Journal of Advances in Science and Technology
Vol. 20, Issue No. 1, March-2023, ISSN 2230-9659

How do Artificial Neural Networks Work

Aryan Rose*

Class 12, Welham Boys School

Abstract - Artificial Neural Networks (ANNs) are computer models inspired by the structure and operation of the human brain. They comprise of linked nodes, called neurons, grouped in layers. Information passes across these neurons, and each connection between neurons is connected with a weight denoting its significance. The network's learning process includes altering these weights depending on input data to increase its capacity to generate correct predictions or classifications. During training, the network compares its output to the intended output, computes the error, and then applies optimization methods to reduce this error. Once trained, ANNs may be used for numerous tasks, including as image identification, natural language processing, and decision-making, making them a formidable tool in the area of artificial intelligence.

Keyword - Artificial, Networks

----------X----------

INTRODUCTION

An artificial neural network is one of the most promising new machine learning technologies that has evolved in recent years as AI has gained acceptance in the academic community. The human nervous system serves as a model for the development of artificial neural networks. When an organism is triggered from the outside, messages are conveyed by neurons, and the organism's awareness is engaged to generate relevant feedback. This occurs because the biological neuronal cells are linked by synaptic connections and constitute complex neural network systems. An artificial neural network mimics this quality by linking together many virtual neurons, exchanging data through the right activation functions, and processing data in a manner similar to that of a human brain. (Hinton G.E., N. Srivastava, A. Krizhevsky, and E. et al. There are three distinct ways of thinking that may occur in a human brain: abstract, picture, and inspiration. Image thinking, in particular, is the exact direction of AI development, as it allows machines to learn knowledge and produce subjective cognition that can replace human work by inspiring creative solutions to problems. To aid or perhaps replace humans in many jobs that require subjective identification or prediction, artificial neural networks have shown to be an excellent tool. Deep neural networks and convolutional neural networks are only two examples of the state-of-the-art models and algorithms that have emerged as a result of the rapid advancement of artificial neural networks in the early 21st century. Simultaneously, there is a clear split between theoretical and practical orientations in terms of the artificial neural network assessment indices. An algorithm's theoretical impact may be measured by how well it performs in pattern recognition or prediction.2015) (C. Szegedy, W. Liu, Y. Jia, et al.) Furthermore, AI is in its formative years, and building a comprehensive theoretical framework is essential for the advancement of artificial neural network theory. As the complexity of the neural network model rises, so does the volume of data it must handle; as a result, researchers are also interested in optimizing the operating overhead and learning speed of neural networks for practical use. Before analyzing the major research outcomes of artificial neural networks, this study examines the fundamentals of neural networks from both a theoretical and an applied perspective.

Basic Theory of Artificial Neural Network

An artificial neural network is a model with several interconnected layers, each representing a different kind of neuron. The output value of a neuron in a layer that is not the input layer is calculated using the activation function of the neuron with which it is linked in the lower layer. The vector of input parameters is read by the neurons in the input layer, which then sends its output to the neurons in the following layer. Neurons employed for processing input and extracting features have extensive interconnections that are hidden from view in "hidden layers." The categorization outcomes or estimated values might be produced by a single or several neurons in the output layer. A connection between input data and output data is established at a single neuron node. The activation function f, the weight parameter W, and the output value hw,b(x) are all defined by the following equation: hw,b(x) = f (WT x) = f (∑ 3i=1 Wi xi + b) where x1, x2, and x3 are the input values and +1 is used to introduce bias value b to set threshold for the activation function. Nonlinear functions, such as the Sigmoid function, the Rectified linear rectifier unit function, etc., are used to activate the input data when confronted with complicated nonlinear classification jobs and the like. Layers of input data, hidden processing layers,

