A.I. #210:

How AI is reshaping manufacturing industry

Wei Zhou Ph.D.

EWOL Presentation Manuscript

## February 19th 2019

## Haier- GE Appliance Park, Louisville KY

## Engineering Week of Learning

# Disclaimer

Before we head into our topics, I have a little disclaimer to make. Base of my experience, I had people asking me why we don’t have any hands on, technical practical example in this class.

Well, A.I.#210 is really an introductory level course, set up to provide an overview on this discipline and to demystify a lot of topics for general audience.

I really respect the fact that we are all from various backgrounds and I very value the interest from our community in this class.

In my other class M.L.#510, we will focus on practical and technical examples and we will have a lot of fun there.

# Self-Introduction by Wei Zhou (Sr. Data Scientist, Haier-GEA)

<Brief go through Timeline>

<1:40s>

Here is a timeline we will follow to for today’s class.

We will spend about 12 minutes on the Historical backgrounds of AI, some concepts, and important individuals.

In the second part, we will walk through 4 application examples of AI in real life.

We will have a 10-minute break after then. And I will throw in a video teaser for our M.L.#510 class.

After we come back, then we will spend about 20 minutes under the hood, looking into 4 major Deep learning neural networks and 2 how to examples from high level.

In the final section, we will talk about Uncertainty, Stability, and Failsafe issues in A.I.

Then we will spend 15 minutes to conclude today’s presentation. Basically , we will have a Q&A session, and I will then briefly talk about our A.I. efforts/implementation in GEA service technology. I will show you 2 flow charts of our systems in design.

I believe that all of you are sharing a genuine interest in advancing your understanding of computational mechanisms underlying thinking and intelligent behavior and their practical embodiment in machine form.

And we are going to adhere to this plan so we don’t run overtime.

<Show page **#2** Main title>

<30s>

Hi everyone again, my name Wei Zhou. I am coming to you from service technology to give this presentation.

It’s my honor and great pleasure today to host this class for my colleagues in appliance park.

This is the title of our class today, how AI is reshaping manufacturing industry. Despite how narrow the title may sound, we do have the repertoire to go through given the fact not everyone is familiar with the topic nor dealing with AI or deep learning techniques, or intense analytical studies on daily basis. So, we will stay on a high level and be inspirational.

<Show page **#3** Section 1>

<30s>

We’re/You are the generation of surfers riding on the fascinating wave of intellectual evolution, and (we) have wondered about the nature of mind over the centuries and, more particularly, (and we are) those who have been optimistic that (in our generation) we can learn new insights about thinking—that we can and likely will comprehend the machinery, the principles underlying thinking and intelligent behavior.

<Show page **#4** Important individuals/ ideas>

<2m:25s>

Speaking of its history, people have witnessed hundreds of years of thinking in reflection, but more recently in the 18th century from Julian Offray de la Mettrie, had a significant work “Man a Machine”, in which he expressed his strong belief that man, body and mind, worked like a machine. The first time in document, a human being firmly believe that a being can be a machine.

And then on to Charles Babbage, who built the first “Mechanical computing machine”, which resembles its basic architecture to a modern computer. It has separate data and program storage (sort of like “memory”), instructional operation, i.e. operation is instruction based, its control unit can make conditional “jumps” (like if, else if , else conditional operators), and it even has a separate I/O.

Then on to Alan Turing, John von Neumann, . In John von Neumann’s famous work “The computer and the brain”, his book discussed how the brain can be viewed as a computing machine. Though the work is speculative in nature, it discussed the differences between brains and computers of his day. Such as processing power, processing speed, parallelism,etc.,. Even today, these topics remain in leading research areas.

And then to Allen Newell and Herbert Simon. They are significant AI theorists of their days. They have proposed the idea to use production rules for simulating human’s problem-solving behavior. Their proposal laid down the bed stones for today’s AI research and directed its development framework.

And there is a larger upswing in interest over the past 60 years.

<Show page **#5** paradigm evolution from computability, decision theory >

<2m:00s>

In the summer of 1956, a defining summer study group at Dartmouth University, led by John McCarthy and attended by an extraordinary set of folks at the first time used and defined the phrase artificial intelligence. The proposal from the meeting is remarkable in its modernity, and could almost describe current research focus. It raised the level of abstraction of thinking away from the kinds of work that John von Neumann and others had been pursuing on optimization and action under uncertainty.

