3 PREDICTION MODELS BASED ON MACHINE LEARNING APPROACH

Researchers have contributed to the development of effective prediction mechanisms over the years. Last two and half decades provided more breakthroughs in deep learning and machine learning methods. These methods have provided new dimension and perspective of developing prediction models. Artificial Neural networks design and development was a key aspect in the evolution of these learning algorithms. Mc-Culloch Pitt, Rosenblat and Kohonen made valuable contributions for the evolution of these neural network based designs and training. This chapter provides insights into these models along with the knowledge related to intelligent algorithms, which are also good enough for the prediction. Information pertinent to modern optimization algorithms and their potential features are provided in this chapter (Alizadehsani et al., 2021).

Eventhough the work is more related to deep learning methods, these optimization algorithms are equally important for the researchers to explore other possibilities. Therefore, short information about few optimization algorithms were provided along with neural network based algorithms. Another reason for including the information of intelligent algorithms is as the future scope of this work. The scope of the extension of this work also lies upon the exploration of these prediction models based on intelligent algorithms, which may help future research to a greater extent.

Earlier days of research in this field suggested using regression analysis, fuzzy logic and other decision making based systems. The decision and prediction are possible outcomes of these logic based systems. But in the context of the health sector, the specifications are ever changing i.e. dynamic conditions. This also would keep the necessity of modifying prediction algorithms as per the modern day demands in health sector (World Health Organization 2020). Section 3.1 and 3.2 provides fundamental understanding about learning algorithms and Section 3.3 onwards the work carried out using machine learning approach to forecast the disease spread was presented.

3.1   ARTIFICIAL NEURAL NETWORKS-DEEP LEARNING ALGORITHMS

Man's development is closely linked with his creation and existence. History and Science together had proven many times that our body and its operation became inspiration to many technological innovations. For example, the structure of eye was inspirational in exploring various techniques of image processing, automated vision as well as new dimensional photography. In the same lines, human brain has become an interesting topic of study to many researchers. This study infact has been an interdisciplinary. Biologists, Physicists, Computing researchers and Electronics experts studied human brain and derived various approaches of learning and training. Human brain also a much explored topic for Psychologists too (Wu and Googan 2019).

Therefore, emulation or simulation of human brain operation has become one of the core researches for many across the globe. The scientific, technology and mathematics based studies are increasing every decade on this topic.

Development of Artificial Neural Network (ANN) was the initial step in this journey. These ANNs are simple computing systems inspired by biological neural networks. The fundamental unit is artificial neuron. Each connection in ANN is a synapse model of a human brain. These connections are called as edges. The weights on these connections are updating cyclically. This is done in every iteration of the computation to reduce the error.

Mc-Culloch and Pitt developed first model in 1940s. This good work is continued by Frank Rosenblatt in late 1950s, which lead to the famous 'PERCEPTRON MODEL'. Ivakhnenko and Lapa in mid 1960s worked towards proposing "Group method of data handling is yet another milestone in this computational research. This is the first functional method, which has many applications in the fields of computing and micro-electronics.

Later Teuvo Kohonen and Timo Honkela infused more breath into these networks. In particular, Kohonen in 1980s and in early years of 21st Century did considerable amount of work in this arena. Kohonen proposed "Self-Organized Maps", which are prominently known as Kohonen Networks. These were developed out of early artificial neural networks. This is the brief history and development of "DEEP LEARNING", as the roots of these algorithms are traced back to 1940s of Mc-Culloch and Pitts' era and the pace for this knowledge is given by Ivankhnenko and Lapa. Therefore, Deep Learning is basically understood

as the method that uses Artificial Neural Networks to perform complex calculations involving huge data. These calculations may support data analysis, data set training and data interpretations.

Deep learning otherwise also shall be understood as a type of Machine Learning that works based on the knowledge related to functions of the human brain. Health, Electronics, e-Commerce industries uses deep learning technics in the early days, now the applications extended to many fields. In this section, keeping the scope of this work in consideration the necessary information to understand deep learning algorithms is presented.

