

# AI for Everyone

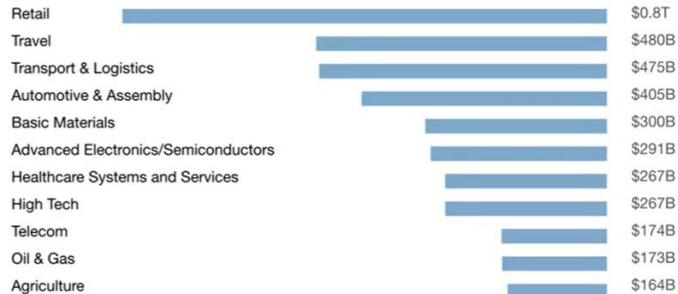
## Week 1

### Introduction

## Introduction

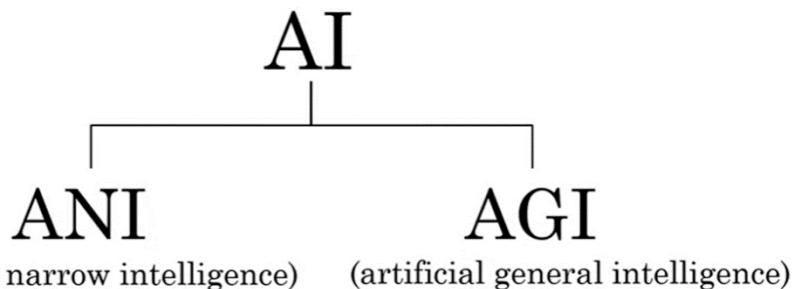
AI value creation  
by 2030

\$13  
trillion



[Source: McKinsey Global Institute.]

## Demystifying AI



- ANI (artificial narrow intelligence –)
  - Smart speaker, self-driving car, web search, AI in farming and factories
  - A lot of progress
  - Case studies in this week
- AGI (artificial general intelligence)
  - Do anything a human can do
  - Not much progress

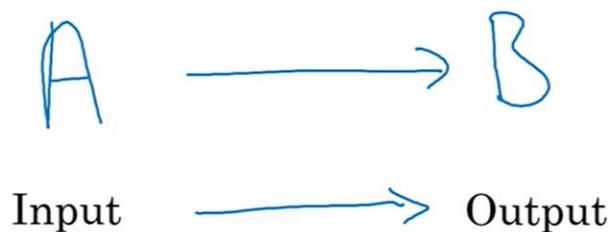
# What you'll learn

- What is AI?
    - Machine Learning
    - Data
    - What makes an AI company
    - What machine learning can and cannot do
    - Optional: Intuitive explanation of Deep Learning
  - Building AI projects
  - Building AI in your company
  - AI and society
    - What is AI?
- 

Machine Learning

*Supervised learning*

## Supervised Learning



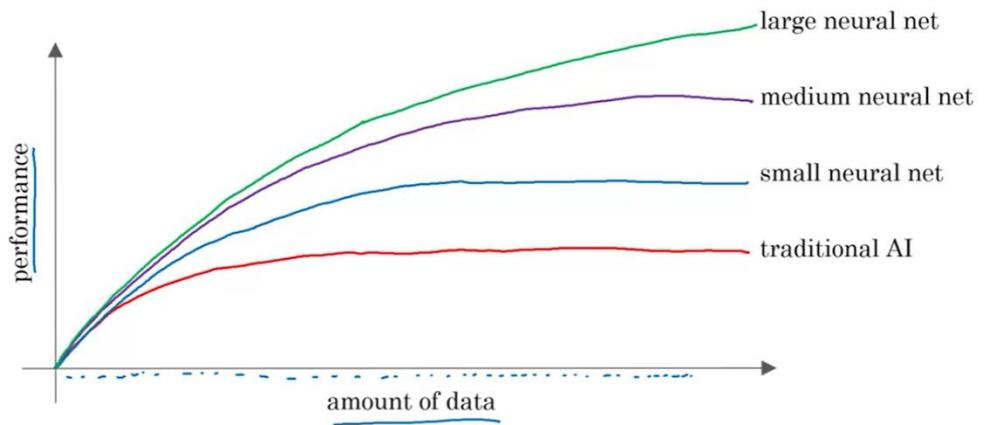
Examples

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

- Supervised learning

*Why Now?*

## Why Now?



What is Data?

## Example of a table of data (dataset)

size of house (square feet)	# of bedrooms	price (1000\$)
523	1	115
645	1	150
708	2	210
1034	3	280
2290	4	355
2545	4	440

image	label
	cat
	not cat
	cat
	not cat

Diagram illustrating the relationship between the two datasets:

- The first dataset (A) has 6 rows, indicated by a bracket labeled "A".
- The second dataset (B) has 4 rows, indicated by a bracket labeled "B".
- A blue arrow points from dataset A to dataset B, labeled "A → B".
- Handwritten annotations show:
  - A checkmark next to the first row of dataset B.
  - A cross next to the second row of dataset B.
  - A checkmark next to the third row of dataset B.
  - A cross next to the fourth row of dataset B.
- Below the second dataset, a bracket labeled "B" spans the last two rows, and a bracket labeled "A" spans the first two rows.

*Example*

## Example of a table of data (dataset)

size of house (square feet)	# of bedrooms	price (1000\$)
523	1	115
645	1	150
708	2	210
1034	3	280
2290	4	355
2545	4	440

Diagram illustrating the relationship between the two datasets:

- The first dataset (A) has 6 rows, indicated by a bracket labeled "A".
- The second dataset (B) has 4 rows, indicated by a bracket labeled "B".
- A blue arrow points from dataset A to dataset B, labeled "A → B".

- Example of a table of data (dataset)

## Acquiring data

# Acquiring data

- Manual labeling



- From observing behaviors

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

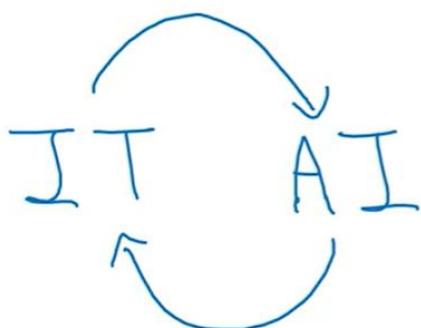
- Download from websites / partnerships

A

B

- Manual labelling
- From observing behaviors
  - Machines might be faulty

## Use and mis-use of data



Don't throw data at an AI team and assume it will be valuable.

## Misconceptions

- a) 3 years of data for IT team, we are collecting so much data. After 3 years, I will have perfect dataset.  
Really bad day strategy.  
Recommendation, once you get data from IT, show to AI team. AI team will feedback to IT.

Get AI team earlier.

- b) I have a lot data and AI Team make sense of it.

More data is usually better than less data, but don't take for granted that just because TBs of data AI team make magically valuable.

AI team can guide you, what is the data that is most valuable.s

*Data is messy*

## Data is messy

- Garbage in, garbage out

- Data problems
  - Incorrect labels
  - Missing values

- Multiple types of data

images, audio, text

*unstructured*

size of house (square feet)	# of bedrooms	price (1000\$)
523	1	115
645	1	0.001
708	unknown	210
1034	3	unknown
unknown	4	355
2545	unknown	440

*structured*

- Garbage in, Garbage out.
- Data problems
- Multiple type of data

The terminology of AI

*ML vs data science*

# Machine learning vs. data science

Home  
prices

size of house (square feet)	# of bedrooms	# of bathrooms	newly renovated	price (1000\$)
523	1	2	N	115
645	1	3	N	150
708	2	1	N	210
1034	3	3	Y	280
2290	4	4	N	355
2545	4	5	Y	440

ML. A → B

Running AI system  
(e.g., websites / mobile app)

A

B

Homes with 3 bedrooms are more expensive than homes with 2 bedrooms of a similar size.

Newly renovated homes have a 15% premium.

# Machine learning vs. data science

Machine learning

“Field of study that gives computers the ability to learn without being explicitly programmed.”

-Arthur Samuel (1959)

Data science

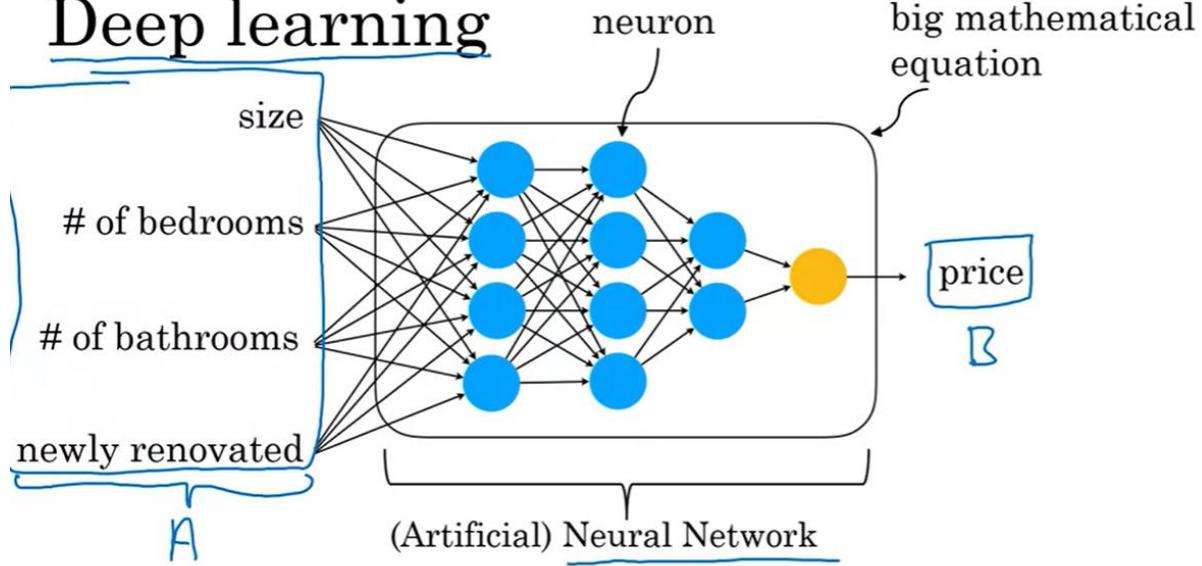
Science of extracting knowledge and insights from data.

software

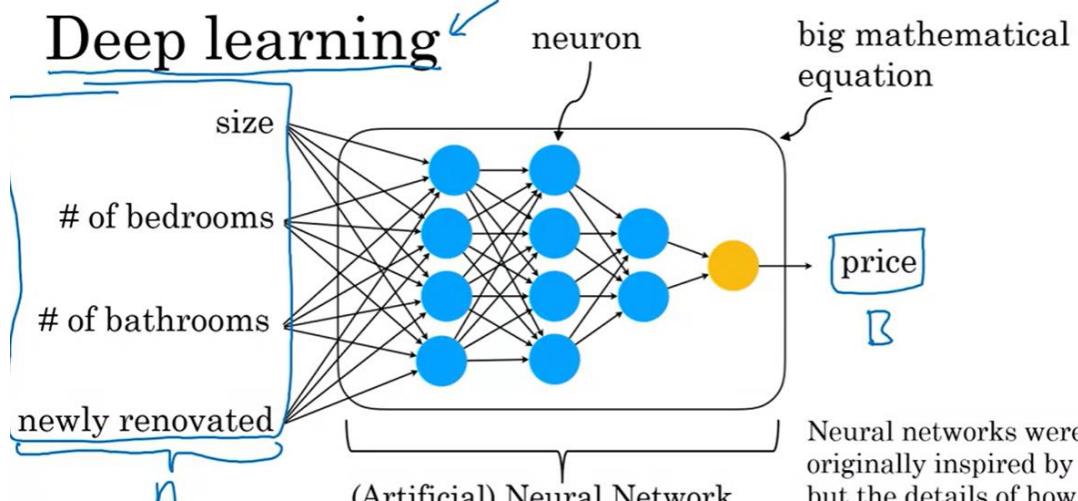
slide deck

Deep learning

# Deep learning



# Deep learning



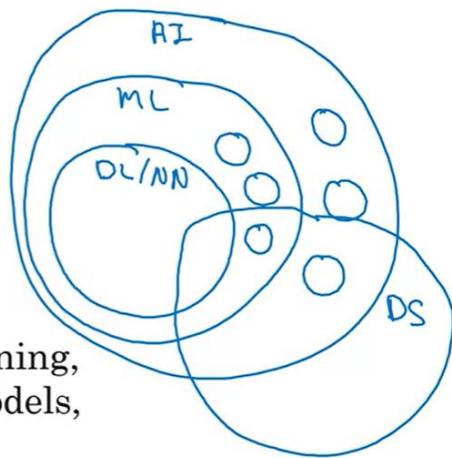
Neural networks were originally inspired by the brain, but the details of how they work are almost completely unrelated to how biological brains work.

