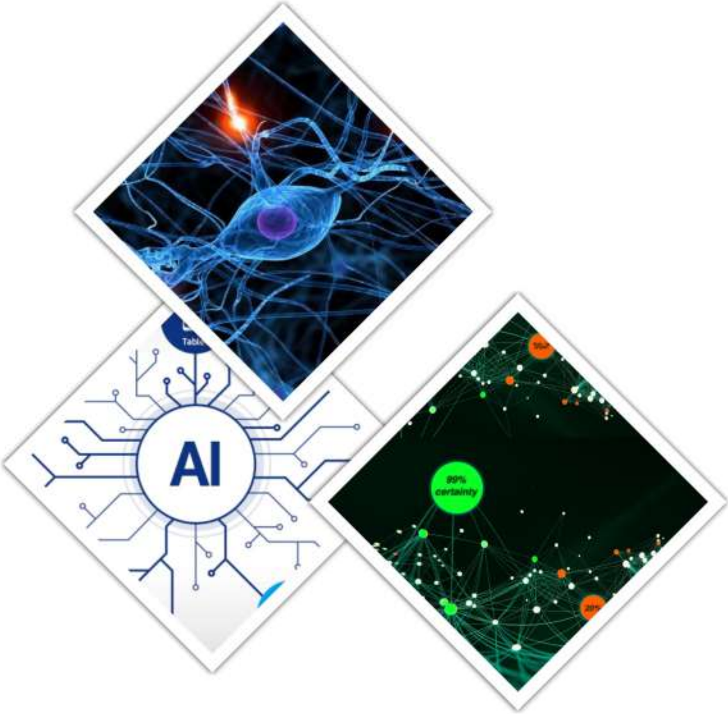
ISBN:978-93-95341-09-7

Introduction to Artificial Neural Networks



Dr. T. Arumuga Maria Devi
Mr. A. Chockalingam | Mrs. P. Thanga Selvi
Dr. M. Santhanakumar | Mrs. R. Hepzibai

INTRODUCTION TO ARTIFICIAL NEURAL NETWORKS

Dr.T.Arumuga Maria Devi
Assistant Professor,
Centre for Information Technology and Engineering,
Manonmaniam Sundaranar University,
Tirunelveli, Tamil Nadu, India

Mr.A.Chockalingam
Assistant Professor (T) & Researcher,
Centre for Information Technology and Engineering,
Manonmaniam Sundaranar University,
Tirunelveli, Tamil Nadu, India.

Mrs.P.Thangaselvi
Researcher,
Centre for Information Technology and Engineering,
Manonmaniam Sundaranar University,
Abhishekapatti, Tirunelveli, Tamil Nadu, India.

Dr.M.Santhanakumar
Assistant Professor (T),
Centre for Information Technology and Engineering,
Manonmaniam Sundaranar University,
Tirunelveli, Tamil Nadu, India.

Mrs.R.Hepzibai
Researcher,
Centre for Information Technology and Engineering,
Manonmaniam Sundaranar University,
Abhishekapatti, Tirunelveli, Tamil Nadu, India.

Published by

SK Research Group of Companies

The International Journal & Book - SKRGC Publication

142, Periyar Nagar, Madakulam

Madurai - 625003, Tamil Nadu, India

skrgc.publisher@gmail.com/+91 9790120237

https://skrgcpublication.org/isbn/

SKRGC Publication
Read | Write | Teach

Title:	INTRODUCTION TO ARTIFICIAL NEURAL NETWORKS
Author's Details:	Dr.T.Arumuga Maria Devi Mr.A.Chockalingam Mrs.P.Thangaselvi Dr.M.Santhanakumar Mrs.R.Hepzibai
Published by:	SK Research Group of Companies, Madurai 625003, Tamil Nadu, India
Publisher's Address:	142, Periyar Nagar, Madakulam, Madurai 625003, Tamil Nadu, India
Edition Details:	I
ISBN:	978-93-95341-09-7
Month & Year	November, 2022
Copyright ©	SK Research Group of Companies
Pages	207
Price	₹450/-

PREFACE

Hello everyone,

We have tried to help you by making a simple book on:

Introduction to Artificial Neural Networks.

