International Journal for Innovative Engineering and Management Research

PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL

www.ijiemr.org

IJIEMR Transactions, online available on 26th Dec 2022. Link :http://www.ijiemr.org/downloads.php?vol=Volume-11&issue=Issue 12

10.48047/IJIEMR/V11/ISSUE 12/132

TITLE: ARTIFICIAL NEURAL NET BASED CLASSIFICATION TOWARDS AUDIO EXTRACTION

Volume 11, ISSUE 12, Pages: 992-998

Paper Authors **APARNA PANIGRAHY, DR. SANJAY KUMAR TIWARI**

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

ARTIFICIAL NEURAL NET BASED CLASSIFICATION TOWARDS AUDIO EXTRACTION

CANDIDATE NAME= APARNA PANIGRAHY
DESIGNATION= RESEARCH SCHOLAR SUNRISE UNIVERSITY ALWAR
GUIDE NAME = DR. SANJAY KUMAR TIWARI
DESIGNATION= PROFESSOR
SUNRISE UNIVERSITY ALWAR

ABSTRACT

The recent interest in artificial neural networks has prompted several implementations across many different subfields of study. Learning in neural networks offers a promising alternative to more traditional approaches to study. Neural networks may be used to a wide variety of issues, including those involving pattern association, pattern classification, regularity detection, image processing, voice analysis, simulation, and so on. In order to accomplish our Content-based audio classification job, a neural classifier was developed in this study employing artificial neural networks. To put it simply, a neural network is a computer model that attempts to mimic the way the human brain works when learning. When we talk about "artificial" neural networks, we're referring to those that have been built into computer programs that can do the massive amounts of computations required for learning.

KEYWORDS: Artificial Neural Net, Audio Extraction, computer model.

INTRODUCTION

The capabilities of the brain are the result of the vast interaction between all cells and their simultaneous processing; individual cells function like basic processors. More than a billion nerve cells do information processing in the human brain. A diagram of a neuron (a kind of nerve cell) is shown in Figure 1. A neuron, as seen in the illustration, has a core, dendrites that receive information, and an axon that has its own dendrites and sends the information along to other neurons.

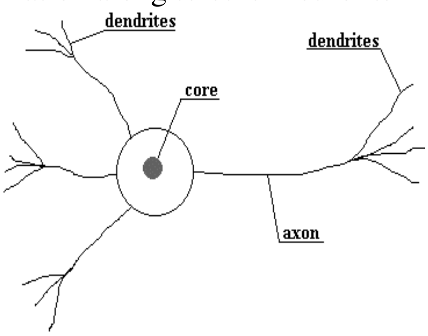

Figure 1: Structure of a neural cell in the human brain

Along the dendrites, neurons transmit information through electrical impulses. Dendrites collect information from outside the neuron, mix it together, and send it down the axon to the terminal dendrites, from whence it may be sent on to other neurons if the stimulus is strong enough to activate the neuron. A neuron is considered to be inhibited if it is unable to transmit information because of insufficient incoming stimulation.

NEURAL NETWORK

A neural net, like the human brain, has neurons and connections between them. Information is being sent from one neuron to another through its outgoing connections. Weights are what link neurons in a neural network. These weights have numerical values that are used to imitate the "electrical" data. Altering these weight values simulates different network architecture. An idealized neural network neuron is shown in Figure 2. The neuron receives data (the input) on its incoming weights in this network. A propagation function operates on this data by adding up the values of all weights that are sent in. The activation function of the neuron then compares the result to a predetermined threshold. When the input level is greater than the threshold, the neuron becomes active; otherwise, it remains inactive. When a neuron is stimulated, it sends a signal through its efferent weights to all of the other neurons in its network.

