

ARTIFICIAL INTELLIGENCE

DS —

Statistics,
Probability,
Linear Algebra

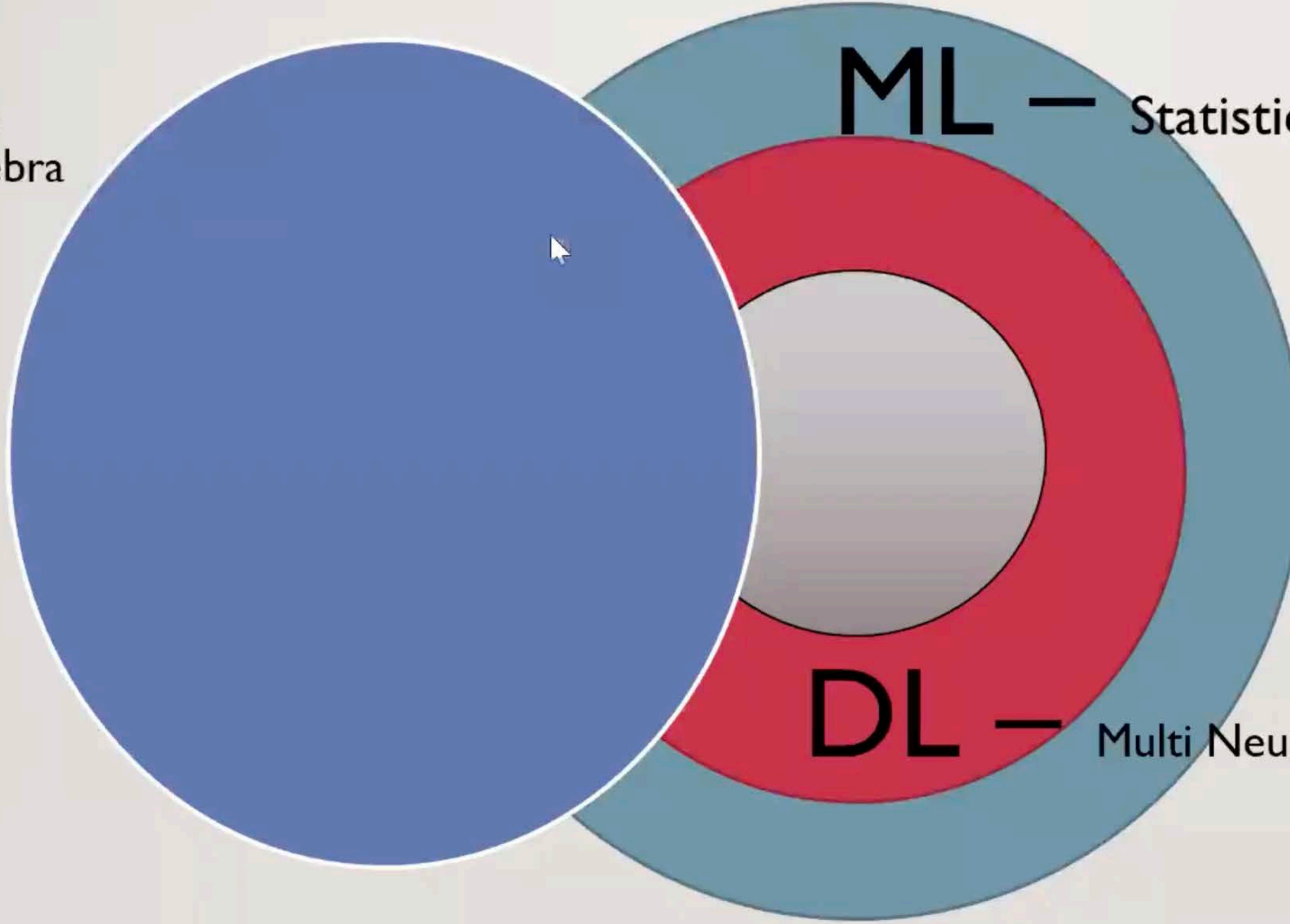
AI — Enables the machine to think

ML — Statistical tools to explore data

- Supervised Learning
- Unsupervised Learning
- Reinforced Learning

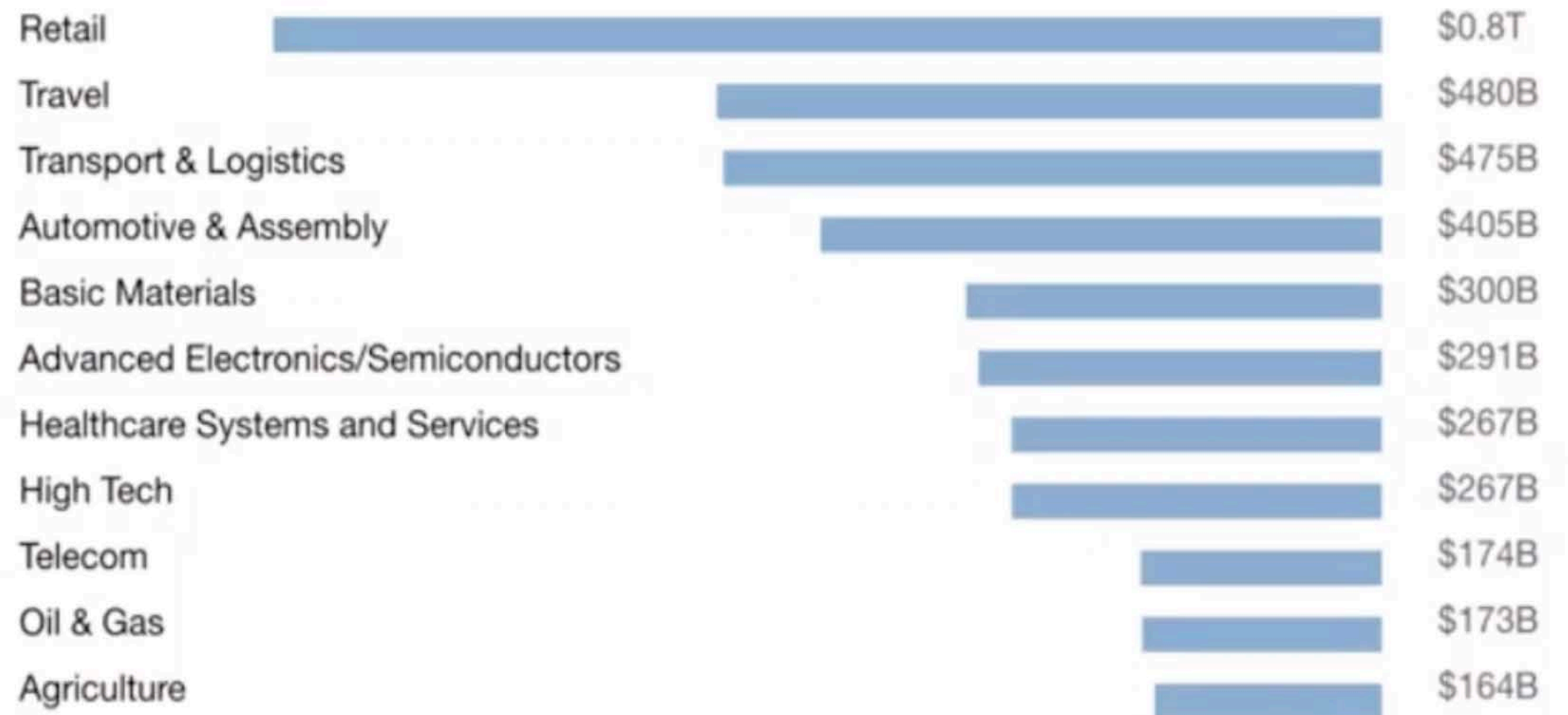
DL — Multi Neural Network Architecture

- ANN
- CNN
- RNN



AI value creation
by 2030

\$13
trillion



[Source: McKinsey Global Institute.]


```
graph TD; AI[AI] --- ANI[ANI]; AI --- AGI[AGI]; ANI --- ANI_desc["(artificial narrow intelligence)"]; ANI --- ANI_examples["E.g., smart speaker, self-driving car, web search, AI in farming and factories"]; AGI --- AGI_desc["(artificial general intelligence)"]; AGI --- AGI_goal["Do anything a human can do"]
```

AI

ANI

(artificial narrow intelligence)

E.g., smart speaker, self-driving car, web search, AI in farming and factories

AGI

(artificial general intelligence)

Do anything a human can do

AGENDA

- What is Machine Learning
- What is Data
- What Machine Learning can and cannot do
- Deep Learning

From Beginning

From Current Slide

Present Online

Custom Slide Show

Start Slide Show

Set Up Slide Show

Hide Slide

Rehearse Timings

Record Slide Show

Set Up

☐ Keep Slides Updated

☒ Use Timings

☒ Play Narrations

☒ Show Media Controls

Monitors

Monitor: Automatic

☒ Use Presenter View

66

67

68

69

70

71

WHAT IS MACHINE LEARNING?

