

# Neural Network Driven Supper Artificial Intelligence Based on Internet of Things and Big Data

\*<sup>1</sup>Bahman Zohuri, <sup>2</sup>Masoud Moghaddam

<sup>1</sup>University of New Mexico, Electrical Engineering and Computer Science Department, Albuquerque, New Mexico USA, and Galaxy Advanced Engineering (CEO), USA

<sup>2</sup>Masoud Moghaddam, Galaxy Advance Engineering Consultant, Albuquerque, New Mexico USA

## Abstract

With today's growing information and overloading of its volume, it is becoming tremendously difficult to analyze the huge amounts of data that contain the information and makes it very strenuous and inconvenient to introduce an appropriate methodology of decision-making fast enough to the point that it can be, considered as real time. The demand for real time processing information and related data both structured and unstructured is on the rise and consequently makes it harder and harder to implement correct decision making at enterprise level to keep the organization robust and resilient against either man made threats or natural disasters.

Today's campaign against any cyber-attack has put a huge demand on cyber security and on information security folks at different levels of any organization. Therefore, processing incoming data as sets of information becomes more and more critical. Furthermore, the data are often, imprecise and will include both quantitative and qualitative elements. For these reasons it is important to extend traditional decision-making processes by adding intuitive reasoning, human subjectivity and imprecision. To enhance this process of decision-making, these authors have taken an unorthodox approach by applying a new growing technology known as neural network as part of driving infrastructure for artificial intelligence system to take over from human being in order to satisfy the demand for real time decision making.

## Introduction

Artificial Neural Networks (ANNs) are the wave of the future, anywhere from being capable of doing everything from translating financial data into financial predictions, to mapping visual images to robotic commands, to classifying medical images for diagnostic test and so on. That why a book like this is essential reading for computer programmers, hardware and software engineers and specialists, managers, biologists, and other professionals who want to stay in the vanguard of their fields.

Artificial Intelligence of today are foundation of future smart robots to carry on whatever tasks could be assigned to them. These robots may be designed around the idea of processing data by far faster than human brain and to be considered as plat form for computational agents of near future going forward in time. Per description by (Poole and Mackworth) [1]. "A computational agent is an agent whose decisions about its actions can be explained in

terms of computation. That is, the decision can be broken down into primitive operation that can be implemented in a physical device. This computation can take many forms. In humans this computation is carried out in "wetware"; in computers it is carried out in "hardware." Although there

\*Corresponding author: Bahman Zohuri, University of New Mexico, Electrical Engineering and Computer Science Department, Albuquerque, New Mexico USA, and Galaxy Advanced Engineering (CEO), USA. E-mail: bahmanz@aol.com

Received June 28, 2018; Accepted July 12, 2018; Published July 26, 2018

Citation: Bahman Zohuri (2018) Neural Network Driven Supper Artificial Intelligence Based on Internet of Things and Big Data. SF J Artificial Intel 1:1.

Copyright: © 2018 Bahman Zohuri. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

are some agents that are arguably not computational, such as the wind and rain eroding a landscape, it is an open question whether all intelligent agents are computational". Additionally, Poole and Mackworth are clamming that an agent acts intelligently, when performs under the following conditions, where our interest falls into how this agent acts as well, so when can judge it by its action:

- What it does is appropriate for its circumstances and its goals,
- It is flexible to changing environments and changing goals,
- It learns from experience, and
- It makes appropriate choices given its perceptual and computational limitations. An agent typically cannot observe the state of the world directly; it has only a finite memory and it does not have unlimited time to act.

Thus, the fundamental and infrastructure of a computational agent is an agent whose decisions about its actions can be explained in terms of computation. That is, the decision can be broken down into primitive operation that can be implemented in a physical device. The central *scientific goal* of Artificial Intelligence (AI) is to understand the principle that make intelligent behavior possible in natural or artificial systems. This is done by:

- The analysis of natural and artificial agents,
- Formulating and testing hypotheses about what it takes to construct intelligent agents, and
- Designing, building, and experimenting with computational systems that perform tasks commonly viewed as requiring intelligence

As part of science, researchers build empirical systems to test hypotheses or to explore the space of possibilities. These are quite distinct from applications that are built to be useful for an application domain.

It is arguable that intelligence is different: you cannot have fake intelligence. If an agent behaves intelligently, it is intelligent. It is only the external behavior that defines intelligence; acting intelligently is being intelligent. Thus, artificial intelligence, if and when it is achieved, will be real intelligence created artificially. However, an agent that is not really intelligent could not fake intelligence for arbitrary topics.

**Note that:** the definition is not for intelligent thought. We

are only interested in thinking intelligently insofar as it leads to better performance. The role of thought is to affect action.

The central engineering goal of AI is the design and synthesis of useful, intelligent artifacts. We actually want to build agents that act intelligently. Such agents are useful in many applications [1].

Although, modern computers, from low-level hardware to high-level software, are more complicated than any human can understand, yet they are manufactured daily by organizations of humans. Human society viewed as an agent is arguably the most intelligent agent known.

### 1.1 Artificial Neural Network (ANN)

Artificial neural networks are a narrow-sensed abstraction of the human brain; thus, the organization of the artificial neural system is very similar to the one of biological neurons. The comprehensive understanding of biological neurons is not complete; however, the basic functionality that contributes to the learning ability of a system is implemented in artificial neural networks. The fundamental element, an artificial neuron, is a model based on known behavior of biological neurons that exhibit most of the characteristics of human brains that we are interested in [Vel98]. This is the most significant difference from conventional computers, which have internal fixed instructions to perform specific functions.

Artificial neural networks can be also described as highly parallel distributed computing models. The fundamental processing units, neurons, are highly connected with strengths, which are dynamically changed during the system's learning process.

The earliest work in neural computing goes back to the 1940's when McCulloch and Pitts [2] introduced the first neural network computing model. Originally in 1943 Warren S. McCulloch, a neuroscientist, Walter Pitts, a logician, published their first paper [2].

The original MCP Neurons had limitations. Additional features were added which allowed them to "learn." The next major development in neural networks was the concept of a perceptron which was introduced by Frank Rosenblatt in 1958. Essentially the perceptron is an MCP neuron where the inputs are first passed through some "preprocessors," which are called association units. These association units detect the presence of certain specific features in the inputs. In fact, as the name suggests, a perceptron was intended to be a pattern recognition device, and the association units correspond to feature or pattern

detectors.

The original MCP Neurons had limitations. Additional features were added which allowed them to “learn.” The next major development in neural networks was the concept of a perceptron which was introduced by Frank Rosenblatt in 1958. Essentially the perceptron is an MCP neuron where the inputs are first passed through some “preprocessors,” which are called association units. These association units detect the presence of certain specific features in the inputs. In fact, as the name suggests, a perceptron was intended to be a pattern recognition device, and the association units correspond to feature or pattern detectors.

The early model of an artificial neuron which was introduced by Warren McCulloch and Walter Pitts is also known as Linear Threshold Gate (LTG) or sometimes it is called, the Threshold Logic Unit (TLU). It is a neuron of a set of inputs  $I_1, I_2, I_3, \dots, I_m$  and one output  $Y$ . The linear threshold gate simply classifies the set of inputs into two different classes. Thus, the output  $Y$  is a binary form. Such a function can be described mathematically using these equations:

$$sum = \sum_{i=1}^N I_i W_i \quad \text{Equation 1}$$

and

$$y = f(Sum) \quad \text{Equation 2}$$

In Equation 1, parameters  $W_1, W_2, W_3, \dots, W_m$  are weight values normalized in the range of either (0,1) or (1,-1) and associated with each input line, and the *Sum* is the weighted sum, while the function  $f$  is a linear step function at threshold  $T$ , as shown in Figure 1 below.

In Figures 1-3: the value of  $T$ , is presentation of a threshold constant. The symbolic representation of the Linear Threshold Gate (LTG) is shown in Figure 2.

The McCulloch-Pitts model of a neuron is simple yet has substantial computing potential. It also has a precise mathematical definition. However, this model is so simplistic that it only generates a binary output and also the weight and threshold values are fixed. The neural computing algorithm has diverse features for various applications [4]. Thus, we need to obtain the neural model with more flexible computational features.

In the 1950's, Rosenblatt's work resulted in a

two-layer network, the Perceptron, which was capable of learning certain classifications by adjusting connection weights. Although the perceptron was successful in classifying certain patterns, it had a number of limitations. The perceptron was not able to solve the classic Exclusive or (XOR) problem (Figure 3). Such limitations led to the decline of the field of neural networks. However, the Perceptron had laid foundations for later work in neural computing.

Figure 1: Linear Threshold Function

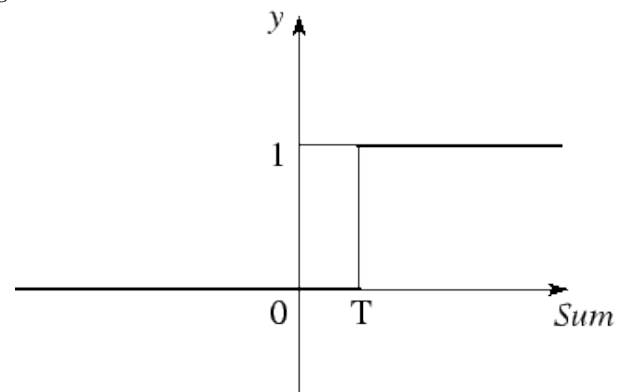


Figure 2: Symbolic Illustration of Linear Threshold Gate [3]

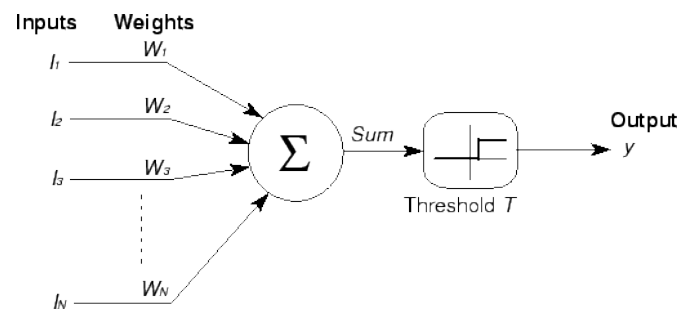
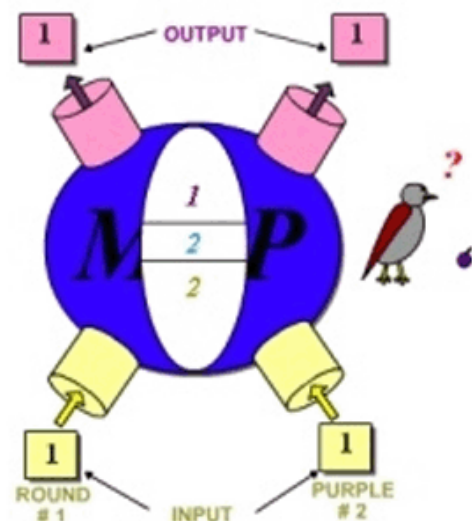


Figure 3: Illustration of Exclusive or (XOR)



**Note that:** In machine learning, the Perceptron is an algorithm for supervised learning of binary classifiers (functions that can decide whether an input, represented by a vector of numbers, belongs to some specific class or not).

