International Journal of Scientific & Engineering Research, Volume 11, Issue 8, August-2020
ISSN 2229-5518

Adaptive Neuro-Fuzzy Inference System with Non-Linear Regression Model for Online Learning Framework

Abdul Joseph Fofanah, Ibrahim Kalokoh, Korjo Tesyon Hwase and Alex Peter Namagonya

Abstract— In this paper, we present a combined hybrid of an adaptive neuro-fuzzy inference system with an adaptive online learning model (ANFIS-AOLM) using a non-linear regression machine learning technique to construct the Neural Network. We proposed ANFIS-AOLM for learning by modeling and controlling imprecise defined, uncertainty system with a significant role of the neuro-fuzzy method. The various simulation exercise was conducted with define input and output values, define rules and the dataset was created and trained the Neural Network. Exploration of the dataset was tested for ANFIS-I using "Grid Partitioning"; ANFIS-II using "Subtractive Clustering"; and ANFIS-III using "Fuzzy Clustering Means". The study was conducted in oder to explore the effectiveness of the combined hybrid learning opportunities using ANFIS and Non-linear Regression models. The study was experimented using MATLAB programming language for the adaptation of the online learning framework with learner's capability to learn from the existing information and adapt accordingly in the learning environment. The results obtained from the analysis shows that with a membership function of 20 and 25 an average of 97.7% accuracy was achieved. However, as the number of membership functions increases the better the performance of the ANFIS and the higher the computational response time. The non-linear regression model also indicated that the relationship of the ANFIS system and the proposed technique is adequate for learner's adaptation and the decision made through learning mechanism.

Index Terms— Online learning, ANFIS, Neural Network, non-linear regression, Adapative learning

1 Introduction

In recent years, artificial intelligence (AI) techniques are contributing significantly in various domain. There are various kinds of intelligent techniques AI has contributed thus includes: genetic fuzzy systems, genetic programming, neural networks, neuro-fuzzy system, particle swamp optimization and many more to name few. Jang [1] contributed to the most significant artificial intelligent (AI) technique titled neuro-fuzzy systems. The neuro-fuzzy system comprises of a combination of fuzzy logic and artificial neural network (ANN). The neuro-fuzzy system follows the standard principles of fuzzy logic based on learning abilities with neural networks. Principally, the IF-THEN statements rules are utilized in the neuro-fuzzy systems to study the perceptibility and imperceptibility values. These fuzzy rules are created based on the training dataset. In order to determine the relevance of the research, this paper reviewed various neuro-fuzzy systems applications in different environments. The development of the internet of things (IoT) has brought about another new era of big data based on deep learning techniques which are widely utilized to resolve different challenges in big data [2], [3], [4], [5], [6].

To improve efficiency and to achieve a learning task, an adaptive online learning system aims to adapt learning content to each learner's framework such as features, interest or needs. Online learning techniques have become a significant method that is contributing greatly to many educational platforms. Computers, networks, personal digital assistants (PDAs), mobile phones and some digital media techniques are being used by several learners in different platforms to improve cooperation and by cross-fertilizing ideas beyond borders.The mindfulness of framework in this regard is highly significant because it enables the cloud environment to be used in a method that upkeeps the learner at any specified location. The learning framework awareness concept is significantly vital that needs to be considered in the strategy of adaptive learning systems for online learning systems to be truly effective [7].

System developments are based on secure, storing and exploration of information about each learner. Essentially, learner preferences can be determined or predicting learner behavior, hence making up a significant concept to make suitable decisions for each learner. The system needs to make inferences about their learner's preferences and expected opportunities based on personal service requirements; thus, this can be accomplished by making postulations about learners based on the interaction interface between the system and learners. Here, machine learning algorithms would be applied to handle those functionalities requirement stetting mechanism.

An adaptive online learning system is a technique that implements adaptation based on a learner model and updates the learner model with newly derived facts. Using imprecise information, an online learning application is basically a system that is interactive. Essentially its interpretation is based on uncertainty and im-pressiveness. Many systems for user modeling employ uncertainty techniques [8]. However, machine learning techniques are generally used for learning modeling due to the complexity of learners' context relationships between them, hence difficult to represent. Webb et al. [9] highlighted four key challenges related to the utilization of machine learning methods for developing modeling criteria: large datasets requirement, labeled data requirement, concept drift, and computational complexity.

- *A. J. Fofanah is with the Department of Mathematics & Comouter science, Faculty of Environmental Science, Milton Margai College of Education and Technology, Sierra Leone. He has Associate Degree in Mathematics, B. Sc (Hons.) in Computer Science, master's Degree in computer science and Master's in Software Engineering. E-mail: abduljoseph.fofanah@gmail.com*
- *I. Kalokoh is with the Department of Computer Science, Faculty of Environmental Science, University of Makeni, Sierra Leone. He has B. Sc Degree in Information Technology and Master's Degree in Software Enigineering, E-mail: ikalokoh@gmail.com*

International Journal of Scientific & Engineering Research, Volume 11, Issue 8, August-2020
ISSN 2229-5518

Largely personalization techniques rely on machine learning techniques that require large labeled data in order to provide better results and avoid over-fitting the dataset. The drift is the continuous change in learner's interests and profiles.

