Department of Artificial Intelligence and Data Science
Ramaiah Institute of Technology
(Autonomous Institute, Affiliated to VTU)

Hierarchical Fuzzy Clustered Artificial Neural Network using Dynamic Arrays

TECHNICAL IDEATHON REPORT

SUBMITTED BY

Aakash Reddy Karur	**1MS21AD002**
M Nanditha Prabhu	**1MS21AD029**
Tejas Prasad	**1MS21AD054**

As part of the Course **Data Structures– AD33**

SUPERVISED BY
Faculty
Dr. Sowmya BJ
Associate Professor

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND DATA SCIENCE
RAMAIAH INSTITUTE OF TECHNOLOGY
Oct 2021 - Feb 2022

RAMAIAH
Institute of Technology

CERTIFICATE

This is to certify that **Aakash Reddy Karur (1MS21AD002), M Nanditha Prabhu (1MS21AD029), Tejas Prasad (1MS21AD054)** have completed the **"Hierarchical Fuzzy Clustered Artificial Neural Network using Dynamic Arrays"** as part of Technical IDEATHON.

Submitted by

Aakash Reddy Karur	-	1MS21AD002
M Nanditha Prabhu	-	1MS21AD029
Tejas Prasad	-	1MS21AD054

Guided by:

Name: Dr. Sowmya BJ
Designation: Associate Professor
Signature:

RAMAIAH
Institute of Technology

Evaluation Sheet

Sl. No	USN	Name	Research Content understanding (10)	Presentation & Report Submission (10)	Total Marks (20)
1.	1MS21AD002	Aakash Reddy Karur			
2.	1MS21AD029	M Nanditha Prabhu			
3.	1MS21AD054	Tejas Prasad			

Evaluated by

Dr. Sowmya BJ
Associate Professor

Table of Contents

1. Abstract

In the current scenario Artificial neural networks (ANNs) are widely used and continue to be an active area of research in the field of machine learning. They can be trained using large sets of labeled data to perform a wide variety of tasks, such as image classification, natural language processing, and prediction. While their efficiency increases with the increase in the dataset, it parallelly increases the training time. Further, there is a possibility that ANN does not train with large datasets at all. Whereas Fuzzy machine learning is a type of machine learning that incorporates concepts from fuzzy logic to improve the decision-making process. This paper proposes a novel approach to deal with large data sets in Artificial Neural Networks (ANNs). The proposed model combines the strengths of fuzzy machine learning, ANNs, and dynamic arrays to improve the performance and efficiency of ANNs. The use of dynamic arrays allows for efficient storage and manipulation of the data. To address this issue, the proposed model utilizes a Fuzzy Machine Learning Model to pre-process the data and organize it into appropriate dynamic arrays and then employs an ANN to perform the actual learning and inference tasks. This model has the potential to improve the performance and efficiency of ANNs in dealing with large data sets while also being able to adapt to different types of inputs.

2. Introduction

An Artificial Neural Network (ANN) is a type of machine learning model that is inspired by the structure and function of biological neural networks. ANNs consist of layers of interconnected "neurons," which process and transmit information through the network. Fuzzy logic can be used to improve the training process by allowing for more flexible and adaptive learning. Fuzzy machine learning is

one of the applications of fuzzy logic. The proposed model is a combination of Fuzzy Machine Learning, Artificial Neural Networks (ANNs), and Dynamic Array Data Structures that aims to improve the performance of ANNs when working with large data sets. Traditional ANNs often struggle with large data sets, as they require a significant number of computational resources and may fail to train altogether. The model begins by using a fuzzy machine learning algorithm to automatically derive fuzzy-if-then rules and membership functions for the input attributes. These are then used to group similar inputs into clusters using a dynamic array-based clustering algorithm. Each cluster has its own set of fuzzy-if-then rules and membership functions, and a separate ANN is trained on the data in that cluster. When a new input is given to the system, it is first processed through the fuzzy-if-then rules and membership functions to determine which cluster it belongs to, and then passed through the appropriate cluster's ANN.