Here is the quote:

"Machine methods for forming abstractions from sensory and other data,"

"carrying out activities, which may be best described as self-improvement,"

"manipulating words according to the rules of reasoning and rules of conjecture,"

and "developing a theory of complexity for various aspects of intelligence."

<Show page #**6** paradigm shift into artificial intelligence>

<3m:10s>

Over the last 50 years, we see the major shift in paradigms.

A new paradigm was forming, explicitly branching AI away from the more numerical decision sciences and operations research…. …into a worldview that included valuable and novel focuses on high-level symbols, logical inference, and cognitive psychology…

Now the Computers are used to explore heuristic models of human cognition, as well as well-defined, structured, puzzle-like problems, game playing, and planning, and reasoning.

So in short, the new paradigm focus on Perception, Learning, Reasoning, and NLP.

So there is a major bifurcation, the new paradigm has totally different mission and vision. Let me replay this page.

<scroll back and re-display this page>

On the right, I used some clips from a 2016 sci-fi movie “Passengers”, I believe many of you have also watched the movie. I love movies, I watched this one with my wife for more than 4 times. It is a love story and I am deeply motivated by the plots, but more so by the background. This starship, named “Avalon”, its mission is to carry 5000 passengers in hibernation state, travel over 120 years from earth to a far-away planet called “Homestead II”.

Even though a fictional plot, the starship “Avalon” presents a perfect example of A.I., which it not only maintains life support for all passengers on-board for 120 years AS A SUPPORTING MACHINE, but also prepare and self-defend alone the trajectory under various scenarios AS A HUMAN COMPANION. As shown, the spaceship has a very long protruding structure on the front, forms a plasma-type energy shield to smash and vaporize incoming meteorites. When it runs through a dense meteorite belt, it foresees this and gradually borrows energy from all 3-major passenger pillars together with its fusion drive to reinforce the protection.

This example exactly tells how the mission and vision of A.I. differ from other paradigm.

<Show page **#7** major areas of efforts>

<2m:00s ~>

Over the last 6 decades, many subfields have greatly evolved often with the whole communities. Lots of efforts going on that make up the broader tapestry of what we today call “The Artificial Intelligence” research.

Focus:

Difference between Speech recognition and NLP.

Deep learning application is a big leap forward in Learning effort.

Planning is based on decision science and dynamic programming, reinforcement learning, etc.

.

<Show page **#8** Section 1 Quiz>

<3m:00s ~>

<Show page **#9** Section 2: Real Life and Industry Applications>

<:15s ~>

In this Section, we will discuss 4 examples. We look into the recent progress in recognition technology.

In our examples, we look into how people put efforts to bring recognition technology into industrial production and how these products impact our day-to-day life.

<Show page **#10** Voice Recognition Splash Page>

<:20s ~>

Let’s look into some recent progress in Speech Recognition.

We see a large upswing over the last 2 decades, in our ability to study perception and reasoning with the growth in computing applications, which today we call “deep learning” applications in the “Learning” effort.

<Show page **#11** Switchboard contest>

<1m:15s ~>

Let’s take a close look at some changes and progress. You can see here, speech recognition error rate on a very famous challenge problem actually several problems here. Say for example, the **switchboard challenge**, you see, for quite a while, around 2000, its hard to get word error rate down, and they could not move forward but plateau for quite a few years. Then with the growth of deep learning applications after 2012, major research groups are able to making some remarkable reductions in error rates on conversational transcriptions.

<Show page **#12** Speech Recognition, it actually outperformed human>

<1m:30s ~>

Why use switchboard and what is it?

The standard test for measuring the accuracy of speech recognition uses a fairly old dataset called <NIST 2000 Switchboard>: recordings of telephone conversations between strangers, about common topics like sports and politics.

“This is the standard speech set used by all researchers for the last 20 or so years”. “It’s natural conversations that people had when they talked over the phone. They’re not talking to their own family, so they’re using standard English.”

It is hard to compare the performance among different companies. Like say, Google uses a different type of sample set.

