**Artificial Intelligence – NOTES**

When we think of AI we think of machines exhibiting behaviors that we associate with humans, like perception

Object Detection/Recognition

Language Translation

**\*SLIDE: What AI can and cannot do**

**What AI can do**:

Anything you can do without much thought, AI can do.

**What AI cannot do**:

It can’t respond by outputting a complex piece of text in context or by empathizing.

**\*SLIDE:** What makes an ML/AI problem easier (more feasible)

1. Learning a “simple” concept (doesn’t require much thought)

-doesn’t do well with subtle ambiguities

2. Lots of (Supervised) data available

**\*SLIDE**: **Table of AI applications**

\***SLIDE:**

Understanding: “Is that your computer?” “Isn’t that your computer?”

**\*SLIDE:** ML requires lots of data

Neural networks and deep learning don’t learn like a human. Humans don’t need thousands of examples to learn.

Recognizing pneumonia from image scans:

A young doctor (humans) can probably look at 10-15 images and understand how to diagnose pneumonia from an image. ML would need like 10,000 images to learn reliably.

**\*SLIDE:** ML performs poorly on new types of data

 Neural Networks: They can’t reason like humans do, they are terrible at applying their knowledge to new areas.

-bone scans of rheumatoid arthritis

-hand radiography/x-ray (left) and bone scintigraphy scans (middle and right)

-different types of images

**\*SLIDE:** new facility scans

Learning on specific types of scans and then going perhaps to a different facility which scans a little different will cause an AI system to perform poorly.

-humans are much better at adapting/transferring their learning from one context to a similar, but not exact, context.

-An AI system is much less robust than humans with generalizing their learning, or figuring out what to do with new data.

-I hope these give you some examples of what might be feasible and worth exploring.

**Deep Learning**

A lot of what Deep learning enables is perception.

-recognizing images

-recognizing what people are saying

Deep learning is a central tool in solving perception problems

\*It’s state of the art on everything having to do with Computer Vision and Speech Recognition

-it’s also very good at discovering new medicines, understanding natural language, and understanding documents.

-Deep Learning shines where there’s lots of data and complex problems to solve

\*Deep Learning is really a family of techniques that adapts to all sorts of data and all sorts of problems

**Neural Network**

**\*SLIDE: On-site Interview using a Neural Network**

Non-technical explanation of an Artificial Neural Network

-A artificial neuron or computational unit

You give it the input (A) and the output (B) and it figures out everything in the middle by itself.

Neurons can only learn simple functions, but stacked together (like Lego bricks) they can compute incredible complicated functions (learn incredibly complicated things) that give you very accurate mappings from the input to the output.

**\*SLIDE:** summary of Neural Network

To build a Neural Network you would feed it lots of data (examples of input labeled with the correct output).

It’s the algorithm’s/software’s job to figure out what the neurons (in the green) should be computing, so that it can completely and automatically learn the most accurate function that maps the input to the output.

If you give it enough data and a large enough neural network it can do an incredibly good job of mapping the input (features) to the output (labels).

**\*SLIDE:** Deep Neural Network

**-----**

**Differentiating cat vs. dog**

With machine learning we would have to define the features, such whiskers, and let our system identify which features are more important in classifying a particular animal.

With deep learning, it automatically finds which features are most important for classifying the animals.

**-----**

**Deep Learning**

-unstructured data

**\*SLIDE: Convolutional Neural Networks**

-they share their parameters across space

**\*SLIDE: Image**

-width, height, depth (RGB)

-that’s our input

**\*SLIDE: Convolutional pyramid**

Now imagine taking a small patch of this image

The convolutions will form kind of a pyramid, with the initial (shallow) image on the bottom.

-you will then apply convolutions that are going to progressively squeeze the spatial dimensions while increasing the depth.

-at the top is the classifier. You have a representation where all of the spatial information has been squeezed out, and only parameters that map to content of the image remain.

-so that’s the general idea

**\*SLIDE: CNN to fully connected layer**

Stack up your convolutions, reducing the dimensionality, and increasing the depth of your network layer after layer

-once you have a deep and narrow representation, you can then connect the whole thing to a normal (fully connected) neural network

-Now you’re ready to train your classifier

**What AI can do**

**COMPUTER VISION**

One of the major successes of Deep Learning has been **Computer Vision**.