DATA SET STATISTICS

TOY PROBLEM

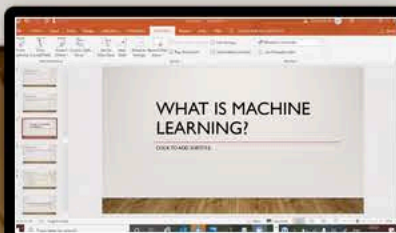
MACHINE LEARNING

WHAT IS MACHINE LEARNING?

CLICK TO ADD SUBTITLE

ARTHUR SAMUEL(1959) :

- ML is a field of study that gives computers the ability to learn without being explicitly programmed,



Tools: Select, Draw with Touch, Eraser, Lasso Select

Pens: Green, Orange, Blue, Grey, Pink, Red, Add Pen

Stencils: Ruler

Convert: Ink to Shape, Ink to Math

Slide thumbnails:

- 66 DIFFERENTIAL STATISTICS
- 67 POINT ESTIMATION
- 68 WHAT IS MACHINE LEARNING?
- 69 DIFFERENTIAL STATISTICS
- 70 TOM MITCHELL (1998) *T = Filtering spam msg*
E = 1K E = 50K
P = 0.2 P = 0.07
- 71 MACHINE LEARNING

TOM MITCHELL (1998)

- A computer is said to learn
- from experience E
- with respect to some task T
- and some performance measure P ,
- if its performance on T ,
- as measured by P ,
- improves with experience E .

