Benefits and Threads of Intelligent Technology

John Sum
Institute of Technology Management, National Chung Hsing Univesity,
Taichung, Taiwan. {pfsum@nchu.edu.tw}

***Abstract*—By surveying on the evolution of the intelligent technology from 1950s to 2010s, the benefits and threads of intelligent technology to every one of us are commented. With the advancement on the technologies for the development of system of higher level of intelligence, the demand for administrative and managerial staffs could be reduced. As a result, it could come to the end of the middle-level management.**

***Index Terms*—Evolution of Technology, Intelligent Technology, Level of Intelligence, Managerial Implications,**

I. Introduction

Evolution of intelligent technology has somehow depended on the evolution of technology. With the advancement of technologies, such as hardware and software technologies, intelligent technology could further been advancing and applying in our works and living. To have a better picture on the evolution of intelligent technology, it is better to have an overview on the evolution of technology around the invention of steam engine in the 18 century.

A. From the Renaissance to the Agricultural Revolution

During the period of **Renaissance**, learning and education had been re-advocated in Europe. Along with the invention of printing press, books were printed. Knowledge and ideas were disseminated rapidly. With an increasing number of studies on natural philosophy, new knowledge in science were aroused and led to the period of **scientific revolution** across Europe in the 16th century. While there were wars in the European continent which caused many casualties, UK on the other hand had a rather peaceful environment. Thus the population of UK grew relatively faster than the other European countries and led to a higher demand on food supply. In this regard, UK had laid new policies to boost the food supply and led to the **British agricultural revolution** in the 17th century in Edinburgh. **Tools and machines** were invented for the farmers and applied in agricultural production. It could thus be marked as the first technological revolution in this history. In this period of time, the major sources of power for driving automation were labors, animals and rivers.

B. Industry 1.0: Steam Engine

In the 17th century, Europeans switched their main source of fuel from wood to coal. Thus, coal mining became important and highly demanded across Europe. Many coal mines were deepened and became flood after penetrating underground water. Thus, powerful water pumps were needed. In the end, steam engine powered water pump was invented.

Making machines and tools are engineering tasks. They require scientific knowledge, a lot of trials and practises. Perhaps, it is one reason why the earliest steam engines were invented by Englishmen and led to the first industrial revolution. Steam engine is an important invention not just it brought to the world the power for automation. It could let a factory locate away from rivers (i.e. water power). Before the invention of steam engines, factories could still use machines and tools to enable their production. However, the factories had to be built along rivers. With steam engine, many factories could be built in other locations. No matter what, a key achievement brought out from the first industrial revolution is production automation. The number of labors was reduced. In addition, redesign production process and the use of new machine could further improve the product quality.

C. Industry 2.0: Electrification

The first industrial revolution marked the change on the source of power, from water power to steam engine. The second industrial revolution marked another major breakthrough – electric power is added on the list. With an increasing number of electricity suppliers and electric motors, together with mechanical engineer like Frederick W. Taylor and brilliant entrepreneur like Henry Ford, various production processes were re-designed based on the idea of assembly line which was driven by a large electric motor. Workers were standing side by side along the assembly line. Each worker only worked for a simple task. The workers were also grouped in various teams. Each team had a supervisor who in charge of the quality of the works in the team. With no doubt, the production cost down and hence the price of a product was then reduced. At the same time, the labors earned more.

D. Industry 3.0: Computerization

World War I and World War II marked another turning point in the history – the invention of electronic computers. To automate a computing task was not a new idea in the 20th century. In the 19th century, Charles Babbage had already designed and built a mechanical computing machine called 'Difference Engine' to perform numerical approximation[1]. With the invention of vacuum tubes, electronic components like diode and triode were made. Eventually, electronic digital computers Z2, Colossus and ENIAC were made and used during the second world war. During the war time, these computers were mainly used as computing machines for scientific researches. The purpose is essentially the same as the purpose of difference machine, for computation.

Until the war was over, digital computers for commercial use had finally been made, LEO from Lyon in England and later the UNIVAC from Remington Rand in US. These computers extended the functionalities of the earlier computers by adding functions, in forms of computer programs, to handle

The work presented in this paper is supported in part by research grants from the Ministry of Science and Technology (MOST) Taiwan numbering 105-2221-E-005-065-MY2, 108-2221-E-005-036, 109-2221-E-005-060 and 110-2221-E-005-053.

[1]https://en.wikipedia.org/wiki/Charles_Babbage.

data processing and data storage. These computer machines together with the corresponding programs formed the earliest *computerized information systems*. With the advancement of computer network technologies, the Internet and then the smartphone (like iPhone), the scale of an information system shifts from a single department-scale information system to a firm-scale network of multiple information systems; and then from a firm-scale enterprise information system to an information system connecting the firm, the suppliers and the customers.

E. Industry 4.0: AI and Others

In some countries, like German and US, have marked the middle of 2010s the starting of the forth industrial revolution. One reason is that AI technologies have been used in the industrial sector. Robotic receptionists have been installed in many service firms to handle customers enquires. By using intelligent software robot, customer service could now be implemented without hiring any human worker. Customer relationship management could be raised to the next level. Market trend could now be conducted and analyzed by an intelligent software program. New products to be developed could then be recommended to the management team. Moreover, marketing executives could use intelligent software to analyze the sentiment of a customer through the posts appeared on the Internet. The executives could also apply AI system to compile marketing strategy and make target marketing more efficient. There are a lot more jobs that AI could do, like production process design and scheduling. With auto-driving vehicles, the efficiency of logistic control could further be improved. Supply chain management could thus be advanced to the next level.

F. Levels of Intelligence

From the first industrial revolution to the forth industrial revolution, one can ready see that technological advancement has brought to the industry from routine labor work replacement to intelligent work replacement, see Table I. With the technological advancement on the production machines and the information systems, it is anticipated that the jobs to be done by human workers could largely be replaced.