1.1 Online Learning in Big-Data Analytics

However, in both private organizations, the public sector, and research institutions Big Data has drawn great attention [10], [11], [12], [13]. Currently, and it is observed that Big data technologies have contributed largely to the greenhouse effects and largely in all fields [14], [15]. It is clearly noticed that big data are likely to grow in the future and will essentially shape our daily environment especially in real-time applications and it is predicated that by 2023 big data will exponentially expand to 175 Zettabytes [16], [17], [18]. The combination of all these data successfully becomes Big data and the provisioning of well-being, effective delivery, and interference [19], [20], [21], [22], [15]. In research papers or books, Big data also provides and save large data. It again uses to save nature and biological applications as profuse data being collected connecting to our ecological, footprint and perceptible impact [23], [24], [15], [25], [26]. Hence, Big data analytics has also impacted significantly on global warming and climate change [20]. In the healthcare domain, Big data has contributed to data related to electronic patient records (EPR), medical imaging data, and genetic data. Researchers using data mining techniques to develop systems for natural language processing (NLP), content-based image retrieval system, and penalized logistic regression are major sources of Big data development [27], [28]. In social networking such as text data, graph data with the adoption of data mining techniques using sentimental analysis, community detection, social influence analysis, and collaborative filtering systems all because of Big data development environments [29], [30], [31], [32], [33]. Furthermore, in CCTV surveillance basically video type of data and utilization of data mining techniques, the labor-based surveillance systems were developed [34]. Sensor data and machine-generated data of both unstructured data and the log file type of data respectively. Using data mining techniques, the contextual anomaly detection and frequent pattern mining tools were developed to enhance and improve fault tolerance mechanisms of big data technologies [34].

1.2 Benefits of Online Learning

The process of enhancing online learning is also referred to as "Asynchronous Learning" which is a student-centered teaching method that practices online learning materials to accelerate communication and cross-fertilize ideas outside the constrictions of time, and a place among a network of people. Asynchronous learning is based on constructivist theory, a student-centered technique that emphasizes the significance of peer-to-peer interactions. This technique is a hybrid of self-study with asynchronous interactions to support learning, and it can be used to facilitate learning in distance education. This hybrid network of learners and the electro-network in which they connect via communication are referred to as an asynchronous "Learning Network". The online learning materials used to provision asynchronous learning include email and online discussion boards. Course management systems such as moodle has been designed to upkeep and protect online interaction. By permitting users to organize discussions, post and reply to messages, and upload and access multimedia to facilitate an effective learning environment.

Asynchronous learning's greatest benefit to students is the freedom it gives them to access the course and its teaching materials within a suitable time by the learner, and from any location, with an Internet connection. This allows accessibility for dif-ferent groups of students, reaching from traditional-setting, on-campus students, to working professionals, to international students in foreign countries. Asynchronous learning situations enhance and facilitate a "high degree of interactivity" between learners who are disconnected both geographically and temporally and afford students many of the social benefits of face-to-face interaction. Since students can express their opinions without interruption, they have more suitable time to reflect on and respond to class materials and their classmates than in a traditional class-room. Additionally, asynchronous learning provides a record of almost everything that occurs in that environment. All educational resources, email history or corre-spondences, and interactions can be electronically archived. Learners can choose an appropriate schedule to review their course materials, lectures, and presentations, as well as correspondence between learners.

This paper presents an online learning reasoning mechanism. The main objective of this reasoning mechanism is to offer an appropriate learning framework format for online learning application, using an Adaptive Neuro-Fuzzy Inference System-Adapted Online Learning Model (ANFIS-AOLM) based on constructed learning theories and acquired a learner profile. Basically, the learner profiles comprise of learner's knowledge, preferences, goals, plan, loca-tion, and probably other significant aspects that are utilized to offer personalized learning frameworks.

2 Related Work

In this section, we explored the literature by reviewing research work on different ANFIS system using modeling techniques in regression, classifications. We thought this could help for thoroughly searching issues and fitting these challenges to intelligently solve problems related to the ANFIS system. Essentially, this provides a benchmark for carrying out this paperwork and making robust predictions that could current challenges in Neural Networks and help future work. This section explains the following: ANFIS with the regression approach, ANFIS with classification approach and ANFIS with multiple kernel learning approaches.