and output layers make up the fundamental building blocks of every artificial neural network. (Madgwick et al. 2011; Harrison & Vaidyanathan) With the exception of the input layer, the activation function is used to determine the output value of any neuron in the network by multiplying its input data by the output data from the connected neuron in the low layer and its weight. Training a neural network mostly entails fine-tuning the network's weight parameters in order to achieve optimum performance. The practical application impact of an artificial neural network requires training and learning using massive data. Back Propagation, also known as the BP algorithm (Glorot Xavier, A. Bordes, and Y. Bengio. 2011), is now the most used training technique for neural networks.The BP algorithm is implemented by first feeding the neural network input from the training set, then using forward propagation to calculate the network's output values, and finally using a mean-square cost function to minimize the difference between the network's output and the expected output value of the training set. Finally, the network's weight parameters are adjusted backwards, layer by layer, using the Gradient Descent Algorithm. After many rounds of training, a neural network is termed convergent if the error between the network output value and the predicted value stays within a limited, stable value range.

LITERATURE REVIEW

Roza Dastres and Mohsen Soori (2021) Using linked connections, Artificial Neural Networks are constructed to do complex calculations. An unspecified number of cells, nodes, units, or neurons form the network and link the input to the output. It's a piece of software that simulates the way the human brain analyzes information. Artificial neural networks are used in many different contexts, including but not limited to: self-driving cars, character recognition, picture compression, stock market prediction, risk analysis systems, drone control, welding quality analysis, computer quality analysis, emergency room testing, oil and gas exploration, and many more. ANN systems have several uses in the field of marketing, including but not limited to: predicting customer behavior; developing and comprehending more nuanced buyer categories; marketing automation; content development; and sales forecasting. of order to advance the study field, this article presents a review of current development and applications of the Artificial Neural Networks by analyzing recent successes in the published papers. This allows for the presentation of the created ANN systems and the introduction of new approaches and applications of the ANN systems.

WANG Xin-gang (2017) with the advancement of AI, there has been a heightened interest in artificial neural network algorithm models that mimic the activity of human brain neurons. Artificial neural networks' feature extraction and classification abilities have found widespread application in areas such as image recognition, speech recognition, natural language processing, trend prediction, and others where human image thinking is either unnecessary or inefficient. Also, as training continues, an artificial neural network improves as a machine learning system. This study presents the analytical thoughts and views by analysing and comparing numerous neural network technology models and algorithms based on the development history and research state of artificial neural networks.

Enzo Grossi, Massimo Buscema (2008) Coupling computer science with other theoretical foundations like nonlinear dynamics and chaos theory paves the way for the development of 'intelligent' agents like ANNs that can dynamically adapt to issues with a high degree of complexity. The study of complexity is made possible by the fact that ANNs can mimic the dynamic interplay of several components concurrently, and they can also make conclusions on an individual basis as opposed to only observing patterns. When compared to traditional statistical methods, these tools may have some distinct benefits. The goal of this paper is to introduce gastroenterologists to ANN-related ideas and paradigms. When applied correctly, the ANN family allows for the greatest possible insight into complex, dynamic, and multidimensional events that are difficult to forecast using the more conventional "cause and effect" framework.

Nitin Malik (2005) The purpose of an ANN, which is a functional imitation of a simplified model of biological neurons, is to build useful 'computers' for solving real-world problems and to reproduce intelligent data evaluation techniques like pattern recognition, classification, and generalization using simple, distributed, and robust processing units called artificial neurons. Non-linear static-dynamic systems are implemented in ANNs in a fine-grained parallel fashion. Because of its massively parallel and dispersed structure, neurons have a great processing capacity, which contributes to ANN's intelligence and allows it to handle challenging tasks. The ability to implement ANN algorithms and architectures with VLSI technology for real time applications is a major factor in the recent resurgence of interest in ANN. Theoretical and practical results in a wide range of fields have fueled a rapid expansion in the number of ANN applications in recent years. This article provides an overview of recent studies and the rapid expansion of several ANN-related applications. This article provides a high-level summary of ANN concepts, models, and uses. We explore future trends and identify potential application areas.