TABLE I
CHANGES IN THE INDUSTRIAL WORKS

Stage	Job Replacement	Level of Intelligence
1st	Rowing sailor	Low
2nd	Production labor	Low
3rd	Production labor	Low
	Administrative staff	Middle
4th	Production labor	Low
	Administrative staff	Middle
	Marketing staff	High
	Customer service	High
	Order placement	Low
	Order fulfillment	Low
	Retailing	Low
	Shop management	High
X	Production labor	Low
	Administrative staff	Middle
	Marketing staff	High
	Customer service	High
	Order placement	Low
	Order fulfillment	Low
	Retailing	Low
	Shop management	High
	Product/Service design	Very high
	Production process design	Very high
	Market survey	Very high
	Research and development	Very high
	Scientific research	Very high
	Social science research	Very high
	Report writing	High
	Story composition	Very high
	Video/Film production	Very High
	Music composition	Very high
	Education	Very high

II. EVOLUTION OF INTELLIGENT TECHNOLOGIES

Making a computer to be intelligent has long been a fascinating topic to scholars. One renowned scholar and the pioneer of AI is definitely the English mathematician Alan Mathison Turing who laid the first reinforcement learning-based framework for learning machines and the proposal of Turing test to examine if a machine is intelligence [1], [2]. Subsequently, the quest of a learning machine has attracted a number of scholars to investigate and develop machines to realize the learning behind a machine [3]–[5]. Their very first purpose is to see if it is possible to make a machine thinks or behaves like human being. Since then, various researchers have joint the effort to investigate on the above issue. Here, I simply outline some major breakthroughs.

A. AI Research in the 1950s – 1960s

After electronic computer had been invented in the early 20 century and then commercial computers had been released in the 1950s, the use of a computer to realize human intelligence have been started. Two approaches to realize such purpose, namely the symbolic logic approach and biological brain approach.

1) Symbolic Logic (Psychological): One group of researchers followed the line of thought of Alan Turing and John von Neumann. We call this line of thought the symbolic logic approach. Every event and object in the world could be represented as a symbol. Each symbol can then be encoded as a binary code. Our human logical thinking is essentially a process of symbolic manipulation. If the process of symbolic manipulation is mapped to a program, a digital computer could mimic human logic thinking. In contemporary terminology, the system built by using this idea is called multi-agent system.

This approach has also attracted many psychologists who are interested in artificial intelligence, especially the behavioral psychologists and cognitive psychologists. They applied the theories to develop learning algorithms for the program to be running in a computer. Symbolic approach has then been one major approach in AI research. Fuzzy set theory, which was introduced by Lotfi Zadeh in 1965 [6], is one example along this approach. Later on, the research works on expert systems in 1970s-1980s, intelligent agent researches starting from 1980s [7] and case-based reasoning starting from 1990s are also along this approach.

2) Neural Network (Biological): In contrast to symbolic logic approach, another approach is to develop an intelligent system based on the finding from biological brain. One key finding in the 1940s was from Warren S McCulloch, Walter Pitts and their co-worker [8], [9] on modeling the behavior of a neuron as a threshold logic unit, which is currently called the McCulloch-Pitts neuron. Neural network approach is referred to the pioneer work by Frank Rosenblatt on his work on Perceptron in the 1950s [5], [10], [11]. Perceptron is a hardware machine mimicking human biological neuronal network with a learning mechanism.

B. AI Research in the 1960s – 1970s

This period of time is also called the AI winter. While Perceptron has drawn public attention in its initial release, it soon encountered its saliency period from 1960s to 1970s. Before the saliency period, neural network approach of learning algorithms development with applications had already attracted many scholars from electrical engineering. Many new models and learning algorithms were proposed and developed, together with vigorous mathematical analysis on the properties of those models and algorithms. These models and algorithms were largely be applied in signal processing and system control. During the saliency period, there were a number of researchers still working on neural network modeling while neural network project did not easily get funded. So, scholars from the major universities (eg. Stanford and UC Berkeley) avoided to put the terms, like 'intelligent', 'neural' and 'learning', on their project titles in order to secure funding from the National Science Foundation, USA. Instead, they use the term 'adaptive system' – the project has nothing to do with neural network.

C. AI Research in the 1970s – 1980s

In this saliency period, many scholars still developed new neural network models and learning algorithms [12]–[14]. Many of them were then applied to solve toy problems and some of them were applied to solve real problems. At the same period of time, some scholars worked on the theoretical aspects of neural networks and AI/ML [15]–[20]. Some scholars worked on a more general framework and theory for AI/ML and brain [21]–[24]. Some scholars investigated the relations between a neural network and a computer from the perspective of theoretical computer science [25]–[30]. Three notable scholars in this period are Shun'ichi Amari, Michael A. Arbib and Stephen Grossberg who laid the earliest theoretical foundation for neural networks, AI and the brain. Along with neural network researches, another milestone in this period is the theory of fuzzy set [6], [31], [32].

1) Neocognitron: Inspired by David Marr's theory on cerebellar cortex and cerebellum neocortex [33]–[37], a Japanese scholar Kunihiko Fukushima developed Cognitron [38] and later Neocognitron [39] computational models for object recognition. These models have later influenced Yann LeCun to develop the so-called *convolution neural network* for character recognition in the late 1980s [40], [41], the AlexNET [42] for image object recognition and today deep neural networks [43].

2) Self-Organizing Map (SOM): Data clustering is an important problem in statistical analysis. Given a set of unlabelled multi-dimensional data, one would need to figure out how many groups of data in the set. Usually, two criteria are defined to measure the goodness of the clustering. First, each group has to be far away enough from other groups. Second, within a group, the data must be close enough among each other. This problem has not just attracted researchers in statistical analysis but also attracted researchers in electrical engineering and computer science. It is because clustering problem is essentially the same as the data compression (equivalently vector quantization) problem in communication engineering.

While various algorithms had already been proposed in 1970s and applied in real applications, those algorithms were unable to get the topographic relations among clusters, i.e. the neighborhood relations among clusters. In the 1970s, researchers in brain science had already found that the neurons on cerebral cortex can self-organize to form topographic map. Inspired by this finding, Teuvo Kohonen proposed a clustering algorithm called self-organizing map (SOM) [44]. This new algorithm is able to do data clustering and at the same time find the topographical relations among clusters. In the 1990s, SOM has largely been applied in data dimension reduction problems, mapping of a set of high dimensional data to a lower dimensional space.