-such as image classification and object recognition

-being able to take an image and classify what’s in it

**Face Recognition**

**\*SLIDE**: **Face ID**

Machines are great at face recognition, often better than humans

**\*SLIDE**: **recognizing the same face**

- Matching up photos of strangers’ faces is surprisingly difficult

-  You have a picture of a person and you just have to tell whether it’s the same person or not

- this is a really demanding task and there’s a lot of error.

- Numerous studies have shown that people will be wrong 10 to 30 percent of the time when asked to determine whether two photos of similar-looking strangers are the same person

- face matching can be tricky, no matter how experienced you are.

**Image Classification/Object recognition**

**\*SLIDE:** Dog

**Object Detection**

**\*SLIDE**: **Object detection**

Instead of analyzing and classifying the whole image, it will tell us where in the image different objects are and what they are.

**Image Segmentation**

**\*SLIDE**: **Image Segmentation**

Image segmentation takes this one step further. Instead of just drawing rectangles around items, it draws precise boundaries. It tells us not just where things are, but also where the pixels are that belong to each item.

In a scan it scan segment out where the liver is, or the heart, or the bone in an image.

**\*SLIDE**: **sheep - Object Detection/Image Segmentation**

**Image Tracking**

**\*SLIDE**: **Image Tracking**

Shows how the algorithm is tracking different people and helps the computer figure out where things are going

-These are some of the major areas of computer vision.

-perhaps some of them will be useful foer some of your projects

**NATURAL LANGUAGE PROCESSING**

**\*SLIDE**: **Natural Language Processing**

AI and Deep Learning in particular is also making a lot of progress in Natural Language Processing

NLP refers to AI understanding natural language

One example:

**\*SLIDE**: Text classification

- Is email Spam or Not

**\*SLIDE**: Sentiment recognition

From an email, tweet or Yelp review, the sentiment recognition algorithm should be able to tell whether the sentiment is positive or negative.

**\*SLIDE**: Information retrieval

For web search, you type in a text query and you want the AI to find relevant documents

**\*SLIDE**: Machine translation

-translate one language into another.

\*These are the major categories of useful NLP applications.

**SPEECH**

**\*SLIDE**: **Speech recognition**

-Speech-to-text

**\*SLIDE**: **Speaker ID**

-Listening to someone speak and figuring out the identity of the speaker

-just as face recognition helps to verify your identity through a picture, speaker ID can also help verify your identity by listening to you speak

**\*SLIDE**: **Speech synthesis** (**Text-to-speech)**

-taking a sentence written in text and turning that into an audio file

-TTS systems are increasingly sounding more natural and human-like

**AI Applications**

**----------------------------------**

**SLIDE**: AI Autopilots

-the average flight of a Boeing plane involves only seven minutes of human-steered flight, which is typically reserved only for takeoff and landing

**AgTech**

**\*SLIDE**: Ag

John Deere bought an AI startup in a quest to automate farming. (Many plants are becoming resistant to RoundUp.) It’s developing cutting-edge machine vision tools that help famers scan fields, assess crops, and get rid of weeds — all at the same time. Its technology makes for more efficient crop spraying.  It’s a set of cameras that fix onto crop sprayers and use deep learning to identify plants (similar to face recognition). If it sees a weed, it’ll hit it with pesticide; if it sees a crop, it’ll drop some fertilizer.

The See & Spray precision weed control technology allows Ag machines to see every plant in the field, and spray very precisely. Can tell the farmer how many weeds and what type of weeds exist in the field so that they can tailor their herbicide programs towards those weeds. They claim that you can save up to 90% of the amount of herbicide that you would spray if you were spraying the entire field.

**---------------------------**

**Fashion**

**\*SLIDE**: Fashion

Webshop visitors often abandon the site when an item is **out of stock, not in their size or** **it is not exactly what they were looking for.**This is when AI comes to the rescue.

The advantage of this kind of recommendations is that, unlike most AI-based solutions, it **doesn’t rely on behavioral or customer data**, as it focuses on the visual similarities of the item the person is looking at. The clothing recognition is carried out using computer vision, and visual features are extracted.

**\*SLIDE**: **Visual search**

Like “Shazam” for fashion.

Asos’ Style Match

**\*SLIDE**: **Visual search**

- the feature lets users take or upload a photo to the app to find visually similar clothing styles that are available to shop.