- Deep learning

*AI has many tools*

## AI has many tools

- Machine learning and data science
- Deep learning / neural network
- Other buzzwords: Unsupervised learning, reinforcement learning, graphical models, planning, knowledge graph, ...



What makes an AI Company

*A lesson from the rise of the Internet*

## A lesson from the rise of the Internet

### Internet Era

Shopping mall + website  
≠ Internet company

- A/B testing
- Short iteration time
- Decision making pushed down to engineers and other specialized roles

### AI era

Any company + deep learning  
≠ AI company

- Strategic data acquisition
- Unified data warehouse
- Pervasive automation
- New roles (e.g., MLE) and division of labor

Internet Era

- A/B testing
- Short iterative time
- Decision making pushed down to engineer and other specialized roles.

AI Era

- Any company + deep learning is NOT AI company
- Strategic data acquisition

- Unified data warehouse
- Pervasive automation
- New roles (MLE) and division of labor

*AI Transformation*

1. Execute pilot projects to gain momentum
2. Build an in-house AI team
3. Provide AI Training
4. Develop an AI strategy
5. Develop internal

## AI Transformation

1. Execute pilot projects to gain momentum
2. Build an in-house AI team
3. Provide broad AI training
4. Develop an AI strategy
5. Develop internal and external communications

What machine learning can do and cannot do

*Can do*

## Supervised Learning

# Supervised Learning

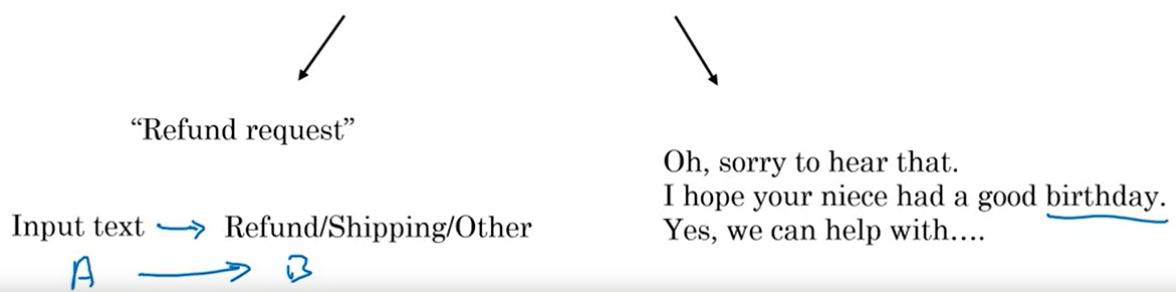
Input (A)	Output (B)	Application
email	spam? (0/1)	spam filtering
→ audio	text transcripts	speech recognition
English	Chinese	machine translation
ad, user info	click? (0/1)	online advertising
→ image, radar info	position of other cars	Self-driving car
→ image of phone	defect? (0/1)	visual inspection

Anything you can do with 1 second of thought,  
we can probably now or soon automate.

## What machine learning today can and cannot do

The toy arrived two days late, so I wasn't able to give it to my niece for her birthday.

Can I return it?



Refund request = can do

Sorry to hear that = cannot do.

*What happens if you can try*

## What happens if you try?

Input (A) → Output (B)  
User email → 2-3 paragraph response

1000 examples

“My box was damaged.” → Thank you for your email.

“Where do I write a review?” → Thank you for your email.

“What’s the return policy?” → Thank you for your email.

“When is my box arriving?” → Thank yes now your....

Ss

*What makes a ML problem easier*

## What makes a ML problem easier

1. Learning a “simple” concept

$\leq 1 \text{ sec}$

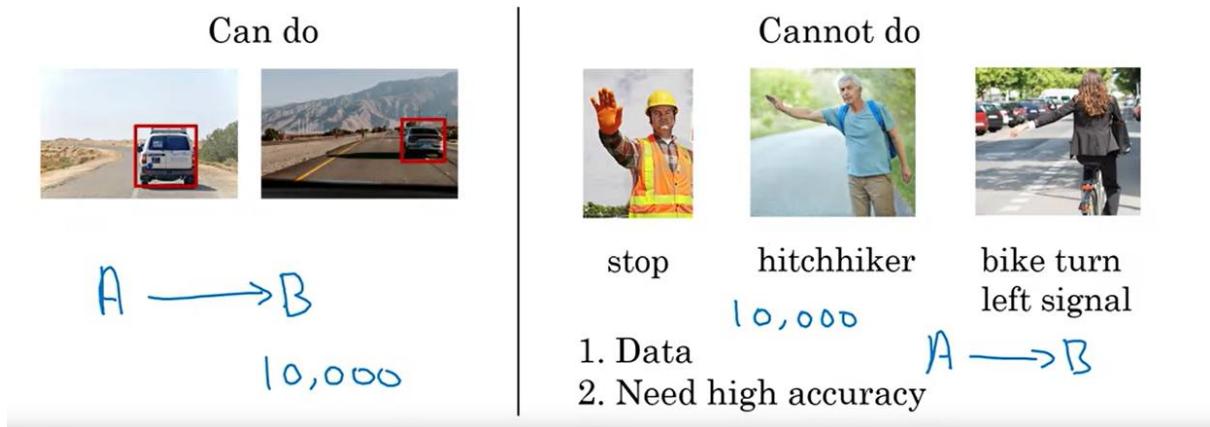
2. Lots of data available

A, B →  $\underbrace{A}_{\sim} \rightarrow B$

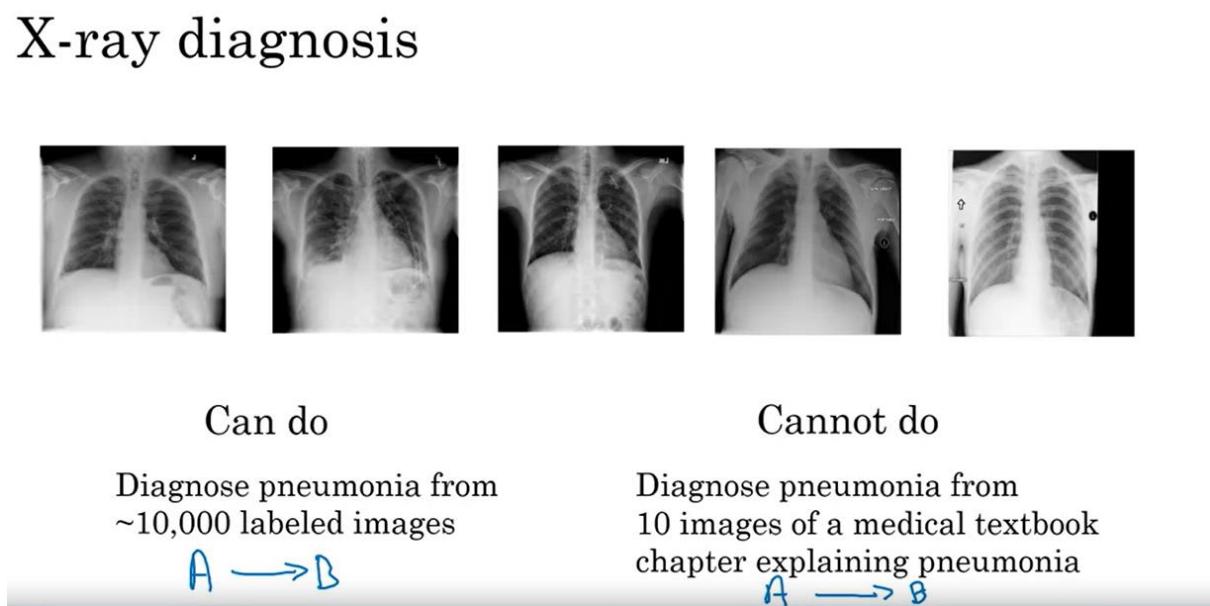
More examples of What machine learning can do and cannot do

*Self-driving car*

# Self-driving car



*X-ray diagnosis*



*Strengths and weaknesses of machine learning*

# Strengths and weaknesses of machine learning

ML tends to work well when:

1. Learning a “simple” concept
2. There is lots of data available

ML tends to work poorly when:

1. Learning complex concepts from small amounts of data
2. It is asked to perform on new types of data

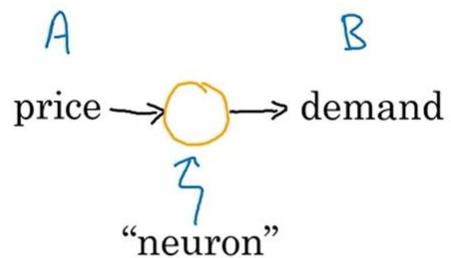
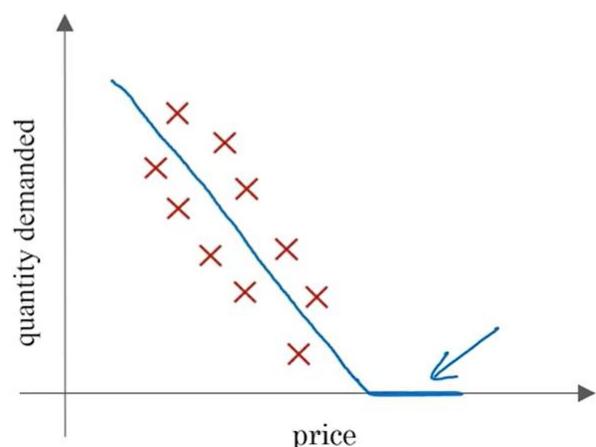