3) Oja PCA: Owing to map a set of data from high dimensional space to single dimension space, Erkki Oja proposed a neural network model and a learning algorithm[2] that can let the neural network to learn from the set of data the dimensional space corresponding to the principle component [45]. As the work done by the neural network is exactly the same as the statistical model principle component analysis (PCA)[3], this Oja model is sometimes called the PCA network or simply PCA.

4) Hopfield Network: Influenced by the binary state of a neuron, John Hopfield showed that a fully connected recurrent neural network with proper design is able to solve travelling salesman problem [46]–[48][4] Later, scholars have found that this neural network can be applied to solve other combinatorial optimization problems. Moreover, scholars have showed that the problem solving ability of Hopfield networks could also be enhanced by using the idea of simulated annealing [49]–[52].

5) Associative Memory: Alongside with the Hopfield network, associative memory models have been introduced in almost the same period of time [14], [53], [54]. These models are mimicking the memory property of our brains. The earliest models of this type are fully connected, like Hopfield network. The learning rule for these models is the *Hebbian learning*, a learning rule postulated by Donald Hedd in his book entitled *The Organization of Behavior* [55]. Sooner, Bart Kosko presented an associative memory model with two layers [56]. For the neuronal nodes of the same layer, there is no connection among each other. Bart Kosko named it bidirectional associative memory (BAM).

Interestingly, Hopfield network or associative memory has the same structure as the Boltzmann machine (to be introduced shortly) which was introduced by Geoffrey E. Hinton in the same period of time [57]. BAM has the same structure as the restricted Boltzmann machine (RBM) [58] and the logistic belief network [59] introduced by Geoffrey E. Hinton in the 2000s. These models are part of the deep neural networks.

6) Multi-layer Perceptron (MLP): In the early 1980s, neural network research came to its second wave in the history.

[2] It is called the Oja rule in some textbooks on neural networks and AI/ML.

[3] If you are familiar with the factor analysis (FA) model, you can consider the PCA model is a special case of FA model, in which the FA model has only one factor.

[4] Today, this recurrent neural network model is commonly called the Hopfield network.

Co-founded by James L. (Jay) McClelland[5] and David E. Rumelhart[6], a group of scholars from various disciplines and different universities across US was formed in UCSD, called the PDP research group, to investigate the psychological basis of human perception and mental phenomena by parallel and distributed processing (PDP) models, i.e. neural network models. The results obtained by the PDP group have later been included in a two volume monographs [60], [61]. By their joint effort, many new neural network models have been proposed and new learning rules are developed. Two notable models are the multilayer perceptron (MLP) and the Boltzmann machine. Instead of using threshold logic unit as the neuronal model, the transfer function of the neurons in the MLP is modelled as a sigmoid function. $z(u) = 1/1 + \exp(-u)$.

The learning algorithm developed for training a MLP is called the backpropagation (BP). It is also called error BP. Basically, backpropagation is derived based on the idea of gradient descent. With loss of generality, it is assumed that the MLP has only one output node. Given a set of data $\mathcal{D} = \{\mathbf{x}_k, y_k\}_{k=1}^{N}$. Let the model of a MLP is simply denoted by $f(\mathbf{x}, \mathbf{w})$, where $\mathbf{x}$ is the input vector and $\mathbf{w}$ is the weight vector (model parameters). The cost function to measure how good the MLP is given by $E(\mathbf{w}) = \frac{1}{N}\sum_{k=1}^{N}\left(y_k - f(\mathbf{x}_k, \mathbf{w})\right)^2$. The learning algorithm is given by $\mathbf{w} \leftarrow \mathbf{w} - \mu\nabla_{\mathbf{w}}E(\mathbf{w})$. MLP has then been applied in solving many engineering problems, like pattern classification, image recognition and system control.

7) Botlzmann Machine: Boltzmann machine is a fully connected stochastic network. That is to say, the output of each neuron will have to feed to all other neurons. Suppose that the network has n neurons. The output of the i^{th} neuron is denoted as z_i and $z_i \in \{0, 1\}$, a two-state neuron. The weight connecting the i^{th} and the j^{th} neurons are denoted as w_{ij}. Besides, $w_{ji} = w_{ij}$. Each neuron is also associated with a bias θ_i. Now, we can describe the working principle of this Boltzmann machine. For the i^{th} neuron, the total signal received is given by $u_i = \sum_{j=1,j\neq i}^{n} w_{ij}z_j + \theta_i$. This value determines the probability that the output z_i is one in accordance with the following probability mass function. $P(z_i = 1|u_i) = \frac{1}{1+\exp(-u_i)}, \quad P(z_i = 0|u_i) = \frac{\exp(-u_i)}{1+\exp(-u_i)}$. Once a neuron's output has been updated, another neuron is randomly picked and its output is updated by the same manner. To understand the idea behind the learning algorithm of Boltzmann machine, knowledge in advanced calculus and statistics is required. I stop short in here. If you are interested in this theoretical background, you can refer to the monographs [60], [61] for detail mathematical derivation of the learning algorithm.

8) Recurrent Neural Network: Owing to understand how speech is perceived, James J. McCleland and Jeffrey L Elman of the PDP research group developed a recurrent neural network model called TRACE [62]. In subsequent decades, various recurrent neural network models were developed and applied in speech recognition. Hopfield network, associative memory, Boltzmann machine and TRACE mark the stage for the early recurrent neural network development.

9) Belief Network: Belief network or Bayesian network was proposed in 1980s [63]. It conceptualizes the probabilistic reasoning for expert system design. It calculates the probability the happening of an event Z given the appearance of events A, B, C and so on.

10) Reinforcement Learning: One more remarkable contribution in this period of time is the reinforcement learning [7], [64], [65], in which the idea is advocated by Andrew Barto and Richard Sutton in the early 1980s. It conceptualizes from psychological theory on human learning – operant conditioning – an agent gets reward if its action can reach a temporal goal and otherwise the agent will be penalized. If the agent has been rewarded, the agent reinforces the association between the stimuli and the action. So, the agent could get another reward if the same stimuli appears again. Otherwise, the agent degrades the association between the stimuli and the action. As a result, the agent could avoid punishment if the stimuli appears again.

