Deep Learning: A New Perspective

Nikita Bhatt[#], Nirav Bhatt*, Purvi Prajapati*

[#]*U & P U Patel Department of Computer Engineering, CSPIT, CHARUSAT, Gujarat, India*
[*]*Department of Information Technology, CSPIT, CHARUSAT, Gujarat, India*

Abstract **- In the last few years, deep neural networks have conquered research area in machine learning and pattern recognition. Deep learning is machine learning techniques that automatically learn hierarchical representations in deep architectures for classification. The goal is to find more important features by using neural networks. In the era of big data where for any real world application, large amount of data need to be processed, deep learning is proven to be the superior to other machine learning techniques. It is applied into the diverse area like speech recognition, natural language processing and collaborative filtering and proved to be superior to other machine learning techniques. It generates large scale neural networks that allow the machine to learn and compute by itself without any human intervention. In this paper we represent the in depth survey on deep learning and its application.**

Keywords **- Deep Learning, Machine Learning, Artificial Intelligence, Big Data, Image processing.**

I. INTRODUCTION

When programmable computers were first conceived, people wondered whether they might become intelligent system that work and react like humans [4]. So researchers start understanding nervous system that inspired the concept of Artificial Neural Network (ANN) which is a subfield of Artificial Intelligence (AI).Artificial Intelligence is an area of computer science that emphasizes the creation of intelligent machines that work and react like humans. It is the field of study looking for ways to create computers that are capable of intelligent behavior [1]. According to John McCarthy (father of AI), AI is the science and engineering of making intelligent machines, especially intelligent computer programs. The technology can broadly be categorized into three groups: Narrow AI, Artificial General Intelligence (AGI), and Super Intelligent AI. In Narrow AI, Systems explicitly designed to solve specific, reasonably well-defined problems. IBM's Deep Blue, which beat chess grand master Garry Kasparov at the game in 1996, or Google DeepMind's AlphaGo, which in 2016 beat Lee Sedol at Go, are examples of narrow AI—AI that is skilled at one specific task. In Artificial general intelligence, systems designed to autonomously learn new tasks and adapt to changing environments. Super intelligent AI takes things a step further. In other words, it's when the machines have outsmarted us. Now the question is where that Intelligence comes from. That's where Machine Learning comes from [2]. Machine learning is one subfield of AI. The core principle of Machine Learning is that machines take data and "learn" for themselves. It's currently the most promising tool in the AI kit for businesses. It explores the development of algorithms that learn from given data [2]. These algorithms are able to learn from past experience (i.e. the given data) and teach themselves to adapt to new circumstances and perform certain tasks. Machine learning allows computers to handle new situations via analysis, self-training, observation and experience. Conventional machine-learning techniques were limited in their ability to process natural data in their raw form. For decades, constructing a pattern-recognition or machine-learning system required careful engineering and considerable domain expertise to design a feature extractor that transformed the raw data (such as the pixel values of an image) into a suitable internal representation or feature vector from which the learning subsystem, often a classifier, could detect or classify patterns in the input [6]. Representation learning is a set of methods that allows a machine to be fed with raw data and to automatically discover the representations needed for detection or classification. Deep-learning methods are representation-learning methods with multiple levels of representation, obtained by composing simple but non-linear modules that each transform the representation at one level (starting with the raw input) into a representation at a higher, slightly more abstract level. With the composition of enough such transformations, very complex functions can be learned. The key aspect of deep learning is that these layers of features are not designed by human engineers: they are learned from data using a general-purpose learning procedure [6]. So Deep learning is a subset of ML. It uses some ML techniques to solve real-world problems by tapping into neural networks that simulate human decision-making. Deep learning can be expensive, and requires massive datasets to train itself on. That's because there are a huge number of parameters that need to be understood by a learning algorithm, which can initially produce a lot of false-positives [1].

