International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

A STUDY OF ARTIFICIAL NETWORK TOWARDS AUDIO ORDERING SYSTEM

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

ABSTRACT
In this research, we investigate the use of convolutional neural networks to create a ranking system for categorizing audio based on its content. Using neural networks, this cutting-edge solution streamlines and personalizes the audio experience for consumers depending on the content's innate features. This research dives deeply into the inner workings of artificial neural networks to clarify their structure and operation. As a result, it's clear that large and varied datasets are crucial for training these networks properly. Research also highlights the need of continuously optimizing neural network topologies to accommodate shifting user tastes and content trends. Examining the practical consequences of the Content-Based Audio Classification Ordering System is a major contribution of this research. It explains how this system might make users' lives easier by speeding up and simplifying tasks like information classification and retrieval. Benefits to content producers and distributors in terms of audience analysis and content improvement are also discussed.

KEYWORDS: Artificial Network, Audio Ordering System, information classification

INTRODUCTION

Media such as speech, music, and ambient noise all carry vital messages. As a result, the challenge of categorizing these various sounds is gaining importance. A human listener can tell the difference between several audio kinds after hearing just a little snippet of the signal. However, employing computers to resolve this issue has proved to be very challenging. However, numerous devices with just mediocre precision might still be deployed. The ability to analyze and categorize audio has several uses. Examples of widespread use include the maintenance of audio archives, the use of commercial music in various contexts, surveillance, and the entertainment business. In order to search and index the various available digital audio databases on the web, audio segmentation and categorization is essential. Classifying speech data according to speaker might aid in effective navigation through broadcast news archives, which has attracted a lot of attention recently.

Signal processing and classification are the two primary pillars of audio classification, as they are of many other pattern classification problems. Features extraction from an audio source is the focus of the signal processing section. Time-frequency analysis techniques, first developed for speech processing, are used here to handle audio information. Data classification is the focus of the classification phase, which uses the statistical information gleaned from the signals to assign classes to the data.

The proliferation of portable digital audio players has led to a rise in the storage capacity of personal computers for digital music files. When you want to listen to music that fits a certain emotion, like cheerful, sad, or furious, it might be challenging to narrow down your options. Both the individual listener and the online music retailer will have to spend time categorizing their music collections.

Music's emotional components are the ones most closely linked to its expressiveness, according to studies in music psychology and music education. Music mood emotion has been recognized as a significant criteria employed by individuals in music searching indexing and storage, according to research of music information behavior. Music mood is subjectively evaluated, making categorization challenging. Although there seems to be a very strong connection between the audio (music) and the emotional state of the listener.

While listening to music, there are numerous external factors that subtly alter it. Some examples of such things include time signatures, musical instruments, and scales. Lyrics are a singularly significant thing because of the impact they have on our thoughts. It's a challenging subject since it involves intricate challenges of Digital Signal Processing to extract audible words from songs and then categorize the music properly. These computations investigated a rather elementary method of deciphering music from sound patterns. There have been attempts to solve this issue by extracting patterns from auditory characteristics, which is similar to the traditional pattern recognition problem.

ORIGIN OF ARTIFICIAL INTELLIGENCE SYSTEM

1956 The phrase "artificial intelligence" was originally used by John McCarthy, who also organized the first meeting entirely dedicated to the issue, the Dartmouth meeting.

The Logic Theorist (LT), created by Allen Newell, J.C. Shaw, and Herbert Simon at what was then Carnegie Institute of Technology and is now Carnegie Mellon University, was shown.

1957 Newell, Shaw, and Simon's GPS, or General Problem Solver.

1952-62 In the game of checkers, Arthur Samuel of IBM created the first computer program that could compete with a human world champion. The excellent skill of the checkers player was the result of Samuel's machine learning systems.

1958 The Lisp language was created by John McCarthy of MIT.

The geometry theorem prover developed by Herb Gelernter and Nathan Rochester (IBM) uses a "typical" case diagram as a semantic description of the domain.

Margaret Masterman and her Cambridge University coworkers develop semantic nets for machine translation in the late 1950s and early 1960s.

1961 First written in Lisp by James Slagle (Ph.D. research, MIT), the symbolic integration software SAINT was used to answer first-year calculus problems.

1962 Unimation was the first firm to specialize on manufacturing robots.

1964 Bert Raphael's work at MIT on the SIR program exemplifies the usefulness of a knowledge-based ontology for FAQ generators.

1965 ELIZA is an English-speaking conversational bot developed by MIT's Joseph Weizenbaum. When a version was created that "simulated" the discussion of a psychiatrist, it became a popular toy at AI labs on the ARPA-net.