So, we hope that our book can be useful to you.

We hope that you like it and thank you.

- Dr.T.Arumuga Maria Devi, Mr.A.Chockalingam,

Mrs.P.ThangaSelvi, Dr.M.Santhanakumar and Mrs.R.Hepzibai

ACKNOWLEDGMENTS

First and foremost, I want to express my gratitude and appreciation to my parents, my family, the Almighty, and all of my well-wishers for their blessings throughout the writing of my book and for helping it to be published successfully.

My heartfelt thanks to Senior Professor Dr.N.Krishnan, Head, Centre for Information Technology and Engineering, Manonmaniam Sundaranar University, Tirunelveli for his continued academic assistance and encouragement to contribute towards intellectual achievements.

My sincere and heartfelt thanks to my husband Dr.P.Kumar, Associate Professor, Centre for Information Technology and Engineering, Manonmaniam Sundaranar University, Tirunelveli for his support and encouragement in successful completion of this book.

I owe a great deal to the co-authors of this book Mr.A.Chockalingam, Mrs.P.ThangaSelvi, Dr.M.Santhanakumar and Mrs.R.Hepzibai for their hard work, helpful suggestions and commitment to complete this book.

- Dr.T.Arumuga Maria Devi

Our special thanks to all the faculty members of Centre for Information Technology and Engineering, Manonmaniam Sundaranar University, Tirunelveli for their support.

Our heartfelt thanks to Dr.M.Sathish Kumar, Publisher, SK Research Group of Companies (SKRGC Publication), Madurai for his effective co-ordination and the publication team for taking this book for publication.

Without the participation and support of several individuals, many of whose names may not be listed here, this undertaking would not have been able to be completed.

We truly appreciate and gladly acknowledge their contributions.

Finally, thanks to all our family and friends for their moral support.

- Dr.T.Arumuga Maria Devi, Mr.A.Chockalingam,

Mrs.P.ThangaSelvi, Dr.M.Santhanakumar and Mrs.R.Hepzibai

CONTENT

CHAPTER I

1.1 A Brief History of Artificial Neural Network (ANN)

The three eras listed below can be used to categorise ANN's history:

1. ANN between the 1940s and the 1960s

The following are some significant advancements from this period:

1943 It has been claimed that the work of biologist Warren McCulloch and mathematician Walter Pitts, who created a straightforward neural network using electrical circuits to represent how neurons in the brain may function, is where the concept of neural networks first emerged.

The Organization of Behavior, written by Donald Hebb in 1949, advanced the idea that repeatedly activating one neuron by another improves its strength each time it is utilised.

Taylor first described an associative memory network in 1956.

Rosenblatt developed the Perceptron learning method for the McCulloch and Pitts neuron model in 1958.

1960 saw the development of the "ADALINE" and "MADALINE" models by Bernard Widrow and Marcian Hoff.

2. ANN from the 1960s to the 1980s

The following are some significant advancements from this period:

Rosenblatt attempted to use the "backpropagation" strategy for multilayer networks in 1961, but was unsuccessful.

Taylor created a winner-take-all circuit with output unit inhibitions in 1964.

By Minsky and Papert, the multilayer perceptron (MLP) was created in 1969.

Kohonen discovered associative memories in 1971.

Adaptive resonance theory was created by Gail Carpenter and Stephen Grossberg in 1976.

3. 1980s to the Present: ANN

The following are some significant advancements from this period:

The main advancement in 1982 was Hopfield's Energy strategy.

1985 saw the creation of the Boltzmann machine by Ackley, Hinton, and Sejnowski.

1986 saw the introduction of the Generalized Delta Rule by Rumelhart, Hinton, and Williams.

In 1988, Kosko created binary associative memory (BAM) and introduced fuzzy logic to artificial neural networks (ANN).

The historical analysis reveals that this field has made substantial advancements. Applications to difficult problems are being developed for neural network-based processors, which are becoming more prevalent. Undoubtedly, the development of neural networks is at a transitional phase right now.