Frank Rosenblatt early 1950 work was consistent of the Perceptron and introduction of a network, which was composed of the units and it was enhanced version of McCulloch-Pitts Threshold Logic Unit (TLU) model. Rosenblatt’s model of neuron, a perceptron, was the result of merger between two concepts from the 1940s, McCulloch-Pitts model of an artificial neuron and Hebbian learning rule of adjusting weights [5]. In addition to the variable weight values, the perceptron model added an extra input that represents bias. Thus, the modified form of Equation 1 is now as follows:

$$sum = \sum_{i=1}^N I_i W_i + b \quad \text{Equation 3}$$

In Equation 3, the constant  $b$  represents the bias value.

In the early 1980’s, researchers showed renewed interest in neural networks. Recent work includes Boltzmann machines, Hopfield nets, competitive learning models, multilayer networks, and adaptive resonance theory models.

## 1.2 Artificial Neuron with Continuous Characteristics

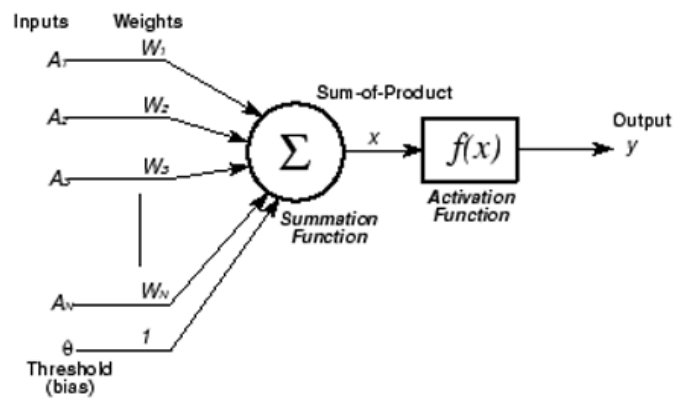
Based on the McCulloch-Pitts model described previously, the general form an artificial neuron can be described in two stages shown in Figures 3-4. In the first stage, the linear combination of inputs is calculated. Each value of input array is associated with its weight value, which is normally between 0 and 1. Also, the summation function often takes an extra input value  $\theta$  with weight value of 1 to represent threshold or *bias* of a neuron. The summation function will be then performed as,

$$x = \sum_{i=1}^N A_i W_i + \theta \quad \text{Equation 4}$$

The sum-of-product value is then passed into the second stage to perform the activation function which generates the output from the neuron. The activation function ‘squashes’ the amplitude the output in the range

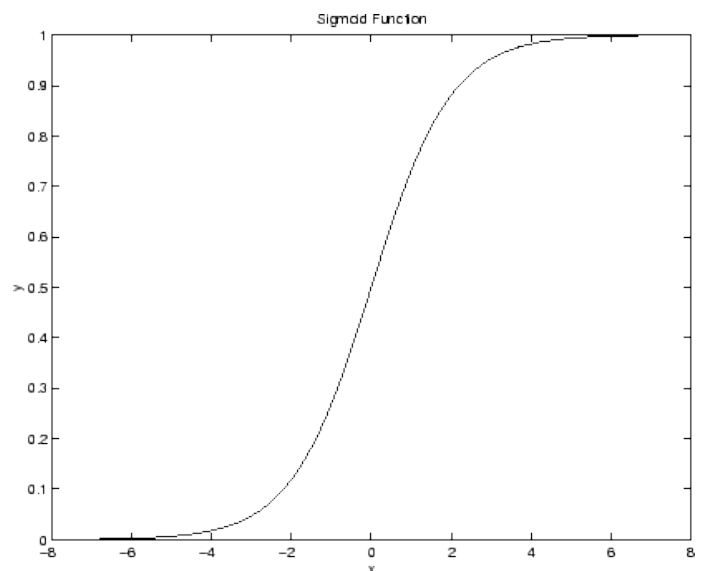
of  $[0,1]$ , or alternately  $[-1,1]$  [6]. The behavior of the activation function will describe the characteristics of an artificial neuron model.

Figure 4: More General Neural Model



The signals generated by actual biological neurons are the action-potential spikes, and the biological neurons are sending the signal in patterns of spikes rather than simple absence or presence of single spike pulse. For example, the signal could be a continuous stream of pulses with various frequencies. With this kind of observation, we should consider a signal to be continuous with bounded range. The linear threshold function should be ‘softened’[5].

Figure 5: Sigmoid Function



One convenient form of such ‘semi-linear’ function is the logistic sigmoid function, or in short, sigmoid function as shown in Figure 5. As the input  $x$  tends to large positive value, the output value  $y$  approaches to 1. Similarly, the output gets close to 0 as

$x$  goes negative. However, the output value is neither close to 0 nor 1 near the threshold point. This function is expressed mathematically as follows:

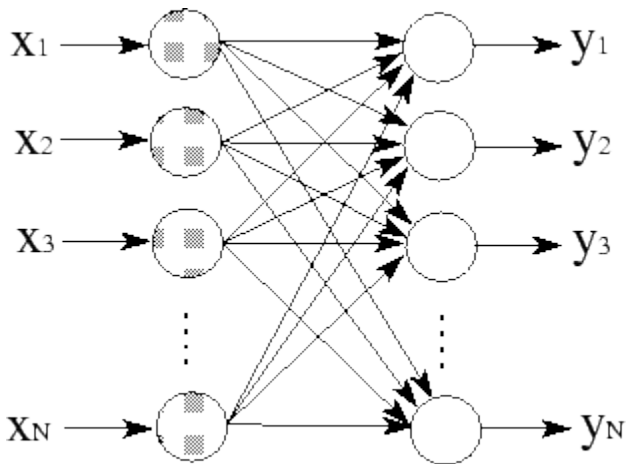
$$y = \frac{1}{1 + \exp(-x)} \quad \text{Equation 4}$$

Additionally, the sigmoid function describes the “closeness” to the threshold point by the slope. As  $x$  approaches to  $-\infty$  or  $\infty$ , the slope is zero; the slope increases as  $x$  approaches to 0. This characteristic often plays an important role in learning of neural networks.

### 1.3 Single Layer Network

By connecting multiple neurons, the true computing power of the neural networks comes, though even a single neuron can perform substantial level of computation [6]. The most common structure of connecting neurons into a network is by layers. The simplest form of layered network is shown in Figure 6. The shaded nodes on the left are in the so-called input layer. The input layer neurons are to only pass and distribute the inputs and perform no computation.

Figure 6: Single Layer Neural Network

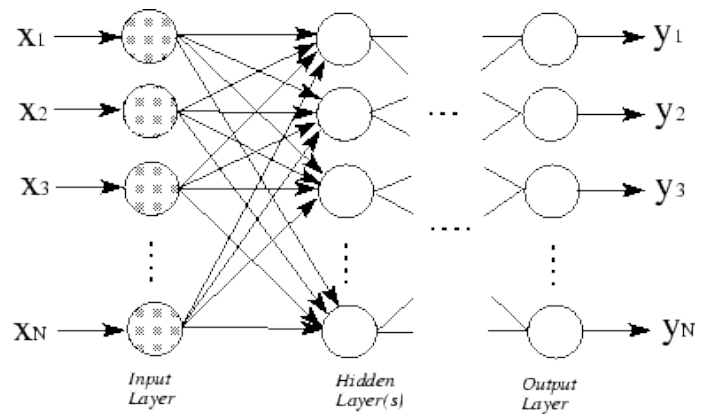


Thus, the only true layer of neurons is the one on the right. Each of the inputs  $x_1, x_2, x_3, \dots, x_N$  is connected to every artificial neuron in the output layer through the connection weight. Since every value of outputs  $y_1, y_2, y_3, \dots, y_N$  is calculated from the same set of input values, each output is varied based on the connection weights. Although the presented network is fully connected, the true biological neural network may not have all possible connections the weight value of zero can be represented as “no connection”.

### 1.4 Multilayer Network

To achieve higher level of computational capabilities, a more complex structure of neural network is required. Figure 7 shows the *multilayer neural network* which distinguishes itself from the single-layer network by having one or more hidden layers. In this multilayer structure, the input nodes pass the information to the units in the first *hidden layer*, then the outputs from the first hidden layer are passed to the next layer, and so on.

Figure 7: Multiple Layer Neural Network



Multilayer network can be also viewed as cascading of groups of single-layer networks. The level of complexity in computing can be seen by the fact that many single-layer networks are combined into this multilayer network. The designer of an artificial neural network should consider how many hidden layers are required, depending on complexity in desired computation.

### 1.5 Learning Process

Perhaps, the most primary significance of a neural network is the ability to learn the incoming information and to improve the performance of processing information. The term learning refers to many concepts by various viewpoints, and it is difficult to agree on a precise definition of the term. In neural networks, we define learning as the following sequence of events:

1. Stimulation by an environment in which the network is embedded.
2. Changes in free parameters of the network as the result of stimulation.
3. Responses in a new way to the environment for improved performance.

A Learning algorithm is a prescribed set of well-defined rules for learning of a neural network. There are

many types of learning algorithms; the common goal of learning is the adjustment of connection weights.

There are two classes of learning: supervised and unsupervised learning. Supervised learning requires an external source of information in order to adjust the network. On the other hand, in unsupervised learning, there is no external agent that overlooks the process of learning. Instead, the network is adjusted through internal monitoring of performance. In this thesis, we mainly deal with supervised learning since understanding the backpropagation network, which focuses on supervised learning, is our goal here in this paper and the Chapter 3 of book by Zohuri and Moghaddam [1].

