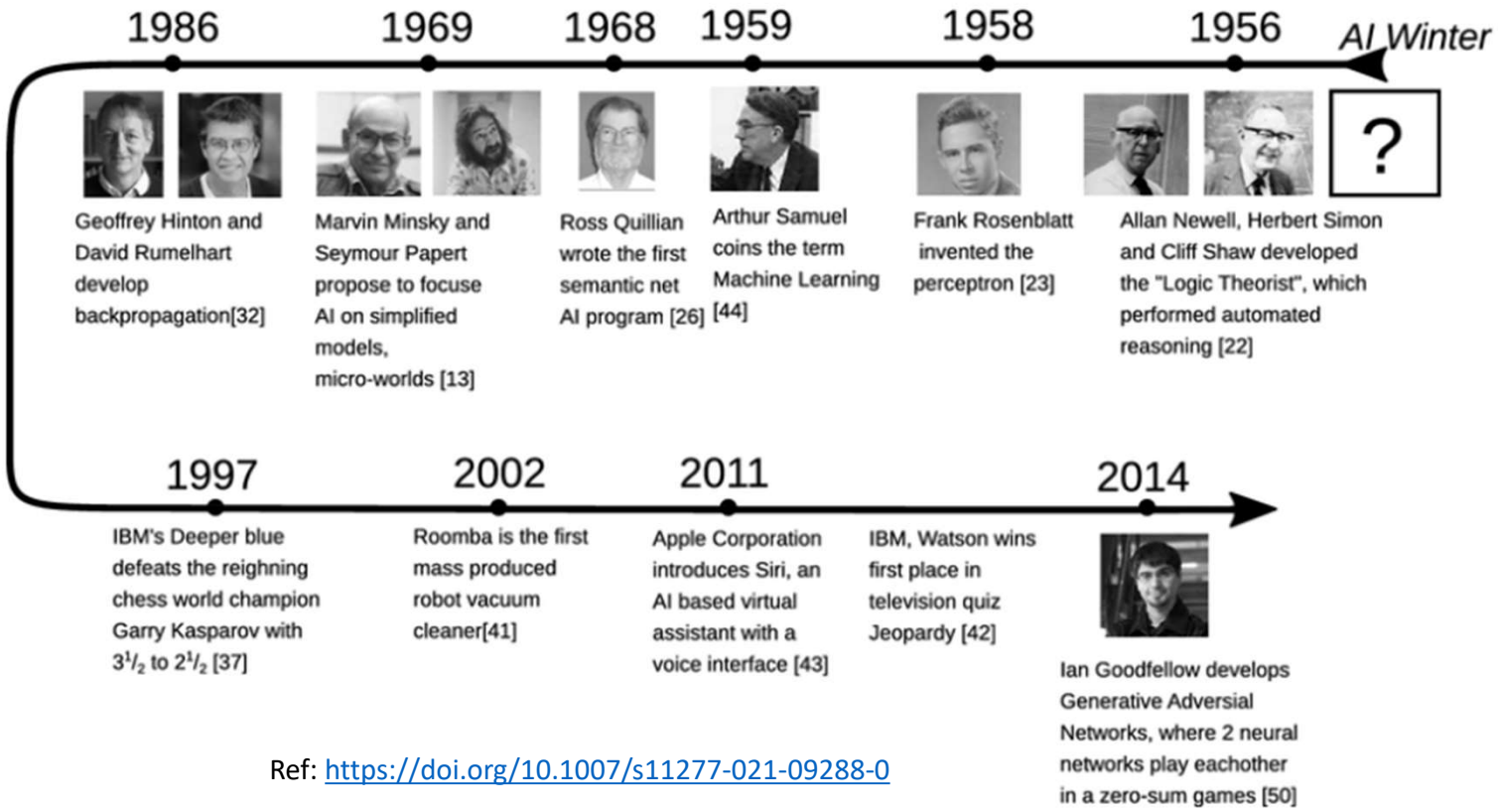
- ✓ Appearance of large, high-quality labeled datasets
- ✓ Massively parallel computing with GPUs
- ✓ Backprop-friendly activation functions, Improved architectures
- ✓ Software platforms, Cloud Compute, APIs, Libraries
- ✓ New regularization techniques, Robust optimizers

Timeline: Is AI, ML, DL really New?



Ref: https://doi.org/10.1007/s11277-021-09288-0

Timeline: Is AI, ML, DL really New?



Ref: https://doi.org/10.1007/s11277-021-09288-0

Where is Machine Learning?

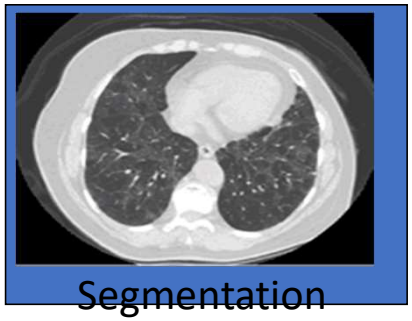
E-Commerce

NETFLIX

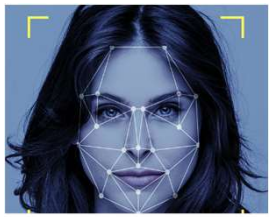
Recommendation Systems

Create Photographs, Paintings

Speech Recognition



Segmentation



Facial Recognition

Virtual Assistants

Chess/ Go Champions

Autonomous Cars/Navigation

Image Courtesy: Google

Other Applications

- Surveillance
- Automated Assembly
- Mail Sorting
- Face detection (photography)
- Robot Navigation
- Content-Based Image Retrieval
- Entertainment
- And many more…

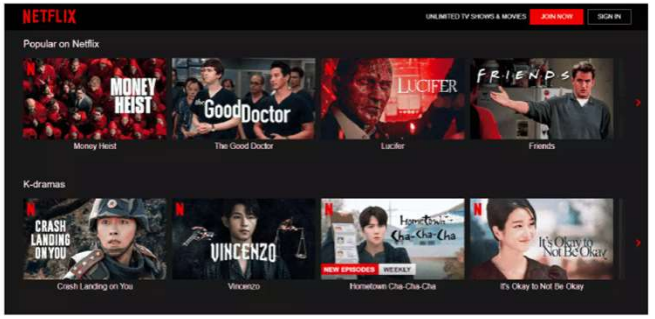


Image Courtesy: Google