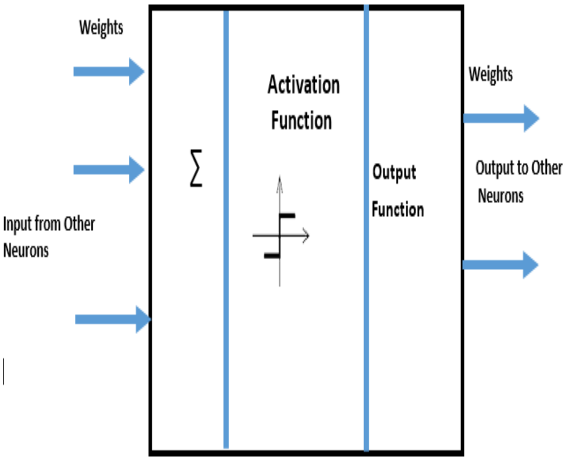

Figure 2: Structure of a neuron in a neural net

The layers of a neural network are the namesake for the groups of neurons they include. Except for the input and output layers, each neuron in a given layer are typically connected to all neurons in the layer above and below it. A neural network may include zero, one, or more hidden layers, which are responsible for relaying data from the input layer to the output layer. Information may also be sent in the other direction, backwards across the network, depending on the learning method in use. A neural network with three layers of neurons is seen in Figure 3. A neural network does not often look like this. The neurons in a layer of certain neural nets are organised as a matrix, whereas in other cases there are no hidden layers at all. There are many different kinds of neural networks, but what they always have in common is at least one weight matrix and connections between at least two layers of neurons.

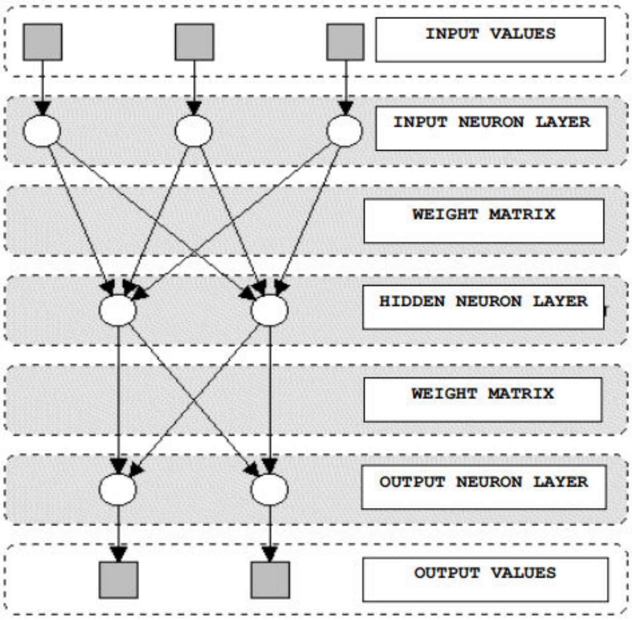

Figure 3: Neural net with three neuron layers

TYPES OF NEURAL NETWORKS

There is a wide variety of neural networks, each with its own set of advantages. Neural networks, in general, are often regarded as highly adaptable problem-solving tools. Ability, specifically the error tolerance of neural networks, should be addressed clearly. That is to say, if a neural net has been taught to handle a certain kind of issue, it will be able to correctly remember its previous solutions even if the new problem being addressed is somewhat different. While neural networks are effective at solving complex issues, there is no assurance that the answers they provide will be right. These solutions are just estimates, and they always include some amount of inaccuracy.

Perceptron

F. Rosenblatt pioneered the use of the perceptron in 1958. It is a very basic neural net that consists of only two layers of neurons and works with binary data (0 or 1). Learning occurs under supervision, and the net is put to work classifying patterns. The basic perceptron may be shown in Figure 4.

Multi-Layer Perceptron

In 1969, M. Minsky and S. Papert presented the first implementation of the Multi-Layer-Perceptron. As can be seen in figure 4.4, this extended Perceptron consists of an input layer, an output layer, and one or more hidden neuron layers in the middle.