Looking at the chart, after a large progress in deep learning applications in 2012, Microsoft and IBM had very remarkable reduction in their word error rate. In 2017, Google’s team announced that they had a 4.9% word error rate. It is lower than human word error rate at 5.3%.

<Show page **#13** Safety Systems that Save Lives SPLASH Page>

<:45s ~>

The latest car-safety technologies have incorporated recognition and tracking technology in the control system. And these technologies have the potential to significantly reduce crashes.

In this part, we discuss how these systems can help keep us on the right track.

The motion pictures on this page give us some ideas. A moving car using camera. The car sees the road and draws marking to show lane boundaries.

<Show page **#14** Two Major Active Safety Systems>

<:25s ~>

Cars with active safety systems nowadays can help us avoid or mitigate a crash in all kinds of situations. There are two active safety systems in our examples. Lane departure warning system and Lane Keep Assist system.

<Show page **#15** LDW Systems>

<1m:15s ~>

Lane departure warning system actively tracks lane boundaries and sends warnings, say a steering wheel vibration, or seat vibration, to the driver when the car crosses lane boundary.

The 3 major features are, continuous monitoring, realtime warning, and some systems watches the drivers’ facial movement and gives warning if it detects that the driver is not paying attention to the road.

There is a balance hard to maintain. According to “Consumer Report”, 12% of the users reported the system avoided at least one crash, and 80% of the users strongly favors the system. However, 41% of users reported at least one false warning, and 19% finds the system annoying.

<Show page **#16** LKA Systems>

<:25s ~>

As a step forward, when you drift away from your lane, lane-keeping assist (LKA) provides steering input or braking to correct the vehicle if it starts to exit the lane.

It also does active monitoring, active steering, and provide correction when oversteering occurs.

According to “Consumer Report”, 7% of the users reported the system avoided at least one crash, and 70% of the users strongly favors the system. However, 9% of users think it is incorrect, and 30% finds the system annoying.

<Show page **#17** Autonomous Vehicle Splash Page >

<:25s ~>

In this part, let’s discuss Autonomous Vehicle, or Driverless car. It is a very complicated system which usually combines

1. Visual recognition/tracking system
2. Speed sensing system
3. Temporal and Spatial Planning
4. Very deep Learning models with numerous rules

<Show page **#18** SAE’s Categorization of Autonomous Vehicle >

<2:25s ~>

Before we jump into an example, let’s first look at some interesting facts.

The SAE International standard that defines the six levels of driving automation. It is a standard adopted by the U.S. Dept. of Transportation and the National Highway Traffic Safety Administration.

**Level 0 (No automation):** This is where the vast majority of cars and trucks are today. The driver handles steering, throttle, and braking (ST&B) monitoring the surroundings, as well as navigating, and determining when to use turn signals, change lanes, and turn. But there can be some warning systems (blind-spot and collision warnings).

**Level 1 (Driver assistance):** Vehicles in this level can handle S or T&B, but not in all circumstances, and the driver must be ready to take over those functions if called upon by the vehicle. That means the driver must remain aware of what the car is doing and be ready to step in if needed.

**Level 2 (Partial assistance):** The car handles ST&B, but immediately lets the driver take over if he detects objects and events the car is not responding to. In these first three levels, the driver is responsible for monitoring the surroundings, traffic, weather, and road conditions.

**Level 3 (Conditional assistance):**The car monitors surroundings and takes care of all ST&B in certain environments, such as freeways. But the driver must be ready to intervene if the car requests it.

**Level 4 (High automation):** The car handles ST&B and monitoring the surroundings in a wider range of environments, but not all, such as severe weather. The driver switches on the automatic driving only when it is safe to do so. After that, the driver is not required.

**Level 5 (Full automation):** Driver only has to set the destination and start the car, the car handles all other tasks. The car can drive to any legal destination and make its own decisions on the way.

<Show page **#19** Waymo Driverless Car, a Google Company >

Ready to playback 3 video clips after brief introduction of Waymo.

<4:25s ~>

Waymo Driverless Car, it’s a Google company since 2009. Waymo started letting its (Level 4) driverless car in the streets of many cities since 2017. Meaning that these vehicles have no human sitting inside and they are running on the streets.

