International Journal of
System Design and Information Processing (SDIP)
http://ijsdip.org
ISSN: (Print): 2319-9288 | (Online): 2321-0591

An Overview of Artificial Neural Network, Working and Training of ANN

[1]Noor Mahboob Subani, [2]J.Raja, [3]K Senthil Vijay, [4]V. Neeraja, [5]Dr.M.Vadivukarassi
[1,2,3,4]Assistant Professor
[2,3,4]Department of CSE , [1]Department of AI&DS
St.Martin's Engineering, Dulapally, Secunderabad-500100 , Telangana, India

Abstract— An Artificial Neural Network (ANN) is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems. ANNs, like people, learn by example. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. Learning in biological systems involves adjustments to the synaptic connections that exist between the neurons. This is true of ANNs as well. This paper gives overview of Artificial Neural Network, working & training of ANN. It also explain the application and advantages of ANN.

Keywords— *Artificial Neural Network, Neurons, Pattern Recognition, Learning Process, Data Classification.*

INTRODUCTION

The study of the human brain is thousands of years old. With the advent of modern electronics, it was only natural to try to harness this thinking process. The first step toward artificial neural networks came in 1943 when Warren McCulloch, a neurophysiologist, and a young mathematician, Walter Pitts, wrote a paper on how neurons might work. They modeled a simple neural network with electrical circuits. Neural networks, with their remarkable ability to derive meaning from complicated or imprecise data, can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. A trained neural network can be thought of as an "expert" in the category of information it has been given to analyse. Other advantages include:

1. Adaptive learning: An ability to learn how to do tasks based on the data given for training or initial experience.
2. Self-Organisation: An ANN can create its own organisation or representation of the information it receives during learning time.
3. Real Time Operation: ANN computations may be carried out in parallel, and special hardware devices are being designed and manufactured which take advantage of this capability.
4. Fault Tolerance via Redundant Information Coding: Partial destruction of a network leads to the corresponding degradation of performance. However, some network capabilities may be retained even with major network damage.

Neural networks take a different approach to problem solving than that of conventional computers. Conventional computers use an algorithmic approach i.e. the computer follows a set of instructions in order to solve a problem. Unless the specific steps that the computer needs to follow are known the computer cannot solve the problem. That restricts the problem solving capability of conventional computers to problems that we already understand and

know how to solve. But computers would be so much more useful if they could do things that we don't exactly know how to do. Neural networks process information in a similar way the human brain does. The network is composed of a large number of highly interconnected processing elements (neurons) working in parallel to solve a specific problem. Neural networks learn by example. They cannot be programmed to perform a specific task. The examples must be selected carefully otherwise useful time is wasted or even worse the network might be functioning incorrectly. The disadvantage is that because the network finds out how to solve the problem by itself, its operation can be unpredictable. On the other hand, conventional computers use a cognitive approach to problem solving; the way the problem is to solved must be known and stated in small unambiguous instructions. These instructions are then converted to a high level language program and then into machine code that the computer can understand. These machines are totally predictable; if anything goes wrong is due to a software or hardware fault. Neural networks and conventional algorithmic computers are not in competition but complement each other. There are tasks are more suited to an algorithmic approach like arithmetic operations and tasks that are more suited to neural networks. Even more, a large number of tasks, require systems that use a combination of the two approaches (normally a conventional computer is used to supervise the neural network) in order to perform at maximum efficiency.