2.1 ANFIS with Regression Approach

Measurement of anthropometric using regression analysis have been utilized for mul-tiple non-linear regression models relative to modeling the practical relationships between different kinds of measurement such as anthropometric and many more utilizes the traditionally non-linear regression model. Consequently, and in recent years, based on artificial intelligent techniques asserting to be universal approxima-tor, were developed as a proposed replacement to statistical techniques, particularly in non-linear models with functional relationships. As a result of both software and hardware computational intelligence, soft-computing in relation to ANFIS ap-proaches have recently drawn the attention of

International Journal of Scientific & Engineering Research, Volume 11, Issue 8, August-2020
ISSN 2229-5518

researchers across the world. This is in parallel with an extraordinary ability of the human mind to reason and learn in an environment of probabilistic uncertainty and imprecision (Zadeh [35]). The main purpose of soft computing is to integrate such influential artificial intelligence methodologies as both neural network and fuzzy inference systems. Therefore, AN-FIS is a new inference system put forward by Jang et al.[19] and Fuller [36]. Their proposed techniques asserted that it is very effective in representing highly non-linear functionalities.

Kaya et al.[37] proposed an intelligent model in which they tried to compare an estimation of anthropometric measurements in adolescents by adopting an hybrid techniques of artificial intelligent method known as "Adaptive neuro-fuzzy inference system" and stepwise regression analysis. Their proposed method demonstrated that ANFIS is an alternative method to regression analysis for the estimation of anthropometric measurements. The best of independent variables for each output have been selected by stepwise regression analysis however, their technique sets are not optimal ones for ANFIS method.

In statistics there are different methods used for analysis, multiple linear re-gression analysis is frequently used to summarize data and also to understanding the relationship between variables (Nourusis [38]). Typically, stepwise regression is an amalgamation of backward and forward techniques or procedures and have ac-counted for the most frequently used method in regression analysis (Newbold [39]; Draper and Smith [40]). The method describes that the first variable is chosen in the same way as in the forward selection. The procedure terminates without independent variables entering into the equation, provided if the variable fails the entry requirements. Upon progressing the criterion, the second variable is selected based on the highest partial correlation. It also enters the equation provided if it passes the entry criterion. The concept and process after the first variable are entered, stepwise selection differs from the forward selection as explain and this includes: as in the backward elimination, the first variable is examined to provide a clue whether it should be eliminated according to the removal criterion. Next step the variable that is not in the equation is determined for removal. Variables can only be removed until none of the remaining variables meet the removal criterion. Finally, the vari-able selection terminates when there are no more variables meeting the entry and removal criteria accordingly.

2.2 ANFIS with Classification Approach

In the early 1990s, there have been many contributions to the study of artificial neural networks (ANNs) and have been proved effective. However, due to inade-quate or inefficient computer network systems, much research was lacking by then. The development of new technologies in computer network systems and the study of ANN has drawn so much attention to many researchers. Artificial neural networks have been used as computational tools to describe pattern classification, particu-larly in the healthcare domain with respect to diseases diagnosis. Additionally, due to its ability of greater predictive power than any signal processing analysis meth-ods ([41],[42],[43],[44]). Again, the significant function of fuzzy-logic and theory with the capabilities to handle uncertainty when making decisions in medical applications. Consequently, fuzzy sets have drawn interest in different domains such as data analysis, decision making, pattern recognition, diagnostics, production tech-niques, modern information technology, cloud computing technologies and many more to name but few [45], [46], [47]. Neuro-fuzzy systems are fuzzy systems, which use ANNs theory in order to determine their properties such as fuzzy sets and fuzzy rules by processing data samples. The significant applications of Neuro-fuzzy sys-tems control the power of the two archetypes namely: fuzzy logic and ANNs by utilization of mathematical characteristics of ANNs in adjusting rule-based fuzzy systems that approximate the way humans process information. The neuro-fuzzy specific technique in neuro-fuzzy development is the adaptive neuro-fuzzy inference system (ANFIS), that has proved to be very powerful and momentous results in modeling non-linear functionalities. The robustness of ANFIS indicates a member-ship function parameter that are removed from the dataset and defines the system behavior. Its learning properties in the dataset provide an ANFIS system to adjusts the system parameters according to a given error criterion [26, 32, 18, 4]. Usher et al. [52],Guler et al.[50] successfully implemented an ANFIS in biomedical engineering was reported for classification problem and data analysis [53].