T = Filtering spam msg

E = 1K

E = 50K

P = 0.2

P = 0.07

MACHINE LEARNING

- Supervised Learning

A -----> B

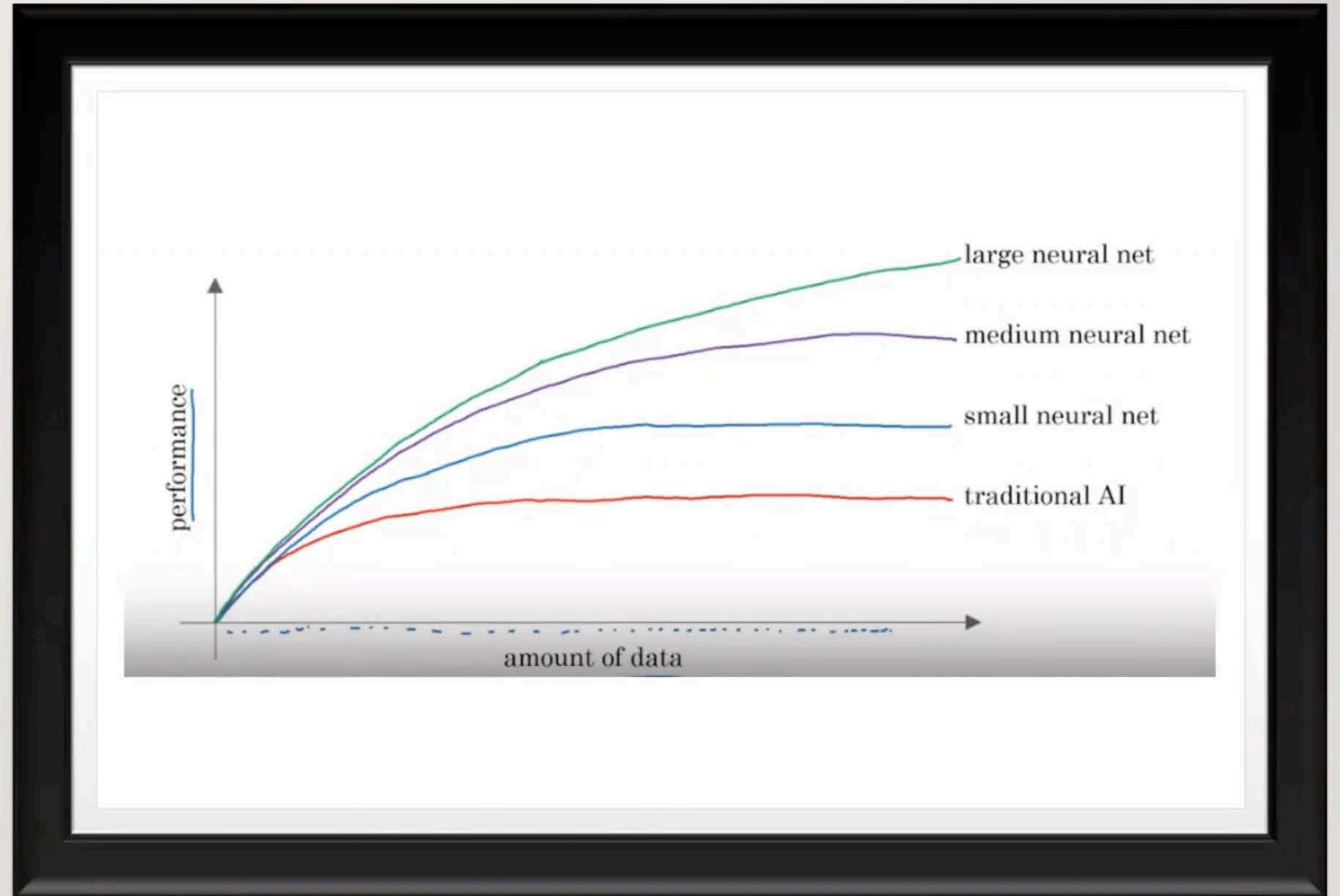
Input -----> Output

SUPERVISED LEARNING

INPUT(A)	OUTPUT(B)	APPLICATION
Email	Spam(0/1)	Spam filtering
Audio	Text Transcript	Speech Recognition
English	Chinese	Machine Translation
Ad, user info	Click(0/1)	Online Advertising
Image, Radar info	Position of other cars	Self-driving car
Image of phone	Defect(0/1)	Visual Inspection

WHY NOW???

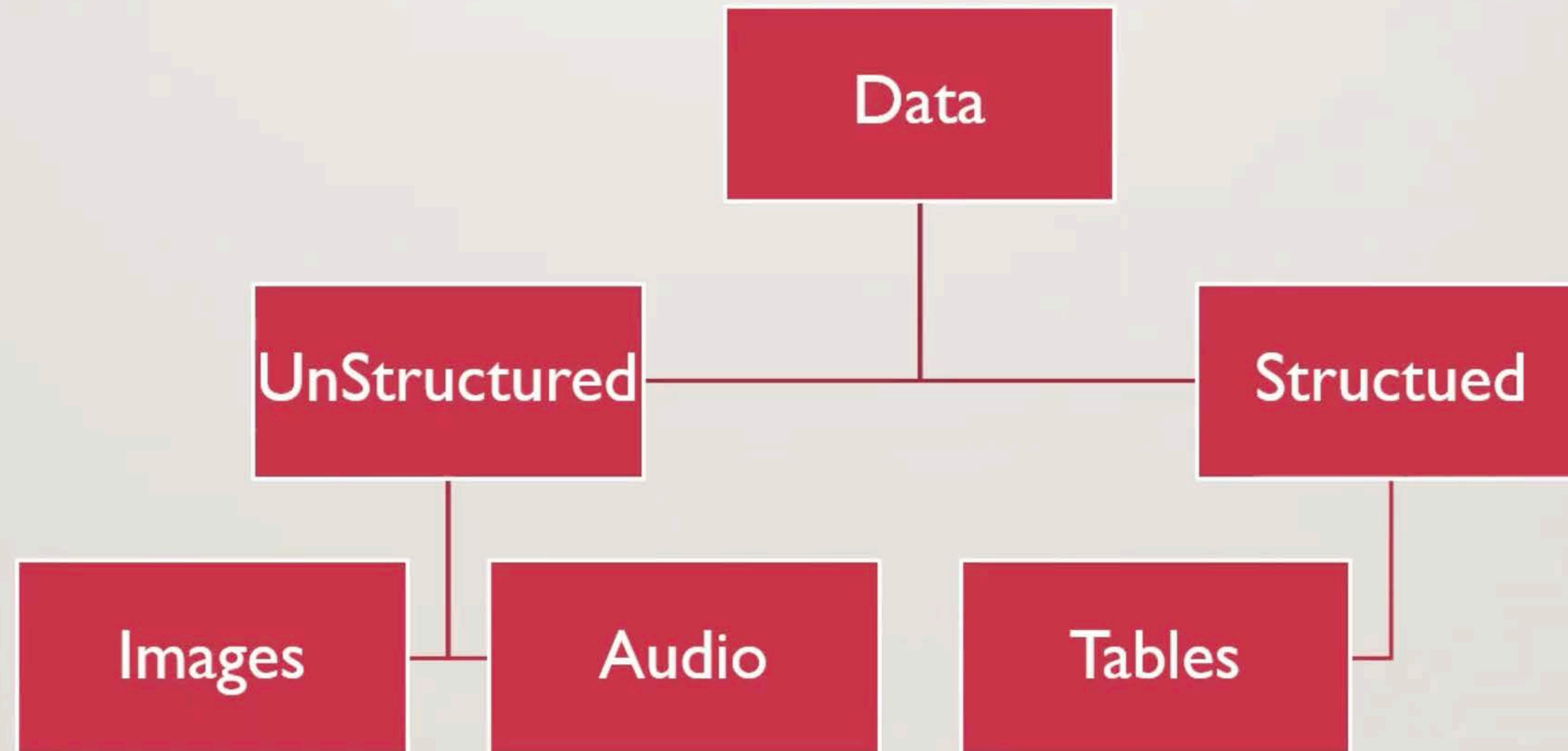
BIG DATA



SUPERVISED LEARNING

INPUT(S)	OUTPUT(S)	APPLICATION
Text	Text (S)	Spam filtering
Audio	Text Transcription	Speech Recognition
Images	Classes	Machine Translation
Ad user info	Click(S)	Online Advertising
Image, Radar info	Position of other cars	Self-driving car
Image of phone	Defect(S)	Visual Inspection