### 1.6 Back-Propagation Neural Networks

Back-Propagation Neural Network (BPNN) employs one of the most popular neural network learning algorithms, the Backpropagation (BP) algorithm. It has been used successfully for wide variety of applications, such as speech or voice recognition, image pattern recognition, medical diagnosis, and automatic controls. One of the most striking early applications was NETTalk by T. J. Sejnowski and C. R. Rosenberg in 1986. [7] The NETTalk was able to learn the rules of phonetics, then the system produced a sound by reading from the sequence of given letters, with a behavior of a child learning to read aloud [8].

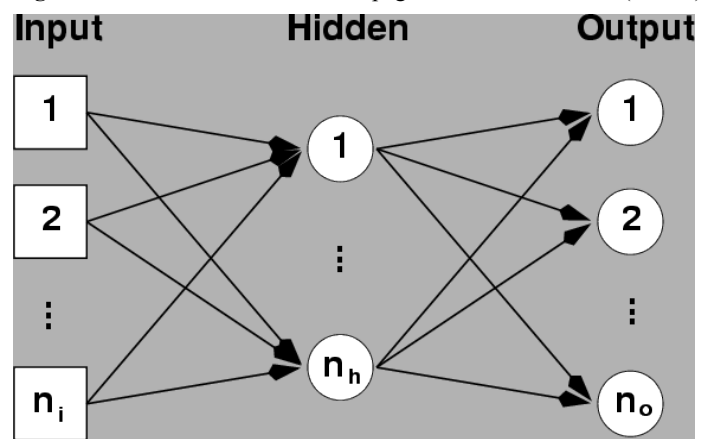
Back-propagation made a tremendous step forward from the single-layer perceptron network. With a more sophisticated learning rule, backpropagation networks overcome the limitations that single-layer networks have. Back-propagation is also the most suitable learning method for multilayer networks. Perhaps, the reason why the backpropagation made the major turning point is because the learning rule has a solid mathematical foundation and it is practical [9].

Overall the BPNN, as it is illustrated in Figure 8 was developed by Rumelhart et al. [10] as a solution to the problem of training multi-layer perceptron. The fundamental advances represented by the BPNN were the inclusion of a differentiable transfer function at each node of the network and the use of error back-propagation to modify the internal network weights after each training epoch.

As it is depicted in the Figure 8, we see a schematic of a back-propagation neural network. The back-propagation neural networks used in this work all have three layers of neurons, or nodes (input, hidden,

and output). Each node in the input and hidden layers is connected to each of the nodes in the next layer (hidden or output). All connections between nodes are directed (i.e., the information flows only one way), and there are no connections between the nodes within a particular layer. Each connection between nodes has a weighting factor associated with it. These weights are modified using the back-propagation algorithm during the training process to produce “learning”.

Figure 8: A Schematic of a Back-Propagation Neural Network (BPNN)



The BPNN was chosen as a classifier primarily because of its ability to generate complex decision boundaries in the feature space [11]. There is even work suggesting that a BPNN, under appropriate circumstances, can approximate Bayesian posterior probabilities at its outputs [12]. This is significant because a Bayesian classifier provides the best performance possible (i.e., lowest error rate) for a given distribution of the feature data. As with other non-parametric approaches to pattern classification, it is not possible to predict the performance of a BPNN a priori. Furthermore, there are several parameters of the BPNN that must be chosen, including the number of training samples, the number of hidden nodes, and the learning rate.

However, since we are interested in a real time processing of input data as fast as they come in and filter this wave data information to a trusted one and stream them to the decision-making authority based on Service Level Agreement (SLA), as it was discussed in previous chapters, it is very important we use appropriate posterior. Therefore, suggestion is leaning toward Fuzzy Logic System (FLS) as futuristic approach, for new generation of Artificial Intelligence System, in particular if we are going to place them as a Business Resilience System as it has been discussed by Zohuri and Moghaddam [12].

Based on the work of Baum and Haussler [13], it is possible to place a bound ( $m$ ) on the number of training samples needed to guarantee a particular level of performance on a set of test samples drawn from the same distribution as the training data. Specifically, if at least  $m$  samples are used to train a network with  $W$  weights and  $N$  nodes such that a fraction equal to  $1 - \varepsilon/2$  of them are classified correctly, then one can be confident that a fraction  $1 - \varepsilon$  of future (test) samples from the same distribution will be classified correctly, where

$$m \geq O\left(\frac{W}{\varepsilon} \log \frac{N}{\varepsilon}\right) \quad \text{Equation 5}$$

As a specific example, to guarantee no more than a 10% error in classifying the test data, the number of training samples should be equal to roughly 10 times the number of weights in the network. For a typical network generated below, this represents a requirement for 5000-10000 training samples. It is simply not tractable to generate that many images. Fortunately, this bound does not preclude the possibility of generating a successful classifier using fewer training samples, as many studies have empirically demonstrated.

The theoretical basis for selecting the number of hidden nodes to use in a single hidden layer network is not well developed. The only general method available to optimize this parameter is to test the network with various numbers of hidden nodes and select the one that performs best.

More details in the subject of Back-Propagation Neural Networks in reference by Zohuri and Moghaddam [1].

However, in today's growing data at Big Data level and coming for every direction, we need to deal both structured and unstructured data processing, which is described in next section here.

## 2.0 Structured and Unstructured Data Processing

The amount of data that is being created and stored on a global level is almost inconceivable, and it just keeps growing. That means there's even more potential to glean key insights from business information yet only a small percentage of data is actually analyzed. What does that mean for businesses? How can they make better use of the raw information that flows into their organizations every day? With respect to the mass categorization that is

central to most computer operations, there are two types of relevant data, which affect speed of assimilation as well as information recall: structured data and unstructured data. Smart robots need both type of data sort and process these data as fast they receive them to level of trusted degree for their processing procedure and set assignment known as service level agreement.

For the most part, structured data refers to information with a high degree of organization, such that inclusion in a relational database is seamless and readily searchable by simple, straightforward search engine algorithms or other search operations; whereas unstructured data is essentially the opposite. The lack of structure makes compilation a time and energy-consuming task. It would be beneficial to a company across all business strata to find a mechanism of data analysis to reduce the costs unstructured data adds to the organization.

Of course; if it was possible or feasible to instantly transform unstructured data to structured data, then creating intelligence from unstructured data would be easy. However, structured data is akin to machine-language, in that it makes information much easier to deal with using computers; whereas unstructured data, loosely speaking is usually for humans, who do not easily interact with information in strict, database format.

Email is an example of unstructured data; because while the busy inbox of a corporate human resources manager might be arranged by date, time or size; if it were truly fully structured, it would also be arranged by exact subject and content, with no deviation or spread which is impractical, because people do not generally speak about precisely one subject even in focused emails.

Spreadsheets, on the other hand, would be considered structured data, which can be quickly scanned for information because it is properly arranged in a relational database system. The problem that unstructured data presents is one of volume; most business interactions are of this kind, requiring a huge investment of resources to sift through and extract the necessary elements, as in a web-based search engine. Because the pool of information is so large, current data mining techniques often miss a substantial amount of the information that's out there, much of which could be game-changing data if efficiently analyzed.

Big data is a term that describes the large volume of data – both structured and unstructured that inundates a business on a day-to-day basis. But it is not the amount of data that's important. It is what organizations do with the

data that matters. Big data can be analyzed for insights that lead to better decisions and strategic business moves.

Harvesting a platform that provides a robust solution for collecting both structured and unstructured data from the Internet (Figure 9). We need to take a unique approach to “connecting” those unconnected strands of information through the use of metadata. We need harvesting technology that employs multiple threads to mass-harvest scalable quantities of unstructured data such as web-based central processing, if you will. Harvests should be based on multiple user-developed queries with results (web pages, PDF’s, XLS, PPT, XML, etc.) qualified through customizable filters. This should be developed four scoring algorithms that index the information based on relevancy to further qualify the documents returned, ensuring the user is seeing only super-relevant content. The final user interface displays the qualified results in a searchable database based on customizable facets (URL, file-type, source category, people mentioned, places mentioned, companies mentioned, custom keywords, etc.) [1].

Figure 9: Presentation of Unstructured Data Similarity



Finding a way to analyze and create intelligence from the wealth of unstructured data available on the Web can be expected to endow an organization with the direct benefit of drastic increases in overall effectiveness and speed of decision-making and implementation.

## 2.1 Big Data Infrastructure Demands

The need for big data velocity imposes unique demands on the underlying compute infrastructure. The computing power required to quickly process huge volumes and varieties of data can overwhelm a single server or server cluster. Organizations must apply adequate compute power to big data tasks to achieve the desired velocity. This can potentially demand hundreds or thousands of servers that can distribute the work and operate collaboratively.

Achieving such velocity in a cost-effective manner is also a headache. Many enterprise leaders are reticent to invest in an extensive server and storage infrastructure that might only be used occasionally to complete big data tasks. As a result, public cloud computing has emerged as a primary vehicle for hosting big data analytics projects. A public cloud provider can store petabytes of data and scale up thousands of servers just long enough to accomplish the big data project. The business only pays for the storage and compute time actually used, and the cloud instances can be turned off until they’re needed again.

To improve service levels even further, some public cloud providers offer big data capabilities, such as highly distributed Hadoop compute instances, data warehouses, databases and other related cloud services. Amazon Web Services Elastic Map Reduce is one example of big data services in a public cloud.

## 2.2 The Human Side of Big Data

However, these tools only address limited use cases. Many other big data tasks, such as determining the effectiveness of a new drug, can require substantial scientific and computational expertise from the analytical staff. There is currently a shortage of data scientists and other analysts who have experience working with big data in a distributed, open source environment.

Big data can be contrasted with small data, another evolving term that’s often used to describe data whose volume and format can be easily used for self-service analytics. A commonly quoted axiom is that “big data is for machines; small data is for people”.

## 2.3 Big Data History and Current Considerations

As we said at the beginning of this chapter, the Big Data is a term that describes the large volume of data both structured and unstructured that inundates a business on a day-to-day basis. But it is not the amount of data that's important. It's what organizations do with the data that matters. Big data can be analyzed for insights that lead to better decisions and strategic business moves.

While the term "big data" is relatively new, the act of gathering and storing large amounts of information for eventual analysis is ages old. The concept gained momentum in the early 2000s when industry analyst Doug Laney articulated the now-mainstream definition of big data as the three Vs as it was described before:

### Volume

Organizations collect data from a variety of sources, including business transactions, social media and information from sensor or machine-to-machine data. In the past, storing it would've been a problem but new technologies (such as Hadoop) have eased the burden.

### Velocity

Data streams in at an unprecedented speed and must be dealt with in a timely manner. RFID tags, sensors and smart metering are driving the need to deal with torrents of data in near-real time.