11) Applications: In this period of time, major research works in neural networks focused on developing models to explain psychological behavior of a human being, like object and speech recognition; and game playing. Only a few models, from the works in fuzzy systems, have been applied in industrial automation and mechanical control.

D. AI Research in the 1990s – 2000s

The UCSD PDP Group has proposed and developed a number of models which could be applied to interpret human reasoning. Since then, explosive number of research works on AI/ML have been conducted. Neural network, fuzzy system and AI/ML had then emerged as hot topics in 1990s and attracted scholars from different disciplines to work together on these areas. Their disciplines include psychology, neuroscience, mathematics, statistics, physics, computer science, electrical engineering, economics, linguistics, philosophy and others.

1) Computational Intelligence: At the same time, fuzzy logic[7] and genetic algorithm[8] came in place. Genetic algorithms refer to the optimization algorithms mimicking genetic mutation and evolution. Strictly speaking, genetic algorithm is not part of AI/ML. But, genetic algorithm could be applied in neural network researches and fuzzy systems researches. Today, the name genetic algorithms has evolved to the name called evolutionary computation. Neural networks and learning systems, fuzzy systems and evolutionary computation become three major topics of interest in the IEEE Computational Intelligence Society.

In the area of neural networks, pioneer researchers in the 1980s continued to develop new models and new algorithms advancing the models and algorithms developed in the 1980s. Other new models and new algorithms from other scholars were proposed in this period of time. Many of them were applied in real-life applications. Thus, some models their designs are no more biological brain-orient. Instead, their designs are application-oriented. For instance, functional-link networks [66], radial basis function (RBF) networks [67]–[70] and NARX/NARMAX models [71], [72] are essentially mathematical models oriented. They have no biological or psychological essence.

[5]http://www.standford.edu/~jlmcc.

[6]https://en.wikipedia.org/wiki/David_Renelhart.

[7]https://en.wikipedia.org/wiki/Fuzzy_logic.

[8]https://en.wikipedia.org/wiki/Genetic_algorithm.

2) Statistical Learning Models: In this period of time, a lot of models were proposed. Theoretical analytical works on the properties of the models were conducted. A notable model called support vector machine (SVM) was introduced in the 1990s [73], [74]. Later support vector regression (SVR) machine [75], [76] was introduced. These models have largely been applied in classification problems.

Moreover, the logistic belief network [77], the long short term memory (LTSM) model [78], the restricted Boltzmann machine (RBM) [58] and the deep belief network [59] were developed. They have thus been embraced in statistical learning models and then machine learning models. These models have laid the foundation for the deep neural networks in the 2010s.

3) Spike Neural Networks: In the 1990s, some scholars in the area of computational neuroscience shifted the model of a neuron from sigmoid function to the biological-inspired neuron models. They include the Lapicque integrate-and-fire model [79], Hodgkin-Huxley model [80]–[84], the FitzHugh–Nagumo model [85], [86], the Morris-Lecar model [87] the Hindmarsh-Rose model [88], [89] and others[9].

These neuron models are applied to model biological neuronal networks. Their behaviors could be investigated by computer simulation. This type of neural network is called the third generation neural network, while the early neural network like Perceptron with threshold logic unit as the neuron is called the first generation neural network and the neural network with sigmoidal neuron like MLP is called the second generation neural network.

In the third generation neural network, many new learning algorithms have been developed. Some followed the idea of backpropagation. Some followed closely to a biological learning rule – Spike-timing-dependent plasticity (STDP). The idea is basically the Hebbian learning with the time difference between the firing time of the postsynaptic neuron and the firing time of the presynaptic neuron. If the postsynaptic neuron fires after the receiving of the firing of the presynaptic neuron, the synaptic strength between the two neurons will be incremented. If the time difference is shorter, the incremental change will be larger. On the other hand, if the postsynaptic neuron fires before the receiving of the firing of the presynaptic neuron, the synaptic strength between the two neurons will be decremented. If the time difference is shorter, the decrement will be larger.

While the spike neural network has captured the most biological brain structure, theoretical analysis on this type of neural network is difficult as the neuron models are complicated. With STDP learning, the analysis is even harder. Therefore, various researches are limited on simulations. While difficult, this area of research has drawn substantial attention in European Union (EU). Thus, a 10-year EU funded *Human Brain Project*[10] has launched in 2013. The project employs some 500 scientists at more than 100 universities, teaching hospitals, and research centres across Europe. Hopefully, more mysteries about our brains could be revealed in the future.

4) Applications: In this period of time, those intelligent technologies have not yet been transferred to commercial products. Large number of application-oriented researches, especially in system control, were conducted [70], [72], [90]–[93]. Various artificial neural network models and fuzzy systems were applied in pattern recognition, automatic target tracking, truck parking, echo noise cancellation, medical diagnosis.

Many intelligent technologies were applied in recommendation systems for product recommendation, search recommendation and target marketing. Some of them were applied in big data analytic for customers preference analysis and consumer behaviorial research.

5) Notable Events: One remarkable event has to be mentioned. This event has brought to the public the attention on the phenomenal ability of an AI machine. IBM Deep Blue defeated the reigning world chess champion Garry Kasparov[11] in 1997. A computer could defeat human brain.

E. AI Research in the 2010s

In this period of time, many notable events had happened. Intelligent products and intelligent services have been developed and available in the marker. As the number of notable events and the number of intelligent products/services are huge, it is not possible to mention them all. Here, I have only selected some of them.