WORKING AN ARTIFICIAL NEURAL NETWORK

The other parts of the art of using neural networks revolve around the myriad of ways these individual neurons can be clustered together. This clustering occurs in the human mind in such a way that information can be processed in a dynamic, interactive, and self-organizing way. Biologically, neural networks are constructed in a three-dimensional world from microscopic components. These neurons seem capable of nearly unrestricted interconnections. That is not true of any proposed, or existing, man-made network. Integrated circuits, using current technology, are two- dimensional devices with a limited number of layers for interconnection. This physical reality restrains the types, and scope, of artificial neural networks that can be implemented in silicon. Currently, neural networks are the simple clustering of the primitive artificial neurons. This clustering occurs by creating layers which are then connected to one another. How these layers connect is the other part of the "art" of engineering networks to resolve real world problems. Basically, all artificial neural networks have a similar structure or topology as shown in Figure1. In that structure some of the neurons interfaces to the real world to receive its inputs. Other neurons provide the real world with the network's outputs. This output might be the particular character that the network thinks that it has scanned or the particular image it thinks is being viewed. All the rest of the neurons are hidden from view. But a neural network is more than a bunch of neurons. Some early researchers tried to simply connect neurons in a random manner, without much success. Now, it is known that even the brains of snails are structured devices. One of the easiest ways to design a structure is to create layers of elements. It is the grouping of these neurons into layers, the connections between these layers, and the summation and transfer functions that comprises a functioning neural network. The general terms used to describe these characteristics are common to all networks. Although there are useful networks which contain only one layer, or even one element, most applications require networks that contain at least the three normal types of
layers - input, hidden, and output. The layer of input neurons receive the data either from input files or directly from electronic sensors in real-time applications. The output layer sends

International Journal of
System Design and Information Processing (SDIP)
http://ijsdip.org
ISSN: (Print): 2319-9288 | (Online): 2321-0591

information directly to the outside world, to a secondary computer process, or to other devices such as a mechanical control system. Between these two layers can be many hidden layers. These internal layers contain many of the neurons in various interconnected structures. The inputs and outputs of each of these hidden neurons simply go to other neurons. In most networks each neuron in a hidden layer receives the signals from all of the neurons in a layer above it, typically an input layer. After a neuron performs its function it passes its output to all of the neurons in the layer below it, providing a feed forward path to the output. the drawings are reversed, inputs come into the bottom and outputs come out the top.)

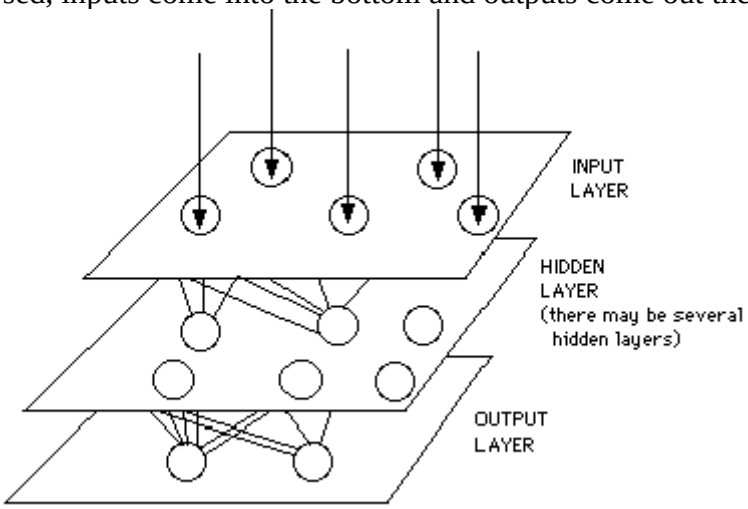

Figure 1- A Simple Neural Network Diagram

TRAINING OF AN ARTIFICIAL NEURAL NETWORK

Once a network has been structured for a particular application, that network is ready to be trained. To start this process the initial weights are chosen randomly. Then, the training, or learning, begins. There are two approaches to training - supervised and unsupervised. Supervised training involves a mechanism of providing the network with the desired output either by manually "grading" the network's performance or by providing the desired outputs with the inputs. Unsupervised training is where the network has to make sense of the inputs without outside help. The vast bulk of networks utilize supervised training. Unsupervised training is used to perform some initial characterization on puts. However, in the full blown sense of being truly self learning, it is still just a shining promise that is not fully understood, does not completely work, and thus is relegated to the lab.