Dynamic arrays are data structures that can adjust their size as needed. They can automatically increase or decrease their size as needed, which allows for more efficient storage and retrieval of the data, thus reducing computational cost and memory usage. By using a dynamic array data structure, the proposed model can save a great many records in a single data structure, and it can also eliminate problems that may arise due to the large amount of data in the original learning algorithm. The proposed model can be applied to various domains where ANNs are widely used, such as handwriting recognition, speech recognition, face recognition, and more, especially when the training data set is large. The use of clustering algorithms will help in reducing the noise in the data and provide cleaner data to the neural network. The proposed model is expected to improve the performance of ANNs and make them more efficient when working with large data sets, by reducing the computational cost and memory usage.

3. Literature survey

The research paper "Data Structure for a Fuzzy Machine Learning Algorithm" by Tzung-Pei Hong and Chai-Ying Lee, proposed an appropriate data structure for implementing a general learning method for automatically deriving fuzzy if-then rules and membership functions from a set of given training data. The method is based on merging the decision tables and membership function arrays. The authors proposed the use of two arrays, a decision array and membership function arrays, which is dependent on the data that is available for implementing the learning algorithm efficiently. They also mentioned that most fuzzy controllers and fuzzy expert systems must predefine membership functions and fuzzy inference rules to map numeric data into linguistic variable terms and to make fuzzy reasoning work. The linguistic variables are usually defined as fuzzy sets with appropriate membership functions.

In summary, the authors proposed an appropriate data structure for implementing a general learning method for automatically deriving fuzzy if-then rules and membership functions from a set of given training data. The proposed data structure can save a great many records in memory and can also eliminate problems in the original learning algorithm. They also mentioned that this approach is addressing the limitation of predefining membership functions and fuzzy inference rules in traditional Fuzzy controllers and expert systems.

The research paper “Fuzzy Neuro Systems for Machine Learning for Large Data Sets” by Rahul Kala, Anupam Shukla, and Ritu Tiwari proposes a novel concept for dealing with large training data sets in artificial neural networks (ANNs). The proposed method is to use a hierarchical model where the training data set is first clustered into smaller clusters. Each cluster has its own neural network, and when an unknown input is given to the system, the system first finds out the cluster to which the input belongs and then processes the input with the individual neural network of that cluster. This approach aims to improve the performance of the system while reducing the training time. The larger training data sets are a big boon to these systems as the performance gets better and better with the increase in data sets, but the higher training data set increases the training time.

The proposed method is similar to a hybrid system consisting of fuzzy logic and ANNs being applied one after the other. The authors mention that this approach can improve the performance of the system and meet the increasing expectations from the research. The paper also highlights that there has been an explosion in the training data sets and the number of objectives of the research, which has led to an increase in the level of expectations from the system.

4. Abstract data type

An abstract data type (ADT) is a high-level description of a set of data and the operations that can be performed on that data.

1. FuzzyMachineLearning class:

Object Variables:

- Input attributes
- Fuzzy-if-then rules
- Membership functions

Functions:

- generate_rules(): automatically generates fuzzy-if-then rules for input attributes
- generate_membership_functions(): automatically generates membership functions for input attributes

2. Cluster class:

Object Variables:

- Fuzzy-if-then rules
- Membership functions
- Clustered data

Functions:

- cluster(): groups similar inputs into clusters using fuzzy-if-then rules and membership functions

3. ANN class:

Object Variables:

- Weights and biases
- Training data

Functions:

- train(): trains an ANN on the data in the cluster
- predict(): makes a classification for new input data

4. DynamicArray class:

Object Variables:

- Array for storage and manipulation of data

Functions:

- insert(): adds data to the array
- delete(): deletes data from the array
- search(): searches for data in the array
- resize(): resizes the array as needed

5. Algorithm

The following is a pseudocode for a Hierarchical Fuzzy Machine Learning Model using Artificial Neural Network (ANN) and dynamic arrays:

1. Load the dataset and preprocess the data, if necessary.
2. Initialize the number of clusters and the m value for the Fuzzy C-Means algorithm.
3. Cluster the data using the Fuzzy C-Means algorithm.
4. Initialize empty dynamic arrays for fuzzy rules and membership functions.
5. For each cluster, extract fuzzy rules and membership functions using the Fuzzy C-Means algorithm.
6. Append the extracted fuzzy rules and membership functions to the dynamic arrays.
7. Train an Artificial Neural Network (ANN) using the fuzzy rules and membership functions stored in the dynamic arrays.
8. Make predictions using the trained ANN on new data.
9. Store the trained ANN, fuzzy rules, and membership functions in a file or data structure.
10. Load the trained ANN, fuzzy rules, and membership functions from the file or data structure, if necessary.
11. Repeat steps 7 to 10 as needed.