METHODOLOGY

The process through which the published works were evaluated and analyzed is described below. Depending on the desired level of detail, several types of documents were gathered and analyzed. The reviews are categorized according to the kind of modifications made, the outcomes, the fields of

application, and the years in which they were published. Journals, conference proceedings, and technical reports were all used in the publishing process. When looking for articles on a certain subject, the term "artificial neural network using conjugate gradient" is often used as a keyword. To further illustrate the breadth of ANN's practical use, a second set of papers was selected. Therefore, the purpose of this work is to zero in on the most promising unexplored element of the development of CG formulation in ANN.

ANN Advancement

Since its discovery in the 1950s, ANN research has seen significant improvement and evolution. There are several methods in ANN that may be used to boost learning performance, so you aren't limited to the CG formulation. Many scientists, nevertheless, use different bio-inspired algorithm methods, such as the genetic algorithm, particle swarm optimization, bee colony optimization, and so on. The following table details the issues faced by ANN researchers, along with the methods they used to address them. Each study's findings are shown as well, demonstrating how the implemented changes enhance the experimental testing data's performance. The table also includes examples of the many different CG that have been used to solve ANN issues. Both feed-forward and back-propagation algorithms are used in this process. The results of using CG demonstrate an increase in precision, reliability, error reduction, and convergence speed.

Table 1: The effectiveness of the ANN-modification method and its results

Problem	Technique	Result
To access CG methods in large-scale typical minimization problems	CGFR, PR, Powell, Shanno-Phua	Result are performed in graph for numerical results on the tested problem. CGR & Powell perform better on 1st test problem. 2nd prob fit better on Shanno-Phuaand Powell.
A study of NN applications in business - The characteristics Feasibility in buss apps & examples ANN drawback	Basic RFN, FFNN	Reviewed of NN common characteristics and feasibilities in business fields
To describe the similarities and differences of ANN and AIS to both nervous and immune system	ART, ART1	ANN show better robustness: – very flexible and damage tolerant , memory:-pattern synaptic strength, learning:- global & local learning rule
To review the ANN applications in forecasting	ANNs approach	ANNs is: Satisfy the performance in forecasting Many factors can affect performance of ANNs
To develop an algorithm for determining the optimal initial weight	FFNN - based on Cauchy's inequality an linear algebraic method	Increase convergence rate and proposed uniform distributed weights performed better than algo in normal distribution. It able to apply on network with different activation function.
To review ANNs utilization in real problems and overview of ANNs theory, learning rules, networks, methodologies and history.	ANNs general theory and BPANNs	Tested on pyrolysis mass spectra for microorganism identification show ANNs more robust and rapid with higher accuracy and cost effectiveness.
To present a realistic guideline of NN in engineering applications	MLP, RBF, NRBS, modified Jenkins 'hypercube'	RBF showed poorer performance. MLP & NRBF equally well but NRBF has faster training
To investigate the use of weather prediction in application of NN to loan forecasting	NN design	The mean of load forecasting using NN was more accurate than traditional procedure. Weather ensemble prediction in NN have strong potential to expand in future.
To introduce a learning method for two-layer NN for sensitivity-based linear learning method	SBLLM, LM,SCG, GDX,GD, SGD	SBLLM – high speed,min error, good performance, homogeneous behavior
To promote a new approach improved training of BP by adapting modify initial search	BFGS, BFGS/AG	The average result for all 4 classification problems (total time convergence) Trainbfgs – 4.465s BFGS – 4.755s BFGS/AG- 3.180s
To improve training efficiency of BPNN by modifying initial search direction with new fast learning algorithm based on FR update with adaptive gain FR	traincgf, CGFR, CG CGFR/AG	Proposed algorithm is generic and easy to implement in all commonly used gradient based optimization processes. It also robust and potentially to enhanced the computational efficiency of the training process.
To model a simulated Water Bath System and a Continually Stirred Tank Heater (CSTH) by using MLP variants for better performance, generalized ability and cheaper	MLP, DFLANN, FLANN, M-FLANN	M-FLANN performed better interpolation (0.03595) and extrapolation (0.2116) ability compare to other NN variants.
To review a purpose and role of ANN in data mining practices	FFNN, BPA	ANN has shown better performances; high accuracy, noise tolerance, ease maintenance, can be implemented in parallel hardware etc
To improve the training efficiency of BP-NN algo by adaptively modifying the initial search direction	CG based NN-algo, FRCG, BP-NN	New technique based on CG updates improve about 10-15% by adaptively modify the search direction. It also robust and potentially enhance the computational efficiency.
To optimize the total structural cost, under constraints related to minimum target for different limit states or performance requirements	NN for mean & standard deviation for dynamic analysis	NN able to represent the relationship between structural responses and intervening random variables also between achieved reliabilities and design parameters
To propose modified algos for FFNN by using gradient of performance function (energy & error)	MMFNN; traincgAJ, traincgll & traincgak	traincgAJ performed better result on 3 tested problem with stable result
To solve imbalance data set problems by focusing on optimizing decision boundary of a step function at ANN and applied PSO to train real predicted output	Std ANN, PSO, Modified ANN	Modified ANN with G-Mean test (Std & Modified ANN) show better result that easy to be implemented in various field. Modified NN : Average : 64.965
To study the minimum temperature for	MLP,BPA	Result shown in graph of accuracy for weather prediction