Let’s take a look at how it works.

As the system drives, this self-driving car uses LIDAR, which sends out millions of laser beams per second. It builds up a detail pictures of world all 360 degrees around it. It also uses short wave radar to detect how far away objects are and their relative speed. And it also uses  high definition cameras to track traffic visual information, like whether the traffic light is red, yellow, or green, or whether there is a different traffic pattern needs to be followed, say, a road block or a repair is right on front. It then combines all that data to understand the world around it . For example, in this split of a second, it knows exactly where it is on the road, it can also identify everything around it in full 360 degrees, then predict what those things might do next. And it doesn't just do that for the objects that you and I can see.

The control system actually has much deeper understanding of the situation, it take all the information the car have seen over time for hundreds of thousands of pedestrians, cyclist, and vehicles that being out there and understand what they look like and use that to infer what other pedestrian look like, and what other vehicles should look like. And then even more importantly, the system can take from a model of how it expects them to move through the world so here the yellow box is a pedestrian crossing in front of this self-driving car, here the blue box is the cyclist the system anticipate they are going to nudge out around the car on the right. here this is cyclist coming down the road, the system knows they are going to continue going down the shape of the road. Here, somebody makes a right turn, and here somebody is making a U turn in front of this self-driving car, and the system anticipate that and response safely.

Here is an example, the system is waiting on a group of cyclists and you probably have never seen and would never expect to see anywhere around you. And here is an example of the system has to deal with drivers, even though they are the small ones. Watch on the right, a person jumped out from the truck at us. And watch the left, a car decides to make a right turn on the last possible moment. Here is an example system makes lane change and the car on the left wants to make that change as well. And here, on the left, the system sees a car run through a red light and yield to it. Similarly, a cyclist run through a red light as well and the system responded safely as well. And, there are people who do we don't know what sometimes on the road, like this car pulling in between two self-driving cars.

<Show page **#20** Quiz Time 3-min>

<3m:00s ~>

<Show page **#21** Intermission >

<6m:00s ~>

<Show page **#22** Intermission Rollout>

<4m:00s ~>

<Show page **#23** Section 3. Neural Network and “How to” Examples >

<:30s ~>

In this section, we are going to be more practical and technical. Though we don’t plan to do any coding, we will walk through the high level view of 4 major deep learning models. We will discuss the terms frequently used and what they are using layman terms.

So we all can understand and some of you in future could brag about it to your friends when they are confused. And that is actually part of my goal.

<Show page **#24** Understand 4 major Neural Networks >

<1m:45s ~>

There are keywords frequently used in deep learning or machine learning, including artificial neural network, Feed forward, back Propagation, and Gradient Descent.

Let’s take a look at what they really are using layman terms.

Artificial neural network is actually a computational model that loosely resembles a biological neural network by concepts. The basic calculation units in Artificial neural network is called neuron and neurons consists of layers, weighted values are passing from neurons to layers, from inputs to outputs, which is called Feed forward.

Sometimes, a deep learning network has the ability to evaluate its output and determine that output is not ideal and it passes the signal backwards and advise the network to adjust the direction and magnitude to reevaluate the problem so that better results can be acquired. This process is called Back Propagation.

The action that the network adjusts the direction and magnitude is called gradient descent.

<Show page **#25** KNN >

<2m:25s ~>

**KNN** is a **non-parametric, lazy**learning algorithm.

KNN can be used for **classification** —An object is classified by a majority vote of its K number of neighbors, with the object being assigned to the class most common among its k nearest neighbors.

Simply put, for two classes, A stand for blue and B stand for red, for an unknown data, the majority class vote of the k nearest data points determines its class.

**Pros**:

No assumptions about data — useful, for example, for nonlinear data

Simple algorithm — to explain and understand/interpret

Versatile — useful for classification or regression

**Cons**:

Computationally expensive — because the algorithm stores all of the training data

Sensitive to outliers

<Show page **#25** DNN >

<2m:25s ~>

"Normal" neural networks usually have one to two hidden layers and are used for SUPERVISED prediction or classification.

Deep learning neural network architectures differ from "normal" neural networks because they have more hidden layers.