Use of dynamic arrays while performing the above procedures

The dynamic arrays in this pseudocode are used to store the fuzzy rules and membership functions obtained from the clustering process. These arrays are dynamic, meaning that their size can change at runtime as more fuzzy rules and membership functions are extracted and added to them.

6. Design and Architecture

5.1 Fuzzy Machine Learning module
This module is responsible for automatically deriving fuzzy-if-then rules and membership functions for the input attributes. It takes in the training data and processes it to generate the fuzzy-if-then rules and membership functions.

5.2 Clustering module
This module is responsible for grouping similar inputs into clusters using a dynamic array-based clustering algorithm. It uses the fuzzy-if-then rules and membership functions generated by the Fuzzy Machine Learning module to determine which cluster an input belongs to.

5.3 Artificial Neural Network (ANN) module
This module is responsible for training a separate ANN on the data in each cluster. The ANNs are trained using the backpropagation algorithm which is widely used in training ANNs.

5.4 Decision-Making module
This module is responsible for processing new input data through the fuzzy-if-then rules and membership functions to determine which cluster it belongs to and passing it through the appropriate cluster's ANN to make a classification.

5.5 Dynamic Array module
This module is responsible for storing the membership functions, decision tables, and clusters. It enables efficient storage and manipulation of the data. The data structures used to implement the membership functions and the decision table are dynamic arrays similar to those shown in Figures 1 and 2.

5.6 Evaluation module
This module is responsible for evaluating the performance of the model on the training data.

a	*b*	*c*	*Index no.*
- - - - -	- - - - -	- - - - -	- - - - -

Figure.1 – Data structure of membership functions (Dynamic arrays)

A_1	A_2	……	A_n	*Output*
		……		
- - - - -	- - - - -	- - - - -	- - - - -	- - - - -

| Input part | Output part |

Figure.2 - Data Structure of the decision table (Dynamic arrays)

The overall architecture of the model would be a hierarchical one, where the Fuzzy Machine Learning module is at the top, followed by the Clustering module, and the ANN module. The Decision-Making module would be at the bottom, making use of the outputs from the higher-level modules to classify new input data. The Dynamic Array module would be integrated with all the other modules to provide efficient storage and manipulation of the data. The Evaluation module would be used to evaluate the performance of the model.