II. MACHINE LEARNING

The general focus of machine learning is the representation of the input data and generalization of the learnt patterns for use on future unseen data [7]. There are three types of machine learning algorithms. In Supervised Learning, a model is prepared through a training process in which it is required to make predictions and is corrected when those predictions are wrong. The training process continues until the model achieves a desired level of accuracy on the training data. In Unsupervised Learning, a model is prepared by deducing structures present in the input data. It may be through a mathematical process to systematically reduce redundancy, or it may be to organize data by similarity. In Semi Supervised Learning, Input data is a mixture of labelled and unlabelled examples. There is a desired prediction problem but the model must learn the structures to organize the data as well as make predictions. Classification and Regression are Example problems in Supervised Learning. Classification using a machine learning algorithm has 2 phases: In Training phase, Machine Learning algorithm is trained using a dataset comprised of the images, text, documents and their corresponding labels and In Prediction phase, trained model is utilized to predict labels of unseen images. Fig 1 shows Machine Learning and Representation Learning. (A) Shows classical machine learning workflow that can be divided into four steps: data pre-processing, feature extraction, model training and model evaluation. (B) Shows Supervised Machine Learning Method in which input feature x is related to an output feature y, whereas unsupervised method learns factors about x without observed labels. (C) Shows Raw input data are often high-dimensional and related to the corresponding label in a complicated way, which is challenging for many classical machine learning algorithms (left plot). Alternatively, higher-level features extracted using a deep model may be able to better discriminate between classes (right plot). (D) Shows Deep networks use a hierarchical structure to learn increasingly abstract feature representations from the raw data [8].

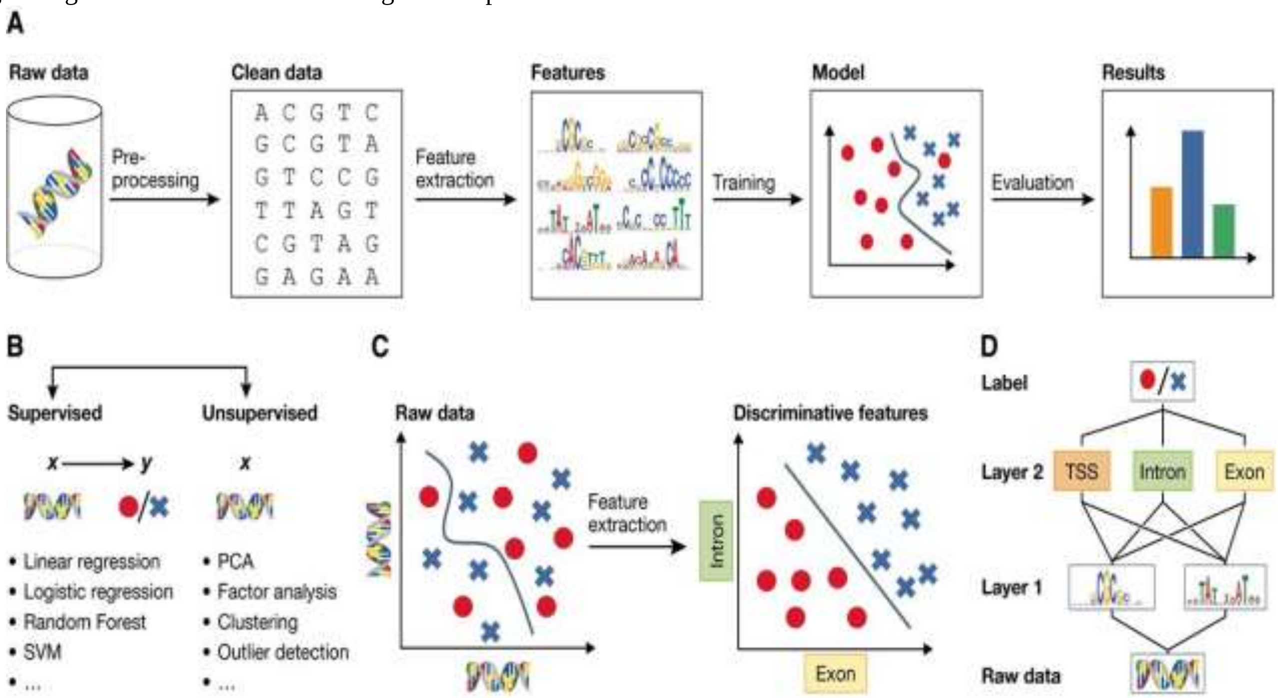

Fig. 1 Machine Learning and Representation Learning [8]

In March 2016, a major AI victory was achieved when DeepMind's AlphaGo program beat world champion Lee Sedol in 4 out of 5 games of Go using deep learning. Deep learning also has business applications. It can take a huge amount of data—millions of images, for example—and recognize certain characteristics. Text-based searches, fraud detection, spam detection, handwriting recognition, image search, speech recognition, Street View detection, and translation are all tasks that can be performed through deep learning [1].