A computational model that replicates how nerve cells in the human brain function is known as an artificial neuron network (or neural network).

Artificial neural networks (ANNs) employ learning algorithms that enable them to autonomously adjust—or, in a sense, learn—as they are presented with fresh data. They become a very powerful tool for non-linear statistical data modelling as a result.

There are three or more interconnected layers in an artificial neural network. Neurons in the input layer make up the first layer. These neurons transmit information to deeper layers, which then transmit the final output information to the final output layer.

CHAPTER II

2.1 Layers used to Build ConvNets

Each tier in a covnets's hierarchy uses a differentiable function to transform one volume into another.

Types of layers:

Let's use an image with the dimensions 32 by 32 by 3 as an example.

1. Input Layer: This layer contains the image's raw input, which is an image with dimensions of 32 by 32 by 3.
2. Convolution Layer: Using the dot product between each filter and each image patch, this layer calculates the output volume. If we apply a total of 12 filters to this layer, the resulting volume will have the dimensions 32 x 32 x 12.
3. The element-wise activation function layer will be applied to the convolution layer's output in this layer. The following activation functions are frequently used: Tanh, Leaky RELU, RELU: max(0, x), Sigmoid: 1/(1+e-x), etc. Since the volume is unchanged, the output volume will have the following dimensions: 32 by 32 by 12.
4. Pool Layer: The main goal of this layer, which is occasionally added to covnets, is to minimise the volume. This expedites computation, conserves memory, and prevents overfitting. Two common types of pooling layers are max pooling and average pooling. In the event that we choose a maximum pool with 2 x 2 filters and stride 2, the resulting volume will have dimensions of 16x16x12.

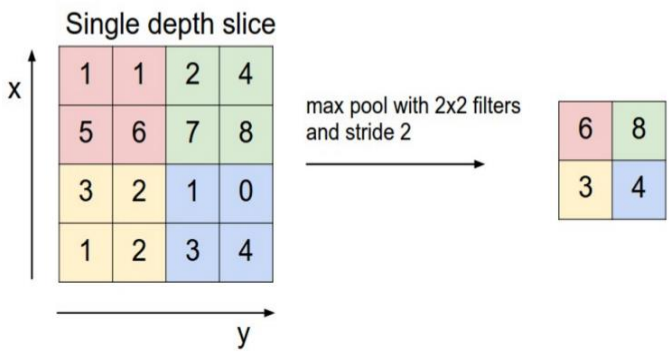


Figure 2.1 Pooling Layers

5. Fully-Connected Layer: This layer, which is a typical neural network layer, calculates the class scores using the input from the layer before it and outputs a 1-D array with a size equal to the number of classes.

2.2 Long-Short Memory (LSTM)

Long Short Term Memory (LSTM) networks are a type of recurrent neural network that may learn order dependence in sequence prediction problems. The current RNN step uses the output from the previous step as its input. The LSTM was developed by Hochreiter & Schmidhuber. It addressed the problem of RNN long-term reliance, in which the RNN can predict words based on recent data but cannot predict words kept in long-term memory. As the gap length increases, RNN's performance is inefficient. By default, the LSTM may hold data for a very long time. It is employed in the processing, prediction, and classification of time-series data.

CHAPTER III

3.1 The Mathematical Representation

1. The Swish activation function has the following benefits over ReLU:

Swish is a smooth function, thus it doesn't change course suddenly like ReLU does close to x = 0. Instead, it gently curves higher from 0 towards values 0, then downward.

- The ReLU activation function cancelled out small negative values. These negative values might be useful for identifying the patterns in the data, though. It is a win-win situation because large negative numbers are wiped out for reasons of sparsity.
- The non-monotonic nature of the swish function improves the representation of input data and weight to be learned.

2. Gaussian Error Linear Unit (GELU)

The top NLP models BERT, ROBERTa, ALBERT, and others are compatible with the Gaussian Error Linear Unit (GELU) activation function. The combination of dropout, zoneout, and ReLUs-related characteristics drives this activation function.

A neuron's output is produced by ReLU and dropout together. ReLU accomplishes this deterministically by dropout stochastically multiplying by zero and multiplying the input by zero or one, depending on whether the input value is positive or negative.