1967 Dr. Joel Moses of MIT The first knowledge-based software to be fully implemented in mathematics.

Doug Engelbart created the mouse at SRI in the late 1960s. Washington, DC was the site of the first International Joint Conference on Artificial Intelligence (IJCAI).

Jane Robinson and Don Walker founded the groundbreaking Natural Language Processing group at SRI in the early 1970s.

In 1971, Terry Winograd's PhD thesis at MIT coupled his language comprehension software, SHRDLU, with a robot arm that carried out orders written in English to show that computers could grasp English phrases within the limited universe of children's blocks.

In a match and replay that attracted a lot of attention (See Deep Blue Wins), the chess software Deep Blue ultimately prevailed against reigning world champion Garry Kasparov. (May 11th, 1997).

Over 5,000 people came out to see the first official Robo-Cup soccer event, which included table-top contests with 40 teams of interacting robots.

In the year 2000, the concept of interactive robot pets (also known as "smart toys") first proposed by novelty toy manufacturers in the 18th century became commercially accessible.

MIT researcher Cynthia Breazeal writes about her emotional robot KISMET in a published paper.

The autonomous Stanford car Stanley was declared the winner of the DARPA Grand Challenge. (2005) October. (See A victor in a grueling desert race, but not behind the wheel.

In search of meteorite samples, the Nomad robot travels to uncharted parts of Antarctica.

Today To see the creation of history, check out AI in the Media.

AREAS OF AI RESEARCH (AI AND RELATED FIELDS) AND ITS APPLICATIONS

Expert System:

- An expert system, often called a knowledge based system, is a computer software that mimics the behavior of a human expert in a certain field.
- Expert systems are computer programs that use encoded information to do tasks that would otherwise need human understanding.
- The inference in this system is carried out by use of symbolic calculations.
- The current generation of expert systems is meant to complement existing expertise, not replace it.
- Medical diagnosis, chemical analysis, geological research, and computer system setup are just a few of the many fields where they have proved valuable.
- Since the future of the expert system sector holds enormous promise for both practical applications and economic gain.The media has started paying a lot of attention to it.

Natural language processing:

- ✓ The purpose of natural language processing is to bridge the gap between humans and machines by facilitating communication in human languages like English rather than via the use of specialized programming languages.

- ✓ There are two subfields within N.L.P. :
- ✓ Natural language comprehension, which explores ways to teach a computer to understand spoken English. This helps computers better understand their human counterparts.
- ✓ To help humans better comprehend computers, "natural-language generation" software attempts to generate text in a natural language, such as English.

Speech recognition:

Reading and writing are secondary to voice as the main interactive means of human communication.

- ✓ In order for computers to hear our voices and comprehend what we're saying, researchers have been hard at work on speech recognition technology.
- ✓ By streamlining human-machine dialogue, speech recognition studies aim to further the field of natural language processing.

Computer vision:

- Connecting a camera to a computer so it can see photographs is a quick and easy process.
- However, it is not an easy process to interpret such visuals so that the computer fully comprehends what it is viewing.
- Vision is the key sense that humans employ to take in information about their surroundings. The sense of sight is far more dominant than the other senses.
- The purpose of computer vision research is to enable computers to see and interpret their environments in the same way that humans do.

Robotics:

- Electromechanical machine that may be instructed to carry out human-like manual labor.
- A robot is "a re programmable multi-functional manipulator designed to move material, parts, roots, or specialized devices through variable programmed motions for the performance of a wide range of tasks," according to the Robotic Industries Association.
- Robotics is not regarded to be a subset of artificial intelligence.
- A 'dumb' robot is one that has no intellect beyond what it has been programmed to carry out.
- An intelligent robot has sensors, such a camera, that enable it to adapt to its surroundings rather than just'mindlessly' carry out commands.

Some Examples

- "Chinook" won the 1994 man-machine checkers global championship. Checkers is a family of two-player board games characterized by the use of uniform pieces that move diagonally and are captured by jumping over their opponents' pieces.