### Variety

Data comes in all types of formats from structured, numeric data in traditional databases to unstructured text documents, email, video, audio, stock ticker data and financial transactions.

At SAS, we consider two additional dimensions when it comes to big data:

### Variability

In addition to the increasing velocities and varieties of data, data flows can be highly inconsistent with periodic peaks. Is something trending in social media? Daily, seasonal and event-triggered peak data loads can be challenging to manage. Even more so with unstructured data.

### Complexity

Today's data comes from multiple sources, which makes it difficult to link, match, cleanse and transform data across systems. However, it's necessary to connect

and correlate relationships, hierarchies and multiple data linkages or your data can quickly spiral out of control.

## 2.4 Big Data's Big Potential

The amount of data that's being created and stored on a global level is almost inconceivable, and it just keeps growing. That means there's even more potential to glean key insights from business information – yet only a small percentage of data is actually analyzed. What does that mean for businesses? How can they make better use of the raw information that flows into their organizations every day?

## 2.5 Why is Big Data Important?

The importance of big data does not revolve around how much data you have, but what you do with it. You can take data from any source and analyze it to find answers that enable:

- 1) Cost reductions,
- 2) Time reductions,
- 3) New product development and optimized offerings, and
- 4) Smart decision making. When you combine big data with high-powered analytics, you can accomplish business-related tasks such as:

- ✓ Determining root causes of failures, issues and defects in near-real time.
- ✓ Generating coupons at the point of sale based on the customer's buying habits.
- ✓ Recalculating entire risk portfolios in minutes.
- ✓ Detecting fraudulent behavior before it affects your organization.

## 2.6 Who Uses Big Data?

Big data affects organizations across practically every industry. See how each industry can benefit from this onslaught of information.

### 1. Banking

With large amounts of information streaming in from countless sources, banks are faced with finding new and innovative ways to manage big data. While it's important to understand customers and boost their satisfaction, it's equally important to minimize risk and fraud while maintaining regulatory compliance. Big data brings big insights, but it also requires financial institutions to stay one step ahead of the game with advanced analytics.

## 2. Government

When government agencies are able to harness and apply analytics to their big data, they gain significant ground when it comes to managing utilities, running agencies, dealing with traffic congestion or preventing crime. But while there are many advantages to big data, governments must also address issues of transparency and privacy.

## 3. Manufacturing

Armed with insight that big data can provide, manufacturers can boost quality and output while minimizing waste processes that are key in today's highly competitive market. More and more manufacturers are working in an analytics-based culture, which means they can solve problems faster and make more agile business decisions.

## 4. Education

Educators armed with data-driven insight can make a significant impact on school systems, students and curriculums. By analyzing big data, they can identify at-risk students, make sure students are making adequate progress, and can implement a better system for evaluation and support of teachers and principals.

## 5. Health Care

Patient records. Treatment plans. Prescription information. When it comes to health care, everything needs to be done quickly, accurately and, in some cases, with enough transparency to satisfy stringent industry regulations. When big data is managed effectively, health care providers can uncover hidden insights that improve patient care.

## 6. Retail

Customer relationship building is critical to the retail industry and the best way to manage that is to manage big data. Retailers need to know the best way to market to customers, the most effective way to handle transactions, and the most strategic way to bring back lapsed business. Big data remains at the heart of all those things.

For example we can see Big Data in action at United Parcel Service (UPS). As a company with many pieces and parts constantly in motion, UPS stores a large amount of data much of which comes from sensors in its vehicles. That data not only monitors daily performance, but also triggered a major redesign of UPS drivers' route

structures. The initiative was called ORION (On-Road Integration Optimization and Navigation), and was arguably the world's largest operations research project. It relied heavily on online map data to reconfigure a driver's pickups and drop-offs in real time. Figure 10 is illustration of typical logistic for United Parcel Service (UPS).

Figure 10: UPS Logistics



The project led to savings of more than 8.4 million gallons of fuel by cutting 85 million miles off of daily routes. UPS estimates that saving only one daily mile per driver saves the company \$30 million, so the overall dollar savings are substantial.

It is important to remember that the primary value from big data comes not from the data in its raw form, but from the processing and analysis of it and the insights, products, and services that emerge from analysis. The sweeping changes in big data technologies and management approaches need to be accompanied by similarly dramatic shifts in how data supports decisions and product/service innovation.

### 2.7 How It Works?

Before discovering how big data can work for your business, you should first understand where it comes from. The sources for big data generally fall into one of three categories:

#### 1. Streaming Data

This category includes data that reaches your IT systems from a web of connected devices. You can analyze this data as it arrives and make decisions on what data to keep, what not to keep and what requires further analysis.

#### 2. Social Media Data

The data on social interactions is an increasingly attractive set of information, particularly for marketing,

sales and support functions. It's often in unstructured or semi-structured forms, so it poses a unique challenge when it comes to consumption and analysis.

### 3. Publicly Available Sources

Massive amounts of data are available through open data sources like the US government's data.gov, the CIA World Fact book or the European Union Open Data Portal.

After identifying all the potential sources for data, consider the decisions you'll need to make once you begin harnessing information. These include:

#### 1. How to Store and Manage it

Whereas storage would have been a problem several years ago, there are now low-cost options for storing data if that's the best strategy for your business.

#### 2. How Much of it to Analyze

Some organizations don't exclude any data from their analyses, which is possible with today's high performance technologies such as grid computing or in-memory analytics. Another approach is to determine upfront which data is relevant before analyzing it.

#### 3. How to Use any Insights You Uncover

The more knowledge you have, the more confident you'll be in making business decisions. It's smart to have a strategy in place once you have an abundance of information at hand.

The final step in making big data work for your business is to research the technologies that help you make the most of big data and big data analytics. Consider:

- Cheap, abundant storage.
- Faster processors.
- Affordable open source distributed big data platforms, such as Hadoop.
- Parallel processing, clustering, MPP, virtualization, large grid environments, high connectivity and high throughputs.
- Cloud computing and other flexible resource allocation arrangements

Training the many layers of virtual neurons in the experiment took 16,000 computer processors the kind of computing infrastructure that Google has developed for its search engine and other services. At least 80 percent of the

recent advances in AI can be attributed to the availability of more computer power, reckons Dileep George, cofounder of the machine-learning startup Vicarious.

There's more to it than the sheer size of Google's data centers, though. Deep learning has also benefited from the company's method of splitting computing tasks among many machines so they can be done much more quickly. That's a technology Dean helped develop earlier in his 14-year career at Google. It vastly speeds up the training of deep-learning neural networks as well, enabling Google to run larger networks and feed a lot more data to them.

Already, deep learning has improved voice search on smartphones. Until last year, Google's Android software used a method that misunderstood many words. But in preparation for a new release of Android last July, Dean and his team helped replace part of the speech system with one based on deep learning. Because the multiple layers of neurons allow for more precise training on the many variants of a sound, the system can recognize scraps of sound more reliably, especially in noisy environments such as subway platforms. Since it's likelier to understand what was actually uttered, the result it returns is likelier to be accurate as well. Almost overnight, the number of errors fell by up to 25 percent—results so good that many reviewers now deem Android's voice search smarter than Apple's more famous Siri voice assistant.

For all the advances, not everyone thinks deep learning can move artificial intelligence toward something rivaling human intelligence. Some critics say deep learning and AI in general ignore too much of the brain's biology in favor of brute-force computing.

One such critic is Jeff Hawkins, founder of Palm Computing, whose latest venture, Numenta, is developing a machine-learning system that is biologically inspired but does not use deep learning. Numenta's system can help predict energy consumption patterns and the likelihood that a machine such as a windmill is about to fail. Hawkins, author of *On Intelligence*, a 2004 book on how the brain works and how it might provide a guide to building intelligent machines, says deep learning fails to account for the concept of time. Brains process streams of sensory data, he says, and human learning depends on our ability to recall sequences of patterns: when you watch a video of a cat doing something funny, it's the motion that matters, not a series of still images like those Google used in its experiment. "Google's attitude is: lots of data makes up for everything," Hawkins says.

But if it does not make up for everything, the

computing resources a company like Google throws at these problems can't be dismissed. They're crucial, say deep-learning advocates, because the brain itself is still so much more complex than any of today's neural networks. "You need lots of computational resources to make the ideas work at all.

## 2.8 Real Time Data Processing and Data Mining

Data mining is the process of selecting and exploring data to discover previously unknown patterns and historical trends that can be used to develop models for predicting future outcomes.

Demystifying Data Mining applies data mining, predictive modeling and real-time analytics in government homeland security, banking, retails, oil and gas operations, etc. Mining large reservoirs of data in oil and gas operations involves committing to key processes and technologies – and embracing new ways of thinking about problem solving. To extract value from vast data stores and change the way decisions are made, many operators have turned to advanced data mining techniques along with real-time analytical and data processing capabilities. This paper explores practical approaches, workflows and techniques that are used in oil and gas operations. It also examines the role of exploratory data analysis; model development and modeling techniques; and approaches to putting models into production.

Data mining is the process of selecting and exploring data to discover previously unknown patterns and historical trends that can be used to develop models for predicting future outcomes. Quantitative techniques uncover patterns and relationships in data that are used to build descriptive and predictive models.

The terms data mining and predictive modeling are often used interchangeably but they are distinct. Data mining is the process of uncovering patterns in a sample set of data and then developing models that find the same desired pattern across a much larger universe of data. Predictive modeling is the process of applying these models during the course of a business process to predict an outcome.

Analytical technologies have made it possible to understand massive amounts of data to assist in decision making across the enterprise. These same technologies advanced data mining techniques and real time analytical and data processing capabilities can change the way organizations make decisions. And when improved decision making is applied in a structured manner, it can

yield significant returns. When these time-tested tools and techniques are adopted by business analysts, they can enable rapid results without requiring the analysts to obtain advanced statistical training.

The scope of activities related to data mining and predictive modeling includes:

- Data preparation to merge multiple data sets, resolve missing values or outliers, and reformat data as needed.
- Exploratory data analysis to discover relationships and anomalies in the data.
- Variable transformation, enrichment and selection to better focus the modeling process.
- Model building using competitive algorithms to search for data combinations that reliably predict the outcome.
- Testing and validation of the champion model to ensure that the model generates output as expected when applied to new data.
- Putting models into production in applications and databases to optimize business processes and improve business decisions.
- Monitoring the model performance to ensure the model is predicting well and does not need to be recalibrated.

## 2.9 Behind the Scenes with Data Mining

Most data aggregation issues are not technical, but rather are related to domain knowledge and data ownership.