Zoneout, an RNN regularizer, stochastically multiplies inputs by one.

By multiplying the input by either zero or one, which is stochastically determined and depends on the input, we combine this functionality. We take the neuron input x and multiply it by m Bernoulli(x), where x is the standard normal distribution's cumulative distribution function and (x) = P(X x).

Since neuron inputs typically follow a normal distribution, especially with batch normalisation, this distribution was used.

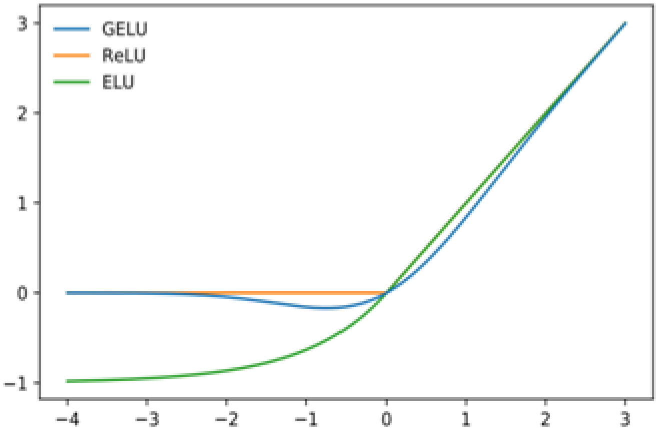


Figure 3.1

Gaussian Error Linear Unit (GELU) Activation Function

Mathematically it can be represented as

GELU

$$f(x) = xP(X \le x) = x\Phi(x)$$
$$= 0.5x\left(1 + tanh\left[\sqrt{2/\pi}\left(x + 0.044715x^3\right)\right]\right)$$

For all tasks in the areas of computer vision, natural language processing, and speech recognition, GELU nonlinearity outperforms ReLU and ELU activations.

The Scaled Exponential Linear Unit (SELU) was developed in self-normalizing networks and handles internal normalisation, ensuring that the mean and variance of each layer are maintained. This normalisation is made possible by SELU by modifying the mean and variance.

CHAPTER IV

4.1 Bidirectional Associative Memory (BAM)

The auto-associative Hopfield network and ANNs, which acquire associations of data via continuous training, have several known limitations. Early in the 1980s, Bart Kosko created the bidirectional associative memory (BAM), a unique type of recurrent neural network (RNN).

The BAM is a hetero-associative memory that offers the capacity to store related material, irrespective of its nature and structure, without the need for ongoing learning. Data associations are only ever stored simultaneously once before being recalled from the BAM's memory. The BAM typically outperforms the current ANNs, offering noticeably greater performance in recalling the associations. Additionally, compared to previous ANNs, using the BAM perfectly resolves the known bitwise XOR problem, making it possible to store the data encoded to binary (bipolar) form in its memory.

The BAM is able to recall the associations for data assigned to its either inputs or outputs, bidirectionally, unlike the unidirectional Hopfield network. Additionally, it enables extracting the accurate associations from input data that is inadequate or corrupted.

When used as a component of an AI-based decision-making process, the BAM models are fairly effective at predicting the answer to a particular data analysis problem based on multiple associations of numerous associated data.

The BAM model can be used for a variety of applications, including the following:

4.2 Classification and Clustering Data

- Incomplete data augmentation
- Recovering damaged or corrupted data

When the several types of information that the ANNs have learned are insufficient to process the data that has been given to an AI for analysis, the BAM models come in very handy.

For instance, the prediction of missing words from incomplete texts using ANN essentially necessitates the storage of word-to-sentence correlations in the ANN's memory. The same missing words can appear in multiple incomplete sentences, so this would result in an erroneous forecast. In this situation, using the BAM enables the storage and recall of all potential relationships between the data elements, including those between words and phrases, sentences and words, and words and sentences, as well as vice versa. In turn, this significantly improves prediction accuracy. It is also advised to multi-directionally associate different data entities in order to either enhance or cluster these data.