DATA



A → B



Size of House

- 500
- 600
- 700
- 1200
- 1800



#of Bedrooms

- 1
- 1
- 3
- 2
- 3



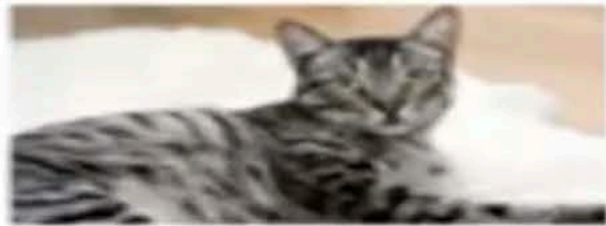



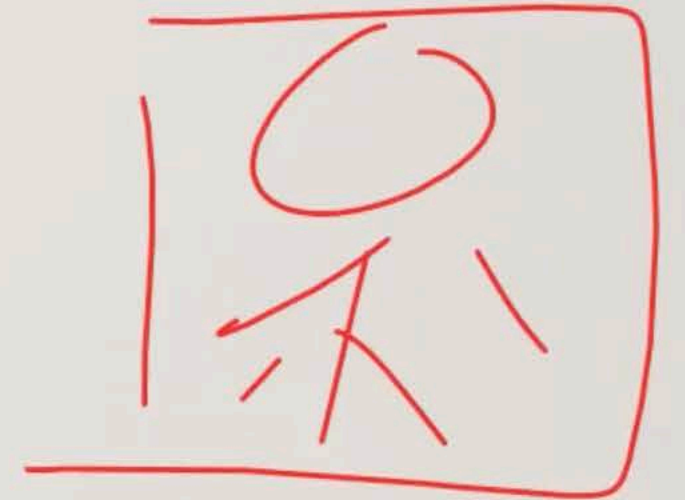
Price(1000\$)

- 1000
- 1400
- 2500
- 2000
- 2400

~~A~~

~~B~~

image	label
	cat
	not cat
	cat
	not cat



A

B

HOW TO COLLECT DATA?

- Manual Labeling
- From Observing behaviours

user ID	time	price (\$)	purchased
4783	Jan 21 08:15.20	7.95	yes
3893	March 3 11:30.15	10.00	yes
8384	June 11 14:15.05	9.50	no
0931	Aug 2 20:30.55	12.90	yes

machine	temperature (°C)	pressure (psi)	machine fault
17987	60	7.65	N
34672	100	25.50	N
08542	140	75.50	Y
98536	165	125.00	Y

- Download from websites/partnerships

DATA IS MEESSY

- Garbage in , garbage out
- Data Problems
 - Incorrect Values
 - Missing Values

MACHINE LEARNING VS DATA SCIENCE

Size of House	#of Bedrooms	Parking Space	Garden	Price
500	1	1	N	10000
600	2	6	Y	50000
700	1	1	N	12000
1000	4	1	Y	80000