A. Machine Learning with Big Data

Machine Learning takes the data and learn from themselves. But today technology based companies like Google, Yahoo, Microsoft and Amazon have collected and maintained data that is measured in Exabyte proportions or larger. Various social media organizations such as Facebook, Twitter, and YouTube have billions of users that constantly generate a very large quantity of data. So we are living in the era of: "Big data". In recent years, Big Data has taken center stage in government and society at large. In 2012, the Obama Administration announced a ''Big Data Research and Development Initiative'' to ''help solve some of the Nation's most pressing challenges'' [11]. Consequently, six Federal departments and agencies (NSF, HHS/NIH, DOD, DOE, DARPA, and USGS) committed more than $200 million to support projects that can transform our ability to harness in novel ways from huge volumes of digital data. In May of the same year, the state of Massachusetts announced the Massachusetts Big Data Initiative that funds a variety of research institutions [12]. In April, 2013, U.S. President Barack Obama announced another federal project, a new brain mapping initiative called the BRAIN (Brain Research Through Advancing Innovative Neuro-technologies) [13] aiming to develop new tools to help map human brain functions, understand the complex links between function and behavior, and treat and cure brain disorders. This initiative might test and extend the current limits of technologies for Big Data collection and analysis, as NIH director Francis Collins stated that collection, storage, and processing of yottabytes (a billion petabytes) of data would eventually be required for this initiative.

Big data generally refers to data that exceeds the typical storage, processing and computing capacity of conventional databases and data analysis techniques [7]. As a resource, Big Data requires tools and methods that can be applied to analyse and extract patterns from large-scale data. The rise of Big Data has been caused by increased data storage capabilities, increased computational processing power, and availability of increased volumes of data, which give organization more data than they have computing resources and technologies to process[7]. In addition to the obvious great volumes of data, Big Data is also associated with other specific complexities, often referred to as the four Vs: Volume, Variety, Velocity, and Veracity [10].

The unmanageable large Volume of data poses an immediate challenge to conventional computing environments and requires scalable storage and a distributed strategy to data querying and analysis. A general theme in Big Data systems is that the raw data is increasingly diverse and complex, consisting of largely un-categorized/unsupervised data along with perhaps a small quantity of categorized/supervised data. Working with the Variety among different data representations in a given repository poses unique challenges with Big Data, which requires Big Data preprocessing of unstructured data in order to extract structured/ordered representations of the data for human and/or downstream consumption. In today's data-intensive technology era, data Velocity – the increasing rate at which data is collected and obtained – is just as important as the Volume and Variety characteristics of Big Data. While the possibility of data loss exists with streaming data if it is generally not immediately processed and analyzed. Veracity in Big Data deals with the trustworthiness or usefulness of results obtained from data analysis. Big Data Analytics faces a number of challenges beyond those implied by the four Vs like format variation of the raw data, fast moving streaming data, trustworthiness of the data analysis, highly distributed input sources, noisy and poor quality data, high dimensionality, scalability of algorithms, imbalanced input data, unsupervised and un-categorized data, limited supervised/labeled data, etc. Adequate data storage, data indexing/tagging, and fast information retrieval are other key problems in Big Data Analytics [7].

III. DEEP LEARNING

The main concept in deep leaning algorithms is automating the extraction of representations (abstractions) from the data [9]. Deep learning algorithms use a huge amount of unsupervised data to automatically extract complex representation. These algorithms are largely motivated by the field of artificial intelligence, which has the general goal of emulating the human brain's ability to observe, analyze, learn, and make decisions, especially for extremely complex problems. Work related to these complex challenges has been a key motivation behind Deep Learning algorithms which strive to emulate the hierarchical learning approach of the human brain. Models based on shallow learning architectures such as decision trees, support vector machines, and case-based reasoning may fall short when attempting to extract useful information from complex structures and relationships in the input corpus. In contrast, Deep Learning architectures have the capability to generalize in non-local and global ways, generating learning patterns and relationships beyond immediate neighbors in the data. Deep learning is in fact an important step toward artificial intelligence. It not only provides complex representations of data which are suitable for AI tasks but also makes the machines independent of human knowledge which is the ultimate goal of AI. It extracts representations directly from unsupervised data without human interference.