OPTIONAL Non-technical explanations of deep learning – Part 1

*What is AI?*

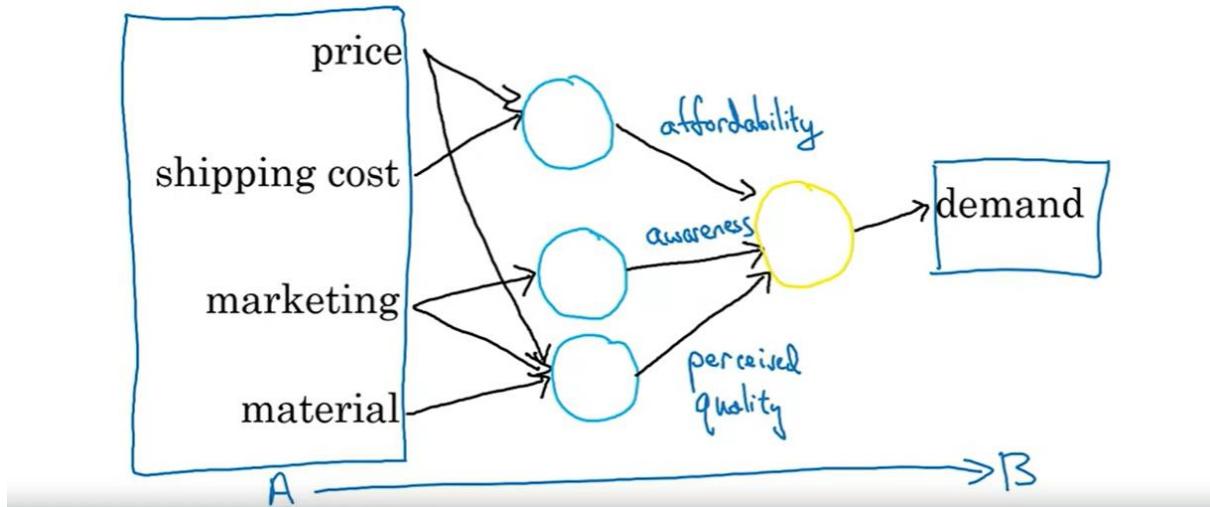
Demand prediction

## Demand prediction



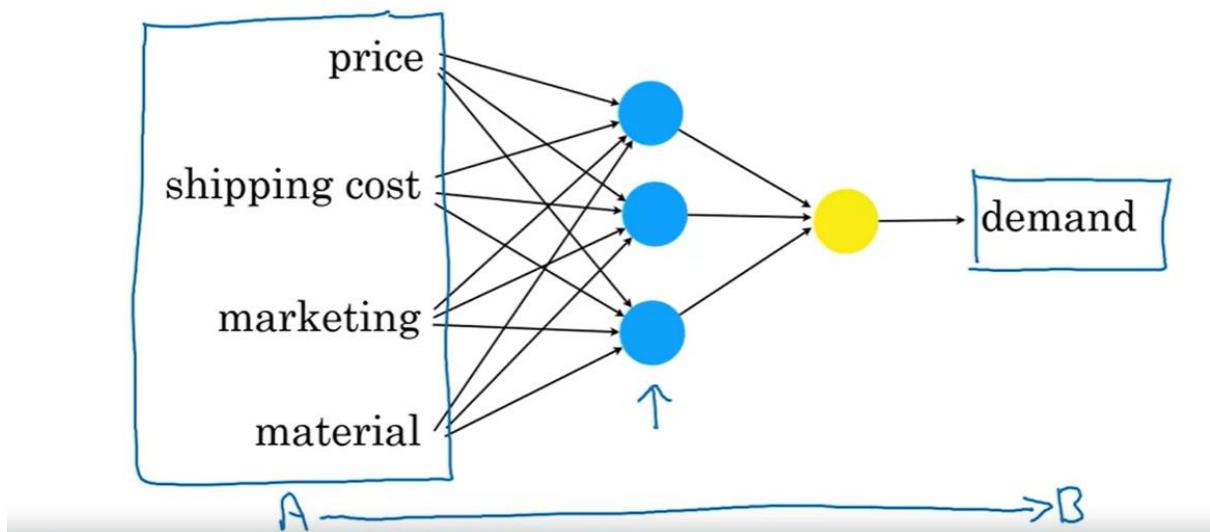
- Artificial neuron

# Demand prediction



- It's just 4 parameters
- In real AI, it would be 10000 parameters

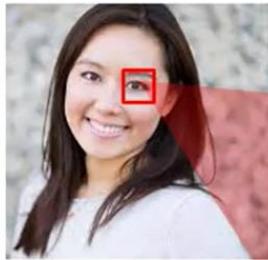
# Demand prediction



OPTIONAL Non-technical explanations of deep learning (more complex) – Part 2

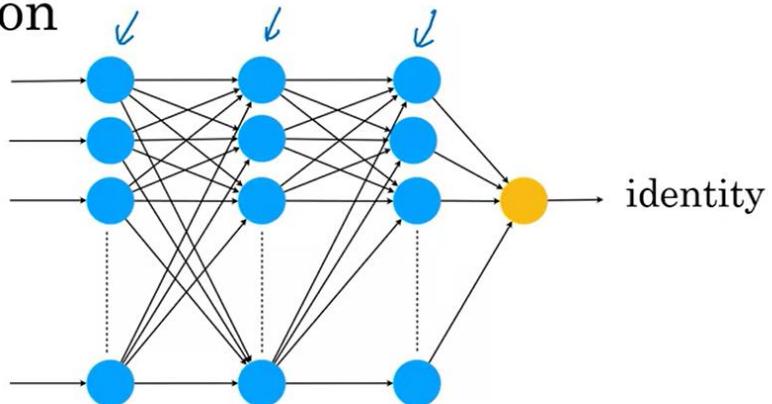
## *Face Recognition*

# 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



1000

1600  
1,000,000  
3,000,000



Lecture notes

Week 1 Quiz

coursera Search in course Search

Hide menu

What is AI?

- Video: Week 1 Introduction 7 min
- Video: Machine Learning 6 min
- Video: What is data? 11 min
- Video: The terminology of AI 9 min
- Video: What makes an AI company? 7 min

Ungraded App Item [IMPORTANT] Have questions, issues or ideas? Join our Community! 10 min

- Video: What machine learning can and cannot do

< Previous Next >

## Week 1 Quiz

Submit your assignment Due Oct 13, 11:59 PM +08 Attempts 3 every 8 hours Start assignment

Receive grade To Pass 80% or higher Your grade -

Like Dislike Report an issue

English Aroh Shukla

coursera Search in course Search

Hide menu

What is AI?

- Video: Week 1 Introduction 7 min
- Video: Machine Learning 6 min
- Video: What is data? 11 min
- Video: The terminology of AI 9 min
- Video: What makes an AI company? 7 min

Ungraded App Item [IMPORTANT] Have questions, issues or ideas? Join our Community! 10 min

- Video: What machine learning can and cannot do

< Previous Next >

## Week 1 Quiz

Submit your assignment Due Oct 13, 11:59 PM +08 Attempts 3 every 8 hours Start assignment

Receive grade To Pass 80% or higher Your grade -

Like Dislike Report an issue

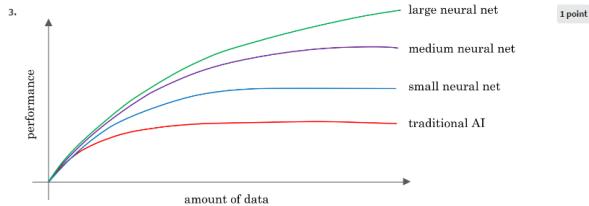
English Aroh Shukla

1. Which of these terms best describes the type of AI used in today's email spam filters, speech recognition, and other specific applications? 1 point

Artificial Narrow Intelligence (ANI)  
 Artificial General Intelligence (AGI)

2. What do you call the commonly used AI technology for learning input (A) to output (B) mappings? 1 point

Supervised learning  
 Reinforcement learning  
 Unsupervised learning  
 Artificial General Intelligence



You want to use supervised learning to build a speech recognition system. The figure above suggests that in order for a neural network (deep learning) to achieve the best performance, you would ideally use: (Select all that apply)

- A large dataset (of audio files and the corresponding text transcript)  
 A small dataset (of audio files and the corresponding text transcript)  
 A large neural network  
 A small neural network

4. The only way to acquire data for a supervised learning algorithm is to manually label it. I.e., given the input A, to ask a human to provide B. 1 point

True  
 False

5. Which of these statements regarding data acquisition do you agree with? 1 point

- It doesn't matter how data is acquired. The more data, the better.  
 Only structured data is valuable; AI cannot process unstructured data.  
 It doesn't help to give data to an AI team, because they can always produce whatever they need by themselves.  
 Some types of data are more valuable than others; working with an AI team can help you figure out what data to acquire.

6. You run a company that manufactures scooters. Which of the following are examples of unstructured data? (Select all that apply.) 1 point

- Audio files of the engine sound of your scooters  
 The maximum speed of each of your scooters  
 The number of scooters sold per week over the past year  
 Pictures of your scooters

7. Suppose you run a website that sells cat food. Which of these might be a good result from a Data Science project? (Select all that apply.) 1 point

- A neural network that closely mimics how cats' brains work.  
 A large dataset of images labeled as "Cat" and "Not Cat"  
 Insights into how to market cat food more effectively, depending on the breed of cat.  
 A slide deck presenting a plan on how to modify pricing in order to improve sales.

8. Based on the terminology defined in Video 4, which of the following statements do you agree with? (Select all that apply.) 1 point

- AI is a type of deep learning. (I.e., all AI algorithms are deep learning algorithms.)  
 Deep learning is a type of machine learning. (I.e., all deep learning algorithms are machine learning algorithms.)  
 The terms "Machine learning" and "data science" are used almost interchangeably.  
 The terms "Deep learning" and "neural network" are used almost interchangeably.

9. Which of these do AI companies do well? 1 point

- Strategic data acquisition  
 Invest in unified data warehouses  
 Spot automation opportunities  
 All of the above

10. Say you want to input a picture of a person's face (A), and output whether or not they are smiling (B). Because this is a task that most humans can do in less than 1 second, supervised learning can probably learn this A-to-B mapping. 1 point

True  
 False

4. The only way to acquire data for a supervised learning algorithm is to manually label it. I.e., given the input A, to ask a human to provide B.

0 / 1 point

- True  
 False

 Incorrect

5. Which of these statements regarding data acquisition do you agree with?

0 / 1 point

- It doesn't matter how data is acquired. The more data, the better.  
 Only structured data is valuable; AI cannot process unstructured data.  
 It doesn't help to give data to an AI team, because they can always produce whatever they need by themselves.  
 Some types of data are more valuable than others; working with an AI team can help you figure out what data to acquire.

 Incorrect

7. Suppose you run a website that sells cat food. Which of these might be a good result from a Data Science project?  
(Select all that apply.)

0.5 / 1 point

- A neural network that closely mimics how cats' brains work.  
 A large dataset of images labeled as "Cat" and "Not Cat"  
  
 This should not be selected
- Insights into how to market cat food more effectively, depending on the breed of cat.  
 A slide deck presenting a plan on how to modify pricing in order to improve sales.

 Correct

8. Based on the terminology defined in Video 4, which of the following statements do you agree with? (Select all that apply.)

0.5 / 1 point

- AI is a type of deep learning. (I.e., all AI algorithms are deep learning algorithms.)  
  
 This should not be selected
- Deep learning is a type of machine learning. (I.e., all deep learning algorithms are machine learning algorithms.)  
 The terms "Machine learning" and "data science" are used almost interchangeably.  
 The terms "Deep learning" and "neural network" are used almost interchangeably.