1) Notable Applications: IBM Watson defeated human players in an Q&A game Jeopardy in 2011[12]. The dataset named ImageNet was introduced in 2009 in the *Conference on Computer Vision and Pattern Recognition (CVPR)* in Florida. In the next year, *ImageNet Large Scale Visual Recognition Challenge (ILSVRC)* was launched[13]. A research team in University of Toronto developed an object recognition system, called Alexnet, won the 2012 championship in the ImageNet Large Scale Visual Recognition Challenge. Google AlphaGo defeated Sedol Lee in a five-game match of Go game[14] in 2016. In 2016, Google has also released a new version of Google Translator which applies a new translation system called Google Neural Machine Translation System[15]. Moreover, various intelligent virtual assistants were developed in this period of time. Apple released Siri for iPhone in 2011. Google developed Google Assistant for use in Android phones and Google Home. Amazon developed Alexa for use in Amazon Echo. Google Home and Amazon Echo are intelligent assistants for use at home.

2) Intelligent Services on Cloud: Today, many cloud platforms (Amazon Web Service, Google Cloud, Tencent Cloud, Alibaba Cloud) have provided intelligence services and APIs for developers to develop new services over these cloud platforms. Besides, some of these platforms have also provided virtual GPU service for those who would like to develop new intelligent technologies.

3) Industrial AI Research: Various tech giants have invested in their AI research programs. Apart from IBM and Microsoft which have a long history in AI research, Google, Amazon, Facebook, Intel, Nvidia, BMW, Tecent, Alibaba, Baidu and many others have allocated budget for AI & Machine Learning research and applications. Moreover, they have formed research teams with particularly focus on the

[9]https://en.wikipedia.org/wiki/Biological_neuron_model.

[10]https://www.humanbrainproject.eu/en/about/overview/.

[11]https://en.wikipedia.org/wiki/Deep_Blue_(chess_computer).

[12]https://en.wikipedia.org/wiki/Watson_(computer).

[13]http://image-net.org/index.

[14]https://en.wikipedia.org/wiki/AlphaGo.

[15]https://en.wikipedia.org/wiki/Google_Neural_Machine_Translation.

theory and applications of AI & Machine Learning. Some of them, like Google and Facebook, have even developed AI libraries for the community to develop advanced applications of AI & Machine Learning.

III. Benefits & Threads

A. Advancing Intelligence of Machines

Another big change will be the intelligent technologies. As mentioned in Section II, IBM Deep Blue, IBM Watson and Google's AlphaGo have already demonstrated their intelligence in many games competing with human contestants. Their successes are due to two major breakthroughs. One is the huge back-end computational resources networked to the front-end terminals. The other is the machine learning (AI) algorithms for analyzing the information to give the best answer. Through the front-end terminal, the engineer could access and control the back-end computational resource to collect huge volume of information over the Internet, analyze the information by some machine learning algorithms and then give the best answer to the engineer. The processing time is almost instantaneous.

On-going researches on artificial intelligence and machine learning have recently conducted intensively in Amazon[16], Facebook (Facebook AI Research[17]), Google (Google Brain[18]), IBM (AI and Cognitive Computing[19]), Microsoft (Machine Learning and Optimization[20]). Even Apple, she has started her AI research in 2016 [94]. One major driver for these researches is NVIDIA's graphical processing units (GPU) [21] and the cloud technologies.

Running a machine learning algorithm is always time consuming if the program is running in any conventional multi-core computer. GPU is basically a CPU specialized design for mathematical computation. Initially, it was designed for processing graphical data. As its computational speed is hundred to thousand time faster than the normal CPU, it has then been applied in running computational intensive machine learning programs for tagging 1.3 millions images [42] and tagging 8 millions video [95]. Another major drive is clearly the cloud, a network of memory and computational resources.

Some of these research results together with their cloud platforms have already been commercialized as intelligent services for anyone who is interested in developing more sophisticated intelligent services for users, like Amazon Machine Learning Services[22], Google Cloud Machine Learning Platform[23] and IBM Watson[24]. Some of these have been commercialized as intelligent business solution for enterprises, like SAP HANA Cloud Platform[25]. Researches in AI for sure will never stop. More intelligent results will show up in the future and more intelligent services will come in the market. More new automation and robotic technologies will be advanced due to the application of AI [96].

[16]https://aws.amazon.com/amazon-ai/.

[17]https://research.fb.com/category/facebook-ai-research-fair/.

[18]https://research.google.com/teams/brain.

[19]http://research.ibm.com/cognitive-computing/.

[20]https://www.microsoft.com/en-us/research/group/machine-learning-and-optimization/.

[21]https://www.nvidia.com/en-us/deep-learning-ai/.

[22]https://aws.amazon.com/machine-learning/.

[23]https://cloud.google.com/products/machine-learning/.

[24]https://www.ibm.com/watson/.

[25]http://www.sap.com/developer/topics/hcp.html.

B. Decline Demand of Administrative and Managerial Staffs

Earlier in the nineteenth century, political economists had already observed and raised the issue on the reduction of wages or quantities of labors due to the employment of machinery. As asserted in the book entitled *On the Principles of Political Economy, and Taxation* written by David Ricardo in 1821 [97], (P.22) *The principles that the quantity of labour bestowed on the production of commodities regulates their relative value, considerably modified by the employment of machinery and other fixed and durable capital*. Ricardo clearly stated that the value of a labor has to be compared with what a machine could do. If a machine could do a better job than a labor, the value of a labor declines.

In the late nineteenth century to the earlier twentieth century, many firms hired a lot of clerks and secretaries to use typewriters for typing paper documents. In the middle of twentieth century, photocopier and computer were invented. Their works were thus replaced by these machines (see P.64-65 in [98]). With further advancement in automation and robotic technologies, demand of human workers in agricultural and manufacturing industries will certainly decline.

Networking technologies led to the reduction of office assistants. Advancement on the intelligent functions of information systems (including management information systems, decision support systems, executive information systems, supply chain management systems, customer relationship management systems and enterprise resources planning systems) not just eases the jobs of administrative and middle management staffs but also reduces the demand on such staffs. With such powerful information systems, the jobs used to be done by these administrative and middle management staffs could now be easily handled by the senior management staffs.

With further advancement in big data, Internet of Things, automation & robotic, and artificial intelligence, (i) a lot more management information systems could be available in the market, (ii) the intelligent levels of such systems would be raised and (iii) factories could be networked together to form a giant autonomous factory producing every thing. As a result, administrative and middle management jobs will also decline.