Chandigarh city by using past of 10 years data and trained its network to analyzed the result by applying NN models.		plotted for 10 years. No solid figure/ value given for this study
To solve imbalanced dataset problems (majority, minority)	BPN, PSO,G-mean Train	Measurement of G-mean Train& Test of proposed ANN show an increment compare to conventional ANN
To improve ABCA based on BPNN for fast and improved convergence rate of hybrid NN learning method	ABCPNN GABPNN	ACBPNN more stable compare to GABPNN with average result 96.145% for all four data set
o review the performance of BPA and several variations to improve the performance assessed by G-mean	BPA, BP-AG, BP-AGMAL	Result for different problems analysis; Accuracy – BPAGMAL -93.1% Training time- BP-AL – 0.00203
To propose a CGNNA in order to achieve a high order accuracy in approximating the second-curvature information of error surface	FNN, CGNNA	Average performance result for every CG Models are: PR – 85.7% PR^{+}- 87% MPR^{+}- 87.78%
To find better accuracy in forecasting by training ANN using several CG models. Numerous of climatic parameter input have been used to test the process.	STD BP, BPCGPB, BPCGFR, BPCGPR	Performance result in % (average) of error prediction: BPCGPR – 3.825% BPCGFR – 2.275% BPCGPB – 4.01% STDBP – 5.47%
To analyze and investigate aspect of QN based on BFGS formula for Hessian approximation with a line search convergent assumption of linearly independent iterates.	BFGS, SRL, DFP updates	Theoretical and numerical aspect of quasi-newton methods are tested based on BFGS update. Modified quasi-Newton are very competitive
To demonstrate better performance of hybrid technique using LM algorithm	ABC-BP, ABC-LM,MABCNN,BPNN, APSO_LM	Proposed APSO_LM is much better compare to other algorithms (BPNN,ABCNN, ABC-BP, ABC-LM) in terms of MSE, SD and accuracy.
To compare the performances of three types of training algorithms for making analysis in medical image of brain hematoma	GD, CG, QN, trainrp, trainscg, trainlm	The convergence speed of trainlm & trainscg higher than other training functions and less number of iterations
To promote a modified PRPCG algorithm for solving nonsmooth unconstrained convex minimization problem	CGPRP	PRP is claimed as most effective CG methods for nonsmooth convex optimization problems.
To compare the minimization error in training the prediction models	GD, Scaled CG, MNN	Scaled CG show faster convergence compared with GD and CG. Able to to train the network to an accuracy level of 10^{-4} and 10^{-5}
To find the best BPA to solve problem of analyzing early detection of DHF	GD, BQN, CGD, RB and LM	LM showed the best outcome compared to other algorithms for solving the earlier problem detection

Study of ANN and MCGs

There has been significant development in several study fields thanks to ANN. ANN is well-suited for use in the fields of medicine and forecasting. More and more fields, from medicine and mathematics to commerce and engineering, are finding usefulness in this expanding field.