Guler et al. [50] conducted a study on a new approach of using ANFIS and was presented for classification of the electroencephalogram (EEG) signals. The tech-nique comprises of five training of ANFIS classifiers to classify the five classes of EEG signals when wavelet coefficients describing the behavior of the EEG signals were designed as inputs. By establishing the technique and a hybrid of the least square techniques the ANFIS classifiers and, were trained with the backpropaga-tion gradient descent [54]. In their submission, a thorough classification between set A of healthy volunteer, eyes open) set B of healthy volunteer, eyes closed, set C of seizure-free intervals of five patients from hippocampal formation of contrasting hemisphere, set D of seizure-free intervals of five patients from epileptogenic zone, and set E of epileptic seizure segments was performed. The predictions of the five ANFIS classifiers were combined against the sixth ANFIS classifier. The accurate classification rates and convergence rates of the ANFIS model were examined and then performance of the ANFIS model was reported. Guler's approach also ex-plained their results: firstly, ANFIS model combined the neural network adaptive capabilities and the fuzzy logic qualitative approach. Secondly some conclusions concerning the saliency of features on classification of the EEG signals were ob-tained through analysis of the ANFIS. Finally, classification results and statistical measures were used for evaluating the ANFIS model. The total classification accu-racy of the ANFIS model was 98.68% which led them to the conclusion that, the ANFIS model can be used in the classification of EEG signals by just considering the misclassification parameter rates.

According to Webb et al.[19], the purposes for which learner models are developed are as follows: underpinning a user's actions can be classified by using cognitive pro-cesses, an exploration of the differences between an expert and individual skills, and identifying the features and behavioral patterns deterministic. In order to provide personal service requirements, learner adaptive systems should be able to deduce learner's behavior. Machine learning models' capabilities can enhance and gener-ate a learner's model and consequently, include all the information

International Journal of Scientific & Engineering Research, Volume 11, Issue 8, August-2020
ISSN 2229-5518

that the system knows about the learner. Generally, the idea starts either by querying the system or by default values. The system reasoning capabilities are like an engine that combines the learner's basic information with other models within the ANFIS system to develop new learning techniques "evidence-based" about an individual learner.

3 Background For Proposed ANFIS-AOLM Model

In this section, we explained the background of the proposed "Adapted Neuro-Fuzzy Inference System-Adapted Online Learning Model (ANFIS-AOLM) frame-work. This is to provide a clear analysis of the significance and impact of an ANFIS system in ANNs. The following are discussed for the proposed ANFIS-AOLM sys-tem: machine learning-based on user profile context, improved learner model anal-ysis, adaptive online learning process, and reasoning methodologies for a learner modeling system.

3.1 Machine Learning Based on User-Profile Context

Guller et al.[19] presented machine learning based on a user profile framework that depicted the process of adapting learning to achieve individual learner features by focusing on the learner's learning technique. **Fig. 1** shows seven adaptation layers with corresponding descriptions that need to be considered as our case of online learning adaptation: *Context Acquisition Layer* is used to collect the information needed for adaptation. It mainly depends on both explicit and implicit information gathering. Explicit information depends on the information provided by the learner and implicit information is collected by supervising the learner's interactions with a system and making assumptions as to the individual. *Information Classification Layer* deals with all data obtained from the previous layer by categorizing the data into several class types. *Learner Model Layer*, Kobsa [63] defines a learner model as a set of information structures designed to represent one or more of the following parameters: goals, plans, and preferences; feature of learners; learner stereotypes and learner behavior. *Information Extraction Layer* evaluates, studies, validates and filters the data based on the current learner's situation. *Learner Profile Repre-sentation* can achieve user requirement services, we should be able to specify learner interests. Constructing a learner profile allows for much more precise outcomes, given an adequately expressive keyword. Accurate information allows the system to more accurately support learner decision making. *Reasoning Layer*, the output of this layer is a set of structural descriptions of what has been learned about learner behavior and interests. *Interface Layer* is formed by the events that are pro-cessed by the adaptation system as well as the questions about the learner that it can answer. This paper proposes a solution to adapt learning framework through the use of online learning technology, based on modelling the learner and all possible frameworks related to individual current situation.

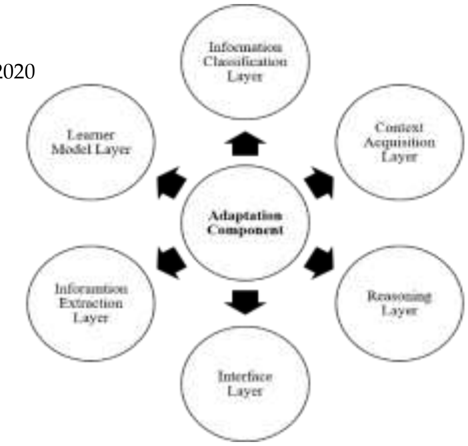


Fig. 1. Adaptation Components

3.2 Improved Learner Model Analysis

Al- Hmouz et al.[22] presented another technique to combine context together to accomplish maximum integration and learner information, hence resulting in insuf-ficient information learner profiles. As indicated in **Fig. 2**, namely the improved learner model comprises of four descriptive components such as a representation of the learner status, the knowledge and shared features status, the situation status, and educational activity status. This component is explained below systematically as depicted in the **Fig. 2** diagram.