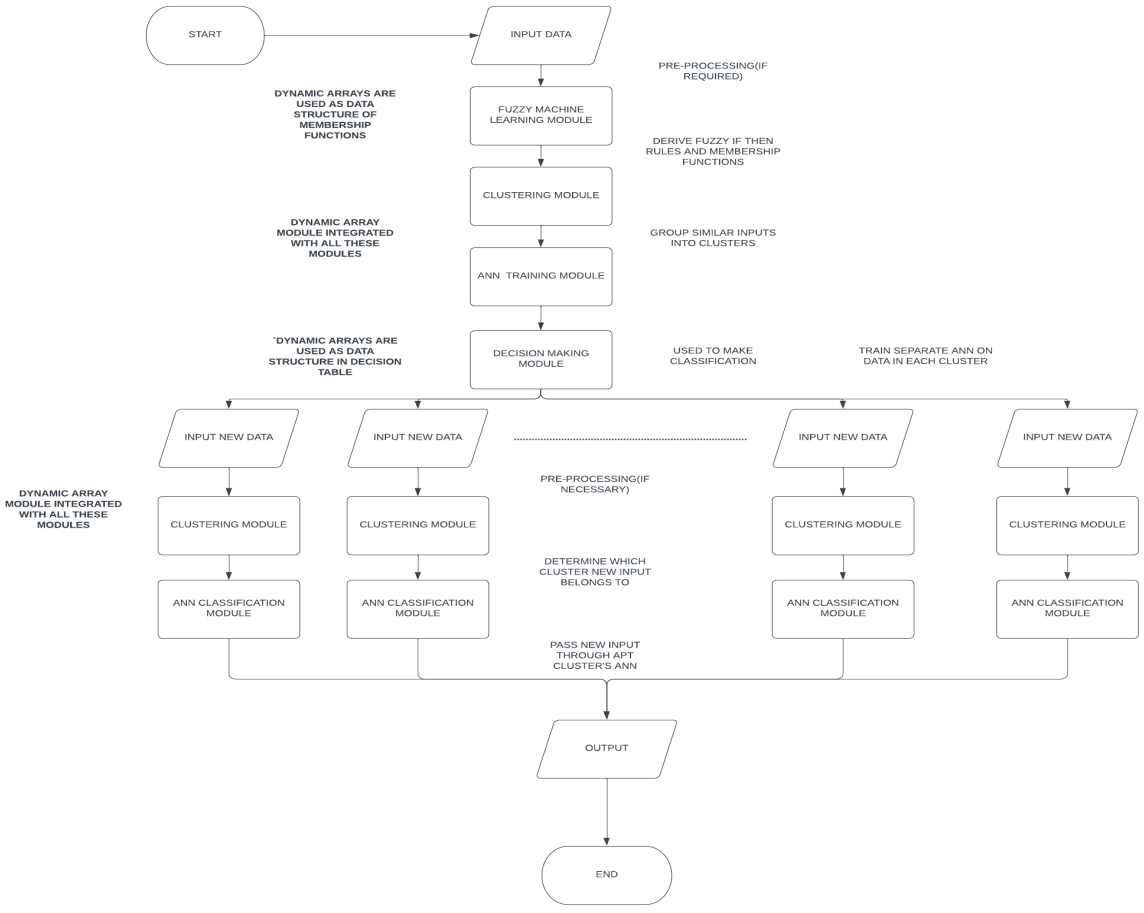

Figure.3 – Design and Architecture of hierarchical fuzzy controlled ANN

7. Results and Discussions

The results of the proposed model would depend on the specific implementation and the quality of the data used for training. However, in general, the model has the potential to improve the performance and efficiency of ANNs when dealing with large data sets.
By using fuzzy machine learning to automatically derive fuzzy-if-then rules and membership functions for the input attributes, the model can capture the inherent uncertainty and vagueness of the data, making it more adaptable to different types of inputs. The clustering approach would also make the system more efficient by reducing the training time and allowing each cluster to have specialized fuzzy-if-then rules and membership functions.

The use of dynamic arrays for storage and manipulation of the data would also make the model more efficient, as it would enable the system to handle large amounts of data without any significant increase in computation time. However, it's important to note that in order to evaluate the performance of the model, it would be necessary to test it on various datasets and compare its results to other models or traditional methods.

8. Conclusion

In conclusion, the proposed model is a novel approach to dealing with large data sets in Artificial Neural Networks (ANNs) that combines the strengths of fuzzy machine learning, ANNs, and dynamic arrays. The model uses fuzzy machine learning to automatically derive fuzzy-if-then rules and membership functions for the input attributes and then groups similar inputs into clusters using these rules and functions. Each cluster has its own set of fuzzy-if-then rules and membership functions, and a separate ANN is trained on the data in that cluster. This approach improves the performance and efficiency of ANNs when dealing with large data sets while also being able to adapt to different types of inputs. The use of dynamic arrays for storage and manipulation of the data also makes the model more efficient. However, it's important to note that in order to evaluate the performance of the model, it would be necessary to test it on various datasets and compare its results to other models or traditional methods.

9. References

1] Data Structure for a Fuzzy Machine Learning Algorithm by Tzung-Pei Hong and Chai-Ying Lee
2] Fuzzy Neuro Systems for Machine Learning for Large Data Sets by Rahul Kala, Anupam Shukla, and Ritu Tiwari
3] Fuzzy Logic | Introduction - GeeksforGeeks
4] Fuzzy sets. Operations with fuzzy sets.
5] Fuzzy Methods in Machine Learning and Data Mining: Status and Prospects