- On February 10, 1996, Deep Blue, a chess-playing computer, defeated Gary Kasporov.
- Industrial control systems have made extensive use of Fuzzy Logic, a method of reasoning in the face of ambiguity.It's a kind of reasoning that takes into account more than just yes/no answers. If there are no clouds, the statement "Today is Sunny..." may be absolutely accurate; if there are a few clouds, it may be 80% true; if it is foggy, it may be 50% true; and if it rains all day, it may be 0% true.
- A machine vision system is a computer vision application used in manufacturing. Machine vision systems employ digital cameras and image processing software to examine and make choices in a manner similar to that of human inspectors. Used for such purpose, too.
- An OCR system is a technique for reading text that has been scanned into a computer from an image. They are proficient at deciphering the European script used in typewritten documents.
- **Hand writing Recognition** Personal digital assistants (PDAs) utilize this. There are millions of them. It's the computer's capacity to take in reasoning.
- **Speech Recognition** is a computer's capacity to understand and obey voice orders.
- **Neural Networks** utilized everywhere from IDSs to gaming engines to bioinformatics An animal brain serves as the inspiration for the structure and operation of a neural network, which is a linked collection of basic processing components or nodes.

ARTIFICIAL NEURAL NETWORKS AND ITS APPLICATIONS

What is a neural network?

A neural network is a versatile system that can pick up on patterns by being shown data over and over again, and then apply those patterns to brand-new situations. In the case of supervised networks, human input is required to specify the goals of training. In contrast, unsupervised networks have their information architectures predetermined.

What do you use a neural network for?

Both regression and classification may be accomplished with the help of neural networks. The results of a regression model are a desirable modification of the input patterns, often expressed as a continuous value. The goal of classification is to place input patterns into one of many classes, with the likelihood of belonging to each class indicated by a numeric value between 0 and 1 on the output.

Why are neural networks so powerful?

Both regression and classification may be accomplished with the help of neural networks. The results of a regression model are a desirable modification of the input patterns, often expressed as a continuous value. The goal of classification is to place input patterns into one of many classes, with the likelihood of belonging to each class indicated by a

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

numeric value between 0 and 1 on the output.

What is the Neural Wizard?

The Neural Wizard is a third-party piece of software that facilitates the process of establishing a neural network. Any of the eight most common neural architectures may be built using this tool, along with the requisite file and probe requirements.

What is Neuro Solutions?

Building, training, and testing neural networks can all be done without leaving Excel thanks to a Neuro solution for Excel. The user may quickly and easily designate columns as input or target and rows as training, testing, or cross-validation by using simple checkboxes. In order to train or test a network, Neuro Solutions for Excel communicates with Neuro Solutions.

How does Neuro Solutions implement neural networks?

The local additive model is followed by Neuro Solutions. In this framework, each element is responsible for its own activations and weights, and may learn from the activations of its neighbors. entity oriented modeling works splendidly here since each part can be treated as a distinct entity that communicates with others through messages. This in turn paves the way for a GUI where networks may be built using icons.

What algorithm does Neuro Solutions use to train recurrent networks?

The back-propagation through time (BPTT) technique is used by Neuro Solutions. This algorithm "unfolds" a dynamic net into a feed-forward net at each time step.

How does Neuro Solutions implement Radial Basis Function (RBF) networks?

Unsupervised clustering algorithms locate the centers of the clusters, from which the centers and widths of the Gaussian axons are calculated. Then, using supervised learning with an objective signal, the weights of the Gaussian axons leading to the output layer are calculated.

Can I implement my own algorithms?

Yes, with the Developer's Lite and Developer's levels, you have access to Dynamic Link Libraries (DLLs), which are the simplest approach to customize Neuro Solutions. The "Engine" property page generates default code for every component, which may be modified and built using MS Visual C++.

Platforms and OS Issues

On which OSes are Neuro Solutions available for use?

Neuro Solutions' complete graphical user interface is compatible with Windows NT and Windows95.

Are there versions of Neuro Solutions for the Macintosh, Sun, etc.?

No, and transferring the GUI environment to other systems is not currently planned (see the following question).

Does Neuro Solutions run under UNIX?

You can use Neuro Solutions on UNIX, but there is no graphical user interface. Users who have purchased a source code license from Neuro Solutions have access to all of the company's core algorithmic code, either as a pre-compiled library or as raw source code. The Professional and Developer editions of the library provide a code generating feature that takes the GUI interface and outputs source code. The user may also create their own code to make library calls.

How can I communicate with Neuro Solutions from another program?

Any OLE client, such as Microsoft Excel

IJARST
International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

or Visual Basic, may interact with Neuro Solutions since it is an OLE compliant server. Neuro Solutions' macro language is leveraged by the OLE commands. This is how our "Neural Wizard" and "Neuro Solutions for Excel" products, which are both separate from Neuro Solutions, operate as clients.

SECURITY OF NEURAL NET ANALYSIS

In general, the user authentication process is either very easy or extremely difficult. Authentication methods have evolved throughout time in fascinating ways. 'Others' may easily forge or steal an identity or breach a password with the many ways technology is evolving. As a result, a plethora of algorithms, each with its own unique take on how to calculate a secret key, have emerged. Random numbers in the range of 106 are generated by the algorithms.