C. The End of Middle Management

In an article in BBC entitled "The end of middle management?" [99], Sydney Finkelstein has pointed out that technology (like computer) is not the only factor leading to the end of middle management. The culture of start-up firms do not like middle managers. The behavior of millennials at work – they believe that they know more than they do. All these factors constitute the down-value of middle management and eventually lead to the end of middle management. Once the demand of the administrative, production and middle management staffs declines, the roles of human resource management will change and the demand of human resource professionals will definitely be reduced.

In the future, the works of senior managers and executives could also be replaced by AI [100]. In the end, management schools would be obsolete. Some of them would have to change their focuses and their roles in the society. Some might have to scale-down the number of undergraduate programs, as the students educated in these programs might not be able to compete with the future AI systems. Management schools should be refashioned to educate graduates with skills that the companies need (P.79 in [101]). But, I like to add a point which is in a spirit originally from Chester Barnard (P.176 in [102]). Management schools should re-focus their roles in helping students to learn how to learn in their life-times.

References

[1] A. M. Turing, “Intelligent Machinery,” National Physics Laboratory. Mathematics Division., Tech. Rep., 1948.

[2] ——, “Computing machinery and intelligence,” *Mind*, vol. 59, no. 236, pp. 433 – 460, 1950.

[3] M. Minsky, “Theory of neural-analog reinforcement system and its applications to the brain-model problem,” Ph.D. dissertation, University of Princeton, Princeton, NJ, 1954.

[4] J. Von Neumann, *The Computer and the Brain*. Yale University Press, 1958.

[5] F. Rosenblatt, “The perceptron: a probabilistic model for information storage and organization in the brain.” *Psychological Review*, vol. 65, no. 6, p. 386, 1958.

[6] L. A. Zadeh, “Fuzzy sets,” *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965.

[7] R. S. Sutton, “Temporal credit assignment in reinforcement learning,” Ph.D. dissertation, University of Massachusetts, Department of Computer and Information Science, 1984.

[8] W. S. McCulloch and W. Pitts, “A logical calculus of the ideas immanent in nervous activity,” *The Bulletin of Mathematical Biophysics*, vol. 5, no. 4, pp. 115–133, 1943.

[9] H. Landahl, W. S. McCulloch, and W. Pitts, “A statistical consequence of the logical calculus of nervous nets,” *The Bulletin of Mathematical Biophysics*, vol. 5, no. 4, pp. 135–137, 1943.

[10] F. Rosenblatt, “Perceptron simulation experiments,” *Proceedings of the IRE*, vol. 48, no. 3, pp. 301–309, 1960.

[11] ——, *Principles of Neurodynamics: Perceptions and the theory of brain mechanisms*. Spartan, 1962.

[12] S. Grossberg, “Neural pattern discrimination,” *Journal of Theoretical Biology*, vol. 27, no. 2, pp. 291–337, 1970.

[13] S.-I. Amari, “Learning patterns and pattern sequences by self-organizing nets of threshold elements,” *IEEE Transactions on Computers*, vol. 100, no. 11, pp. 1197–1206, 1972.

[14] K. Nakano, “Associatron – A model of associative memory,” *IEEE Transactions on Systems, Man, and Cybernetics*, no. 3, pp. 380–388, 1972.

[15] H. Scudder, “Probability of error of some adaptive pattern-recognition machines,” *IEEE Transactions on Information Theory*, vol. 11, no. 3, pp. 363–371, 1965.

[16] S. Grossberg, “Some networks that can learn, remember, and reproduce any number of complicated space-time patterns I,” *Journal of Mathematics and Mechanics*, vol. 19, no. 1, pp. 53–91, 1969.

[17] S.-I. Amari, “Characteristics of randomly connected threshold-element networks and network systems,” *Proceedings of the IEEE*, vol. 59, no. 1, pp. 35–47, 1971.

[18] ——, “Characteristics of random nets of analog neuron-like elements,” *IEEE Transactions on Systems, Man, and Cybernetics*, no. 5, pp. 643–657, 1972.

[19] ——, “Neural theory of association and concept-formation,” *Biological Cybernetics*, vol. 26, no. 3, pp. 175–185, 1977.

[20] M. A. Arbib, “On the convergence of grossberg’s learning equations,” *Journal of Mathematical Analysis and Applications*, vol. 31, no. 3, pp. 545–553, 1970.

[21] S. Amari, “A theory of adaptive pattern classifiers,” *IEEE Transactions on Electronic Computers*, no. 3, pp. 299–307, 1967.

[22] M. A. Arbib, “Artificial intelligence and brain theory: Unities and diversities,” *Annals of Biomedical Engineering*, vol. 3, no. 3, pp. 238–274, 1975.

[23] ——, “Cybernetics after 25 years: A personal view of system theory and brain theory,” *IEEE Transactions on Systems, Man, and Cybernetics*, no. 3, pp. 359–363, 1975.

[24] ——, “Artificial Intelligence: Cooperative Computation and Man? Machine Symbiosis,” *IEEE Transactions on Computers*, no. 12, pp. 1346–1352, 1976.

[25] M. Arbib, “Turing machines, finite automata and neural nets,” *Journal of the ACM (JACM)*, vol. 8, no. 4, pp. 467–475, 1961.

[26] M. A. Arbib, “Automata theory and development: Part I,” *Journal of Theoretical Biology*, vol. 14, no. 2, pp. 131–156, 1967.

[27] S. Grossberg, “Embedding fields: A theory of learning with physiological implications,” *Journal of Mathematical Psychology*, vol. 6, no. 2, pp. 209–239, 1969.

[28] M. A. Arbib, “Toward an automata theory of brains,” *Communications of the ACM*, vol. 15, no. 7, pp. 521–527, 1972.

[29] ——, “From automata theory to brain theory,” *International Journal of Man-Machine Studies*, vol. 7, no. 3, pp. 279–295, 1975.