The process of data mining includes several key activities. These include data aggregation and preparation, exploratory data analysis, modeling and deployment of models in production environments.

Data preparation is typically 80 percent of the effort of an analytical project. One reason is that many organizations lack a single source of complete, high-quality data required for an analytical exercise. Some choose to wait for the completion of a corporate data warehouse that promises to organize, arrange, standardize and clean the data. Unfortunately, these warehouses seldom address all of the relevant data that may be critical to the analytical problem.

Preparing data for data mining should result in an analytical base table, or data set, that has variables associated with the problem being modeled. This data is prepared very differently than a warehouse for historical reporting because it is gathered from many more databases and stored in very different subsets that are tailored for the analysis at hand.

Aggregation and preparation of data for an analytical project must be conducted by members of the team who have the relevant skill set. Most data aggregation issues are not technical, but rather are related to domain knowledge and data ownership. Once the business logic is provided by the domain experts, the process of accessing, joining and reformatting data can be performed by technical staff.

Data standardization and data quality also require collaboration between domain and technical staff. For instance, standardizing a supplier name across multiple systems is not likely to affect the result of an analytical exercise. However, decisions about how to fill missing range values in sensor tag information or how to collapse time measurement intervals should be made by the domain expert advising the analyst.

Once the workflow for aggregating and cleansing the relevant data is determined, the processes and rules related to conducting these tasks can be automated. Automation of data preparation can be done in real time when the data is created, or at intervals that are appropriate to the data's specific, time-sensitive requirements.

### 2.10 Modeling Techniques

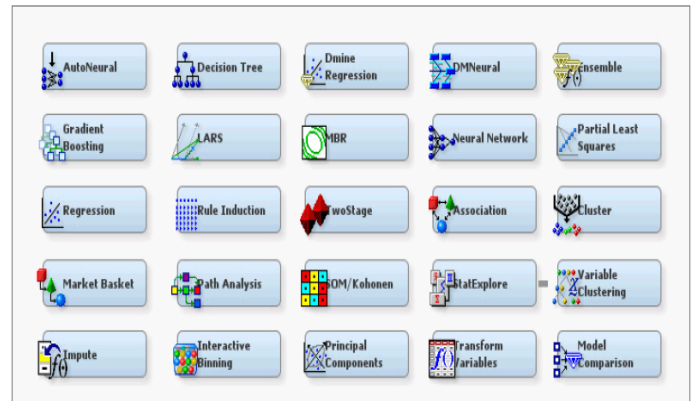
Which technique will produce the best model and therefore the best prediction. Some of these modeling techniques include decision trees, neural networks, least-angle regressions, logistic regressions, memory-based reasoning and rule induction.

A software solution with multiple modeling techniques allows an analyst to quickly and easily apply a particular modeling technique to a data set and interactively work with the parameters to try different configurations. This capability allows modelers to test many different approaches while relying on the software to pick the most accurate model, based on a set of user-selected statistical diagnostic tests. It is also important to be able to build an ensemble model by combining techniques if the combination of two models is more effective than a single model. Advanced data mining software that provides these capabilities allows an analyst to spend time developing true insights as opposed to programming analytical models.

Figure 11 is illustration of analytical methods that may apply to oil field as an example.

Data mining is most effective when deployed as part of an integrated information delivery strategy that is supported by strong business domain specialists, Information Technology (IT) and skilled analysts.

Figure 11: Examples of Analytical Methods that Might be Applied to Oilfield Data



### 2.11 Deploying Models in a Scalable Environment

Significant business value is lost without the ability to put analytical models into production at an enterprise scale. But putting models into production involves key enabling technologies that must be flexible and robust enough to support the various technical requirements dictated by the business problem. The technical solution must support data access regardless of size and source and must provide flexible scalability regardless of computational complexity or the time window established to return a computed value.

Data mining is most effective when deployed as part of an integrated information delivery strategy that is supported by strong business domain specialists, IT and skilled analysts. Determining who leads for each data mining exercise requires consideration of the target audience, the timing of the results, and the anticipated action as a result of the analytical insight. This multidisciplinary approach ensures that the project can have technical success and also generate business value.

The technical infrastructure selected for both modeling and production must be able to support modeling at scale. When modeling large data sets, it is important to be able to process large data sets with computationally intensive tasks and to visualize large data sets that support exploratory analysis. In the case of very large data sets or complex computational processes, it may be necessary to use high-performance computing to return value within a tight time frame.

Just as business outcomes shift with the economy, data constantly changes. As a result, companies should establish models that natively adapt to changes in the data without significant human intervention. It is possible for key variables in a multivariate predictive model to shift

in a statistically significant way not anticipated by the model. The ability to monitor and alert key stakeholders to ongoing model performance is the final step in deploying advanced analytical models to scale.

In summary, Data mining technologies can be utilized to exploit vast amounts of data to yield significant results in short time frames. SAS has helped many organizations apply advanced analytics to achieve significant benefits in the digital oilfield. This journey from data exploration to optimized decisions is achievable and can be deployed at scale in oilfield operations.

### 3.0 Improving Big Data Analytics with Machine Learning-as-a-Service

Machine Learning-as-a-Service (MLaaS) exists as the nexus point for some of the most promising technologies and applications of Big Data analytics. The availability of sophisticated Machine Learning algorithms on demand via the Cloud has significant ramifications for [1]:

**Big Data:** In many ways, the predictive capabilities of Machine Learning are the only means by which the enterprise can make use of all its data. MLaaS can enable organizations to combine structured data with unstructured, external data to automate analytics processes that would otherwise take too long to parse with Big Data sets.

**Cognitive Computing:** As one of the central components in Cognitive Computing, Machine Learning helps organizations take a more cognitive approach to their Data Management by producing applications with cognitive capabilities.

**Data Science:** Machine Learning-as-a-Service has the propensity to both reduce the need for scarcely found Data Scientists and substantially assist them in basing models and business objectives on an organization's data. In either case it eliminates the need to create each individual algorithm.

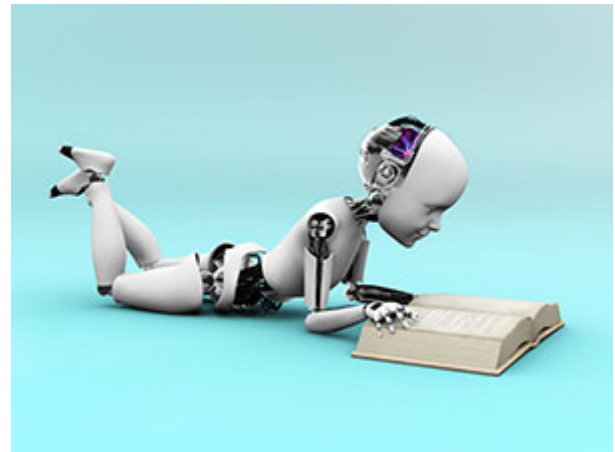
**Application Development:** The capability of MLaaS to accelerate the Data Science process ensures that developers can create better applications more expeditiously to derive near real-time action from analytics. Figure 12 is illustration of Smart Robot as learning machine

**Data Modeling:** Once initial algorithms are construed;

Machine Learning automates the Data Modeling process by producing models on both present and future data to expedite what otherwise would be a time-consuming affair.

**Cloud Computing:** The speed and convenience of MLaaS are accessed through the Cloud, which reinforces this medium as the architecture of choice for Big Data.

Figure 12: The Smart Robot



The cumulative effect is that Machine Learning-as-a-Service expands on the possibilities of Big Data analytics while making them more accessible than ever before. James Kobiulus of IBM recently indicated that:

“These value points derive from Machine Learning’s core function: enabling analytics algorithms to learn from fresh feeds of data without constant human intervention and without explicit programming.”

Machine learning is a type of Artificial Intelligence (AI) that provides computers with the ability to learn without being explicitly programmed. Machine learning focuses on the development of computer programs that can change when exposed to new data.

The process of machine learning is similar to that of data mining. Both systems search through data to look for patterns. However, instead of extracting data for human comprehension as is the case in data mining applications machine learning uses that data to detect patterns in data and adjust program actions accordingly. Machine learning algorithms are often categorized as being supervised or unsupervised. Supervised algorithms can apply what has been learned in the past to new data. Unsupervised algorithms can draw inferences from datasets.

For example, Facebook’s News Feed uses machine learning to personalize each member’s feed. If a member frequently stops scrolling in order to read or “like” a

particular friend's posts, the News Feed will start to show more of that friend's activity earlier in the feed. Behind the scenes, the software is simply using statistical analysis and predictive analytics to identify patterns in the user's data and use to patterns to populate the News Feed. Should the member no longer stop to read, like or comment on the friend's posts, that new data will be included in the data set and the News Feed will adjust accordingly.

### 3.1 Application Development

Some of the most eminent Cloud service providers (Amazon, Google, Microsoft, IBM) are offering MLaaS either independently or as part of other platforms. Twitter recently underscored the importance of Machine Learning by acquiring Whet lab, a Machine Learning startup. Perhaps the most immediate of the aforementioned ramifications of MLaaS is the fact that it enables developers to readily incorporate Machine Learning into their applications. Most providers deliver MLaaS as an offering used expressly in conjunction with their Clouds, which underpins the need to facilitate Big Data applications off premises. More importantly, the ability of developers to utilize Machine Learning algorithms in their applications reduces the reliance on Data Scientists and the complexity associated with creating these algorithms. Applications involving Machine Learning include any assortment of uses from fraud detection and recommender engines to pattern mining, clustering, and other aspects of e-commerce that hinge on Big Data analytics. MLaaS enables developers to take on more responsibility for Data Modeling and Data Science.

### 3.2 Data Science

In addition to reducing the need for Data Scientists, Machine Learning-as-a-Service (MLaaS) can significantly assist these professionals by increasing the complexity of the underlying analytics algorithms that empower the business. One of the key repercussions of MLaaS is that the increased availability of Machine Learning makes it and Big Data less individual concerns, and more integral to overall processes in Data Management. Without having to create each and every algorithm for each and every application, Data Scientists can tackle more profound aspects of *Machine Learning* such as *Deep Learning* and neural networks. In this way, MLaaS becomes as valuable a tool for Data Scientists as it is for organizations without Data Scientists since it enables the former to expand on the capabilities of this discipline within Data Management.