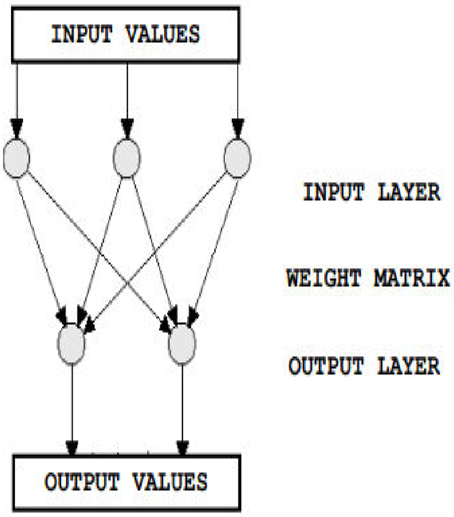

Figure 4: Multi-Layer-Perceptron

Back Propagation Network

One of the most effective neural network types is the Back Propagation Network, which was developed by G.E. Hinton, E. Rumelhart, and R.J. Williams in 1986. Figure 5 depicts the back propagation learning technique being applied to a network with a structure identical to that of the Multi-Layer-Perceptron.

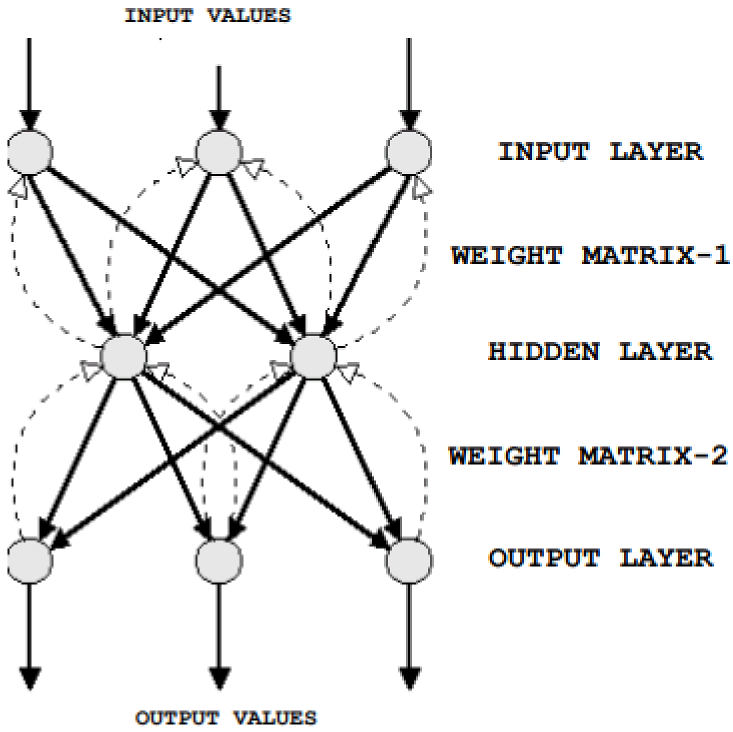

Figure 5: Back propagation Net

LEARNING

The dendrites of neurons in the human brain are the pathways through which information is communicated. When a neuron receives enough stimulation, it sends a signal to all of its neighbors. From there, the data travels to its final destination, where some action is triggered. If the incoming stimulus is too weak, the neuron will not produce an output, and the transmission of the information will be halted. It's not easy to explain how the human brain learns, and nobody understands the precise process. It is thought that during learning, the structure of the connections between neurons changes, allowing only specific neurons to respond to certain stimulations. This implies that, after first learning a fact, strong connections establish between the relevant brain cells, facilitating instant recall.

When these identical brain cells are activated again after learning new information, their connection structure will change to reflect the new data. On the other hand, if it has been a while since a person has remembered certain knowledge, the "weak" connection structure between the relevant brain cells will have developed. It has occurred if the learner "forgot" or dimly remembered a previously acquired knowledge.