Learner Status: This is the first and main component of the improved learner model. It encapsulates the learner profile, learner history, general situation, learner knowledge, and educational activity. Essentially, it indicates and describes the learner adoption of assumptions about knowledge and preferences, the interaction history, and a description of his general condition. The preferences represent the learner's interest in certain subjects, and thus category captures all the common information that can comprise a learner profile.

Situation Status: This is the second and main component of the improved learner model. The learner's present state and general situation are part of the situation sta-tus. This component depicts the learner's current situation with its various features in the real world. The overall situation describes the general status of the system and consists of the following: devices, networks, hardware, and software resources and perhaps other online parameters required to enhance learning.

Knowledge and Shared Properties Status: This is the third main component of the improved learner model. The learner experience is based on assumptions about the learner's expe-rience of the system, associations between inputs, and evidences and procedures concerning the online learning application. The shared features depict the available tools that help learners to achieve an activity if that activity includes coordination.

International Journal of Scientific & Engineering Research, Volume 11, Issue 8, August-2020
ISSN 2229-5518

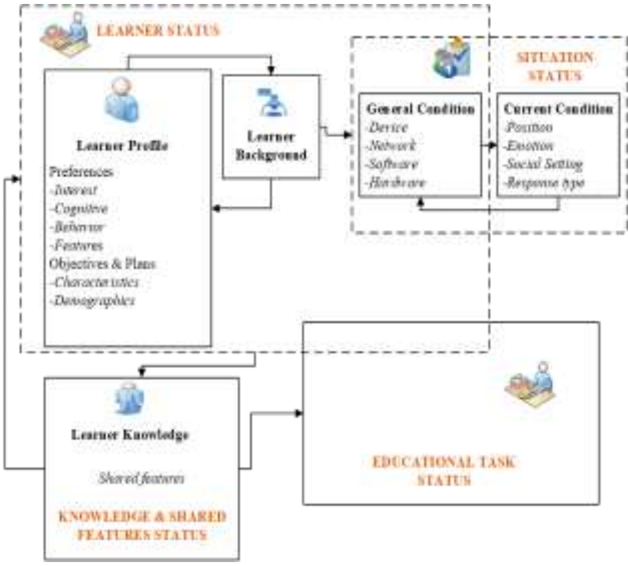


Fig. 2. Learner Model Components

Educational Activity Status: This is the fourth main component of the improved learner model, which comprises of "any service or facility that supplies a learner with general electronic information and educational content that aids in acquisition of knowledge regardless of location and time" [22]. Therefore, the research focused on modeling the learner and all possible contexts in an extensible way that can be used for individualization method in online learning model.

3.3 Adaptive Online Learning Process

Apparently, the reasoning mechanism comprises two main stages namely: fuzzy logic and neural network. The inputs to the adaptation mechanism involve the following: the status of the shared feature, the educational task status in the form of learner profile representation, the situation status and the knowledge-based status (**see Fig. 1**). In order to explain the adaptive online learning application, this section summarized the following stages suitable for the online learning process (**see Fig. 3**).

The process starts when the learner brings out the learning task by interacting with the application by choosing some suitable actions presented on the learner's device interface. Probably by using the application for the first time, the learner will be requested to fill and complete some forms related to personal information and other information through device sensors or some related online technology.

Secondly, if the learner has registered or has used the application before, the learner would be prompted or willing to partake and be willing to initiate the task in hand. The system will deliver a start-up application to show that the learner needs to participate in a learning task. In the process, the learner data or information are collected to create the learner model paradigm.

Thirdly, the reasoning mechanism transforms the framework from one state to another in order to achieve the challenges of the learner framework. Through ma-chine learning algorithms, the system observes learner behavior and preferences and recommends educational resources that are alike to choose in the past evaluation. The reasoning mechanism selects the most adequate media type for the learning context based on the learner challenges and convey the adapted framework content to the learner's interface.