### 3.3 Cloud Vitality

All of the typical benefits of utilizing the Cloud apply to MLaaS: reduced cost, less infrastructure, increased time to value. Additionally, MLaaS offers the sort of boons that other specialized analytics services such as Graph Analytics-as-a-Service (See next section) provide. Those include a conservation of resources dedicated to hiring specialized personnel and the ability to simply outsource difficult analytics work and its analysis. Certain MLaaS vendors, for instance, will not only construct predictive models and algorithms, but also provide analysis of an organization's data as part of their services. Although each provider's process and capabilities differ, most base charges on the amount of data and the length of time service is used some have specific prices for different Machine Learning functions. Virtually all of them involve APIs to make data machine readable and provide a framework for basic algorithms and models that organizations can tailor to varying degrees. An assortment of visualization tools and programming languages is supported, while there is a minimal reliance on code.

### 3.4 The Point of Big Data: The Internet of Things

Machine Learning and MLaaS are projected to play an integral role in the facilitation of the Internet of Things (IoT). The IoT is arguably the ultimate expression of Big Data. It enables perpetual connectivity and constant streaming of data from any variety of gadgets, from the industrial to the personal. It would be virtually impossible for a team of Data Scientists to continually refine the algorithms and models required for real-time and predictive analytics of the immense quantities of data involved in the IoT. The problem with the IoT that Machine Learning ameliorates is the need to not only account for potentially billions of sensors and their constant streaming, but to analyze them in way that produces timely action. Virtually the only way to do so is to build future models and algorithms from historic and real-time data, which is what Machine Learning does. MLaaS expedites that process.

A thing, in the Internet of Things, can be a person with a heart monitor implant, a farm animal with a biochip transponder, and automobile that has built-in sensors to alert the driver when the tire pressure is low--or any other natural or manmade object that can be assigned an IP address and provided with the ability to transfer data over a network.

### 3.5 The Internet of Things

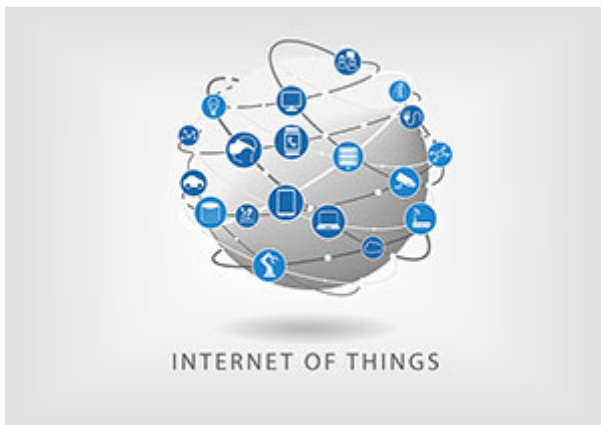
The Internet of Things (IoT) is slated to transform the nature of transactional data from a rigid, performance optimized process, to a dynamic, on-the-fly zone capable of handling the pace and variety of Big Data. There are several reasons to believe Apache Spark will play a formidable role in this transformation:

**In-Memory analytics:** Spark utilizes an in-memory analytics approach that is typically faster than most disk-based methods so much so that it has been reported as significantly outperforming Hadoop.

**Machine learning iterations:** This Big Data parallel-processing framework is primed for Machine Learning algorithms that utilize its swiftness for the sort of rapid iterations required to draw conclusions from and make improvements with immense data sets.

**Hadoop compatible:** Although some view it as competitive with Hadoop, Spark was actually designed to run on HDFS (but it does not have to). In doing so it can add substantial value to one of the most established and widely used Big Data platforms Figure 13.

Figure 13: Internet of Things



With vendor support for Spark including Map R, Cloudera, IBM, Data Stax, Intel, Hortonworks, and other Big Data platforms, its effect on Big Data, the IoT, and transactional data is arguably burgeoning throughout the data sphere.

As IoT becomes more wide spread it will attract more potential for cyber-attacks and fraud. The vast quantity of data that will flow between the connected vehicle, connected home and the insurance company is vulnerable to interception. The new IoT products are also likely to lead to new types of application and claims fraud. Insurers will need to invest more heavily in data security and fraud protection.

While realizing the full potential of IoT for insurance will not be without its challenges, its early exploitation is already producing positive results. IoT undoubtedly makes losses easier to predict and prevent. Smart home devices, wearables and the imminent arrival of the driverless car will usher in a shift toward a new type of customer relationship where insurance will become less reactive and more preventative. The winners will be organizations that overcome today's obstacles to embrace change and capitalize on uncertainty.

The Internet of Things (IoT) is a collection of network-connected physical objects and machines. They have embedded identifications, sensors, and software that can provide an understanding of where they are, what they're doing, and what's going on around them. These devices can communicate with each other and share their data via a network or a cloud-based platform. Examples we hear about most often include a power company's "smart grid" with sensors collectively managing the flow of and demand for electricity, and an individual's "smart home" with climate control, lighting, and security adjusted automatically and remotely.

"The IoT does three things for retailers," said Lori Schafer, Executive Adviser for Retail and CPG at SAS, speaking recently at SAS® Global Forum. "It helps you sense who the customers are and what they're doing; it helps you better understand customer behavior and buying practices; and it allows you to then act upon those insights to create a better overall experience for the customer."

We are moving from a more connected world to a more informed world. And if you're a retailer that isn't informed, you're going to be at a big disadvantage.

### 3.6 Additional Benefits

Additionally, Spark has attracted so much attention in the Big Data space because of its benefits:

**Reduced infrastructure:** Part of the utility of the high-performance speeds of this platform is that they are generated while requiring fewer nodes and clusters than those of more traditional Big Data platforms.

**Open Source aspect:** The open source nature of the solution yields up front cost advantages.

**Application building:** The diversity of code options greatly assists with application building and decreasing time to value.

Accelerated performance: Spark's tremendous speed not only facilitates real time analytics but enables users to incorporate many more data sources than they otherwise would be able to.

Data Science and developer friendly nature: Although Spark was written in Scala, it contains APIs for popular programming languages such as SQL, Java, and Python. The assortment of code languages is exploited by the platform's underlying engine, which can determine how to complete jobs on a cluster, thereby 'hiding' some of the complexity associated with writing code.

### 3.7 Potential Drawbacks

Part of the reason that Spark is unlikely to supplant Hadoop and its ecosystem for the present includes the fact that it:

Lacks autonomy: Although additional components of the Spark platform include a Semantic Graph analytics engine, a Machine Learning library, and tools for data streaming, its autonomy is circumscribed by a lack of a resource manager or file system, which can simply increase its dependency on Hadoop.

Is still immature: Having influenced the data landscape over the past five years, Spark still remains a technology that is in the process of maturing.

Lacks use cases: Enterprise-level deployments of Spark that have demonstrated consistency and performance strengths over time are still in the process of materializing.

Is not primed for operations: Primed for analytics, Spark is less effective in operations due to rigidity in its core abstraction that produces difficulty in making a single change in a data set, and which frequently requires copying the entire set to do so.

### 3.8 Internet of Things Transactions

Current real-time Big Data applications that incorporate elements of both analytics and transactions include standards such as fraud detection, recommender engines, or any variety of Industrial Internet applications pertaining to equipment asset management and preventative maintenance. These applications hint at the value that Spark can provide and the way that the IoT will affect transactions, because they offer a starting point for how the former can already be used.

What the Internet of Things is helping to herald is sophistication in the nature of transactions in which any assortment of analytics (real-time, predictive, and even historic) is required to not only provide insight, but also

actually trigger action that expands upon the possibilities of most current transactions. For instance, while most real-time fraud detection systems simply provide approval or denial based on a finite number of factors, the shift in transactions the IoT is causing can greatly increase the number of those factors. For the consumer, that might mean options for a smart vehicle to obtain gas or for a smart home to reduce power consumption. Regardless of the individual circumstance, the rapidly iterative nature of Spark's extreme expedience to facilitate Machine Learning algorithms will be well served.

### 3.9 Internet on Things Applications Retailers

With IoT, we can now understand the context (the time and the place of the customer) to identify when we are certain the customer needs help or an incentive to purchase, and we can respond proactively.

Key applications of IoT for retailers include supply chain, connected consumer and smart-store applications. In particular, let's look at five areas where retailers are taking advantage of IoT:

1. Predictive equipment maintenance is used for managing energy, predicting equipment failure or detecting other issues. For example, every grocery store has a lot of complex equipment – most people recognize refrigeration units. When these units are instrumented with sensors, we can predict maintenance issues that might affect power consumption for savings or monitor temperature fluctuations to ensure food safety.

2. Moving merchandise more efficiently is one of the goals of smart transportation applications in retail, and IoT can come into play with the maintenance of transport, tracking and route optimization. We know many retailers have been using GPS to track and route trucks in the last couple of years. With IoT, we are able to understand to a much higher degree of accuracy how close a pallet of merchandise is to a given store.

3. When it comes to demand-aware warehouse fulfillment, we're talking about warehouse automation and robotics driven by online and in-store shopping demand. IoT allows us to monitor sales opportunities in real time and track missed in-store sales. It is important to remember that Radio-Frequency Identification (RFID) is a well-tested part of IoT that can be used for inventory management and more accurate service-level optimization. Currently, a typical

distribution center or warehouse is organized by aisles and shelves based on a fixed schematic. The warehouse of the future will be open space where automated pallets self-organize based on real-time demand.

4. Increasingly, the connected consumer is having an impact on brick-and-mortar locations. Retailers understand that customers are able to check in-store pricing and local inventory levels from their mobile devices. Imagine if we could make a customized best-price offer or provide location-based services right in the store. What if we could target our high-value, loyal customers with concierge services? In the past, it was accepted as the norm that we would send mass promotions to customers with the expectation that some acceptable percentage might be interested in that promotion. With IoT, we can now understand the context (the time and the place of the customer) to identify when we are certain the customer needs help or an incentive to purchase, and we can respond proactively.

5. In a smart store, mall traffic can be analyzed across several retailers, so we understand the entire shopping journey. In the past, we had to run expensive survey projects to understand if store associates were being responsive to customer service needs and then enact elaborate staff training programs. Now, within smart stores, we will be able to use video or Wi-Fi foot-traffic monitoring to see if customers dwell over a product area. Then, in real time, direct an associate to help that customer or analyze that information later to adjust store layouts for more efficient customer visits. In addition, by monitoring store traffic and customer demand in real time, we can customize the current in-store shopping experience. That gives us the opportunity to implement rich digital marketing inside the store or announce events to customers via their mobile devices.