(i) Supervised Training- In supervised training, both the inputs and the outputs are provided. The network then processes the inputs and compares its resulting outputs against the desired outputs. Errors are then propagated back through the system, causing the system to adjust the weights which control the network. This process occurs over and over as the weights are continually tweaked. The set of data which enables the
training is called the "training set." During the training of a network the same set of data is processed many times as the connection weights are ever refined. The current commercial network development packages provide tools to monitor how well an artificial neural network is converging on the ability to predict the right answer. These tools allow the training process to go on for days, stopping only when the system reaches some statistically desired point, or accuracy. However, some networks never learn. This could be because the input data does not contain the specific information from which the desired output is derived. Networks also

don't converge if there is not enough data to enable complete learning. Ideally, there should be enough data so that part of the data can be held back as a test. Many layered networks with multiple nodes are capable of memorizing data. To monitor the network to determine if the system is simply memorizing its data in some non significant way, supervised training needs to hold back a set of data to be used to test the system after it has undergone its training. If a network simply can't solve the problem, the designer then has to review the input and outputs, the number of layers, the number of elements per layer, the connections between the layers, the summation, transfer, and training functions, and even the initial weights themselves. Those changes required to create a successful network constitute a process wherein the "art" of neural networking occurs. Another part of the designer's creativity governs the rules of training. There are many laws (algorithms) used to implement the adaptive feedback required to adjust the weights during training. The most common technique is backward-error propagation, more commonly known as back-propagation. These various learning techniques are explored in greater depth later in this report. Yet, training is not just a technique. It involves a "feel," and conscious analysis, to insure that the network is not over trained. Initially, an artificial neural network configures itself with the general statistical trends of the data. Later, it continues to "learn" about other aspects of the data which may be spurious from a general viewpoint. When finally the system has been correctly trained, and no further learning is needed, the weights can, if desired, be "frozen." In some systems this finalized network is then turned into hardware so that it can be fast. Other systems don't lock themselves in but continue to learn while in production use.

(ii) Unsupervised, or Adaptive Training- The other type of training is called unsupervised training. In unsupervised training, the network is provided with inputs but not with desired outputs. The system itself must then decide what features it will use to group the input data. This is often referred to as self-organization or adaption. At the present time, unsupervised learning is not well understood. This adaption to the environment is the promise which would enable science fiction types of robots to continually learn on their own as they encounter new situations and new environments. Life is filled with situations where exact training sets do not exist. Some of these situations involve military action where new combat techniques and new weapons might be encountered. Because of this unexpected aspect to life and the human desire to be prepared, there continues to be research into, and hope for, this field. Yet, at the present time, the vast bulk of neural network work is in systems with supervised learning. Supervised learning is achieving results

APPLICATION OF AN ARTIFICIAL NEURAL NETWORK

The various real time application of Artificial Neural Network are as follows:

- Function approximation, or regression analysis, including time series prediction and modelling.
- Call control- answer an incoming call (speaker-ON) with a wave of the hand while driving.
- Classification, including pattern and sequence recognition, novelty detection and sequential decision making.
- Skip tracks or control volume on your media player using simple hand motions- lean back, and with no need to shift to the device- control what you watch/ listen to.
- Data processing, including filtering, clustering, blind signal separation and compression.

- Scroll Web Pages, or within an eBook with simple left and right hand gestures, this is ideal when touching the device is a barrier such as wet hands are wet, with gloves, dirty etc.
- Application areas of ANNs include system identification and control (vehicle control, process control), game-playing and decision making (backgammon, chess, racing), pattern recognition (radar systems, face identification, object recognition, etc.), sequence recognition (gesture, speech, handwritten text recognition), medical diagnosis, financial applications, data mining (or knowledge discovery in databases, "KDD").
- Another interesting use case is when using the Smartphone as a media hub, a user can dock the device to the TV and watch content from the device- while controlling the content in a touch-free manner from a far.
- If your hands are dirty or a person hates smudges, touch-free controls are a benefit.

DEVELOPMENT OF ARTIFICIAL NEURAL NETWORK

From the idea of artificial neural network to the present high-speed development, the development of artificial neural network has mainly gone through three stages, the first representation of neural network prototype is the perceptron model, neurons are linearly connected, but there are almost no practical application values. After introducing nonlinear connection, the back propagation algorithm is invented, which laid the foundation for the training of neural network. After entering the 21st century, deep learning has become the mainstream of the neural network ideas, and the application of neural network is greatly improved.