The network is a directed graph, meaning that each hidden unit is connected to many other hidden units below it. So each hidden layer going further into the network is a NON-LINEAR combination of the layers below it, because of all the combining and recombining of the outputs from all the previous units in combination with their activation functions.

When the OPTIMIZATION routine is applied to the network, each hidden layer then becomes an OPTIMALLY WEIGHTED, NON-LINEAR combination of the layer below it.

When each sequential hidden layer has less units than the one below it, each hidden layer becomes a LOWER DIMENSIONAL PROJECTION of the layer below it as well. So the information from the layer below is nicely summarized by a NON-LINEAR, OPTIMALLY WEIGHTED, LOWER DIMENSIONAL PROJECTION in each subsequent layer of the deep network.

<Show page **#26** CNN >

<2m:25s ~>

A convolutional neural network is a technique in computer vision to make the algorithm “see” the picture at a deeper level as a composition of various edges, lines, corners to identify the *contents* of the image.

A Convolutional Neural Network is a modification of the Neural Network algorithm that allows a network to search for shapes across a whole image with only a few learnable parameters. The network learns what shapes to look for.

every network layer acts as a detection filter for the presence of specific features or patterns present in the original data. The first layers in a CNN detect (large) features that can be recognized and interpreted relatively easy. Later layers detect increasingly  (smaller) features that are more abstract (and are usually present in many of the larger features detected by earlier layers). The last layer of the CNN is able to make an ultra-specific classification by combining all the specific features detected by the previous layers in the input data.

Convolutional neural networks have layers and layers of "filters". Each of these filters can only see small neighborhood at one time and gets excited when it sees a certain set of edges or shape of its liking, in that small neighborhood.

Now you can see how this next layer is using basic information from the previous layer to make more complex deduction. Since, this layer has access not-only what different filters are excited about in previous layer but also of much larger neighborhood.

The likes and dislikes of this layer will be fed to the next layer, and theirs to next one and so on. With each layer getting excited about more complex and much larger shapes.

<Show page **#26** RNN >

<2m:25s ~>

A recurrent neural network (RNN) is a class of artificial neural network where connections between units form a directed graph along a sequence. This allows it to exhibit dynamic temporal behavior for a time sequence. Unlike feedforward neural networks, RNNs can use their internal state (memory) to process sequences of inputs. This makes them applicable to tasks such as unsegmented, connected handwriting recognition or speech recognition.

Recurrent Neural Network comes into the picture when any model needs context to be able to provide the output based on the input.

Sometimes the context is the single most important thing for the model to predict the most appropriate output.

You can think of the hidden state  as the memory of the network.  captures information about what happened in all the previous time steps. The output at step  is calculated solely based on the memory at time . As briefly mentioned above, it’s a bit more complicated  in practice because  typically can’t capture information from too many time steps ago.

Unlike a traditional deep neural network, which uses different parameters at each layer, a RNN shares the same parameters (above) across all steps. This reflects the fact that we are performing the same task at each step, just with different inputs. This greatly reduces the total number of parameters we need to learn.

<Show page **#27** Voice and Visual Recog. How To Examples >

<2m:25s ~>

<Show page **#34** CNN For Visual Recognition>

<2m:25s ~>

Demonstrates a convolutional neural network (CNN) example with the use of convolution, ReLU activation, pooling and fully-connected functions.

The neural network consists of 3 convolution layers interspersed by ReLU activation and max pooling layers, followed by a fully-connected layer at the end. The input to the network is a 32x32 pixel color image, which will be classified into one of the 10 output classes.

<Show page **#35** CNN For Visual Recognition>

<2m:25s ~>

This is a different Visual Recognition training view. Different feature maps from hidden layers are shown.

And the final fully-connected layer gives classification results.

<Show page **#37** Frame problem for A.I.>

<0:25s ~>

We need to understand our limited systems is facing the large complexity of the open world.

There are several limitations in representation and reasoning. One is qualification, one is Ramification.

<Show page **#38** Know the unknowns>

<0:25s ~>

Scope? Model bank

Enough insight? Data? More insights,

Consideration of all possibilities? Extended real-world experiments

Behavioral constraints? Failsafe design

A.I. machines models with human intervention.