Week 2  
Introduction

# Starting an AI project

- Workflow of projects
- Selecting AI projects
- Organizing data and team for the projects

## Starting an AI Project

- Workflow of projects
- Select AI projects
- Organizing data and teams for projects

## Building AI projects

*Example: Speech recognition*

### Example: Speech recognition



Amazon  
*Echo / Alexa*



Google  
*Home*



Apple  
*Siri*



Baidu  
*DuerOS*

*Key ML project – Speakers*

# Key steps of a machine learning project

## *Echo / Alexa*

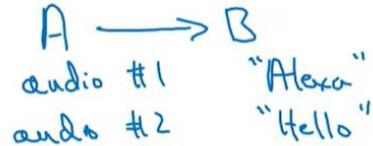
→ 1. Collect data

→ 2. Train model

Iterate many times until  
good enough

→ 3. Deploy model

Get data back  
Maintain / update model



- Collect Data (Audio clips)
- Train model
  - Iterate many times
- Deploy model
  - Get data back
  - Maintain / update model

## *Key ML project – Self-driving*

# Key steps of a machine learning project

## *Self-driving car*

1. Collect data

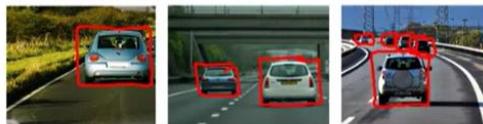


image → position of other cars

2. Train model

Iterate many times until  
good enough



3. Deploy model

Get data back  
Maintain / update model

- Collect data
  - Image => position of other cars
- Train model
  - Iterate many times
- Deploy model

- Get data back
- Maintain / update model

Workflow of a data science project

*Building Data Science project*

Example : Optimizing a sales funnel

## Example: Optimizing a sales funnel



*Key Steps of data science project*

## Key steps of a data science project

### *Optimizing a sales funnel*

→ 1. Collect data

User ID	Country	Time	Webpage
2009	Spain	08:34:30 Jan 5	home.html
2897	USA	13:20:22 May 18	redmug.html
4893	Philippines	22:45:16 Jun 11	mug.html

→ 2. Analyze data

Iterate many times to get good insights

→ 3. Suggest hypotheses/actions

Deploy changes

Re-analyze new data periodically

- Collect Data
- Analyze data
  - Iterate many times to get good insights
- Suggest hypotheses / actions
  - Deploy changes

- Re-analyze new data periodically

## Key steps of a data science project

### *Manufacturing line*



Clay Batch #	Supplier	Mixing time (minutes)
001	ClayCo	35
034	GooClay	22
109	BrownStuff	28

1. Collect data
2. Analyze data  
Iterate many times to get good insights
3. Suggest hypotheses/actions

Mug Batch #	Humidity	Temperature in kiln (F)	Duration in kiln (hours)
301	0.002%	1410°	22
302	0.003%	1520°	24
303	0.002%	1420°	22

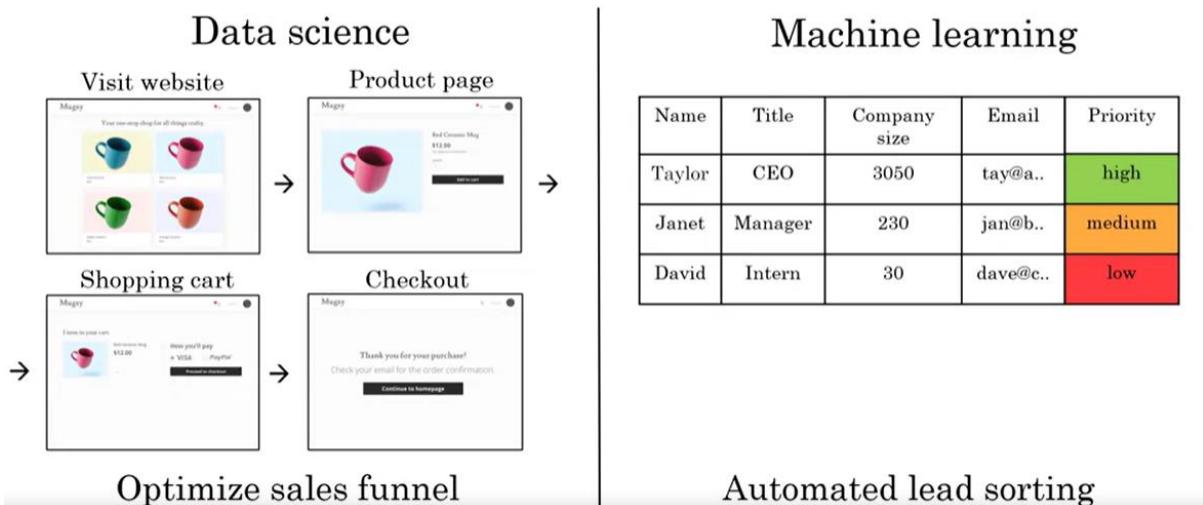
- Collect Data
- Analyze data
  - Iterate many times to get good insights
- Suggest hypotheses / actions
  - Deploy changes
  - Re-analyze new data periodically

Every job function needs to learn how to use data

- Data is transforming.
- Digital survey into paper to digital record
- Your job function or AI or Data science

*Sales*

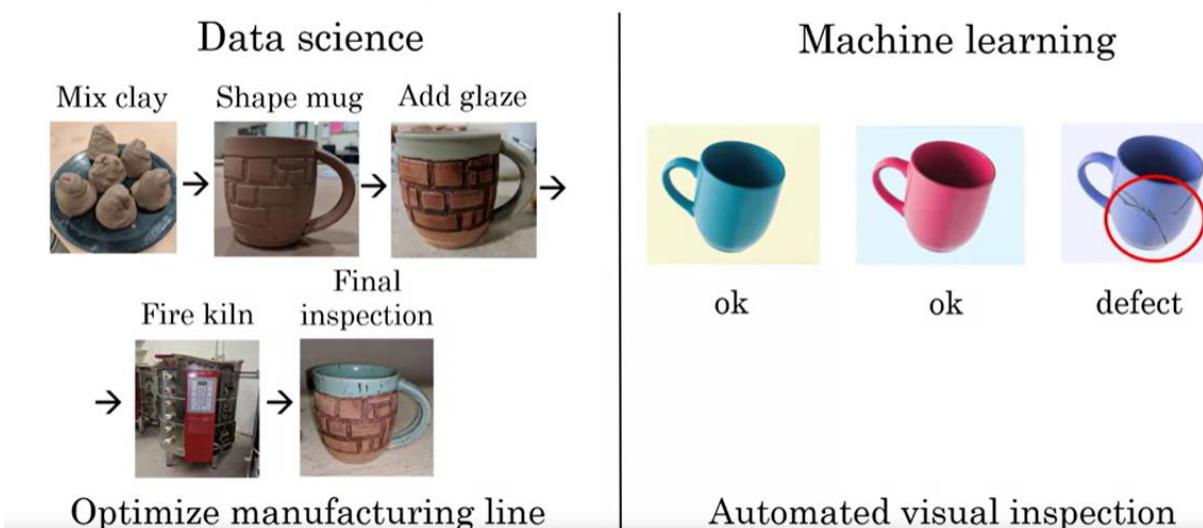
# Sales



- Data Science
  - Optimized sales funnel
- Machine learning
  - Prioritize - Automated lead sorting

*Manufacturing line manager*

## Manufacturing line manager

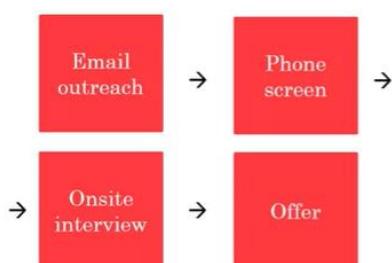


- Line manager of manufacturing
- Data Science
  - Optimized manufacturing line
- Machine learning
  - Automated visual inspection
  - Final inspection – 1000s to check
  - Automate mug,
  - Reduce labor cost
  - Improve quality

## Recruiting

# Recruiting

### Data science



Optimize recruiting funnel

### Machine learning

Jane Doe
Personal Info
Education
Professional
Employment

→ Yes

Tiffany Doe
Personal Info
Education
Professional
Employment

→ No

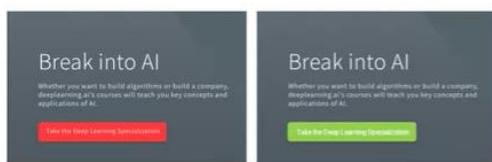
Automated resume screening

- Data Science
  - Optimized recruiting funnel
- Machine learning
  - Automated resume screening
  - Recruitment AI should not be biased

## Marketing

# Marketing

### Data science

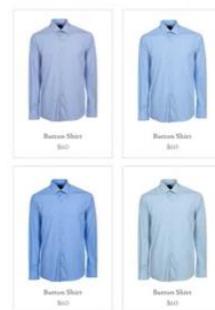


B

A/B testing

### Machine learning

Recommended for you



Customized product recommendation

- Data Science
  - A/B testing
  - 2 versions of website
- Machine learning
  - Customized product recommendation
  - Recommendations

## Agriculture

Data science



Crop analytics

Machine learning



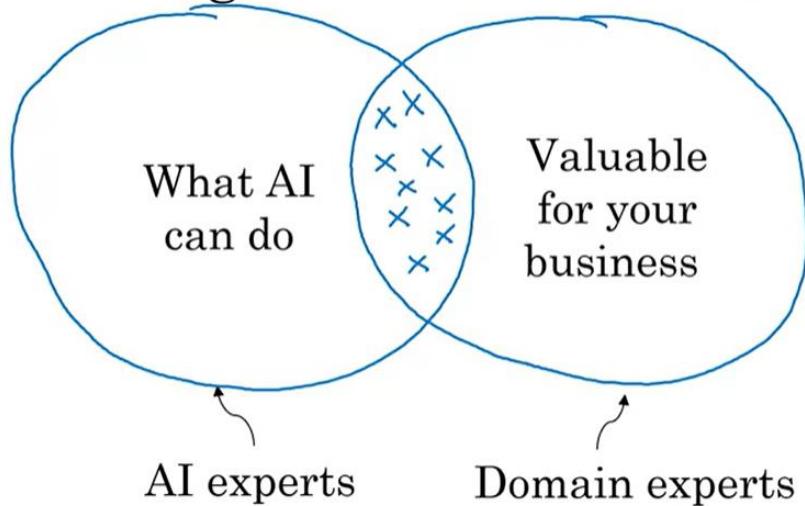
Precision weed killing

- Data Science
  - Crop Analytics
- Machine learning
  - Precision weed killing

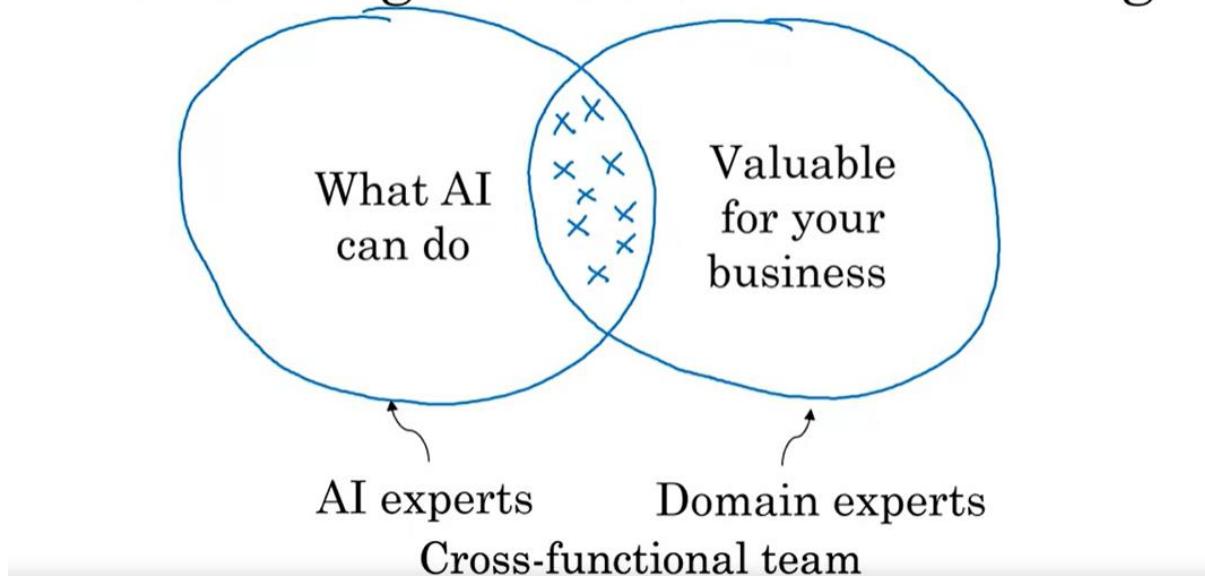
How to choose an AI project (Part 1)

*AI knowledge and domain knowledge*

## AI knowledge and domain knowledge



# AI knowledge and domain knowledge



*Brainstorming framework*

## Brainstorming framework

- Think about automating tasks rather than automating jobs. E.g., call center routing, radiologists.
- What are the main drivers of business value?
- What are the main pain points in your business?