(i) Perceptron model -In the field of machine learning, perceptron is a supervised learning algorithm used for binary classifiers (it can determine whether the input represented by a numeric vector belongs to a particular class of function). It is a linear classifier, which uses a classification algorithm based on the linear prediction function combining the weight set and the feature vector to predict. Perceptron algorithm can be traced back to 1950s, and it is one of the first implementation of artificial neural networks in customized hardware.352Advances in Social Science, Education and Humanities Research (ASSEHR), volume 130 Before the birth of perceptron model, experts in neuroscience and cybernetics and etc. began to think about the possibility of intelligent computing tools, logician Walter Pitts and cybernetics expert Warren McCulloch designed a neural network computing prototype based on mathematics and threshold logic algorithm, called MP model. This linear model classify the input by judging the positive and negative values of the activation function, which is the earliest linear classification model. Since then, the research of neural network has been extended from the study of biological nervous system to the study of artificial neural network model. Inspired by the MP model, psychologist Frank Rosenblatt invented the perceptron model in 1957. The perceptron is designed for character image recognition. In the background of neural networks, the perceptron is an artificial neuron that uses the Heaviside step function as the activation function. The perceptron algorithm is also called a single-layer perceptron, and as a linear classifier, a single-layer perceptron is the simplest feed forward neural network, and it uses the idea of hidden neurons, which is referred to many times by later algorithms. Although the perceptron was initially promising, it was soon proven unable to be trained to identify many types of patterns and could not solve the problem of XOR circuits, which also led to the study of neural networks stagnating for many years.

(ii) Back propagation algorithm- Back propagation is a method used in artificial neural network to calculate each neuron's error contribution after a batch of data processing. It is used to adjust the weight parameters of each neuron and complete the learning process. Technically, this algorithm calculates the gradient of the loss function and uses the gradient descent optimization algorithm, so it is also called the error back propagation. In the training process of the neural network, each input value and its expected output value are required, and the whole algorithm repeats the two-stage loop: forward propagation and weight update. When the input vector is presented to the neural network, it propagates forward through the network layer by layer until it goes out from the output layer. The output value of the neural network is then compared with the expected output value using the cost function, and calculate the error value for each neuron in the output layer. The error value is then passed backward from the output layer through the neural network until each neuron has an associated error value that reflects its contribution to the original output. Back propagation algorithm uses these error values to compute the gradient of the loss function and uses the gradient descent algorithm to optimize the cost function, in order to try to minimize the cost function. Finally, after adjusting the weights and threshold parameters between the neurons in the network layer by layer, one training process of the back propagation algorithm is completed.

CONCLUSION

In this paper we discussed about the Artificial neural network, working of ANN. Also training phases of an ANN. There are various advantages of ANN over conventional approaches. Depending on the nature of the application and the strength of the internal data patterns you can generally expect a network to train quite well. This applies to problems where the relationships may be quite dynamic or non-linear. ANNs provide an analytical alternative to conventional techniques which are often limited by strict assumptions of normality, linearity, variable independence etc. Because an ANN can capture many kinds of relationships it allows the user to quickly and relatively easily model phenomena which otherwise may have been very difficult or imposible to explain otherwise. Today, neural networks discussions are occurring everywhere. Their promise seems very bright as nature itself is the proof that this kind of thing works. Yet, its future, indeed the very key to the whole technology, lies in hardware development. Currently most neural network development is simply proving that the principal works