With the rapid growth of online shopping, retailers are very keen to bring the frictionless customer experience of online shopping into the store wherever they can. They want access to the same type of rich data and high-performance analytics that retailers use to drive websites and mobile shopping trips. Their goal is to have that same limitless control to craft a customer experience and collect detailed data to help them predict how customers will shop.

The differentiation with IoT will come from a retailer's ability to sense, understand and act on IoT data

with analytics. It won't be in the technology, the devices or the IoT plumbing. To take advantage of this new promising area, retailers should focus on IoT applications that better serve customers and create value.

Note that RFID works better than Barcodes. A significant advantage of RFID devices over the others mentioned above is that the RFID device does not need to be positioned precisely relative to the scanner. We're all familiar with the difficulty that store checkout clerks sometimes have in making sure that a barcode can be read. And obviously, credit cards and ATM cards must be swiped through a special reader.

In contrast, RFID devices will work within a few feet (up to 20 feet for high-frequency devices) of the scanner. For example, you could just put all of your groceries or purchases in a bag, and set the bag on the scanner. It would be able to query all of the RFID devices and total your purchase immediately. (Read a more detailed article on RFID compared to barcodes.)

RFID technology has been available for more than fifty years. It has only been recently that the ability to manufacture the RFID devices has fallen to the point where they can be used as a "throwaway" inventory or control device. Alien Technologies recently sold 500 million RFID tags to Gillette at a cost of about ten cents per tag.

One reason that it has taken so long for RFID to come into common use is the lack of standards in the industry. Most companies invested in RFID technology only use the tags to track items within their control; many of the benefits of RFID come when items are tracked from company to company or from country to country.

However, common problems with RFID that are worth to pay attention to are *reader collision* and *tag collision*. Reader collision occurs when the signals from two or more readers overlap. The tag is unable to respond to simultaneous queries. Systems must be carefully set up to avoid this problem. Tag collision occurs when many tags are present in a small area; but since the read time is very fast, it is easier for vendors to develop systems that ensure that tags respond one at a time.

### 3.10 Cloud Ramifications

Spark's utility extends beyond the Internet of Things and its sophisticated transactions and includes a number of repercussions for the Cloud as well, which has become the de facto home for Big Data applications. Most of its IoT transaction capabilities can apply to general Cloud

deployments where its advantages of combining graph analytics, SQL, and sensor data processing or streaming are most useful especially with time-sensitive data requirements. The nexus points between IoT transactions and the Cloud that Spark can play an integral role in is the expansion of Clouds to what amount to IoT Clouds. Labeled as ‘fogs’, IoT clouds help to create a situation in which a Cloud’s core resources (such as bandwidth and storage) are decentralized and pushed to its extremities in order to aggregate these resources across Clouds. In the case of the IoT, all of the interconnected devices will have their own Clouds that are effectively aggregated to conserve resources and reinforce interconnectivity which will require the sort of Machine Learning algorithms and graph analytics that Spark can readily provision.

### 3.11 Data Streaming

Another vital aspect of Spark is its tools for data streaming and accommodating continuously generated machine or sensor data. Although there is no shortage of options for streaming data, one of the distinguishing points of Spark’s tool is that it processes events in a batch-like method during brief intervals of time in which events are continuously collected. Although Spark’s streaming capabilities represent just one of many options of this technology that is still maturing, it helps to provide a gestalt of sorts with its other non-streaming capabilities (discussed above). Thus, the platform ensures that it can still enable the sort of speed necessary for real-time analytics on Big Data sets that are nimble enough for transactions in an Internet of Things world.

### 3.12 Under the Hood

At the abstraction level, Spark consists of a number of collections known as Resilient Distributed Datasets (RDD), which are intractable and based on local files (such as HDFS or others). There are several collection operations common to solutions written in Scala that apply to RDD, including parallelized versions of for each and map. Other functions include capabilities for reduce-by which consolidates entries according to a certain function pertaining to a specific key and join capabilities that collect entries from a pair of RDD based on a shared key. Of particular use is the fact that Spark internalizes the operation sequence that resulted in a particular set of data and can reconstruct that sequence in the event of node failure.

### 3.13 Dynamic Transactions

Given its propensity for enhancing Hadoop and the ubiquity of this Big Data platform, it is unlikely that the upstart Spark will replace Hadoop in the near future. Its assortment of tools, however, will help to facilitate an environment in Data Management where vendors continue to push the envelope for real-time analytics that exploit many of the aspects of Big Data (Machine Learning, graph analytics, streaming, continued SQL support, multiple development languages, and in-memory technologies) that will ensure that this technology has continued relevance. Redefining the nature of transactions, particularly in the impending wake of the Internet of Things, may be the best way for ensuring that failing Big Data initiatives do not contribute to this technology’s untimely demise. Dynamic transactions based on all data, real time and otherwise, can impact the enterprise’s ROI in a number of ways from marketing to improved customer relations.

### 3.14 Natural Language Process Involvement

There are pivotal aspects of both Machine Learning and Cognitive Computing that are predicated on Natural Language Processing (NLP) Although IBM’s Machine Learning APIs are available exclusively through Bluemix and various offerings related to Watson, they help to highlight some of the critical ways in which text analytics via NLP can create a cognitive focus for application building. Some of the more utilitarian Machine Learning services involving NLP include capabilities for relationship extraction and user modeling. The former plies through sentences to identify various points of significance (subjects, actions, places); the latter creates predictions based on text and language analysis about social traits for specific people. Other NLP applications translate different languages and make sense of colloquial expressions. Although these particular features are associated with Watson, additional MLaaS providers have NLP features to enhance their offerings as well.

### 3.15 The Epicenter

In many ways, Machine Learning functions at the epicenter for a number of different facets of Big Data analytics. Its pivotal role only increases with the availability of MLaaS, which helps to democratize this subset of predictive analytics and enhance the roles of laymen and experts alike. As one of the enablers of the IoT, Machine Learning has a secure place in the future of Big Data. Its capacity to create timely action from analytics makes it

essential to Big Data applications. Machine Learning's ultimate value for Big Data and analytics is alluded to in the aforementioned IBM blog:

"In many ways, MLaaS can be the Return-On-Investment (ROI) capstone of Big Data initiatives because Machine Learning algorithms can grow to be highly effective at data scales in volume, velocity and variety. Without MLaaS capabilities that can dynamically respond to myriad concurrent data streams in the cloud, the human race risks drowning in its own Big Data".

### 3.16 The Mathematics of Data: Graph Analytics-as-a-Service

Graph databases are optimal for running advanced analytics because they indicate the relationship between data elements and allow for readily discernible inferences between them yielding answers to questions that users never thought to ask. Leveraging the prowess of graph analytics (especially on Big Data sets), however, has traditionally been hampered by:

A general lack of skills on any number of Semantics technologies (Sparql, RDF, ontologies, etc.) required to get the sort of performance needed for business cases.

A lack of financial resources to provide the sort of heavy spending on infrastructure that some of the most eminent IT departments such as those at Google and Facebook readily utilize for their lucrative deployments.

A dearth of analytics expertise regarding the most viable means of producing the sort of results needed to derive business value from data.

Graph Analytics as a Service (GAaaS), however, can rectify virtually all of these impediments while contributing these additional boons:

Data integration in a single repository that involves all types of data regardless of models and schema, yet which is able to best other Data Lake options by providing much valued semantic consistency.

Reinforcement of the self-service movement in Data Management that simplifies many processes (especially pertaining to analytics) and empowers business users.

Democratizing semantic technologies while spurring adoption rates in a way that begets in earnest the trend towards ubiquitous computing and its importance in the coming years.

According to Algebraix Data CEO Charlie Silver, modestly priced Graph Analytics as a Cloud service

(as compared to on-premise deployments and their considerable infrastructure requirements) ultimately helps level the playing field and gives small- and mid-sized businesses the same resources and advantages of large enterprises:

"Larger companies have tens of thousands of databases whereas small companies have dozens of databases and they are all siloed," Silver said. "Setting up a linked data graph helps get rid of these silos and simultaneously allows for you to do deep reasoning inference analytics". See Figure 14 as illustration of Network of Graph Database

Figure 14: Network of Graph Database



### 3.17 Deep Reasoning Inference Analytics

Graph databases provide an environment in which users can determine logical inferences about their data in ways that relational databases cannot. Graph analytics then exploit those inferences to make predictions and prescriptions about data elements in a reasonable way that acknowledges relatively simple relationships between data, which can make for much more profound analyses. Those simple relationships can become more revealing the more one utilizes this form of analytics.

"As that graph has more data elements on it, the more inferencing and reasoning it can do" Silver said. "It's so rich and so full of insight that that's why we're taking this to the Business Intelligent (BI) analytics world directly".

### 3.18 Data Algebra

While the logical inferences of graph databases help to provision analytics based on overall views of data and their relationships to one another, those databases and analytics options are also immensely improved by

a mathematical representation of data that makes graph analytics much more robust than other options. Algebraix Data is largely able to issue its service through the Cloud due to Data Algebra, in which both data and queries are represented as equations. By applying these mathematical representations to a triple store, the analytics vendor was able to achieve:

Improved scalability to accommodate Big Data sets.

Performance based rates of speed that issue results in close to real time and considerably expedite the deployment of critical technologies such as RDF and Sparql. The underlying engine for Algebraix Data can eschew conventional computations in any number of ways by utilizing algebraic representations and recognizable query patterns, or speed through them by re-using results and selecting optimum algorithms and data structures to perform them.

Data integration regardless of sources, structure, data type, and format.

Although the aforementioned benefits of Data Algebra are significant, its potential to render graph databases as integration hubs may be the most salient advantage it offers as well as the source of those other advantages:

“We’re going after this new market of graph analytics, but the ultimate play is going to be in data integration, where all data can integrate and officially apply to all enterprise data and the Internet,” Silver remarked. “All data can be integrated irregardless of their model, which is a visual artifice because it can all be implemented and represented”.

### **3.19 Integration, Data Modeling, and Data Lake Architecture**

At the core of the integration capabilities that Algebraic Data drives and the celeritous, reliable results it delivers for graph analytics is an environment in which all data appears as mathematics regardless of modeling specifications. This fact is attributed to its Data Lake architecture (which is typical of NoSQL options), and which actually encourages and rewards enterprises for the incorporation of as much data as possible on the linked enterprise database graph. Silver noted:

“Algebra for data...means that data, regardless of its structure or format meaning it could be in a relational table, a graph, it could be hierarchical, voice, video, it doesn’t matter can be represented mathematically or algebraically, which is a huge innovation in the history of software development because what it really allows for is data to arc integrative”.