Machine Learning

$A \rightarrow B$

A

B

Data Science

Houses with garden are more expensive

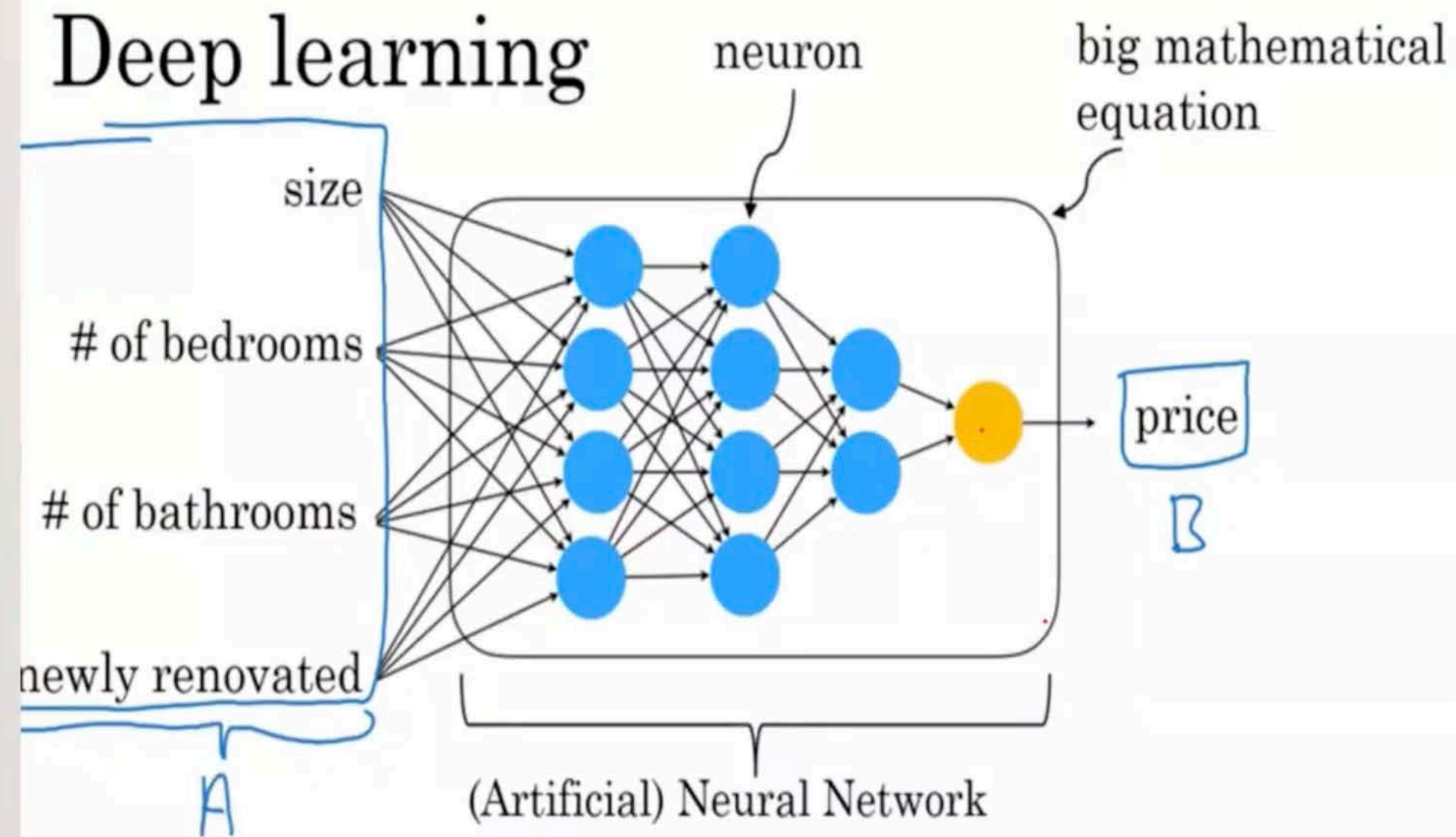
Houses with parking space are 30% more expensive than houses without parking space.

MACHINE LEARNING VS DATA SCIENCE

- Machine Learning
 - “Field of study that gives computers the ability to learn without being explicitly programmed.”
- Data Science
 - Science of extracting knowledge and insights from data

-Arthur Samuel(1959)

DEEP LEARNING



WHAT ML CAN AND CANNOT DO?

Anything that can be done with the 1 second of thought can be automated(probably now or else sooner)

EXAMPLE :- THE SHIRT I HAVE ORDERED IS LOOSE FOR ME. CAN I EXCHANGE IT FOR ONE SMALLER SIZE?

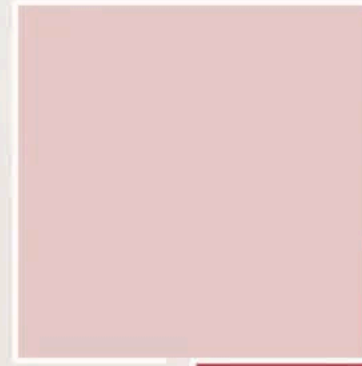
- It can figure out that it is the Exchange Request
- Input Text ---→ Exchange Request
- But It can't say that : "Congratulations !! You might have reduced your weight. So, the shirt is loose to you."