3.2   DEEP LEARNING ALGORITHMS

Deep learning is simple and closely resembles the concept of Artificial Neural Network. It has got an input layer, a hidden layer and an output layer. Hidden layer is the key as it reflects the complexity of the method. Deep learning algorithms use several methods. The most frequently used deep learning algorithms are Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short Term Memory Networks (LSTMs), Multilayer Perceptrons (MLPs), Self-Organizing Maps (SOMs) and Deep Belief Networks (DBN). Deep learning can also be treated as the cutting edge knowledge related to Artificial Intelligence (AI). Swarm inspired AI is known as Artificial Swarm Intelligence, this has advanced AI to the next level of implementation, thereby the deep learning methods (Cohen and Normile 2020).

Artificial Intelligence is narrowing down the gap between the human and machine capacities. Amazing technology is emerging in this knowledge hunt and one such innovation is deep learning. Figure 3.1 represents the relation between AI, ML and DL.

The agenda has been to make the machines empowered to work like humans. Perception of vision, cognitive functions and limbic activities everywhere the human and machine competencies are trying to be matched by this new computational methods supported by the technological advancements. Image processing, Natural language processing, Computer vision with deep learning has become an interesting topic for researchers.

3.2.1   Convolutional Neural Networks

A Convolutional Neural Network (ConvNet/CNN) is one of the prominent Deep Learning algorithms. This was developed by Yann LeCunin 1988, and it was called as LeNet. This is mainly used for image
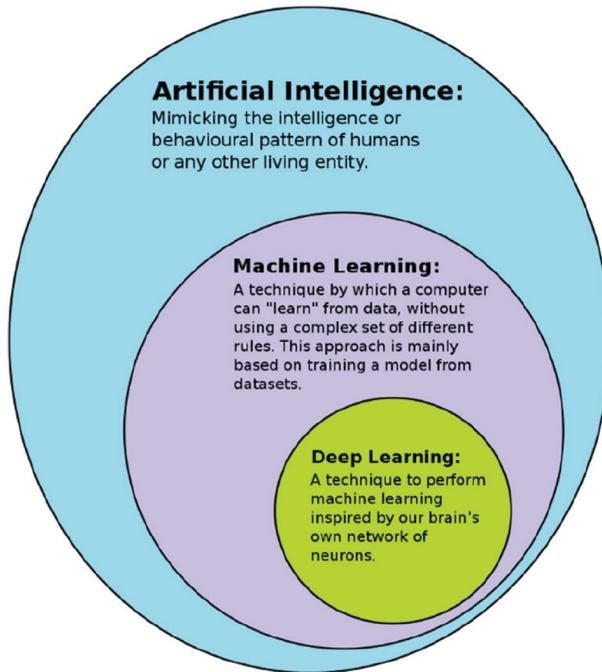
Artificial Intelligence:
Mimicking the intelligence or behavioural pattern of humans or any other living entity.

Machine Learning:
A technique by which a computer can "learn" from data, without using a complex set of different rules. This approach is mainly based on training a model from datasets.

Deep Learning:
A technique to perform machine learning inspired by our brain's own network of neurons.

Figure 3.1. AL-ML-DL (*Source*: Wikipedia).

processing and digits recognition applications. Satellite imaging, medical images, forecasting and to identify anomalies this CNN is normally used.

The input layer shall take an image; assign some weights to objects in image. These weights are assigned in such a way that each object is differentiable from other. CNN's have a Rectified Linear Unit (ReLU) layer to do operations on the elements. The output is generally a rectified feature map. The rectified feature map further feeds into a pooling layer. This pooling layer reduces the feature map dimensions. This also converts pooled feature map into a single dimensional linear vector through flattening. When the flattened matrix is fed to the input, the image classification or identification will be carried out. The pre-procession of CNN is much lower than other algorithms. Figure 3.2 is a depicted version of the aforementioned process.

The Figure 3.2 also resembles the pattern of neurons connectivity in human brain.