### **3.20 Governance and Self-Service Graph Analytics**

As is the case with any Data Lake and the rapid integration of disparate sources, there are very real governance concerns which Silver notes are “customer specific” and must be addressed in order to sustain such an option. However, by offering graph analytics as a Cloud-based service, service providers can potentially mitigate these concerns by leveraging the knowledge and experience of their staffs to help ameliorate this (and most other) aspects of this analytics process. Customers can simply pay to have professionals provide semantic consistency for their various data types.

In fact, there are a number of similarities between GAaaS

and developments in self-service Business Intelligence and analytics options in which laymen end users do little more than apply the results of their analyses to specific business processes. According to Silver, customers simply need to give their data (which is likely in a relational format) to the company which in turn handles the lengthy process of implementing it into a linked data graph. That process typically includes conducting Extract Transfer and Load (ETL) or requisite data transformations, putting data into a graph, and educating users about how to issue queries while applying ontologies. “Our key value proposition is this,” Silver asserted. “We use our mathematics to make this stuff all easy to use; we’re essentially bringing that to the customers making semantic technologies or graph analytics easy to use”.

### **3.21 Ubiquitous Computing**

The reduced ease which GAaaS allows for those attempting to utilize semantic technologies should play a critical role in democratizing semantics and helping to level the playing field in the capabilities of large enterprises and small- to mid-sized operations. Nonetheless, Silver believes the future of semantics and graph databases firmly lies in the latter’s integration capabilities particularly as augmented

by the mathematical representations that Algebraix Data utilizes. The forms those possibilities take vastly exceeds providing analytics for individual customers and pertains to a future in which perhaps all data the world's data is united in a single repository. "The world's data is billions of silos; everything is a silo," Silver acknowledged. "How do you bring those data to be integrated? Now, when you have those common mathematics, that's going to be a big change. That's going to happen over 20, 30 years; its' not going to happen tomorrow."

Such integration would account for ubiquitous computing, which would enable access from anywhere with any variety of devices and provide a degree of interoperability (between data types, organizations, countries, use cases) that is currently not available. Silver is convinced that the method of mathematically representing data that his company employs will ultimately be the enabler of such computing:

"The mathematics of data has implications everywhere in the data world; everywhere in IT," Silver

said. "However, we are narrowly focused right now on the Graph Analytics as a Service. Over time, our intellectual property will be the foundation for many, many applications that I can't even begin to describe".

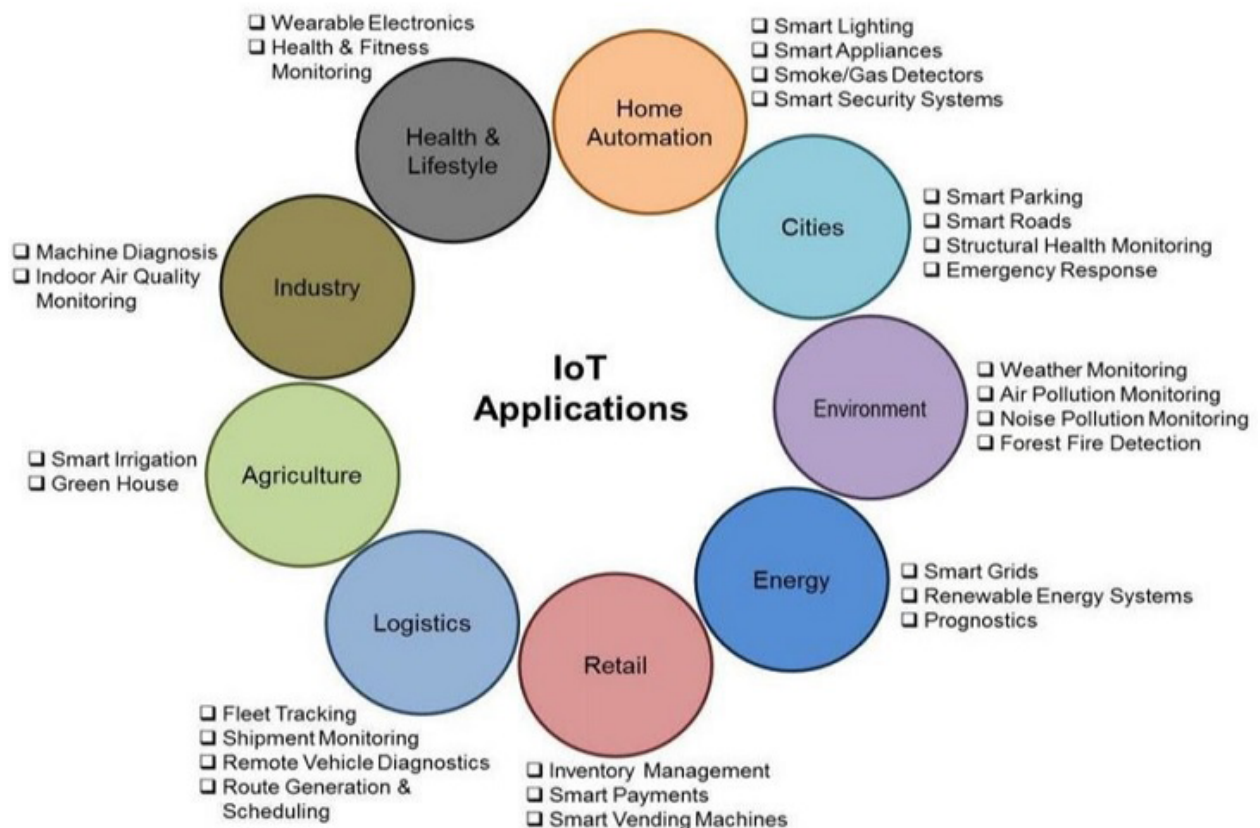
#### 4.0 Conclusions

The 21<sup>st</sup> Century has opened up as well as recognized another dimension to our industry of past and present as renaissance of industrial revolution, where the Information and Communication Technology (ICT) plays as a driving factor of Internet of Things (IoT).

This industrial revolution will bring big changes in all business segments, enormously impacting people's lives as it is presented by Figure 15 here.

Companies within today's industrial environment from ICT business segments in this new market, according to some approximations, will be worth over \$152 Billion Dollars by year 2022 per I-Scoop 2016 prediction. They will have crucial roles in this new industrial period in changing and adjusting their revenue-earning projections

Figure 15: Business Segments for IoT/IIoT Implementation (Courtesy of What is next for IoT and IIoT, Australia 2015 by Ahmed Banafa)



accompanied by investment in relevant resources. instance, mobile telecom operators will have to reorganize and reposition themselves to become ICT operators, and companies in all business segments will have to do the same in changing and adjusting their business activities to include their organization, products and services.

Internet of Things (IoT); Industrial Internet of Things (IIoT); Over the Top (OTT) Applications; Artificial Intelligence (AI); Cloud Services (CS) and many other items are considered as part of ICT going forward. These changes for example will bring new approaches for creating smarter telecommunication products and services in mobile industry and consequently, customers will be able to use completely new types and kinds of products and services in telecommunication markets as a simple aspect of IoT along with growth of data to the level of Big Data and implementation of Super Artificial Intelligence (SAI) as a processing tool by filtering these data in a right direction.

Information and communication technologies have had a profound effect in lots of marketing as a tool.

Although artificial neural networks have simpler processing units than typical Central Processing Units (CPUs), their unique form of parallel processing and tremendous number of interconnections make them incredibly versatile problem solvers. This far-reaching source gets you in on the ground floor of this state-of-art technology, giving you a complete overview of neural network capabilities that drives artificial intelligence, limitations, components, and applications in diverse fields.

An important aspect of this approach by us is to examine biological neural systems and how, artificial neural network is based on them, and, driven by them as well. Key areas discussed include structural diversity... temporal...origins of artificial neural systems...brain structure and function...biological nerve cells... synapses... what is random and what is fixed in the brain's neural networks...and how biological systems really compare to computational neural networks.

For us to deliver such momentum of knowledge in Neural Networking to our reader, we have reached to experts in the field and with their permission we have quoted their notes, lectures and their presentations from their web site at various universities and industries. Our many thanks go to experts like Professor Ingrid Russell of the University of Hartford and permission from Collegiate Microcomputer, Lectures by Kiyoshi Kawaguchi and Professor David Beeman of University of Colorado, Department of Electrical and Computer Engineering

on Computational Neuroscience. In addition, some information that are used was permitted by Dr. David Leech Anderson, Professor of Philosophy at Illinois State University and art works are provided by, “*graphic artists* at The Mind Project”

## References

1. Zohuri B, Moghaddam M (2017) Neural Network Driven Artificial Intelligence: Decision Making Based on Fuzzy Logic Computer Science, Technology and Applications: Mathematics Research Developments, Nova Science Publishers Inc.
2. McCulloch, Pitts (1943) A logical calculus of the ideas immanent in nervous activity in the Bulletin of Mathematical Biophysics 5:115-133.
3. Hassoun, Mohamad H (1995) Fundamentals of Artificial Neural Networks'. The MIT Press.
4. Zurada, Jacek M (1992) Introduction to Artificial Neural System. West Publishing Company.
5. Bose NK, Liang P (1996) Neural Network Fundamentals with Graphs, Algorithms, and Applications.
6. Haykin, Simon (1999) Neural Networks: A Comprehensive Foundation, second edition.
7. Terrence JS, Charles RR (1986) NET talk: a parallel network that learns to read aloud. The Johns Hopkins University Electrical Engineering and Computer Science Technical Report JHU/EECS-86/01, 32 pp.
8. Lerma S, Leonardo O (1991) A Neural Network Approach to a Dimensionality Reduction Problem. Master's thesis, The University of Texas at El Paso, El Paso, TX.
9. Dayhoff JE (1990) Neural Network Architecture: An Introduction. Van Nostrand Reinhold.
10. Rumelhart D, Hinton G, Williams R (1986) Learning representations by back-propagating errors. Nature 323: 533-536.
11. Hornik K, Maxwell S, Halbert W (1989) Multilayer feed forward networks are universal approximators. Neural Networks 2: 359-366.
12. Zohuri B, Moghaddam M (2017) Business Resilience System (BRS): Driven Through Boolean, Fuzzy Logics and Cloud Computation: Real and Near Real Time Analysis and Decision-Making System.

13. Erick B.Baum, David Haussler (1989) What size net gives valid generalization? Neural Computation. 1: 151-160.

**Citation:** Bahman Zohuri (2018) Neural Network Driven Supper Artificial Intelligence Based on Internet of Things and Big Data. SF J Artificial Intel 1:1.