Figure 1 shows that the fields of medicine and prediction and forecasting have led the way in ANN papers. One reason why NN has been so successful in both areas is because the method's structure lends itself well to application in either. Implementing a neural network algorithm as a problem solver may improve results in the field of prediction across a wide range of applications, from weather forecasting and agriculture to stock market research. In addition to their usefulness in fields like mathematics and engineering, neural networks may be put to use to assess an organization's needs and then create a customized framework or structure to meet those demands. Research into neural networks also draws from other areas, including business, image processing, neuro-computing, meteorology, and the testing of many types of data. Neural networks, which have recently risen to prominence in the field of image processing, have been extensively included into this investigation. In contrast, image processing emphasizes novel perspectives on data categorization and pattern recognition as essential elements of network design. In image processing, the Self Organizing Mapping (SOM) network is a heritable neural network subtype. Optimizing neural networks using multi-criteria gradients (MCGs) has moved beyond the realm of image processing and into the realms of medicine and forecasting, and will be the focus of future research.

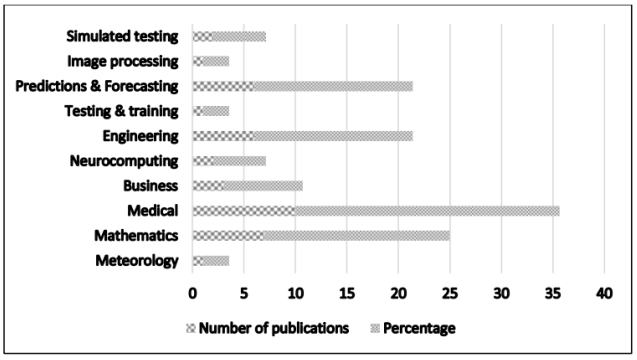

Figure 1: Neural networks have many potential uses

DATA ANALYSIS

MCGs in ANN contributions

Multiple research have shown that neural networks may benefit from a modified version of the conjugate gradient technique. In order to discover the most precise result, the lowest cost function, and the least error square (MSE), the field of optimization is widely employed, beginning with classical CG and basic Gradient Descend (GD).

Table 2: The NN Algorithm Type That Makes Use of

MCGs

MCGs	Description
GD	Gradient Descent – basic and initial CG formulation
CGFR	Fletcher-Reeves improve the GD
CGBFGS	Broyden-Fletcher-Goldfarb-Shanno extended formulation of FR
CGPB	CG formula with Broyden-Powell
CGPR	CG Polak-Ribière improve the previous formula
CGPRP	CG Polak-Ribière-Polyak update the previous PR
CGDP	CG Descent – Powell alike as Quasi-Newton method
trainCG	trainedCG with modification of formulation based on combination of standard CG algorithm
CGNNA	CG with combination of NN functions

Variations on Common ANNSS Alteration Methods

While several studies make use of the concrete MCG formulation, others have found a more effective method of enhancing the neural network learning algorithm. A novel finding with improved performance compared to earlier discoveries was achieved by combining different methodologies. As a result, the following table provides a concise summary of recent trends in neural network tweaking that use not just the conjugate gradient method, but also a hybrid learning algorithm drawn from a wide range of optimization methods.