3.2.2   Recurrent Neural Networks

Recurrent neural networks represent a type of neural networks. RNNs were based on David Rumelhart's work (1986). These are useful in the
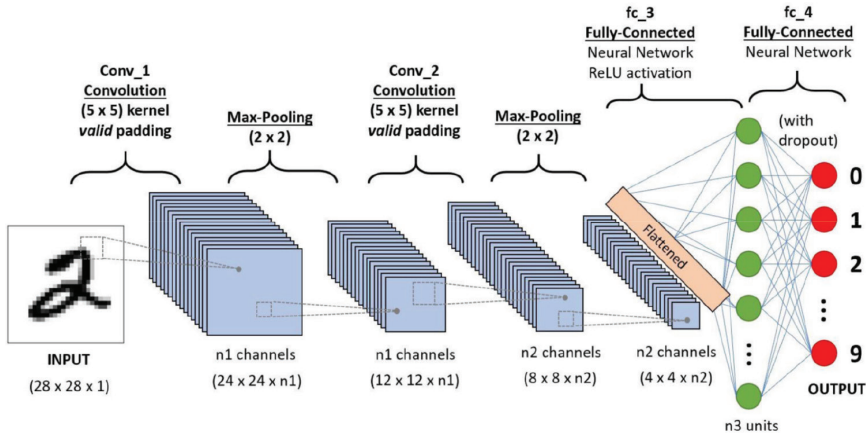
Figure 3.2. A CNN sequence to classify handwritten digits.

modelling of sequential data. These are derived from the feed forward networks of ANN. These are used by "Siri" and "Voice Search" of Google.

This algorithm has a special capacity of "Memory". This is because of the internal loop or memory. This feature makes this RNN has the best contender for the ML, particularly, in sequence data processing. This RNN is responsible for major achievements emerged due to deep learning.

The difference between RNN and Feed forward neural network is shown in the Figure 3.3.

In feed forward the information transfer takes place in only one direction, whereas in RNN the information is transmitted in loops also. Feed forward networks have no memory, RNN has an internal memory. RNNs are used in Natural language processing, machine translation, time series analysis and in Prediction. An Unfolded RNN network is as shown in the Figure 3.4.

The operation is simple. The output at time t − 1 is given to the input at time t, similarly other cycles follows, RNNs process inputs of any type of length. RNN has memory, so that all the parameters of each input it does on a task or memorized, processed by hidden layers to give the output. This reduces additional complexity; these are robust and much promising than many algorithms because of internal memory.

3.2.3   Bi-directional Recurrent Neural Networks

Bi-directional Recurrent Neural Network (BRNN) was proposed in 1997 by Schuster and Paliwal. BRNNs have 2 hidden layers in opposite
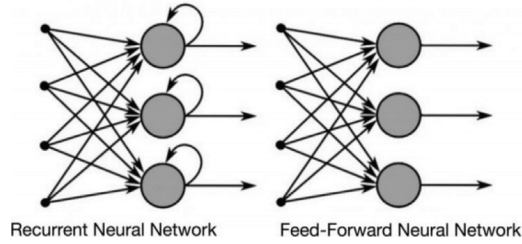
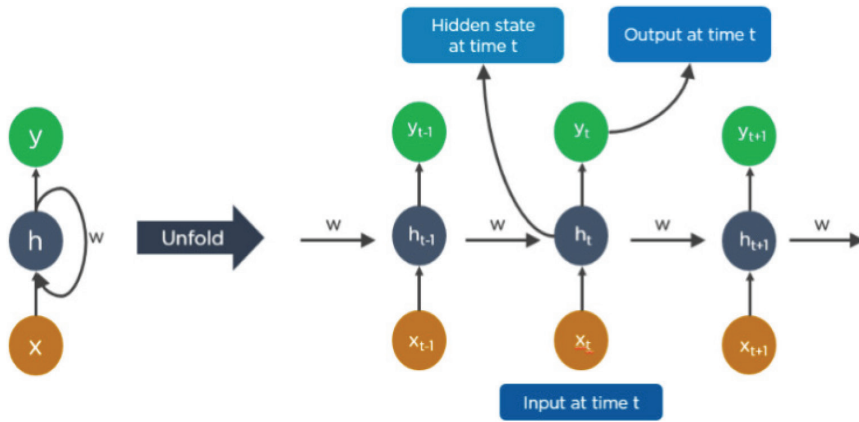**Figure 3.3.** RNN and Feed forward neural network.