3 principles

- Automating tasks then jobs
- Main drivers of business value.
- Main pain points in your business

# You can make progress even without big data

- Having more data almost never hurts.
- Data makes some businesses (like web search) defensible.
- But with small datasets, you can still make progress.

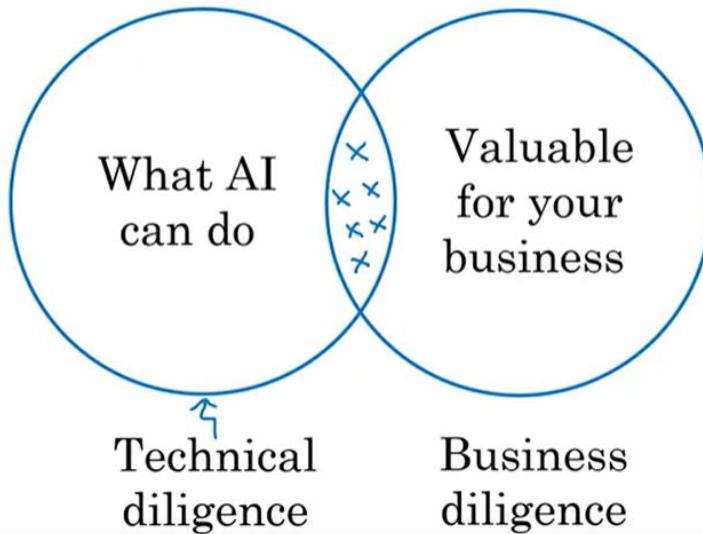


- Having more data almost never hurts
- Data makes some business (like web search) defensible.
- Smaller datasets, you can still make progress

How to choose an AI project (Part 2)

*Due diligence on project*

## Due diligence on project



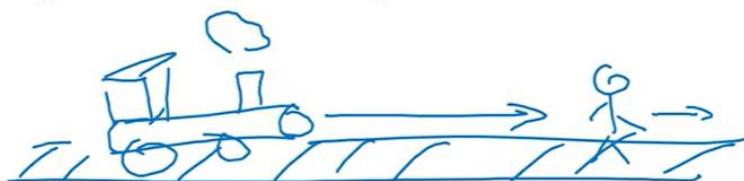
# Due diligence on project

Technical diligence	Business diligence
<ul style="list-style-type: none"><li>• Can AI system meet desired performance</li><li>• How much data is needed</li><li>• Engineering timeline</li></ul>	<ul style="list-style-type: none"><li>• Lower costs</li><li>• Increase revenue</li><li>• Launch new product or business</li></ul>
<ul style="list-style-type: none"><li>• Technical diligence<ul style="list-style-type: none"><li>◦ Can AI system meet your desired performance</li><li>◦ How much data is needed</li><li>◦ Engineering timeline</li></ul></li><li>• Business diligence<ul style="list-style-type: none"><li>◦ Low costs</li><li>◦ Increase revenue</li><li>◦ Launch product or business</li></ul></li><li>• Ethical diligence</li></ul>	<p>} current business</p> <p>} new business</p>

*Build vs buy*

## Build vs. buy

- ML projects can be in-house or outsourced
- DS projects are more commonly in-house
- Some things will be industry standard – avoid building those.

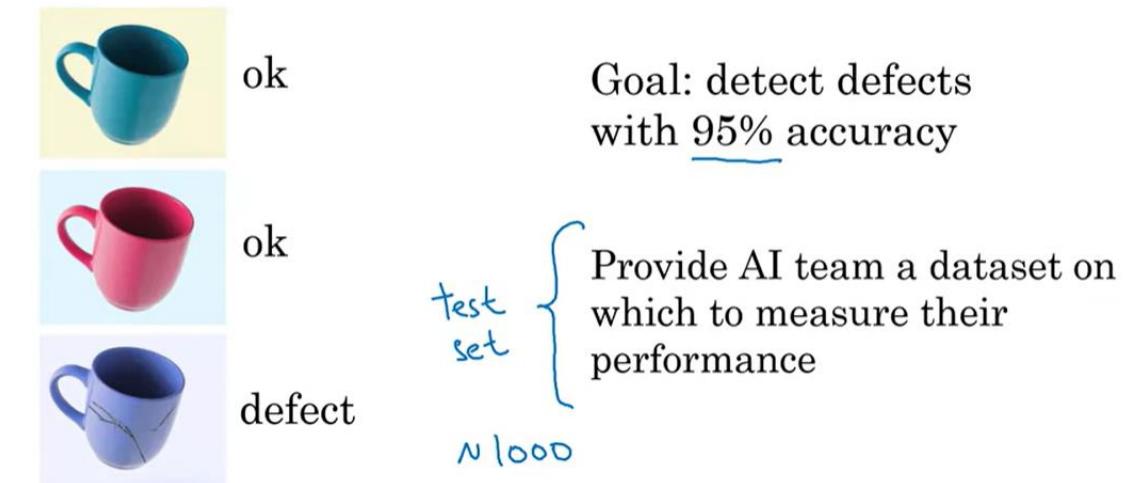


- Buy vs build
- ML projects can be in-house or outsourced
- DS projects are more in-house
- Some things will be industry standard, avoid building those

## Working with an AI Team

*Specific your acceptance criteria*

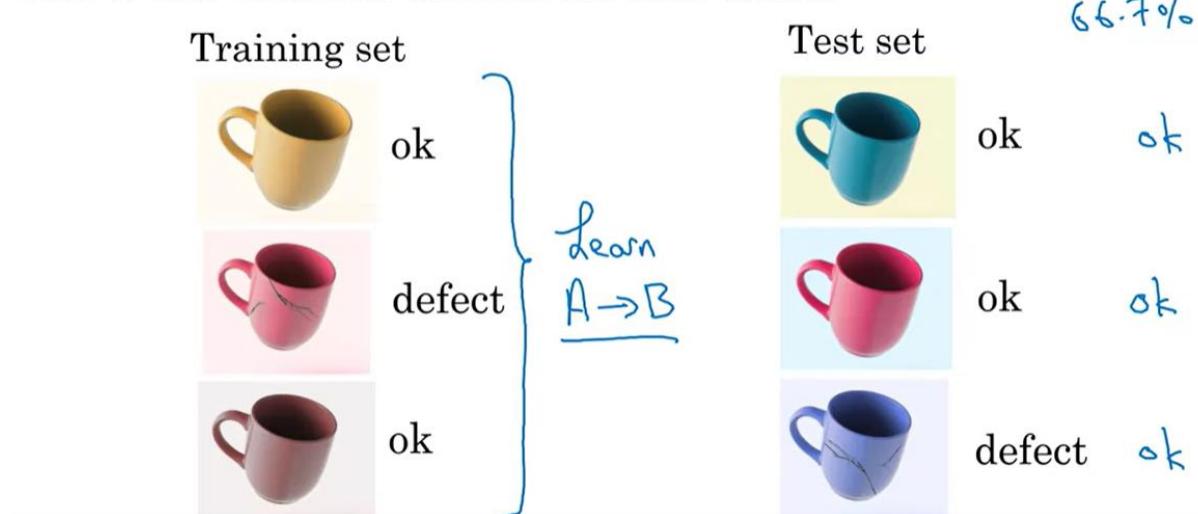
## Specify your acceptance criteria



- Specific your acceptance criteria
- Goal: detects defects with 95% accuracy.
- Need a data for AI Team
- Example of OK and NO okay with defects of coffee mugs
- Called as Testset (1000 pics is fine)

*How AI Team think about data*

## How AI teams think about data



- Training set | Test set

*Pitfall: Expecting 100% accuracy*

## Pitfall: Expecting 100% accuracy

Test set		
	ok	 - defect ok
	ok	 ok defect?
	defect	 ok

- Test set
- Limitations of ML
- Insufficient data
- Data is messy Mislabeled data
- Ambiguous labels

OPTIONAL: Technical tools for AI Teams

*AI Technical tools*

## AI technical tools

### Machine learning frameworks:

- TensorFlow
- PyTorch
- Keras
- MXNet
- CNTK
- Caffe
- PaddlePaddle
- Scikit-learn
- R
- Weka

- ML Frameworks
- Research publications
- Open source repos
  - GitHub

### Research publications:

- Arxiv

### Open source repositories:

- GitHub

CPU VS GPU

# CPU vs. GPU

CPU: Computer processor (Central Processing Unit)



GPU: Graphics Processing Unit



## Cloud vs. On-premises

- CPU – Computer processor (Central processing unit)
- GPU – Graphics Processing Unit
- Cloud vs On-premises
- Edge

## Week 2 Quiz

Menu > Week 2 > Week 2 Quiz < Previous Next >

### Week 2 Quiz

Submit your assignment  
Due: Oct 14, 11:59 PM +08 Attempts: 3 every 8 hours Start assignment

Receive grade  
To Pass: 80% or higher Your grade -

Like Dislike Report an issue

**Congratulations! You passed!**

Grade received 80%  
Latest Submission Grade 80%

To pass 80% or higher

[Go to next item](#)

1. Machine learning is an "iterative" process, meaning that an AI team often has to try many ideas before coming up with something that's good enough, rather than have the first thing they try work.

True  
 False  
 Correct

1 / 1 point

2. Say you want to use Machine Learning to help your sales team with automatic lead sorting. I.e., Input A (a sales prospect) and output B (whether your sales team should prioritize them). The 3 steps of the workflow, in scrambled order, are:

Deploy a trained model and get data back from users  
 Collect data with both A and B  
 Train a machine learning system to input A and output B

What is the correct ordering of these steps?

(i) (ii) (iii)  
 (i) (iii) (ii)  
 (ii) (iii) (i)  
 (ii) (i) (iii)

Incorrect

0 / 1 point

3. What are the key steps of a Data Science project?

Collect data  
 Analyze the data  
 Suggest hypothesis or actions  
 All of the above

Correct

1 / 1 point

4. Machine Learning programs can help: (select all that apply)

Automate visual inspection in a manufacturing line  
 Correct  
 Automate lead sorting in sales  
 Correct  
 Automate resume screening  
 Correct  
 Customize product recommendations  
 Correct

1 / 1 point

5. Unless you have a huge dataset ("Big Data"), it is generally not worth attempting machine learning or data science projects on your problem.

True  
 False  
 Correct

1 / 1 point

6. Say you want to build an AI system to help recruiters with automated resume screening. Which of these steps might be involved in "technical diligence" process? (Select all that apply.)

Making sure that an AI system can meet the desired performance  
 Correct  
 Defining an engineering timeline  
 Correct  
 Ensuring that this is valuable for your business (e.g., estimating the project ROI)  
 Making sure you can get enough data for this project  
 Correct

1 / 1 point

7. Which of these statements about "business diligence" do you agree with?

Business diligence can typically be completed in less than a day.  
 Business diligence is the process of ensuring that the AI technology, if it is built, is valuable for your business.  
 Business diligence is the process of ensuring that the envisioned AI technology is feasible.  
 Business diligence applies only if you're launching new product lines or businesses.