Textual passwords, biometric scanning, tokens or cards (like at an ATM), and other key password paradigms are now supplied to users.As was noted above, an encryption technique is often used for text-based passwords. Cards or tokens serve as proof of identity, while biometric scanning serves as a "natural" signature. However, there are many who despise having to remember their cards, and others who object to the high IR exposure required for Biometric scanning (in order to scan their retinas). Most people nowadays use dictionary terms, pet names, girlfriend names, etc. as their textual passwords. Klein ran these experiments back in the day and found that he could break between 10 and 15 passwords in a single day. Due to advancements in technology, including faster computers and more online resources, this is now an easy task.

Our current plan for securing this system relies on 3D passwords, an adaptable and novel approach to authentication. Passwords now take into account the limitations of human memory. Passwords are usually kept simple so that users may easily remember them. In our method, remembering and recognizing things, as well as using biometrics or tokens as authentication, are all necessary. The 3D password user interface appears when it has been installed and a secure website is entered. This is an optional second password that the user may provide in text form. After the first verification, he will enter a 3D virtual room. Let's use the example of a digital garage to illustrate. The typical garage has a wide variety of tools, machines, and other items, each of which has its own special qualities. The user's actions will then be based on these characteristics. The x, y, and z axes may be used to freely relocate any item in the three-dimensional space. That's the thing with everything: everything can move. All of the things have this characteristic. Let's pretend a person has logged in and is now in the carport. He notices a screwdriver and selects it (xyz coordinates: 5, 5, 5); then he shifts it five positions to the right (xyz coordinates: 10, 5, 5). That is clearly a genuine verification. The genuine user is the only one who can identify the item from among countless options and know exactly what it is for. This is the process of remembering and recognizing something. To establish a password, all one has to do is go up to a radio and change the frequency to a secret one. Card readers and biometric scanners are two examples of input methods that might improve security. A user's identity may be verified on several occasions.

CONCLUSION

This research aims to merge the many elements acquired from audio samples for segmenting and classifying into several categories of audio using an artificial neural network. 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. In light of the above, a multi-layer feed forward neural network using a supervised back propagation learning technique has been used to develop an audio classifier. The network learns from audio characteristics that have been retrieved. The software has the capability to label the audio clip as either music, news, or sports. The technique is also used to categorize musical emotions, such as anger, sadness, and happiness. Despite the fact that our training set is relatively small and consists of audio clips with varying semantic structures, the results obtained for the classification of audio categories are quite encouraging, lending credence to the decision to use these features for audio classification. These studies prepared our software for a more normalized data set and longer audio samples. For it to have economic value, it must be able to scale to large datasets and full-length audios. The current setup allows us to classify sound into three categories: sports, news, and music. Additional audio classification categories may be added to the current system in a future deployment. Our audio classification method has an efficiency of almost 90%. However, there are still significant opportunities to enhance the system's performance. This study's long-term goal is to eventually make use of artificial neural networks in tandem with MFCC methods to guarantee flawless audio categorization.

REFERENCES

1. Aleksander, I. and Morton, “An introduction to neural computing”, H. 2nd edition, International Thomson Computer Press, London 1995, PP.284 ISBN No: 85032-167-1.
2. Potamianos and P. Maragos. Speech analysis and synthesis using an AM-FM modulation model. Speech Communication, 28:195–209, 1999.
3. Mehrabian. Pleasure-arousal-dominance: A general framework for describing and measuring individual. Current Psychology, 14(4):261–292, 1996.
4. Samal and P. A. Iyengar, “Automatic recognition and analysis of human faces and facial expressions: A survey," Pattern Recognition, vol. 25, no. 1, pp. 65-77, 1992.
5. Wieczorkowska, P. Synak, R. Lewis, and Z. Ras. Extracting emotions from music data. In Proceedings of the 15th International Symposium on Methodologies for Intelligent Systems, Saratoga Springs, USA, 2005
6. Boashash and A. M. Zoubir, “Digital signal processing”, Brisbane: Queensland University of Technology, Signal Processing Research Centre, 1995.
7. C.C. Liu, Y.H. Yang, P.H. Wu, and H.H. Chen.Detecting and classifying emotion in popular music. In JCIS, 2006.
8. C.H. Chen, M.F. Weng, S.K. Jeng, and Y.Y. Chuang.Emotion-Based Music Visualization Using Photos.LNCS, 4903:358–368, 2008