X-RAY DAIGNOSIS



Can Do

- Diagnose bone fracture from ~100000 labeled images



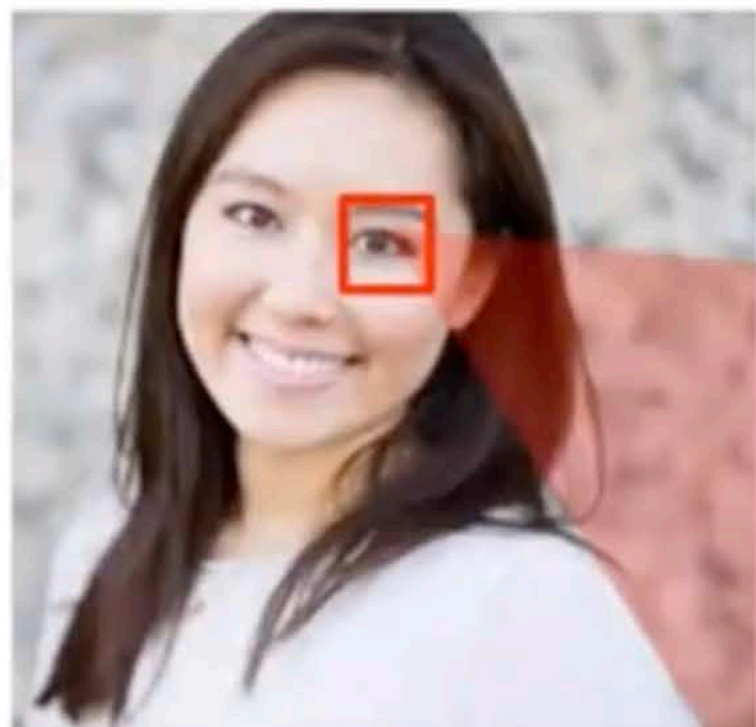
Cannot Do

- Diagnose bone fracture from 10 images of textbook explaining the same.

STRENGTHS AND WEAKNESSES OF ML

- It works well when:
 - Learning a “simple: concept
 - Lots of data is available
- It won't work when:
 - Learning complex concepts from small amount of data
 - When asked to perform on new types of data

Face recognition



30	32	22	12	10	10	12	33	35	30
12	11	12	234	170	176	13	15	12	12
234	222	220	230	200	222	230	234	56	78
190	220	186	112	110	110	112	180	30	32
49	250	250	250	4	2	254	200	44	6
55	250	250	250	3	1	250	245	25	3
189	195	199	150	110	110	182	190	199	55
200	202	218	222	203	200	200	208	215	222
219	215	220	220	222	214	215	210	220	220
220	220	220	220	221	220	221	220	220	222

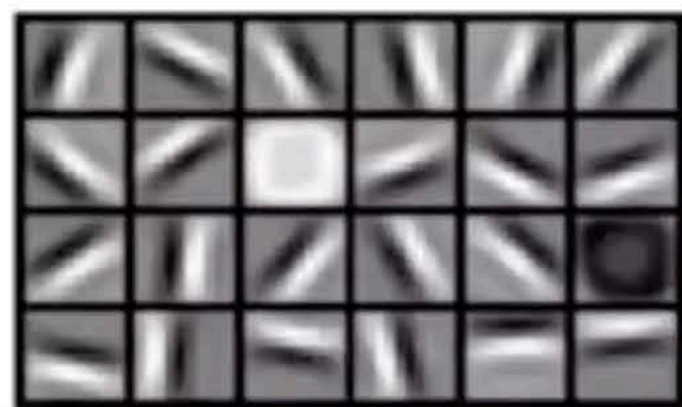
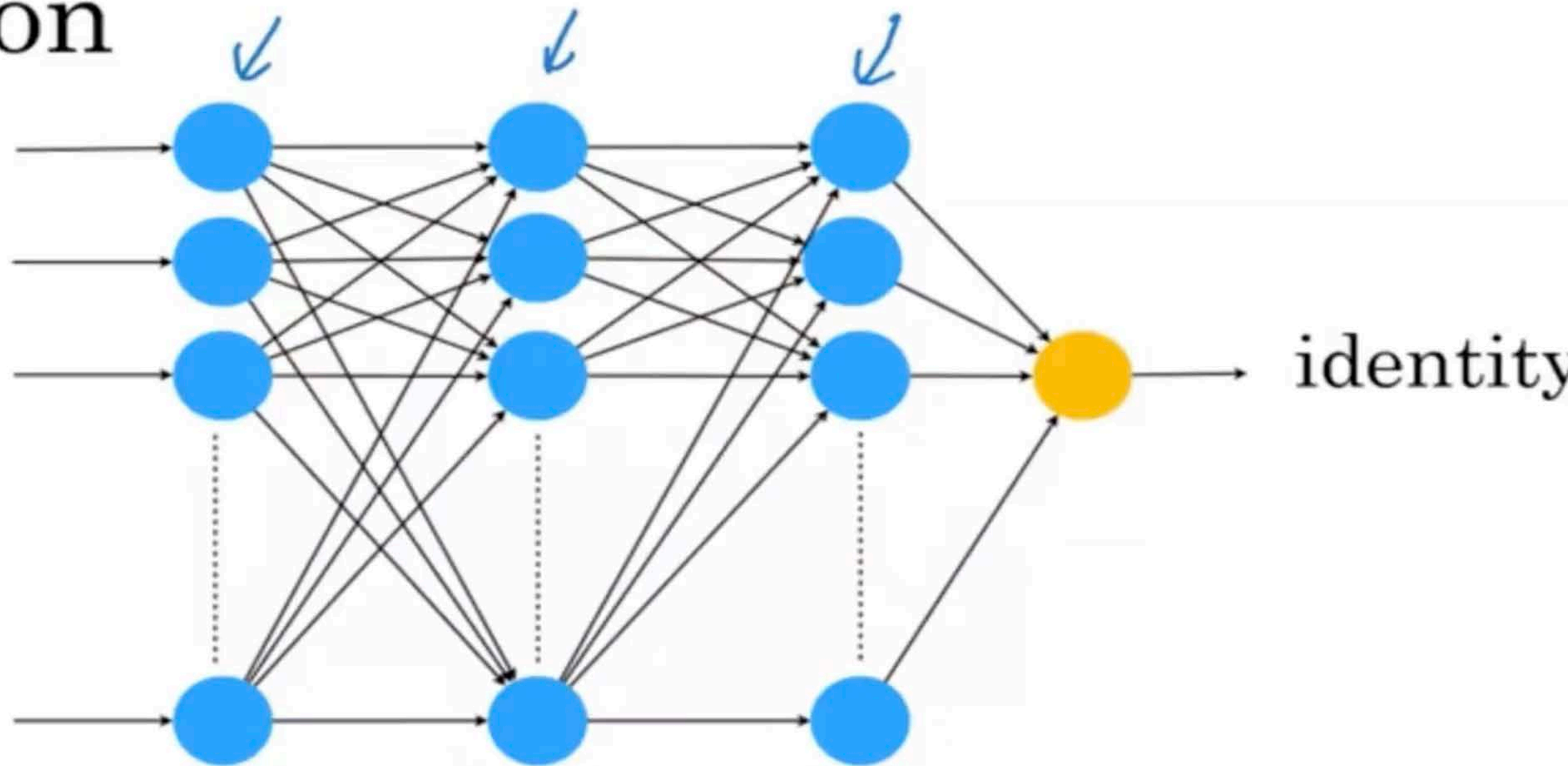
Face recognition