REFERENCES

[1]. Bradshaw, J.A., Carden, K.J., Riordan, D., 1991. Ecological ―Applications Using a Novel Expert System Shell‖. Comp. Appl. Biosci. 7, 79–83.
[2]. Lippmann, R.P., 1987. An introduction to computing with neural nets. IEEE Accost. Speech Signal Process. Mag., April: 4-22.
[3]. N. Murata, S. Yoshizawa, and S. Amari, ―Learning curves, model selection and complexity of neural
[4]. networks,‖ in Advances in Neural Information Processing Systems 5, S. Jose Hanson, J. D. Cowan, and C. Lee Giles, ed. San Mateo, CA: Morgan Kaufmann, 1993, pp. 607-614.
[5]. Bradshaw, J.A., Carden, K.J., Riordan, D., 1991. Ecological ―Applications Using a Novel Expert System Shell.
[6]. Rich ,Knight and B Nair, “Artificial Intelligence”, TMH Publication.
[7]. Jacek M . Zurada, Introduction to Artificial Neural System”, Jaico publishing house.
[8]. Ajith Abraham, “Artificial Neural Networks”, Stillwater,OK, USA, 2005.
[9]. Carlos Gershenson, “Artificial Neural Networks for Beginners”, United kingdom.
[10]. Ke Xu. Research on application of convolutional neural network in image recognition[D]. Zhejiang University:2012.
[11]. P. Werbos. Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences.[Dissertation]. Harvard University:1974.
[12]. D. E. Rumelhart, G. E. Hinton and R. J. Williams. Parallel Distributed Processing: Exploration in the Microstructure of Cognition. MIT Press. 1986.

[13]. T. Lee and D. Mumford. Hierarchical Bayesian Inference in the Visual Cortex. Journal of the Optical Society of America A. vol. 20. 2003:1434-1448.
[14]. G. Hinton, S. Osindero. A Fast Learning Algorithm for Deep Belief Nets. Neural Computation. vol. 18. 2006:1527-1554.
[15]. Yann Lecun, Leon Bottou, Yoshua Bengio. Gradient Based Learning Applied to Document Recognition. Proceedings of IEEE. 86(11). 1998:2278-2324.
[16]. Sanjay Kumar Suman, Dhananjay Kumar and L. Bhagyalakshmi, "SINR pricing in non- cooperative power control game for wireless ad hoc network", KSII Transactions on Internet and Information Systems, KSII TIIS, vol. 8, no. 7, pp. 2281-2301, 2014. https://dio.org/10.3837/tiis.2014.07.005
[17]. L. Bhagyalakshmi, Sanjay Kumar Suman, Sujeetha Devi, "Joint Routing and Resource Allocation for Cluster Based Isolated Nodes in Cognitive Radio Wireless Sensor Networks", Wireless Personal Communication, Springer, vol. 114, issue 4, pp. 3477- 3488, Oct. 2020. https://doi.org/10.1007/s11277-020-07543-4
[18]. K. Mahalakshmi, K. Kousalya, Himanshu Shekhar, Aby K. Thomas, L. Bhagyalakshmi, Sanjay Kumar Suman et. al., "Public Auditing Scheme for Integrity Verification in Distributed Cloud Storage System", Scientific Programming, Hindawi, vol. 2021, Article ID 8533995, Dec. 2021. https://doi.org/10.1155/2021/8533995
[19]. Sanjay Kumar Suman et. al., Detection and prediction of HMS from drinking water by analysing the adsorbents from residuals using deep learning, Hindawi (SAGE Journal) Adsorption Science & Technology, vol. 2022, Article id 3265366, March 2022. https://doi.org/10.1155/2022/3265366
[20]. Bhagyalakshmi and K. Murugan, "Avoiding Energy Holes Problem using Load Balancing Approach in Wireless Sensor Network", KSII Transaction on Internet and Information Systems, vol. 8 , no. 5, pp. 1618- 1637, 2014. https://dio.org/10.3837/tiis.2014.05.007.
[21]. Satyanand Singh, Sajai Vir Singh, Dinesh Yadav, Sanjay Kumar Suman, Bhagyalakshmi Lakshminarayanan, Ghanshyam Singh, "Discrete interferences optimum beamformer in correlated signal and interfering noise, International Journal of Electrical and Computer Engineering, vol 12, no. 2, pp. 1732-1743, April 2022. http://doi.org/10.11591/ijece.v12i2.pp1732-1743