[30] S.-I. Amari, K. Yoshida, and K.-I. Kanatani, “A mathematical foundation for statistical neurodynamics,” *SIAM Journal on Applied Mathematics*, vol. 33, no. 1, pp. 95–126, 1977.

[31] R. E. Bellman and L. A. Zadeh, “Decision-making in a fuzzy environment,” *Management Science*, vol. 17, no. 4, pp. B–141, 1970.

[32] L. A. Zadeh, “Outline of a new approach to the analysis of complex systems and decision processes,” *IEEE Transactions on Systems, Man, and Cybernetics*, no. 1, pp. 28–44, 1973.

[33] D. Marr, “A theory of cerebellar cortex,” *Journal of Physiolgy*, vol. 202, no. 2, pp. 437–470, 1969.

[34] ——, “A theory for cerebral neocortex,” *Proceedings of the Royal Society of London. Series B. Biological Sciences*, vol. 176, no. 1043, pp. 161–234, 1970.

[35] D. Marr and T. Poggio, “A computational theory of human stereo vision,” *Proceedings of the Royal Society of London. Series B. Biological Sciences*, vol. 204, no. 1156, pp. 301–328, 1979.

[36] D. Marr and E. Hildreth, “Theory of edge detection,” *Proceedings of the Royal Society of London. Series B. Biological Sciences*, vol. 207, no. 1167, pp. 187–217, 1980.

[37] D. Marr, *Vision: A computational investigation into the human representation and processing of visual information*. CUMINCAD, 1982.

[38] K. Fukushima, “Cognitron: A self-organizing multilayered neural network,” *Biological Cybernetics*, vol. 20, no. 3-4, pp. 121–136, 1975.

[39] ——, “Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position,” *Biological Cybernetics*, vol. 36, pp. 193–202, 1980.

[40] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, “Backpropagation applied to handwritten zip code recognition,” *Neural Computation*, vol. 1, no. 4, pp. 541–551, 1989.

[41] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[42] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Advances in Neural Information Processing Systems*, 2012, pp. 1097–1105.

[43] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[44] T. Kohonen, “Self-organized formation of topologically correct feature maps,” *Biological Cybernetics*, vol. 43, no. 1, pp. 59–69, 1982.

[45] E. Oja, “Simplified neuron model as a principal component analyzer,” *Journal of Mathematical Biology*, vol. 15, no. 3, pp. 267–273, 1982.

[46] J. J. Hopfield, “Neural networks and physical systems with emergent collective computational abilities,” *Proceedings of the National Academy of Sciences*, vol. 79, no. 8, pp. 2554–2558, 1982.

[47] J. J. Hopfield and D. W. Tank, ““Neural” computation of decisions in optimization problems,” *Biological Cybernetics*, vol. 52, no. 3, pp. 141–152, 1985.

[48] ——, “Computing with neural circuits: A model,” *Science*, vol. 233, no. 4764, pp. 625–633, 1986.

[49] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, “Optimization by simulated annealing,” *Science*, vol. 220, no. 4598, pp. 671–680, 1983.

[50] S. Kirkpatrick, “Optimization by simulated annealing: Quantitative studies,” *Journal of Statistical Physics*, vol. 34, no. 5-6, pp. 975–986, 1984.

[51] H. Szu and R. Hartley, “Fast simulated annealing,” *Physics Letters A*, vol. 122, no. 3-4, pp. 157–162, 1987.

[52] P. J. Van Laarhoven and E. H. Aarts, “Simulated annealing,” in *Simulated Annealing: Theory and Applications*. Springer, 1987, pp. 7–15.

[53] G. Palm, “On associative memory,” *Biological Cybernetics*, vol. 36, no. 1, pp. 19–31, 1980.

[54] G. A. Carpenter, “Neural network models for pattern recognition and associative memory,” *Neural Networks*, vol. 2, no. 4, pp. 243–257, 1989.

[55] D. Hebb, *The Organization of Behavior*. Wiley, New York, 1949.

[56] B. Kosko, “Bidirectional associative memories,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 18, no. 1, pp. 49–60, 1988.

[57] D. H. Ackley, G. E. Hinton, and T. J. Sejnowski, “A learning algorithm for Boltzmann machines,” *Cognitive Science*, vol. 9, no. 1, pp. 147–169, 1985.

[58] G. E. Hinton, “Training products of experts by minimizing contrastive divergence,” *Neural Computation*, vol. 14, no. 8, pp. 1771–1800, 2002.

[59] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, no. 7, pp. 1527–1554, 2006.

[60] D. E. Rumelhart, J. L. McClelland, and P. R. Group, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volumn 1: Foundations*. MIT Press, 1986.

[61] ——, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volumn 2: Psychological and Biological Models*. MIT Press, 1986.

[62] J. L. McClelland and J. L. Elman, "The TRACE model of speech perception," *Cognitive Psychology*, vol. 18, no. 1, pp. 1–86, 1986.

[63] J. Pearl, "Fusion, propagation, and structuring in belief networks," *Artificial Intelligence*, vol. 29, no. 3, pp. 241–288, 1986.

[64] A. G. Barto, R. S. Sutton, and P. S. Brouwer, "Associative search network: A reinforcement learning associative memory," *Biological Cybernetics*, vol. 40, no. 3, pp. 201–211, 1981.

[65] A. G. Barto, R. S. Sutton, and C. W. Anderson, "Neuronlike adaptive elements that can solve difficult learning control problems," *IEEE Transactions on Systems, Man, and Cybernetics*, no. 5, pp. 834–846, 1983.

[66] Y. Pao, *Adaptive Pattern Recognition and Neural Networks*. Addison-Wesley, 1989.

[67] W. Beastall, "Recognition of radar signals by neural network," in *1989 First IEE International Conference on Artificial Neural Networks,(Conf. Publ. No. 313)*. IET, 1989, pp. 139–142.

[68] S. Renals, "Radial basis function network for speech pattern classification," *Electronics Letters*, vol. 25, no. 7, pp. 437–439, 1989.

[69] F. Girosi and T. Poggio, "Networks and the best approximation property," *Biological Cybernetics*, vol. 63, no. 3, pp. 169–176, 1990.