Table 3: Discovering and Changing Techniques in ANN

Modification Technique	Problem specifications	Analysis
Adaptive gain -modify on the parameter setting. modifies the gradient based search direction by changing the gain value for each node	To improve the training efficiency of gradient descent method To modify the gradient based search direction	Robust, easy to compute and easy to implement for most nonlinear CG algorithm.
Adaptive regulation -iteratively adapts regulation parameters by minimizing validation error using conjugate gradient	To minimize generalization error, the sum of the bias, the variance and inherent noise	Improved algorithm for adaptation of regularization parameters.
Cuckoo Search (CS) & Cuckoo Optimization Algorithm (COA) -minimize a cost function defined as MSE between actual and target output by adjusting weight and biases.	To minimize the cost function of MSE between actual and target input	ANN-COA produced slightly better result for predicting students' performance. COA and CS are able to search optimal or near optimal solution in terms of RMSE, MAPE and R.
Three-term CGA -apply a vector based training algorithm established the convex function.	To overcome problems of NN search directions if the objective function is scaled	Promote the sufficient descent and global convergent.

CONCLUSION

The majority of ANN studies have been on developing more effective learning algorithms for training. In order to accomplish this primary objective, appropriate formulations and algorithms will be used to boost the overall performance of testing data. Otherwise, CG is used as a primary factor in the review's back-propagation error correction. Although extensive study with experimental testing is needed to prove the performance, efficacy, and efficiency of the chosen approach, this is not the case at now. Multiple variants of NN agents and performance factors that might enhance or perhaps be the best answer in that sector can be investigated by expanding the scope of the investigation. Since there are many aspects of neural network learning algorithms to consider, this study may be expanded to include such topics in the future. The network's parameters influencer, the number of neurons and hidden layers, the weight adjustment, and the learning rate may all be tweaked further to obtain the best possible outcome. Therefore, the next step in future research will be to evaluate the new formulation, methodologies, and models with an appropriate data set across a wide range of neural network problem domains.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to Mr. Prashant Arora, Head of Department Computers and Technology, Welham Boys School for his invaluable support throughout the research.

REFERENCES

1. Hinton G E, Srivastava N, Krizhevsky A, et al.(2012) Improving neural networks by preventing co-adaptation of feature detectors. Computer Science, 2012, 3(4): 212-223.

2. C. Szegedy. W. Liu, Y. Jia, et al.(2015) Going Deeper with Convolutions. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015:1-9.

3. Madgwick S O, Harrison A J, Vaidyanathan A.(2011) Estimation of IMU and MARG orientation using a gradient descent algorithm.[C]// IEEE International Conference on Rehabilitation Robotics. IEEE, 2011:1-7

4. Glorot Xavier, A. Bordes and Y. Bengio.(2011) Deep Sparse Rectifier Neural Networks. International Conference on Artificial Intelligence and Statistics.

5. Shuguang Liu, Chongxun Zheng, Mingyuan Liu.(1996) Back Propagation algorithm in feedforward neural networks and its improvement: progress and prospect [J]. Computer Science.23(1).

6. Ke Xu.(2012) Research on application of convolutional neural network in image recognition[D]. Zhejiang University. Proceedings of IEEE. 86(11). 2012:2278-2324.

7. Roza Dastres and Mohsen Soori (2021) "Artificial Neural Network Systems". International Journal of Imaging and Robotics (IJIR), 2021, 21 (2), pp.13-25. ffhal-0334954

8. WANG Xin-gang (2017) "Overview of the Research Status on Artificial Neural Networks" 2nd International Forum on Management, Education and Information Technology Application (IFMEITA 2017) Advances in Social Science, Education and Humanities Research (ASSEHR), volume 130.

9. Enzo Grossi, Massimo Buscema (2008) Introduction to artificial neural network, Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

10. Nitin Malik (2005) National Conference on 'Unearthing Technological Developments & their Transfer for Serving Masses' GLA ITM, Mathura, India 17-18 April 2005.

Corresponding Author

Aryan Rose*

Class 12, Welham Boys School