Finally, the size of the greatest floating-point type, float64, is used to constrain each ANN layer's memory capacity. For instance, only 1024 bytes are available in a single ANN layer of the shape (100x100). It goes without saying that in order for the ANN to be able to store the associations and recall them in several directions, the number of layers and neurons inside each layer must be increased. This has an adverse effect on ANN-based memory latency since the demands for learning and making predictions are increasing proportionally to the size of the ANN.

Despite this, the BAM's outstanding learning and prediction performance allows for allocation of as much memory as needed, beyond the memory capacity restrictions of conventional ANNs. Additionally, numerous BAMs with layered structures may be aggregated into memory.

4.3 Associations Stored In Memory

The BAM learns the bipolar patterns that can be created by associating various pieces of data. A binary vector, whose

CHAPTER V

5.1 Self Organizing Maps – Kohonen Maps

is a form of artificial neural network that was also influenced by 1970s biological models of brain systems. It employs an unsupervised learning methodology and used a competitive learning algorithm to train its network. In order to minimise complex issues for straightforward interpretation, SOM is utilised for clustering and mapping (or dimensionality reduction) procedures to map multidimensional data onto lower-dimensional spaces. The input layer and the output layer are the two layers that make up SOM. Below is a description of the Self-Organizing Map's architecture with two clusters and n input features for any sample:

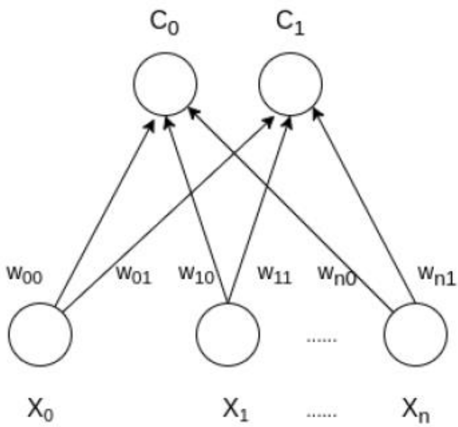


Figure 5.1 Kohonen Maps

How do SOM works?

Consider an input set with the dimensions (m, n), where m represents the number of training examples and n represents the number of features present in each example. The weights of size (n, C), where C is the number of clusters, are first initialised. The winning vector (the weight vector with the shortest distance from the training example, for example, the Euclidean distance) is then updated after iterating over the input data for each training example. Wij = wij(old) + alpha(t) * (xik - wij(old)) is the formula for the weight updating algorithm, where alpha signifies the learning rate at time t, j stands for the winning vector, I for the ith feature of the training example, and k for the kth training example from the input data. The SOM network is trained, and trained weights are utilised to cluster new examples. A new example is included in the collection of successful vectors.

5.2 Neighbor Topologies in Kohonen SOM

The following two topologies, though there are more, are most frequently used:

Topology of a Rectangular Grid

In this topology, the difference between each rectangular grid is 8 nodes, with 24 nodes in the distance-2 grid, 16 nodes in the distance-1 grid, and 8 nodes in the distance-0 grid. The number denotes the winning unit.