Correct

1 / 1 point

8. You want to use supervised learning for automated resume screening, as in the example above. Which of the following statements about the Training Set are true? (Select all that apply.)

It should give examples of both the input A (resume) and the desired output B (whether to move forward with a candidate).  
 The Training set and Test set can be the same dataset.  
 It should give examples of the input A (resume) but not necessarily the desired output B (whether to move forward with a candidate).  
 This should not be selected

It will be used by the AI team to train the supervised learning algorithm.

Correct

0.5 / 1 point

9. For your automated resume screening application, you are now providing a Test Set to the AI team. Which of the following statements about the Test Set are true? (Select all that apply.)

It should give examples of the input A (resume) but not necessarily the desired output B (whether to move forward with a candidate).  
 This should not be selected  
 It will be used by the AI team to evaluate the performance of the algorithm.  
 The Test Set should ideally be identical to the Training Set.  
 It should give examples of both the input A (resume) and the desired output B (whether to move forward with a candidate)

Correct

0.5 / 1 point

10. Which of these are reasons that it's often unrealistic to expect an ML system to be 100% accurate?

You might not have enough data  
 Data can be mislabeled  
 Data can be ambiguous  
 All of the above

Correct

1 / 1 point

## Week 3

Building AI in your company

# Building AI in your company

- Case studies of complex AI products
  - Roles in an AI team
  - AI Transformation Playbook
  - Taking your first step
- 
- Case studies of a complex AI products
  - Role in an AI team
  - AO Transformation playbook
  - Taking your first step

Case study: Smart speaker

*Smart speaker*

## Smart speaker



Amazon  
Echo / Alexa



Google  
Home



Apple  
Siri



Baidu  
DuerOS

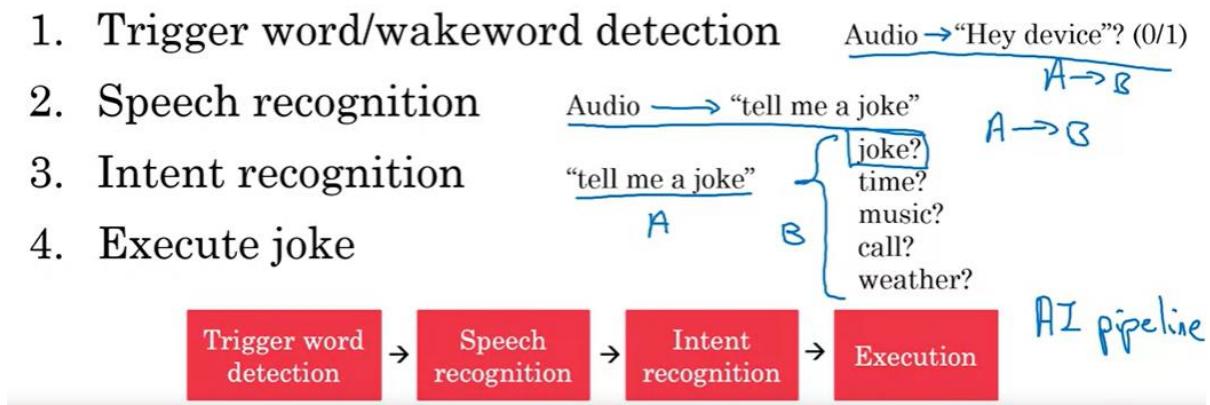
“Hey device, tell me a joke”

*Hey device, tell me a joke.*

## *“Hey device, tell me a joke”*

Steps to process the command:

1. Trigger word/wakeword detection
2. Speech recognition
3. Intent recognition
4. Execute joke



Steps to process the command

1. Trigger work / wakeup detection
2. Speech recognition
3. Intent recognition
4. Execute the joke

AI Pipeline

*Hey device, set timer for 10 mins*

## *“Hey device, set timer for 10 minutes”*

Steps to process the command:

1. Trigger word/wakeword detection      Audio → “Hey device”? (0/1)
2. Speech recognition      Audio → “set timer for 10 minutes”
3. Intent recognition      “set timer for 10 minutes” → timer
4. a) Extract duration  
    “Set timer for **10 minutes**”  
    “Let me know when **10 minutes** is up”
- b) Start timer with set duration

Steps to process the command

1. Trigger work / wakeup detection
2. Speech recognition

3. Intent recognition
4. A) Execute the joke
  - a. Execute the joke  
Let me know when 10 mins is up
  - b. Start timer with set duration

*Other functions*

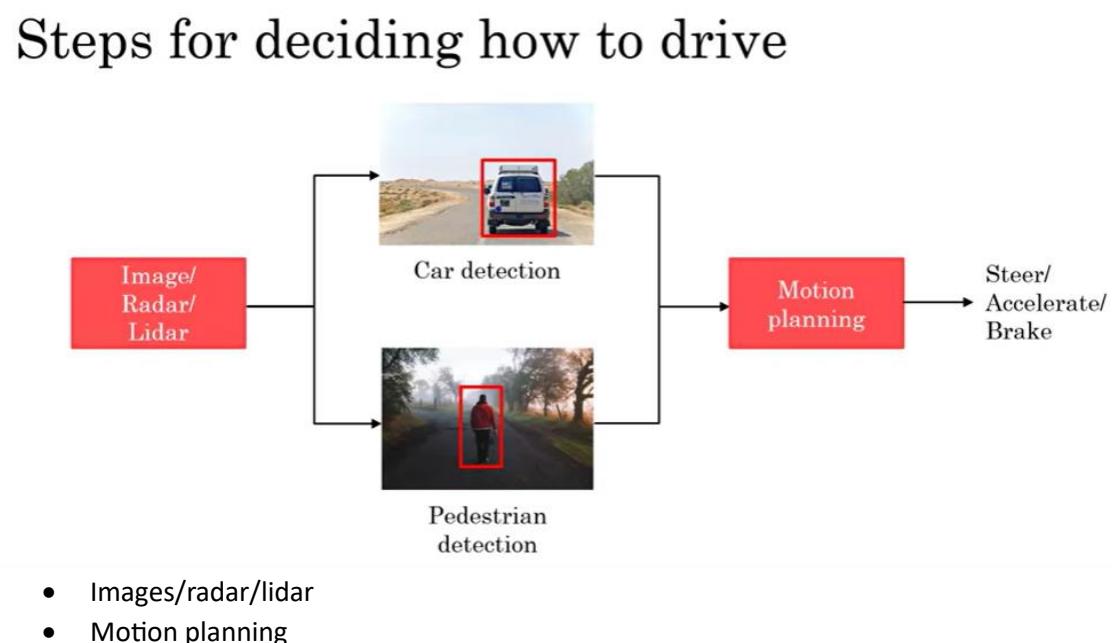
## Other functions

- Play music
- Volume up/down
- Make call
- Current time
- Units conversion
- Simple question
- ...

- Key steps:
1. Trigger/wakeword detection
  2. Speech recognition
  3. Intent recognition
  4. Specialized program to execute command

Case study: Self-driving car

*Steps for deciding how to drive*



## *Key Steps*

### Key steps:

#### 1. Car detection



#### 2. Pedestrian detection

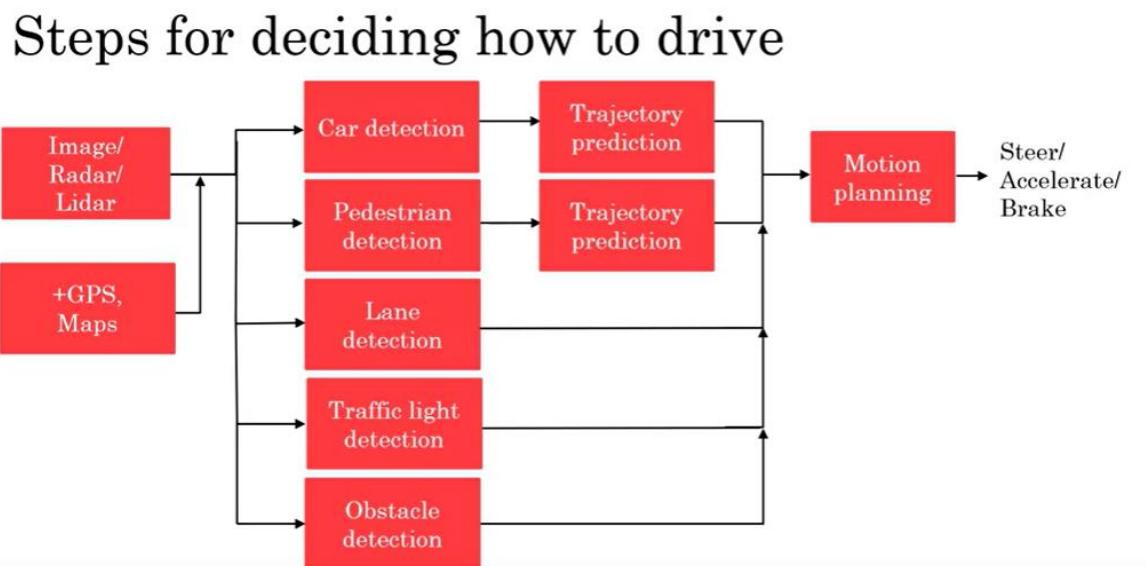


#### 3. Motion planning



1. Car detection
2. Pedestrian detection
3. Motion planning – path and speed

## *Steps for deciding how to drive*



- Images/radar/lidar
  - +GPS, Maps, Tr
- Car detection
  - Trajectory prediction

- Pedestrian detection
  - Trajectory prediction
- Lane detection
- Traffic light detection
- Obstacle detection
- Motion planning
  - Steer / Accelerate / brake

Example roles of an AI team

*Example roles*

## Example roles

- Software Engineer
    - E.g., joke execution, ensure self-driving reliability, ...
  - Machine Learning Engineer
    - A → B
  - Machine Learning Researcher
    - Extend state-of-the-art in ML
- } Applied ML Scientist
- Software Engineer
    - E.g. Joke execution, ensure self-driving reliability
  - ML Engineer
  - ML Researcher
    - Extend state-of-the-art in ML

*Example role 2*

## Example roles

- Data Scientist
    - Examine data and provide insights
    - Make presentation to team/executive
  - Data Engineer
    - Organize data
    - Make sure data is saved in an easily accessible, secure and cost effective way
  - AI Product Manager
    - Help decide what to build; what's feasible and valuable
- 
- Data Scientist
    - Examine data and provide insights
    - Make presentation to team/executive
  - Data Engineer
    - Organize data
    - Ensure it saves in an accessible, secure and cost effective way
  - AI Product Manager
    - Help decide what to build; what's feasible and valuable

*Gettings started with a small team*

- ### Getting started with a small team
- 1 Software Engineer , or
  - 1 Machine Learning Engineer/Data Scientist , or
  - Nobody but yourself
- 
- 1 Software engineer
  - 1 ML / Data Scientists
  - Nobody but yourself

AI Transformation Playbook (Part 1)

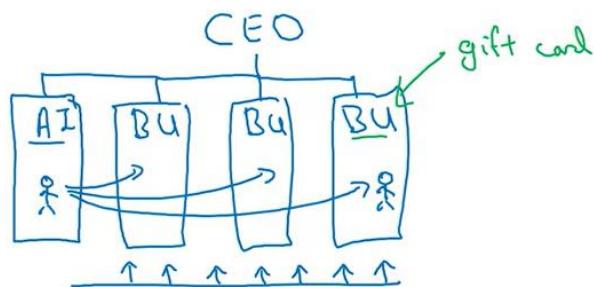
## AI Transformation Playbook

1. Execute pilot projects to gain momentum
2. Build an in-house AI team
3. Provide broad AI training
4. Develop an AI strategy
5. Develop internal and external communications

- Execute pilot projects to gain momentum.
- Build an in-house AI team
- Provide broad AI training.
- Develop an AI strategy.
- Develop internal and external communications.