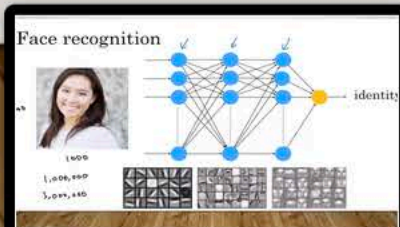
1600

1,000,000

3,000,000

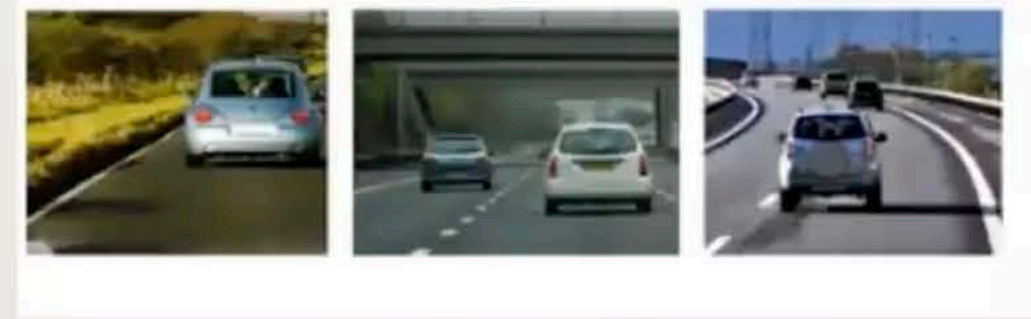


WORKFLOW OF MACHINE LEARNING PROJECT



STEPS FOR SELF

- 1. Collect data
 - Image \rightarrow position of other cars



2.. Train Model

3. Deploy Model

WORKFLOW OF DATA SCIENCE PROJECT

- 1. Collect data
- 2. Analyze data
 - Iterate many times to get good insights

3. Suggest hypotheses/actions

Deploy changes

Re-analyze new data periodically

CASE STUDY : SMART SPEAKER

- Steps

- 1. Trigger word/wake word detection

Audio --→ "Hello device(0/1)

- 2. Speech Recognition

Audio----"Play a song"

- 3. Intent Recognition

song

- 4. Execute song

AI PITFALLS

- Expect Ai to solve everything
- Hiring 2-3 ML engineers and depend purely on them
- Expecting AI to work the first time