Figure 3.4. An Unfolded Recurrent Neural Network.

directions for the same output. This is a type of generative deep learning as mentioned in the Chapter 1. Here the reverse engineering takes place. The output layer can get information backwards and forward states at the same time. This mechanism is followed to get more input to the network. This also infuses more flexibility into the network. Sometimes RNN input data is fixed, whereas in BRNN this is not a requirement. Here the future input is available to the current state. Hand writing recognition is the best specimen for implementing BRNN. This is in other words, splitting the neurons of RNN into 2 directions. One is forward and other is backward. The difference in operation is evident from the Figure 3.5.

In RNN delay is incurred in getting future information, whereas in BRNN it is not the case. Other applications of BRRN include Speech recognition, Translation and dependency parsing.
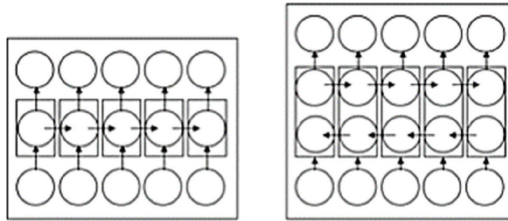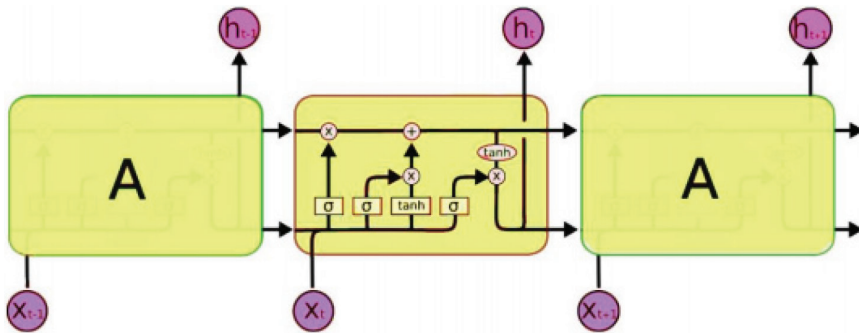
Figure 3.5. RNN and BRNN.

Figure 3.6. Operational diagram of LSTM.

3.2.4   Long Short-Term Memory Networks

The acronym of Long Short-Term Memory Networks is LSTM. This is a category of RNN. This is very much suitable to time series data. For example, the text data. LSTMs are more advantageous than RNN in more than one way. The reason for this edge is the "Selective remembrance of patterns for long time". LSTM has the feature of data retention for long time. So, LSTM has extended feature of memorizing for long duration. Typical LSTM architecture and operation is explained in the Figure 3.6.

LSTMs have memory blocks as represented in the Figure 3.6. Two states are shared with next cell or block. One is cell state and other is hidden state. Memory blocks work on the basis of gates. The functioning of the output gates takes place in three stages. (A) Creation of a vector; by applying tanh function to the cell. (B) Preparing a filter that filters values to the output. (C) Multiplying the vector with filter value. This will be thrown as output.

As the data retention is longer period, they are useful in time series prediction. This is the reason to select LSTM network for this work.

Here 4 interacting layers exchange information in a special manner. These LSTMs are also having applications in composition of music and drug discovery.

3.2.5   Bidirectional Long Short- Term Memory and Deep Belief Networks

Bi-LSTM is the acronym of Bidirectional long short- term memory. This is a combination of forward and backward LSTM. The most prominent application is natural language processing. Long distance dependencies are covered using this model. Training can be given to these Bi-LSTM networks to remember and forget the specific information. This is the key feature of these networks. So, selective memory concept is further extended in these networks. There is a challenge in these networks; the information cannot be encoded from back to front.

Deep Belief Networks (DBNs) consists of multiple layers. They are random and have many variables. These values are often known as hidden units. These are Boltzmann machines. The connections are in between the layers; here also the communication is between previous and next layers. These are having applications in video and image processing. The next section introduces ASI concept, which supports the knowledge base creation for the future scope of this work (Chen 2015).