A neural net, in contrast to the biological model, has a fixed structure consisting of a fixed number of neurons and fixed numbers of connections between them (called "weights"), each of which has a fixed value. The weights themselves are what evolve with training. This signifies, in comparison to the original: When new information "stimulates" (exceeds a threshold value of) certain neurons, those neurons either send the data via their weighted connections to other neurons or block the data from traveling any further. A weight's value increases when data has to be moved and decreases otherwise. To ensure that all inputs result in the intended output, the weight values are adjusted dynamically as the network learns new ones. A matrix containing the weight values between the neurons is generated during the training of a neural net. After proper training, a neural network should be able to use these matrix values to determine the appropriate output for a particular input. After the learning process, there is always some residual error; therefore the produced output is usually just a decent approximation of the ideal output.

Supervised Learning

Supervised learning is a popular training approach in which the system attempts to make predictions based on existing samples. It "learns" from its errors by comparing its predictions to the desired response. The information is first sent to the neurons in the input layer. The information is sent from neuron to neuron. The inputs are given a weighting at each link before being sent to the next node, where the weightings are added together and either strengthened or reduced. This process is repeated until the data are at the output layer, when the model makes a prediction. Predicted results are evaluated against the actual results in a supervised learning system. No adjustment is made to the system weights if the expected and actual outputs are the same. However, the inaccuracy is sent back into the system and the weights are changed if the anticipated output is either greater or lower than the actual result in the data. We refer to the process by which a mistake is sent back into a network as "back-propagation." Supervised learning approaches like the Multi-Layer Perceptron and the Radial Basis Function are discussed. Back-propagation is used by the Multi-Layer Perceptron, whereas the feed-forward strategy used by the Radial Basis Function only requires a single pass of the data during training.

Unsupervised Learning

As opposed to making predictions, unsupervised learning neural networks excel in describing data. During training, the neural network is not presented with any outputs or replies; indeed, in a neural network, the idea of output fields does not even exist. Kohonen networks are the

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

major unsupervised method. Cluster analysis, in which "like" examples are grouped together, is where Kohonen and other unsupervised neural networks shine.

Back Propagation Learning Algorithm

Multi-Layer Perceptron relies heavily on back propagation, a supervised learning process, to update the weights of the network's hidden neuron layer(s). The weight values are adjusted retroactively through a determined output error in the back propagation process. This overall slip must have originated from prior forward propagation work. In forward propagation, the sigmoid activation function is used to stimulate the neurons.

Designing of Audio Classifier

The Audio classifier developed in this study, which is a Multi-Layer Perceptron (MLP) layered Feed forward network. There are fourteen input neurons in the network's single layer, seven hidden neurons in the hidden layer, and three output classes in the network's single layer of output neurons. First, the audio is sorted into the categories of news, sports, and music. There are three subgenres of upbeat music: cheerful, furious, and sorrowful. The supervised back propagation learning technique is used to train the classifier. The classifier takes the retrieved low-level audio attributes as input data. The information is separated into training data and test data. Attribute values in the database must be normalized such that they fall inside the range [0,1] or [-1,1] internally. There are two primary methods for standardizing data:

(i) Max- Min normalization: The data is linearly transformed using the Max-Min normalization method. Let's assume that the lowest and maximum values of attribute A are mina and maxA respectively.

(ii) Decimal Scaling Normalization: To normalize the values of characteristic A, decimal scaling simply shifts the decimal point. MLP is fed information about extracted, normalized audio characteristics. The processing units in an MLP are arranged in layers and linked together through weighted connections. There is a secret intermediate layer between the input and output layers.

Input Layer: The input layer showcased the audio features, which totaled fourteen. Data between 0 and 1 might be normalized and sent to the input layer. The Sigmoid Function is used by the input layer to transform the data it receives into an output, which is often a number between 0 and 1. Each input layer cell has a weighted relationship to every hidden layer cell. Input layer bias and weights are often set to zero as a modeling or simplification shortcut.

Hidden Layer: The network has a single hidden layer made up of seven neurons. The hidden layer's cells are completely coupled to their counterparts in the input and output layers. During training, a Neural Network's overfitting may be prevented by changing the number of cells in a hidden layer. The hidden layer's weights and bias are each given a random value between -1.0 and 1.0 or -0.5 and 0.5 as a starting point. Each cell in the hidden layer uses these starting values to calculate its output, which is then passed on to the output layer.