-it can distinguish between different styles of clothing

It will return items in a similar style, not just similar colors and patterns

**\*SLIDE**: **StyleSnap**

Amazon just released Style Snap which enables you to upload a photograph or screenshot of an outfit that you like. StyleSnap then will present you with recommendations for similar items on Amazon that match the look in the photo. It uses ML to find similar clothes and styles.

**Face recognition and face search**

**\*SLIDE**: **Face Recognition**

Facebook’s facial recognition technology (DeepFace) is so effective that it can identify you in a photo even if it’s not tagged.

If you’re on Facebook, then there’s a model of your face for recognizing you in photos.

 You can opt-out of facial recognition, and Facebook says it will delete the face template (model) used to find (identify) you in photos.

If you’re in public and someone takes a photo, there’s a good chance that Facebook can identify you in it.

**\*SLIDE**: **Face Search**

A child lost in a mall…

-Instead of monitoring the videowall (cameras) and looking through video manually, AI enables us to search for a person across all selected camera streams in a fraction of a second, using only one photo of that person.

AI can find her and match her face with the hundreds or thousands of faces in the locations of interest. The search result is obtained in nearly real time as she passes by a certain camera.

-upload an image of the person of interest, the AI system identifies similar faces, and shows us where the person has appeared, and when.

**Face Recognition**: Match a face in a video stream to a database

**Face Search**: Match a face in a video stream to the target face in uploaded image

**\*SLIDE:** **Faces drawn from DNA**

Each rendering was created by plugging an individual genetic profile.

-predictive tool created by a professor of anthropology and genetics at Penn State University

-studied the ways that genes influence facial development.

-yields an image, drawing connections between genetic markers and points on the face.

-a sketch emerges inferred solely from DNA.

**What AI cannot do (yet)**

------------------------------

**\*SLIDE**:  **Understanding gestures**

**Fingers: peace, victory, two**

AI doesn’t do well with ambiguity and subtlety, nor things that change with context

**Self-driving car**

One of the most exciting and pursued products of the AI era is the self-driving car

Some of the biggest companies are working on this problem (Google, Tesla, Uber, Apple)

**\*SLIDE**: **Self-driving car**

A self-driving car has to be able to avoid other cars as well as pedestrians.

DETECTION:

-Both car detection and pedestrian detection can be done with ML with input to output mapping (takes as input pictures, and tells us where the other cars and pedestrians are)

MOTION PLANNING:

Detection influences motion planning, which translates into steering, accelerating, and braking to get the car to move in the planned direction at a desired speed.

Detection is fed into Motion Planning software that plans the desired path that the car will take, in order to make progress toward the destination without making any collisions.

**\*SLIDE**: Steps (simplified)

Car Detection and Pedestrian detection:

-uses computer vision (sensors) and supervised learning

- takes as input pictures like these and outputs the detected cars and pedestrians

-most self-driving cars have additional cameras on each side and a rear-facing camera so that it can detect cars all around it. It also uses detectors other than cameras.

Motion planning:

-tells you the path and speed that you should drive in order to follow the road and avoid the detected objects.

-Requires steering angle prediction and lane detection.

Additional input:

Cameras, Radar, GPS, Maps data

Trajectory detection, lane, traffic light and obstacle detection (traffic cones, debris)

**\*SLIDE**: **UBER autonomous vehicle**

**\*SLIDE**: **Toyota autonomous vehicle**

**\*SLIDE**: **Waymo’s system (a Level 4 autonomous system)**

- Waymo’s infrastructure includes a variety of sensor, radar, and camera systems.

**\*SLIDE**: **LiDAR view**

(Light Detection and Ranging)

-uses a laser (light in the form of a pulsed laser) to measure ranges or distance to a target.

LiDAR creates a 3D map of the surroundings, identifying mobile and immobile objects, including other vehicles, cyclists, pedestrians, traffic lights, and a variety of road features.

-figuring out what’s in the world around you

-the system provides a 360 degree model of the vehicle’s environment

**\*SLIDE**: **Vision system view (object detection and tracking)**

-**Vehicle-to-Vehicle communication**: if a self-driving car encounters an accident or high-volume but slow-moving traffic, it is capable of relaying the information to other self-driving cars so that they can potentially avoid accidents and traffic.

++++

**\*SLIDE**: **Levels of Autonomy**

Automated vehicle technologies are still in the research and development phase.

This is how the government categorizes automated driving systems. Most in production systems today are at Level 1 or 2.