1. *Execute pilot projects to gain momentum.*
  - More imp for the initial project to succeed rather than most valuable
  - Show transaction within 6-12 months
  - Can be in-hour or outsources
2. *Build in a in-house AI Team*

### 2. Build an in-house AI team



BU= Business Unit

AI function can be under CTO, CIO, CDO, etc. or a new CAIO

- CEO – should fund rather then BUs

- AI | BU1 | BU2| BU2

3. *Provide Broad AI Training*

### 3. Provide broad AI training

Role	What they should learn
Executives and senior business leaders	<ul style="list-style-type: none"> <li>• What AI can do for your enterprise</li> <li>• AI strategy</li> <li>• Resource allocation</li> </ul>
Leaders of divisions working on AI projects	<ul style="list-style-type: none"> <li>• Set project direction (technical and business diligence)</li> <li>• Resource allocation</li> <li>• Monitor progress</li> </ul>
AI engineer trainees	<ul style="list-style-type: none"> <li>• Build and ship AI software</li> <li>• Gather data</li> <li>• Execute on specific AI projects</li> </ul>

The smart CLO knows they should *curate* rather than *create* content

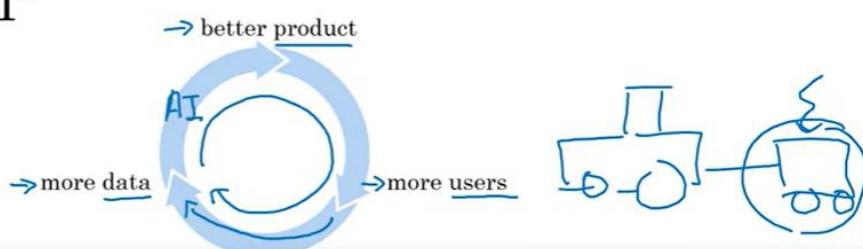
- Roles and what they should learn

AI Transformation Playbook (Part 2)

4. *Develop an AI Strategy – Part 1*

### 4. Develop an AI strategy

- Leverage AI to create an advantage specific to your industry sector
- Design strategy aligned with the “Virtuous Cycle of AI”



- Leverage AI to create an advantage specific to your industry sector.
- Design strategy aligned with the “Virtuous Cycle of AI”

*4. Develop an AI Strategy – Part 2*

## 4. Develop an AI strategy

- Consider creating a data strategy
    - Strategic data acquisition
    - Unified data warehouse
  - Create network effects and platform advantages
    - In industries with “winner take all” dynamics, AI can be an accelerator
- 
- Consider creating data strategy
    - Strategic data acquisition
    - Unified data warehouse
  - Create network effects and platform advantages
    - -in industries with “winner take all” dynamics, AI can be an accelerator

*5. Develop internal and external communications*

## 5. Develop internal and external communications

- Investor relations
- Government relations
- Consumer/user education
- Talent/recruitment
- Internal communications

Detailed AI Transformation Playbook: <https://landing.ai/ai-transformation-playbook/>

- Investor relations
- Government relation
  - Healthcare is highly regulated.
  - Self-driving cars
- Consumer / user education
- Talent / recruitment
- Internal communications

AI pitfalls to avoid

*Do and Donts*

## AI pitfalls to avoid

### Don't:

- Expect AI to solve everything
- Hire 2-3 ML engineers and count solely on them to come up with use cases

### Do:

- Be realistic about what AI can and cannot do given limitations of technology, data, and engineering resources
- Pair engineering talent with business talent and work cross-functionally to find feasible and valuable projects

## AI pitfalls to avoid

### Don't:

- Expect the AI project to work the first time
- Expect traditional planning processes to apply without changes
- Think you need superstar AI engineers before you can do anything

### Do:

- Plan for AI development to be an iterative process, with multiple attempts needed to succeed
- Work with AI team to establish timeline estimates, milestones, KPIs, etc.
- Keep building the team, but get going with the team you have

---

Taking your first step in AI

Some initial steps you can take

# Some initial steps you can take

- Get friends to learn about AI
    - This course
    - Reading group
  - Start brainstorming projects
    - No project is too small
  - Hire a few ML/DS people to help
  - Hire or appoint an AI leader (VP AI, CAIO, etc.)
  - Discuss with CEO/Board possibilities of AI Transformation
    - Will your company be much more valuable and/or more effective if it were good at AI?
- 
- Get friends to learn about AI
    - This course
    - Reading groups
  - Strat brainstorming projects
    - -No project is too small
  - Hire a few ML/DS people to help.
  - Hire or appronment an AI Leader (VP AI, CAIO)
  - Discuss wth CEO/Broad possibilities of AI Transformation

## Week 3 Quiz

Menu > Week 3 > Week 3 Quiz < Previous Next >

Week 3 Quiz

Submit your assignment  
Due Oct 18, 11:59 PM +08 Attempts 3 every 6 hours Start assignment

Receive grade  
To Pass 80% or higher

Your grade -

Like Dislike Report an issue

1. Because smart speakers can carry out multiple functions (such as tell a joke, play music, etc.) it is an example of Artificial General Intelligence (AGI).

1 / 1 point

- True  
 False

**Correct**

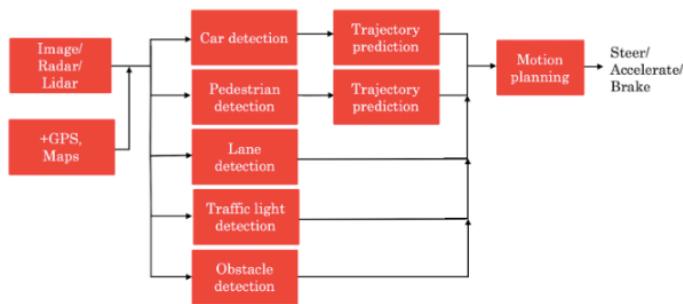
2. What are the key steps to a smart speaker function?

1 / 1 point

- Trigger detection -> intent recognition -> speech recognition -> command execution.  
 Speech recognition → Trigger word detection -> intent recognition -> command execution.  
 Trigger word detection -> intent recognition -> speech recognition -> command execution.  
 Trigger word detection -> speech recognition -> intent recognition -> command execution.

**Correct**

1.



The component for pedestrian detection is usually built using:

- A motion planning algorithm  
 Supervised learning  
 GANs  
 Reinforcement learning

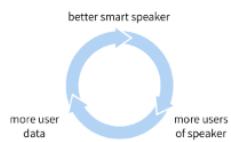
**Incorrect**

2.

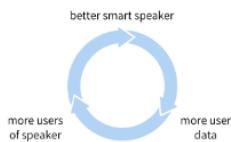
7. Say you are building a smart speaker, and want to accumulate data for your product through having many users. Which of these represents the "Virtuous circle of AI" for this product?

1 / 1 point

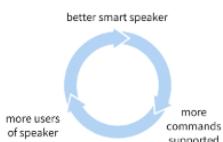
(A)



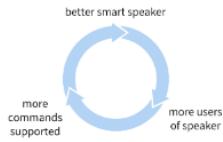
(B)



(C)



(D)



**Correct**

3.

8. Why is developing an AI strategy NOT the first step in the AI Transformation Playbook?

0 / 1 point

- The strategy should be to use the Virtuous Circle of AI, which comes after building a product.
- Without having some practical AI experience and knowing what it feels like to build an AI project, a company usually does not know enough to formulate a sound strategy.
- When transforming a company into an AI company, one does not need a strategy, therefore it can't be the first step.
- There is no reason. Developing an AI strategy IS the first step in the AI Transformation Playbook.

**Incorrect**

4.

8. Why is developing an AI strategy NOT the first step in the AI Transformation Playbook?

0 / 1 point

- The strategy should be to use the Virtuous Circle of AI, which comes after building a product.
- Without having some practical AI experience and knowing what it feels like to build an AI project, a company usually does not know enough to formulate a sound strategy.
- When transforming a company into an AI company, one does not need a strategy, therefore it can't be the first step.
- There is no reason. Developing an AI strategy IS the first step in the AI Transformation Playbook.

 **Incorrect**

9. According to the AI Transformation Playbook, broad AI training needs to be provided not only to engineers, but also to executives/senior business leaders and to leaders of divisions working on AI projects.

1 / 1 point

- True
- False

 **Correct**

5.

10. Which of the following are AI pitfalls to avoid? (Select all that apply)

0.75 / 1 point

- Expecting AI based projects to work the first time
-  **Correct**
- Expecting traditional planning processes to apply without changes
-  **Correct**
- Pairing engineering talent with business talent to identify feasible and valuable projects.

 **This should not be selected**

- Expecting AI to solve everything
-  **Correct**

6.

Week 4

## AI and society

- AI and hype
  - Limitations of AI
    - Bias
    - Adversarial attacks
  - AI, developing economies, and jobs
- 
- AI and hype
  - Limitations
    - Bias
    - Adversarial attacks
  - AI, developing economies and Jobs

A realistic view of AI

*Goldilocks rule for AI*

- ## Goldilocks rule for AI
- Too optimistic: Sentient / super-intelligent AI killer robots coming soon
  - Too pessimistic: AI cannot do everything, so an AI winter is coming
  - Just right: AI can't do everything, but will transform industries
- Too Optimistic
  - Too Pessimistic
  - Just right

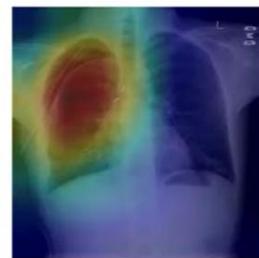
*Limitations of AI – Part 1*

## Limitations of AI

- Performance limitations
- Explainability is hard (but sometimes doable)



Right-sided  
Pneumothorax  
(collapsed lung)



[Rajpurkar et al. (2018). CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning.]  
[Wang et al. (2017). ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases. IEEE CVPR]  
[Images source: NIH Clinical Center Image dataset: <https://nihcc.app.box.com/v/ChestXray-NIHCC>]

- Performance limitations
- Explicability is hard (but sometime doable)

*Limitations of AI – Part 2*

## Limitations of AI

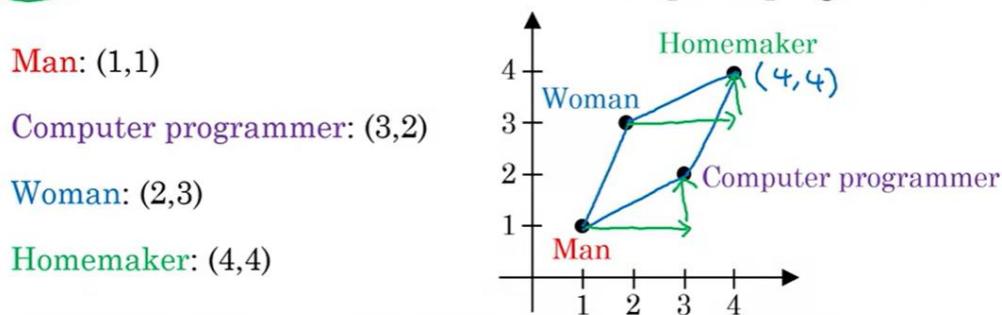
- Biased AI through biased data
- Adversarial attacks on AI

- Biased AI through biased data
- Adversarial attack on AI

Discrimination / Bias

## AI learning unhealthy stereotypes

- Man : Woman as Father : Mother
- Man : Woman as King : Queen
- Man : Computer programmer as Woman : ~~Homemaker~~  
Computer programmer



[Bolukbasi et al. (2016). Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings.]