Output Layer: Three neurons make up the output layer. Each neuron stands for a distinct genre category in the final output. Each output layer cell is directly linked to every hidden layer cell through a random-valued bias and weighted connections. Every hidden layer cell feeds information into every cell in the output layer. Cells in the output layer, like those in the hidden layer, perform the same calculations on input and output.

CONCLUSION

The goal of this study was to use an artificial neural network to classify audio based on its content, by combining the numerous elements retrieved from audio snippets. These methods have allowed us to refine our approach to low-level audio feature-based characterisation of audio. This methodology uses ANN-MFCC to extract audio properties such as time domain, pitch, frequency, and sub band energy. The aforesaid data led to the development of a supervised back propagation learning technique inside a multi-layer feed forward neural network audio classifier. The network learns from audio characteristics that have been retrieved. The algorithm was able to determine if the audio sample was related to sports, the news, or music. Additionally, the technique is used to categorize the music into moods like angry, sad, and cheerful. Even though our training set is relatively small, it contains audio clips with different semantic structures, and this work validates the choice of audio classification features, the results obtained for the classification of audio categories are quite encouraging. Artificial neural networks are one of the most notable areas of invention in the science of artificial intelligence, which has benefited greatly from the fast expansion of technology. These very complex computational models have found widespread use in many fields due to their resemblance to the human brain in areas such as computer vision, natural language processing, and audio categorization. The purpose of this extensive research was to investigate the inner workings of artificial neural networks and their application to the creation of a Content-Based Audio Classification Ordering System. By analyzing this system in detail, we have shown how AI is boosting user experiences and simplifying content management by revolutionizing the way we engage with audio.

REFERENCES

1. Li, I. K. Sethi, and N. Dimitrova et al., “Classification of general audio data for contentbased retrieval,” Pattern Recog. Lett., vol. 22, no. 5, pp. 533-544, April 2001.
2. D. Manoussaki, R.S. Chadwick, D.R. Ketten, J. Arruda, E. Dimitriadis, and J.T. O’Malley. The influence of cochlear shape on low-frequency hearing. ProcNatlAcadSci USA, 105(16):6162–6166, 2008.
3. D. Manoussaki, E. Dimitriadis, and R.S. Chadwick. The cochlea’s graded curvature effect on low frequency waves. Physical Review Letters, 96(8), 2006.
4. D.YangandW.Lee. Disambiguating music emotion using software agents. In Proceedings of the 5th International Conference on Music Information Retrieval, Barcelona, Spain, 2004.
5. Didiot, I. Illina, and D. Fohr et al., “A wavelet-based parameterization for speech/music discrimination,”

Comput. Speech Lang., vol. 24, no. 2, pp. 341-357, Apr. 2010.

6. E. Pampalk, A. Flexer, and G. Widmer, “Improvements of Audio-Based Music Similarity and Genre Classification,” in Proc. of the 6th Int. Conference on Digital Audio Effects, London, UK, Sept. 2003.
7. E. Scheirer and M. Slaney, “Construction and evaluation of a robust multifeature speech/music discriminator,” in Proc. Int. Conf. Acoustics, Speech, Signal Processing ’97, Munich, Germany, 1997.
8. J. Lu, “Multimedia database management systems”, Boston; London: Artech House, 1999.
9. Tzanetakis and P. Cook, “Musical Genre Classification of Audio Signals,” IEEE Trans. Speech Audio Process., vol. 10, no. 5, pp. 293-302, Jul. 2002.
10. G. Tzanetakis and P. Cook. Marsyas: a framework for audio analysis. Organised Sound, 4(3):169–175, 1999.
11. I Nelken, Y Rotman, and O Bar Yosef. Responses of auditory-cortex neurons to structural features of natural sounds. Nature, 397:154–7, 1999.