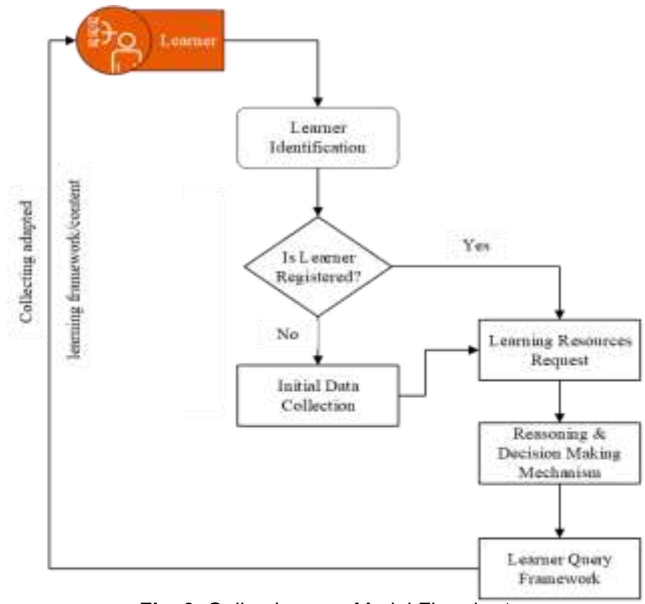


Fig. 3. Online Learner Model Flowchart

3.4 Reasoning Methodologies for Learner Modeling System

Reasoning approaches try to make assumptions about individual learners based on each user's interaction with the system. There have been some methods used to model the learner contexts particularly K-Nearest Neighbors (K-NN), Bayesian Net-work, Genetic Programming, Fuzzy-Logic, and Neural Networks. Hybrid systems, consisting of various machine learning techniques, have also been used. Numerous projects attempting to implement machine learning methods for learner modeling have been modularized in some research works to suit the learner's context in recent years. In the search for adequate answers to some of these research projects have had a variability of objectives and levels of adaptation for individual learners. An exploration of some of these concepts are discussed subsequently.

Reasoning techniques try to develop assumptions regarding user learners based on each individual's interaction with the system. These approaches have contributed greatly in data science and online learning systems and are the basis for carrying out the research. Some of these algorithms includes: Neural Network, K-Nearest Neighbor (K-NN), Bayesian Network, Genetic and Fuzzy algorithms are basically some of the methods utilized to model the learner context framework. More so, hybrid systems comprise of combinations of diverse machines learning schemes that have been highly used. Over the years, there has been different projects attempting to implement machine learning techniques for learner modelling. Some of these projects have had different aims and levels of adaptation for user learners. In order to make this research worthwhile some of these researches carried out are discuss shortly.

Tsiriga and Virvous proposed the K-Nearest Neighbor (K-NN) model for the actualization of the student model in web-based educational applications using a web-based passive tutor mechanism. The main purpose of this system was initial-izing and up-

International Journal of Scientific & Engineering Research, Volume 11, Issue 8, August-2020
ISSN 2229-5518

dating the learner model with the combination of stereotypes. The proposed system model was implemented with 117 learners belonging to diverse stereotypes.

Garcia et al.[17] proposed a Bayesian Network model for detecting learning styles as defined by Felder and Silverman [12]. The proposed system was implemented with 27 computer science engineering learners taking an Artificial Intelligence course. The results were compared by Garcia et al. through a learning style questionnaire completed by the same 27 learners. The results show that the Bayesian network was effective for predicting the learning style of the learners with high accuracy.

A Bayesian network model to support educational instructors in making decisions under uncertain situations was proposed by Varatharajan [51]. The proposed algorithm was implemented with 800 learners acquiring knowledge in an informatics course. The main objective of this algorithm was to archetypal the learner's behavior in order to make predictions about the success of instructor decisions. Xenos found that the proposed algorithm could be a valuable tool in decision-making under uncertain circumstances.

Frias et al. proposed a Neuro-Fuzzy learner model system to diagnose the errors of high school learners by collecting data in courses related to vectors in physics and mathematics with simulation tools. Through some simulation exercises, the system was evaluated by adopting learner data with different knowledge-based and level categories with respect to behavior corresponding to fuzzy values. Furthermore, the error classification of a feed-forward neural network was used to train the data. Fundamentally, most of the projects carried out for the implementation of learning techniques have been tested and adaptive for online learning systems.

4 THEORETICAL CONCEPTS OF ANFIS-AOLM SYSTEM

One of the purposes of carrying out this research is to ensure that we explore some of the theoretical concepts and its contributions to online learning theories using the ANFIS system. In this section we looked at some of these concepts in learning theories. The following represents some theories that could help explain the importance of using suitable online learning applications to learning theories: drawback setting and general framework, application of online learning (first order, second order and online learning with expert advice).

4.1 Drawback Setting and General Framework

Supervised machine learning task such as classification and regression is considered as generic. Making training samples such as $D = \{(x_1, y_2), \dots\dots (x_r, y_1)\}$ *where* $x_p \in \mathbb{R}^d$ is a D-dimensional instance representing the features. And for a mathematical relation of $y^p \in |y| = \{-1, +1\}$ represent a binary classification, $y^p \in \{0,1\}$ represent a multi-class classification with C classes and $y^p \in \mathbb{R}$ for regression tasks, the target label assigned to x_p. The objective of the online learning is to learn a function $F: R^d \to R^{|y|}$. The prediction is represented by a function $\bar{y}_p$. The prediction model is determined by the performance of the learning function which evaluates based on an accumulated drawback offered. The effect can be the error rate in a squared loss for regression model. In order to learn from a prediction function that mitigate this effect over T instances, a loss function basically called Ling-Loss, Squared Loss, Cross-Entropy is chosen for mitigation. We represent this loss function as $\mathcal{L}(F(x), y)$. The instances $X_p, p = 1 \dots\dots P$ arrive chronologically and a prediction $\overline{y_p}$ is made on each instance. Consequently, the setting resulting in the true target of y_p by using this reference, the parameters of the nonlinear function $F(x)$ are updated.