**Level 1**: the car controls the acceleration and braking while the driver would control the steering. Level 1 systems may warrant for full human control anytime.

**Level 2**: the car controls the acceleration, braking and steering, the driver must constantly monitor the system. Some Level 2 systems require that the driver ‘s hands be on the steering wheel to continuously operate.

**Level 3**: Level 3 systems allow the driver to perform other tasks (like texting or watching a movie) while the system takes control of most vehicle operations. However, the system requires driver intervention (within a limited time) for some operations (as specified by the vehicle manufacturer).

**Level 4**: This level supports self-driving with minimal driver intervention but it does so only in select, mapped locations called **geofenced areas**.  (Drivers might even be able to sleep.) (estimated by 2021)

- Some firms believe that level 3 and level 4 autonomy are too dangerous, because the hand-off from machine to human could be unpredictable and dangerous (switching from texting or watching a movie to swerving away from an accident might be an unrealistic expectation).

**Level 5**: No human intervention is required. (The driver becomes another passenger and won’t even be required to know how to drive.)

(estimated by 2025-2030)

-the result should be fewer cars on the road, fewer accidents and shorter traffic lights.

\*But, even with all of that, the issue with AI is the human actor, understanding human gestures and intent (the limiting issue so often for AI is understanding human intent)

**\*SLIDE**: **Stop vs. Wave gestures**

**\*SLIDE**: **Other gestures**

- It would be very hard to build an AI system to recognize human intention from gestures at a high level of accuracy needed for a task such as a self-driving car.

As long as humans are in the equation there area a lot of subtleties.

Self-driving cars will probably work best when all cars are self-driving.

Human Examples:

-someone putting makeup on or checking a text at a stop sign (we will assume the right of way)

-seeing children playing in the yard with a ball up the street. We can anticipate that the ball might go into the street.

DISCUSSION (?)

**How to choose an AI project**

**\*SLIDE**: the best projects are feasible and valuable

Select projects at the intersection of what AI can do and what’s interesting to you and/or (valuable to) your business.

**\*SLIDE**: **FEASIBLE**:

-is it a task that AI can do?

-are you collecting the data?

-how much data will you need?

**\*SLIDE**: **INTERESTING/VALUABLE**

-you can work with your company to brainstorm what might be valuable to the business (requires domain knowledge) that AI can do.

**BRAINSTORMING**: “adding intelligence”

-What are the main pain points in your business?

-some of them main be able to be solved with AI

-this can be a useful starting point for brainstorming AI projects

-What are the main drivers of business value?

-finding AI solutions that augment this can be very valuable

-lower costs

-increase revenue

-launch a new product or business

-Think about automating tasks

-think of all of the tasks that employees do, there may be one that is particularly amenable to ML automation; select the most fruitful option for ML automation

-resumes

-sentiment analysis for email routing

\*\*\* Are you collecting the data? Just start. Some of you may have heard of Big Data. How much data is needed is very problem dependent and some problems may only need a relatively few examples. Although, more data almost never hurts.

**\*SLIDE: Ethical**

-make sure that there’s no bias in resume selection

-recognizing race by voice is an ethical concern

**Case Study**

**Smart speaker**

**\*SLIDE**: Smart speaker (voice-activated device, or voice remote control)

**\*SLIDE**:

1. Activated by wake word

-Just an A to B mapping

- Once it hears the wake word: it outputs a 0/1

2. The software has to take the audio of what came after the wake word and map it to commands it understands using Machine Learning (Speech-to-Text algorithm)

-(e.g., “What is the weather tomorrow”, “Give me the sports scores”, “tell me a joke”, “Call Mom”).

3. Intent recognition

-the algorithm takes what you said and has to figure out what you want it to do

-it has a limited vocabulary (set of commands)

-it takes the string of words from the speech recognition algorithm and maps them to a given intent (out of 10 or 100 that it knows how to execute).

-several strings should be mapped to the same intent

-a well-designed intent recognition system should recognize most ways to ask for the time, etc.

4. Intent execution

**\*SLIDE**: **AI Pipeline**

-This process is called the AI Pipeline

-you have multiple components which process data one step after another.

-each step uses a different ML component

\* The first 2 steps are easily solved. The intent and execution can be more complex depending upon the application.