- Man: woman as father?
- Man: woman as king: queen
- Man: Computer programmer as : ~~Homemaker~~ (should be Computer programmer)
- Man (1,1) – through stats
- Computer programmer (3,2)
- Women (2,3)

### Why bias matters

- Hiring tool that discriminated against women
  - Facial recognition working better for light-skinned than dark-skinned individuals
  - Bank loan approvals
  - Toxic effect of reinforcing unhealthy stereotypes
- 
- Hiring tool that discriminate against company
  - Face recognition
  - Bank loan approvals
  - Toxic effect of reinforcing unhealthy stereotypes

*Combating bias*

## Combating bias

- Technical solutions:
    - E.g., “zero out” the bias in words
    - Use less biased and/or more inclusive data
  - Transparency and/or auditing processes
  - Diverse workforce
    - Creates less biased applications
- 
- Technical solutions
    - E.g. “zero out”
    - User less biased and / or more inclusive
  - Transparency and/or auditing
  - Diverse workforce

Adversarial attacks on AI

## Adversarial attacks on AI



Hummingbird

Minor perturbation →



Hammer



Hare

Minor perturbation →



Desk

- Hummingbird => Hammer
- Hare => Desk

*Physical attacks*

- A different person

- Physical attacks



“Milla Jovovich”

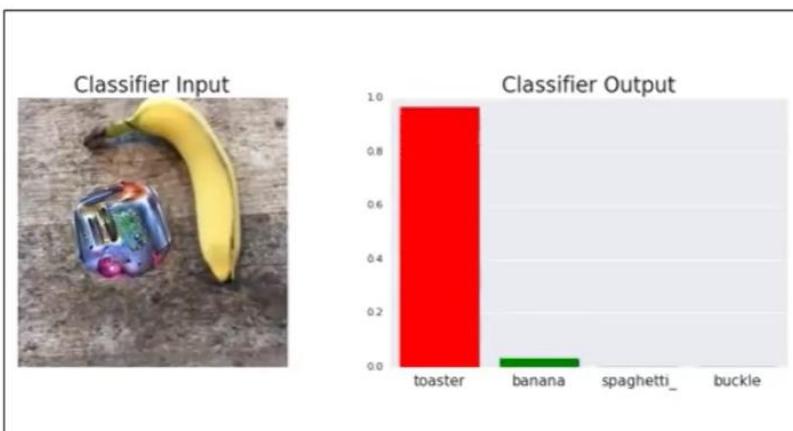


Fails to see stop sign



Banana

## Physical attacks



[Sharif et al. (2016). Accessorize to a Crime: Real and Stealthy Attacks on State-of-the-Art Face Recognition.]

[Eykholt et al. (2018). Physical Adversarial Examples for Object Detectors.]

[Brown et al. (2018). Adversarial Patch.]

### Adversarial defenses

- Defense do exist, but incur some cost
- Similar to spam vs anti-spam

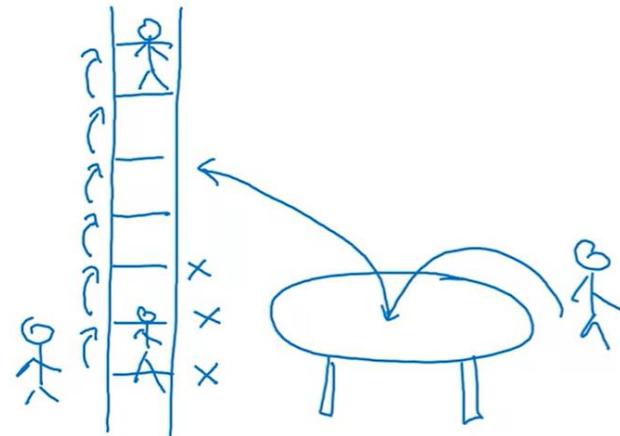
### Adverse uses of AI

# Adverse uses of AI

- DeepFakes
  - Synthesize video of people doing things they never did
- Undermining of democracy and privacy
  - Oppressive surveillance
- Generating fake comments
- Spam vs. anti-spam and fraud vs. anti-fraud
  - Deepfakes
  - Undermining of democracy and privacy
  - Generative Fake comments
  - Spam vs anti-spam

AI and developing economies

*Developing economies*



“Leapfrog”  
-Mobile phones  
-Mobile payments  
-Online education

*How developing economies can build AI*

## How developing economies can build AI

- US and China are leading, but all AI communities are still immature
  - Focus on AI to strengthen a country's vertical industries
  - Public-private partnerships to accelerate development
  - Invest in education
- 
- US and China
  - Still immature
  - Focus on AI
  - Public-private partnership to accelerate development
  - Invest in education

AI and jobs

*AI's impact on job worldwide*

## AI's impact on jobs worldwide

Jobs displaced  
by 2030

**400-800 mil**

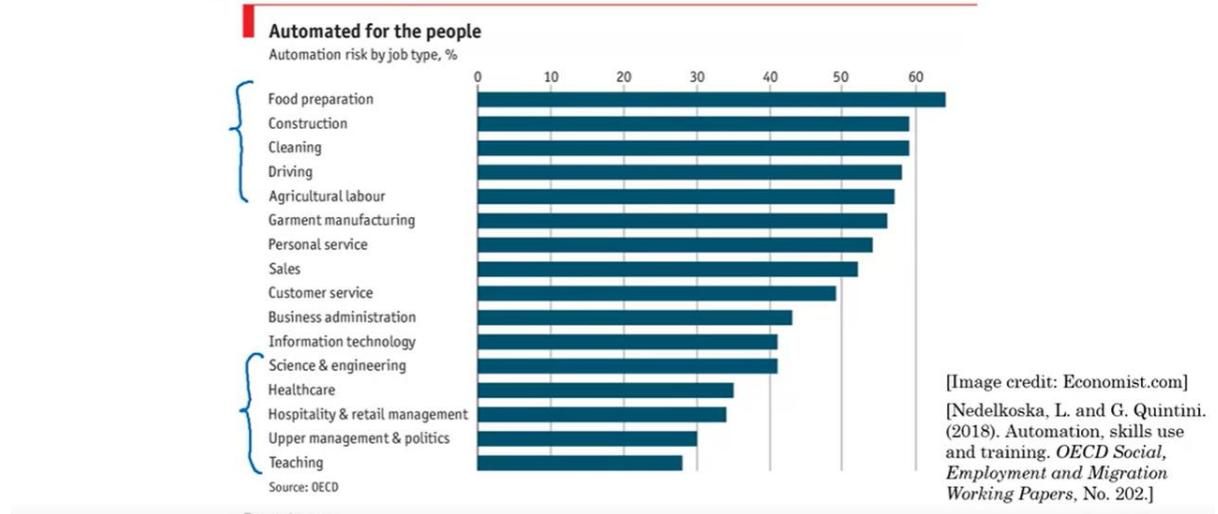
Jobs created  
by 2030

**555-890 mil**

- Jobs displaced by 2030 – 400-800 M
- Jobs created by 2030 – 555 – 890 M

*AI's Impact on jobs worldwide*

# AI's impact on jobs worldwide



## Some solutions

- Conditional basic income: provide a safety net but incentivize learning
- Lifelong learning
- Political solutions
  - Conditional basic income
  - Lifelong learning
  - Political solutions

## Conclusion

### What you have learned?

- What is AI?
- Build AI Projects
- Building AI in your company
- AI and society
- Keep learning.

## Week 4 Quiz

⚠ Try again once you are ready

Grade received **75%**

Latest Submission Grade 75%

To pass 80% or higher

Try again

1. What are the current limitations of AI technology? (Select all that apply)

1 / 1 point

Explainability is hard

 Correct

AI technology can be biased

 Correct

AI technology can discriminate

 Correct

AI technology is susceptible to adversarial attacks

 Correct

There are no limitations to AI technology

2. What is the Goldilocks Rule of AI?

1 / 1 point

One shouldn't be too optimistic or too pessimistic about AI technology

AI's technology will continue to grow and can only benefit society

One should allocate many resources to defend the world from giant killer robots

An AI winter is coming

 Correct

3. Say you are building an AI system to help make diagnoses from X-ray scans. Which of the following statements about explainability of AI do you agree with?

0 / 1 point

1.

- One shouldn't be too optimistic or too pessimistic about AI technology
- AI's technology will continue to grow and can only benefit society
- One should allocate many resources to defend the world from giant killer robots
- An AI winter is coming

 **Correct**

3. Say you are building an AI system to help make diagnoses from X-ray scans. Which of the following statements about explainability of AI do you agree with?

0 / 1 point

- AI systems are intrinsically "black box" and cannot give any explanation for their outputs.
- Explainability is usually achieved through building a chatbot to talk to the user to explain its outputs.
- Lack of explainability can hamper users' willingness to trust and adopt an AI system.
- Most AI systems are highly explainable, meaning that it's easy for a doctor to figure out why an AI system gave a particular diagnosis.

 **Incorrect**

4. Using current AI technology, if a machine learning system learns from text that reflects unhealthy biases/stereotypes, then the resulting AI software may also exhibit similarly unhealthy biases/stereotypes.

1 / 1 point

- True
- False

 **Correct**

5. Using current AI technology, if a machine learning system learns only from text that is completely neutral and does not reflect any gender biases, then we would expect it to exhibit no, or at most minimal, gender bias.

0 / 1 point

- True
- False

 **Incorrect**

2.

Incorrect

6. Which of these are good practices for addressing bias in AI? (Select all that apply)

0.75 / 1 point

Systematic auditing processes to check for bias

Correct

Technical solution such as "zeroing out" bias

Correct

Using an adversarial attack on the AI system to change its outputs to be less biased

This should not be selected

Using more inclusive/less biased data

Correct

7. Which of these are examples of adversarial attacks on an AI system? (Select all that apply)

0.75 / 1 point

Subtly changing an image to make an AI system mistakenly recognize a dog as a cat.

Correct

Adding a sticker to a stop sign to make an AI system fail to detect it.

Correct

Using AI to synthesize a fake video of a politician saying something they never actually said.

This should not be selected

Subtly modifying an audio clip to make a speech recognition system think someone said "Yes, authorized" when they actually said "No, reject."

Correct

3.

 **Correct**

Using AI to synthesize a fake video of a politician saying something they never actually said.

 **This should not be selected**

Subtly modifying an audio clip to make a speech recognition system think someone said "Yes, authorized" when they actually said "No, reject."

 **Correct**

8. If a developing economy has a strong and thriving coffee bean manufacturing industry (or some other vertical industry), then it has an advantage in applying AI to coffee bean manufacturing (or other vertical industry).

**1 / 1 point**

True

False

 **Correct**

9. What are the jobs that AI is most likely to displace over the next several years?

**1 / 1 point**

All jobs will be displaced

Most jobs involving office work (white collar jobs)

Jobs that comprise primarily of routine, repetitive work

Jobs that comprise primarily of non-routine, non-repetitive work

 **Correct**

10. Congratulations! You deserve a pat on the back for finishing this course.

**1 / 1 point**

True

False

 **Correct**

## References:

### 1. detailed ai transformation playbook

[LandingAI\\_Transformation\\_Playbook\\_11-19.pdf](#)