4.2 Application of Online Learning

The application of online learning can be found in many domains. For instance, in Cybersecurity, Finance, and Recommendation Systems specifically problem related to environment of online learning. In finance, a popular application is Online Portfolio Selection. In this model, the function F to be learned is the portfolio vector, such as identification of how to distribute wealth in a set of assets with specific objectives. However, an online learning setting in every time iteration, the environment shows a corresponding instance of vector X. Hence the current portfolio vectors the algorithm realizes some profit or loss. By utilization of this loss information, the algorithm updating the portfolio vector X through some technique, thus modifying the prediction function F. Similarly, in time-series modeling settings are observed on high frequency trading, where decisions are to made automatically depending on the streaming setting factors. Conversely in anomaly detection [9], because the tasks are being monitored continuously. In every time sequence, the environment shows some task is not utilizing the prediction function $F(x)$. Consequently, the model may detect some response from the user or new statistics about the data distribution as a result to which the model would update its prediction function. Though the algorithm of the learning algorithm shows a specific example of supervised learning, certainly online learning may also be unsupervised. Furthermore, very common application of online learning is demonstrated in Recommendation System [19], where the data in the form of individual rating arrive sequentially and promptly in accordance with the recommendation system needed to be enhanced. User preferences could progress over-time, and consequently, the models should be able to outline adaptable to such temporal patterns. Therefore, application of online learning can be categories into three main learning theories: first order, second order, and online learning with expert advice.

First Order Online Learning: This algorithm learns by updating the weight vector w for classification task chronologically by using only the first order information with training data. The concept of first order online learning algorithm is describe below:

Perceptron [19] is the initial learning algorithm. In each iteration, for any mistake by the prediction model, perceptron makes an update as indicated in the relation in (1).

$$\boldsymbol{w_{p+1} \leftarrow w_p + y_p x_p} \tag{1}$$

Online Gradient Descent (OGD) [23] updates the vector w by employing the stochastic technique of gradient descent principle to a single training instance arriving chronologically. Specifically, OGD makes an online update iteratively in the function (2).

International Journal of Scientific & Engineering Research, Volume 11, Issue 8, August-2020
ISSN 2229-5518

$$\boldsymbol{w_{p+1} \leftarrow w_p - \eta \Delta \ell(w_p, x_p; y_p)} \tag{2}$$

$$where\ \eta\ is\ the\ step\ size\ parameter, \ell(w_p, x_p; y_p)\ is\ some\ predefined\ loss\ function$$

Passive Aggressive Learning [36] is an online learning technique that has two main concerns: to avoid the new model deviating too much from the existing one and aggressiveness by updating the model and correcting the prediction mistake as much as possible. The optimization of passive aggressive learning is mathematically defined in (3).

$$\boldsymbol{w_{p+1} \leftarrow \underset{w}{argmin} \frac{1}{2} \left\| w_p - w \right\|^2 subject\ to\ y_p(w * x_p) \geq 1} \tag{3}$$

The derived solution is the update rule as defined below:

$$\boldsymbol{w_{p+1} \leftarrow w_p + T_p y_p x_p} \tag{4}$$

$$\boldsymbol{Where\ T_p = max\left(\frac{1 - y_p(w_p * x_p)}{\left\| x_p \right\|^2}, 0\right)}$$

The model above assumes a hard margin exists, that is data can be linearly separable which may not be necessarily true always, especially noisy data. To overcome this shortcomings, soft-margin passive aggressive variants such as PA-I and PA-II are commonly used with closed-form solutions.

Second Order Online Learning: this online learning objectives is to boost the learning ability by exploiting second-order information through underlying statistical distributions. Some of the most popular second-order algorithm are discussed below:

Confidence-Weighted Learning (CW) [10] is parallel to the aggressiveness tradeoff, accepts that CW utilizes the second order information. CW learning maintains a diverse confidence measures for each user feature, such that weights of lower confidence will be updated more aggressively than those of higher confidence. Example by modelling the weight vector as a Gaussian distribution, CW tradeoffs between passiveness and aggressiveness. CW learning can be cast into the optimization function as in (5).

$$\boldsymbol{(\mu_{p+1}, \Sigma_{p+1}) \leftarrow \underset{\mu, \Sigma}{argmin}\ D_{KL}(N(\mu, \Sigma) || N(\mu_p, \Sigma_p))} \tag{5}$$

subject to $y_p(\mu, x_p) \geq \emptyset^{-1}(\eta)\sqrt{x_p^T \Sigma x_p}$

Here the CW optimization objective is modified for linearly non-separable data as defined in (6).