**Gait recognition**

**\*SLIDE**: Gait recognition

AI Pipeline: **Gait Analysis**

**Machine vs. Man**

While AI is trying to replicate our mental functions, hundreds of jobs such as drivers, radiologists and cashiers are facing substitution with machines in the next 5 years.

**Becoming Human**

In which domains is AI stronger than humans?

Main advantages of **Artificial Intelligence** over Natural Intelligence are:

* **Incredible speed**. While one doctor can make a diagnosis in ~10 minutes, AI system can make a million in the same time.
* **Continuous operation**. Computers don’t need vacations or sleep, they can work without stopping.
* **Less biased**. In realms of Law and Medicine it’s particularly important to have as much data as possible. Because AI can be trained on millions and billions of examples, they have more information to make an important decision.
* **Accuracy**. Whether it is a house or stock price prediction, AI systems can do the job much more precise.

Artificial Intelligence has significant advantages in many tasks.

Main advantages of **Natural Intelligence** over Artificial Intelligence are:

* **Generalizing knowledge**. Humans are much better at transferring previously learned knowledge to novel situations.
* **Multi-tasking**. Regular worker may have dozens of responsibilities, but to teach a computer even 1 task may take months.
* **Differentiating subtleties of context**. Humans can quickly take in a lot of information from multiple senses and differentiate subtleties in ambiguous contexts.
* **Complex movements.** After 60 years of research and development, the most advanced robots can’t compete with the mobility of even a 6 years old child.

**\*SLIDE: Build vs. Buy**

-projects can be in-house or outsourced

-Data Science projects are more often done in-house because they usually require day-to-day knowledge about the business

**\*SLIDE: Automated ML**

-ML/AI projects can be done using an automated service

For many projects ML/AI can be automated

-Common use cases should use pre-trained models

-(Recognizing a dog in a photo vs. your dog in a photo)

-use Google’s expertise

-For something very unique to your business that gives you a defensible advantage, you can use a custom model.

Tech companies are amassing wealth and power, creating a new order in business where the most valuable resource is no longer oil, but data.

Tech companies are destroying the economy, creating a downturn in business where the most valuable resource is no longer available.

Apple today introduced Core ML 3, the latest iteration of its machine learning model framework for iOS developers, bringing machine intelligence to smartphone apps. On-device machine learning is growing in popularity as a way to deploy quickly.

Apple today delayed its announcement of Core ML 3. After many delays, it still does not offer a machine learning model framework for iOS developers. Additionally, on-device machine learning has decreased in popularity.

**\*SLIDE: Custom vs. Pre-trained**

**Next Steps**

Think about ideas that you would like to implement or perhaps volunteer to help someone with their project to motivate your learning of many of these skills.

We’ll walk through how to do ML/AI and start solving some cool or interesting problems.

The goal is for you to be able to do ML/AI on your own problems.

Some initial steps you can take:

-get friends to learn with you (e.g., Python, ML)

-start brainstorming projects (no project is too small; start small and succeed)

**Next Meeting**

How many of you would like help on a project idea that you have?

How many of you would be interested in working with someone on their project idea?

Next time, be brave and share your project idea with the group

(Who will be willing next time?)

The goal is to encourage you to begin!

-Start on a small project to gain momentum.

-it is more important for the initial project to be successful than to be valuable.

-I hope that you will use ML/AI to build exciting and valuable projects either for yourself or for your company.

-Make life better for yourself and for others

Ask yourself, will you or your company be much more valuable and/or effective if you were good at ML/AI?

**Neural Networks – rules of thumb**

With Deep Learning, we can build extremely accurate predictive models.

**1. Number of Layers**: Start with two hidden layers

**2. Number of nodes (size) of intermediate layers**: a number from the geometric progression of 2, e.g., 4, 8, 16, 32, … . The **first layer should be around half of the number of input** data features. **The next layer size as half of the previous**.

**3. Number of nodes (size) of output layer for Classification:**If binary classification then the size is one. For a multi-class classifier, the size is the number of classes.

**4. Activation for intermediate layers**: Use relu activation.

**5. Activation for output layer:**Use sigmoid for binary classification, softmax for multi-class classifier, and linear for regression.

**6. Epochs**: Start with 20 to see if the model training shows decreasing loss and any improvement in accuracy. If there is no minimal success with 20 epochs, move on. If you get some minimal success, make epoch as 100.

**7. Batch size**: Choose the batch size from the geometric progression of 2. For unbalanced datasets have larger value, like 128, otherwise **start with 16**.