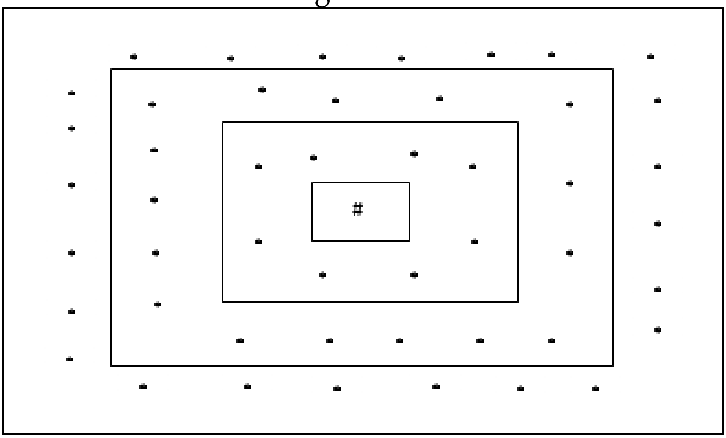


Figure 5.2 Kohonen SOM

For Full Book Contact

Dr.T.Arumuga Maria Devi

Assistant Professor,

Centre for Information Technology and Engineering,

Manonmaniam Sundaranar University,

Tirunelveli, Tamil Nadu, India

Email: deviececit@gmail.com

Email: arumugamariadevi@msuniv.ac.in

Mobile: +91 8667899606

About the Authors

Dr. T. Arumuga Maria Devi B.E., M.Tech., Ph.D., Received B.E degree in Electronics & Communication Engineering from Manonmaniam Sundaranar University, Trinelveli, Tamil Nadu, India in 2003. Completed M.Tech degree in computer & Information Technology from Manonmaniam Sundaranar University, Tirunelveli, Tamil Nadu, India in 2005. Received Ph.D Degree in Information Technology Computer Science and Engineering from Manonmaniam Sundaranar University, Tirunelveli, Tamil Nadu, India in 2012. She is the Assistant Professor of Centre for Information Technology and Engineering, Manonmaniam Sundaranar University, Tirunelveli since 2005. Her Research Includes Signal Processing, Remote Communication, Multimedia and Mobile Computing.

Mr. A. Chockalingam, B.Tech., M.Tech., is a Researcher and Assistant Professor (T) in Centre for Information Technology and Engineering, Manonmaniam Sundaranar University, Tirunelveli, Tamil Nadu, India. He has three years of teaching experience and five years of Research experience in areas like Data Analytics, Deep Learning, Machine Learning and Artificial Neural Networks. He has contributed in a research project "Automatic License Plate Recognition" granted by Centre for Development of Advanced Computing (C-DAC), Trivandrum, India. He has completed '4' edX Verified Certifications in Data Analytics and Cyber Security domains. He has served as a resource person in two Faculty Development Programs, attended 25 FDPs, Workshops, Conferences and Webinars and co-authored a book.

Mrs. P. Thanga Selvi, Research Scholar, Centre for Information Technology and Engineering, Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli - 627012, Tamilnadu, India. She received her Master of Computer Application in Anna University and received Bachelor of Computer Science in Manonmaniam Sundaranar University, Tirunelveli. Her research interests include Networking, Wireless Communication and Hyperspectral Image Processing.

Dr. M. Santhanakumar M.E., Ph.D., Received his BE degree in Computer Science and Engineering from Manonmaniam Sundaranar University, Tirunelveli in 2004, Completed M.E. degree in Computer Science and Engineering from Anna University, Chennai, Tamilnadu, India in 2007 and Received Ph.D. degree in Information and Communication Engineering at Anna University, Chennai, Tamilnadu, India in 2018. Currently he is working as a Assistant Professor (T) in Centre for Information Technology and Engineering, Manonmaniam Sundaranar University, Tirunelveli, Tamilnadu, India. He is working in the field of web content mining and web usage mining. His research papers have been published in International Journal of Business Information Systems, Asian Journal of Information Technology, WSEAS Transactions on Computers and International Journal of Applied Engineering, etc.

Mrs. R. Hepzibai, M.Sc., (Ph.D), is a Researcher in Centre for Information Technology and Engineering, Manonmaniam Sundaranar University, Tirunelveli, Tamil Nadu, India. She has five years of Research experience in areas like Hyper spectral imaging, Deep Learning, and Machine Learning. She has completed Professional Certification Program from IBM Web Sphere IIS Data stage Enterprise Edition V7.5. She also has Professional membership in International Association of Computer Science and Information Technology (IACSIT) and International Association of Engineers (IAENG).She served as a resource person for Three days Training programme on "Basics of Computer and its applications" at Manonmaniam Sundaranar University and also she attended Workshops, Conferences and Webinars.

Published by

SK Research Group of Companies

The International Journal & Book - SKRGC Publication

142, Periyar Nagar, Madakulam
Madurai - 625003, Tamil Nadu, India
skrgc.publisher@gmail.com/+91 9790120237
https://skrgcpublication.org/isbn/

SKRGC Publication
Read | Write | Teach

ISBN 978-93-95341-09-7