In traditional machine learning algorithms, we need to hand-craft the features. By contrast, in deep learning algorithms feature engineering is done automatically by the algorithm. Feature engineering is difficult, time-consuming and requires domain expertise. The promise of deep learning is more accurate machine learning algorithms compared to traditional machine learning with less or no feature engineering [3].

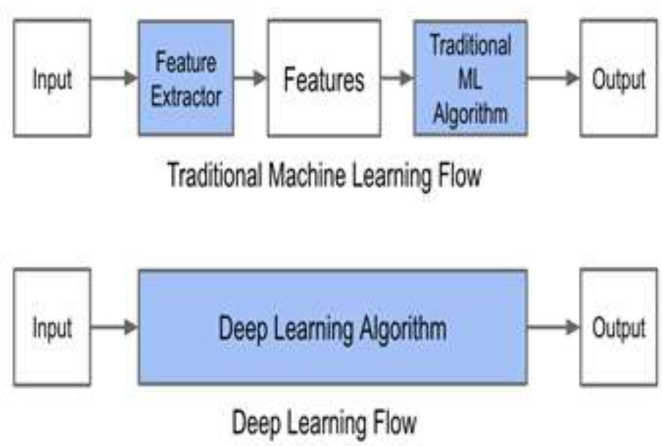

Fig. 2 Feature Engineering in Machine Learning and Deep Learning [5]

Fig 2 shows feature engineering in Machine Learning and Deep Learning [5].Deep Learning is an approach to AI. Specifically it is a type of machine learning, a technique that allows computer systems to improve with experience and data. Machine learning is the only viable approach to building AI systems that can operate in complicated, real-world environments. Deep Learning is a kind of machine learning that achieves great power and flexibility by learning to represent the world as a nested hierarchy of concepts, with each concept defined in relation to simpler concepts, and more abstract representations computed in terms of less abstract ones. Deep Learning only appears to be new, because it was relatively unpopular for several years preceding its current popularity. The field has been rebranded many times, reflecting the influence of different researchers and different perspectives. Broadly speaking, there have been three waves of development of deep learning: deep learning known as cybernetics in the 1940s-1960s, deep learning known as connectionism in the 1980s-1990s, and current rebirth under the name deep learning beginning in 2006. It is motivated by biological nervous systems inspired the concept of Artificial Neural Networks (ANN), a form of information processing. Just as a biological nervous system is a massive interconnection of nodes (called neurons) located within the brain, ANN consists of numerous interconnected processing elements that simultaneously work to solve specific problems. Neural Networks attempt to bring computers a little closer to the brain's capabilities by imitating aspects of information in the brain in a highly simplified way.

IV. APPLICATIONS OF DEEP LEARNING

Following are the various applications of Deep Learning.

Automatic Colorization of Black and White Images:

Image colorization is the problem of adding colour to black and white photographs. Traditionally this was done by hand with human effort because it is such a difficult task. Deep learning can be used to use the objects and their context within the photograph to colour the image, much like a human operator might approach the problem. Generally the approach involves the use of very large convolutional neural networks and supervised layers that recreate the image with the addition of color[14].

Automatically Adding Sounds to Silent Movies:

In this task the system must synthesize sounds to match a silent video. The system is trained using 1000 examples of video with sound of a drum stick striking different surfaces and creating different sounds. A deep learning model associates the video frames with a database of pre-rerecorded sounds in order to select a sound to play that best matches what is happening in the scene. The system was then evaluated using a Turing-test like setup where humans had to determine which video had the real or the fake (synthesized) sounds[5,14].