[70] S. Chen, C. F. Cowan, and P. M. Grant, "Orthogonal least squares learning algorithm for radial basis function networks," *IEEE Transactions on Neural Networks*, vol. 2, no. 2, pp. 302–309, 1991.

[71] S. Chen, S. A. Billings, C. F. Cowan, and P. M. Grant, "Practical identification of NARMAX models using radial basis functions," *International Journal of Control*, vol. 52, no. 6, pp. 1327–1350, 1990.

[72] K. S. Narendra and K. Parthasarathy, "Identification and control of dynamical systems using neural networks," *IEEE Transactions on Neural Networks*, vol. 1, no. 1, pp. 4–27, 1990.

[73] V. Vapnik, I. Guyon, and T. Hastie, "Support vector machines," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

[74] C. Cortes and V. Vapnik, "Support vector machine," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

[75] A. J. Smola *et al.*, "Regression estimation with support vector learning machines," Master's thesis, Technische Universität München, 1996.

[76] H. Drucker, C. J. Burges, L. Kaufman, A. J. Smola, and V. Vapnik, "Support vector regression machines," in *Advances in Neural Information Processing Systems*, 1997, pp. 155–161.

[77] R. M. Neal, "Connectionist learning of belief networks," *Artificial Intelligence*, vol. 56, no. 1, pp. 71–113, 1992.

[78] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[79] L. Lapicque, "Recherches quantitatives sur l'excitation electrique des nerfs traitee comme une polarization," *Journal de Physiologie et de Pathologie Generalej*, vol. 9, pp. 620–635, 1907.

[80] A. L. Hodgkin, A. F. Huxley, and B. Katz, "Measurement of current-voltage relations in the membrane of the giant axon of Loligo," *The Journal of Physiology*, vol. 116, no. 4, p. 424, 1952.

[81] A. L. Hodgkin and A. F. Huxley, "Currents carried by sodium and potassium ions through the membrane of the giant axon of Loligo," *The Journal of Physiology*, vol. 116, no. 4, p. 449, 1952.

[82] ——, "A quantitative description of membrane current and its application to conduction and excitation in nerve," *The Journal of Physiology*, vol. 117, no. 4, p. 500, 1952.

[83] ——, "The dual effect of membrane potential on sodium conductance in the giant axon of Loligo," *The Journal of Physiology*, vol. 116, no. 4, p. 497, 1952.

[84] ——, "Propagation of electrical signals along giant nerve fibres," *Proceedings of the Royal Society of London. Series B-Biological Sciences*, vol. 140, no. 899, pp. 177–183, 1952.

[85] R. FitzHugh, "Mathematical models of threshold phenomena in the nerve membrane," *The Bulletin of Mathematical Biophysics*, vol. 17, no. 4, pp. 257–278, 1955.

[86] J. Nagumo, S. Arimoto, and S. Yoshizawa, "An active pulse transmission line simulating nerve axon," *Proceedings of the IRE*, vol. 50, no. 10, pp. 2061–2070, 1962.

[87] C. Morris and H. Lecar, "Voltage oscillations in the barnacle giant muscle fiber," *Biophysical Journal*, vol. 35, no. 1, pp. 193–213, 1981.

[88] J. Hindmarsh and R. Rose, "A model of the nerve impulse using two first-order differential equations," *Nature*, vol. 296, no. 5853, pp. 162–164, 1982.

[89] J. L. Hindmarsh and R. Rose, "A model of neuronal bursting using three coupled first order differential equations," *Proceedings of the Royal society of London. Series B. Biological sciences*, vol. 221, no. 1222, pp. 87–102, 1984.

[90] K. S. Narendra and K. Parthasarathy, "Gradient methods for the optimization of dynamical systems containing neural networks," *IEEE Transactions on Neural Networks*, vol. 2, no. 2, pp. 252–262, 1991.

[91] K. S. Narendra, "Neural networks for control theory and practice," *Proceedings of the IEEE*, vol. 84, no. 10, pp. 1385–1406, 1996.

[92] K. S. Narendra and S. Mukhopadhyay, "Adaptive control using neural networks and approximate models," *IEEE Transactions on Neural Networks*, vol. 8, no. 3, pp. 475–485, 1997.

[93] L. Chen and K. S. Narendra, "Identification and control of a nonlinear discrete-time system based on its linearization: A unified framework," *IEEE Transactions on Neural Networks*, vol. 15, no. 3, pp. 663–673, 2004.

[94] A. Shrivastava, T. Pfister, O. Tuzel, J. Susskind, W. Wang, and R. Webb, "Learning from simulated and unsupervised images through adversarial training," *arXiv preprint arXiv:1612.07828*, 2016.

[95] S. Abu-El-Haija, N. Kothari, J. Lee, P. Natsev, G. Toderici, B. Varadarajan, and S. Vijayanarasimhan, "Youtube-8m: A large-scale video classification benchmark," *arXiv preprint arXiv:1609.08675*, 2016.

[96] M. Purdy and P. Daugherty, "Why artificial intelligence is the future of growth," Available at http//www.accenture.com/futureofai., Accenture Institute for High Performance, Tech. Rep., September 2016.

[97] D. Ricardo, *On the Principles of Political Economy, and Taxation*. John Murray, 1821.

[98] J. Diebold, "Automation - The new technology," *Harvard Business Review*, vol. 31, no. 6, pp. 63–71, 1953.

[99] S. Finkelstein, "The end of middle management?" Available at http://www.bbc.com/capital/story/20150624-the-end-of-middle-management, June 2015.

[100] D. Sharma, S. Yen, M. Noga, E. Marcade, C. Saravana, and D. Beurteaux, "An AI shares my office," Available at http://www.digitalistmag.com/executive-research/an-ai-shares-my-office., January 2017.

[101] A. Bernstein and J. Joerres, "Globalization, robots, and the future of work," *Harvard Business Review*, vol. 94, no. 10, pp. 74–79, 2016.

[102] C. I. Barnard, "Education for executives," *The Journal of Business of the University of Chicago*, vol. 18, no. 4, pp. 175–182, 1945.