Automatic Machine Translation:

This is a task where given words, phrase or sentence in one language, automatically translate it into another language. Automatic machine translation has been around for a long time, but deep learning is achieving top results in two specific areas: Automatic Translation of Text. And Automatic Translation of Images. Text translation can be performed without any pre-processing of the sequence, allowing the algorithm to learn the dependencies between words and their mapping to a new language. Stacked networks of large LSTM recurrent neural networks are used to perform this translation. Convolutional neural networks are used to identify images that have letters and where the letters are in the scene. Once identified, they can be turned into text, translated and the image recreated with the translated text. This is often called instant visual translation [14].

Automatic Handwriting Generation:

This is a task where given a corpus of handwriting examples, generate new handwriting for a given word or phrase. The handwriting is provided as a sequence of coordinates used by a pen when the handwriting samples were created. From this corpus the relationship between the pen movement and the letters is learned and new examples can be generated ad hoc. What is fascinating is that different styles can be learned and then mimicked [14].

Automatic Text Generation:

This is an interesting task, where a corpus of text is learned and from this model new text is generated, word-by-word or character-by-character. The model is capable of learning how to spell, punctuate, form sentences and even capture the style of the text in the corpus. Large recurrent neural networks are used to learn the relationship between items in the sequences of input strings and then generate text. More recently LSTM recurrent neural networks are demonstrating great success on this problem using a character-based model, generating one character at time [14].

Automatic Image Caption Generation:

Automatic image captioning is the task where given an image the system must generate a caption that describes the contents of the image. In 2014, there were an explosion of deep learning algorithms achieving very impressive results on this problem, leveraging the work from top models for object classification and object detection in photographs. Once you can detect objects in photographs and generate labels for those objects, you can see that the next step is to turn those labels into a coherent sentence description. This is one of those results that knocked my socks off and still does. Very impressive indeed. Generally, the systems involve the use of very large convolutional neural networks for the object detection in the photographs and then a recurrent neural network like an LSTM to turn the labels into a coherent sentence [14].

V. CONCLUSION

We discover the new perspective of deep learning, Representation learning. Deep learning algorithms with unsupervised data, automate the process of extracting complex representation. These algorithms are highly inspired by the field of artificial intelligence, which has the ability to observe, analyze, learn, and make decisions, especially for extremely complex problems like pattern recognition and speech recognition.

REFERENCES

[1]. Hope Reese, "Understanding the differences between AI, machine learning and deep learning", Feb 2017.
[2]. Parina Hassani, "Artificial Learning, Machine Learning and Deep Learning: Know The Difference", Dec 2016.
[3]. Adil Moujahid, "A Practical Introduction to Deep Learning with Caffe and Python", June 2016.
[4]. Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. Deep learning. MIT Press, 2016.
[5]. faizan Shaikh , "Deep Learning vs. Machine Learning – the essential differences you need to know!", April 2017.
[6]. LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton, "Deep learning." Nature 521.7553 (2015): 436-444.
[7]. Najafabadi, Maryam M., et al. "Deep learning applications and challenges in big data analytics." Journal of Big Data 2.1 (2015): 1.
[8]. Angermueller, Christof, et al. "Deep learning for computational biology." Molecular systems biology 12.7 (2016): 878.
[9]. Bengio, Yoshua, Aaron Courville, and Pascal Vincent. "Representation learning: A review and new perspectives." IEEE transactions on pattern analysis and machine intelligence 35.8 (2013): 1798-1828.
[10]. Dumbill, Edd. "What is big data." (2012).
[11]. Weiss, Rick, and Lisa-Joy Zgorski. "Obama administration unveils "big data" initiative: Announces $200 million in new R&D investments." Office of Science and Technology Policy Executive Office of the President (2012).
[12]. K. Haberlin, B. McGilpin, and C. Ouellette. Governor Patrick Announces New Initiative to Strengthen Massachusetts' Position as a World Leader in Big Data. Commonwealth of Massachusetts [Online].Available:http://www.mass.gov/governor/pressoffice/pressreleases/ 2012/2012530-governor-announces-big-data-initiative.html
[13]. Fact Sheet: Brain Initiative, Office of the Press Secretary, The White House, Washington, DC, USA, 2013.
[14]. Jason Brownlee, "8 Inspirational Applications of Deep Learning", July 2016.