$$\boldsymbol{(\mu_{p+1}, \Sigma_{p+1}) \leftarrow \underset{\mu, \Sigma}{argmin}\ D_{KL}(N(\mu, \Sigma) || N(\mu_p, \Sigma_p)) + \lambda_1 \ell(y_1, \mu * x_p) + \lambda_2 x_p^T \Sigma x_p} \tag{6}$$

Online Learning with Expert Advice

Expert advice with online learning predictions entails with creating a concluding prediction based on the prediction made by a set of experts. For instance consider N experts, at each time t step: $t = 1, 2, \ldots \ldots P$, the algorithm decides on probability distribution $P_t$over the experts, where $P_t, i \geq 0$ the weight allocated to expert i, and $\sum_{i=1}^{N} P_t, i = 1$. For each expert i suffers some loss ℓ_t, i determined by the general online learning settings. The loss impacted by the algorithm is then computed as stated in (7).

$$\boldsymbol{\sum_{i=1}^{N} P_t, i\ell_t, i = P_t^T * \ell_t} \tag{7}$$

The average loss of the algorithm without loss of generality, thus assumed that $\ell_t, i \in [0,1]$. This algorithm is called Hedge [21] for the online prediction with expert advice. This scheme is direct simplification of Littlestone and Warmuth's weighted majority algorithm [14]. The scheme maintains non-negative weight vector whose value at time t is represented as $w_t = (w_t, 1, \ldots\ldots w_t, N)$. The scheme uses the normalized distribution to make prediction as indicated in (8).

$$\boldsymbol{P_t = \frac{w_t}{\sum_{i=1}^{N} w_t, i}} \tag{8}$$

After the loss ℓ_t is disclosed, the weight vector w_t is updated using multiplicative rule in (9) to achieve the function.

$$\boldsymbol{w_{t+1}, i = w_t, i\beta^{\ell_{t,i}}, \beta \in [0, 1]} \tag{9}$$

Thus, implies that the weight of expert i will exponentially decrease with the loss $\ell_{t,i}$. This permit Hedge to track the performance of the best expert.

5 ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM ARCHITECTURE EXPLAINED

The adaptive neuro-fuzzy inference system (ANFIS) method was presented by Jang [48]. An ANFIS technique consist of Fuzzy Logic that learns from simple data given inputs into an anticipated out through extremely interconnected Neural Network processing parameters and data connections and subsequently weighted to map the numerical inputs into a desired output.

Technically, ANFIS is a hybrid that has both advantages of machine learning techniques of Fuzzy Logic and Neural Network to form one technique [20]. However, there are some characteristics that enhance ANFIS algorithm to achieve effective and efficient results [25, 26]: it improves fuzzy IF-THEN rules to describe the behavior of multifaceted system, does not require human expertise, enhance fast and precise learning, easy to implement, easy to integrate both linguistic and numeric knowledge for problem-solving, offers anticipated dataset, greater choice of membership functions to utilize, robust generalization capabilities, and exceptional explanation abilities through fuzzy rules.

However, various rules cannot segment the same output membership function in diverse forms. The number of rules settings must be equal to the number of membership functions. In the architecture of ANFIS representation, there are basically two fuzzy rules called IF-THEN on the first-order Sugeno model are measured as:

Rule I: $IF\ x\ is\ M_1\ AND\ y\ is\ N_1, THEN$
$f_1 = p_1 x + q_1 y + r_1.$
Rule II: $IF\ x\ is\ M_2\ AND\ y\ is\ N_2, THEN$
$f_2 = p_2 x + q_2 + r_2.$

International Journal of Scientific & Engineering Research, Volume 11, Issue 8, August-2020
ISSN 2229-5518

$$Where: x \; and \; y \; are \; the \; inputs,$$
$$M_i \; and \; N_i are \; the \; fuzzy \; sets,$$
$$f_i \; are \; the \; expected \; outputs \; within \; the \; fuzzy \; region \; determined \; by \; the \; fuzzy \; rules, p_i, q_i, and \; r_i \; are \; the \; design \; paramters \; that \; are \; determined \; during \; the \; training \; process.$$

The architecture of ANFIS is illustrated in **Fig. 4** which is the reasoning engine for this Sugeno model. This architecture is used to implement two rules (see **Fig. 4**). The circle represents a fixed node and on the other hand, the square represents an adaptive node. We explained this model by exploring on each layer and denotes description of the ANFIS model as follows:

Layer I: In layer I, all the nodes are adaptive nodes, the outputs of layer I are fuzzy membership grade of the inputs which are determined by the (10) below:

$$O_{1,i} = \mu M_i(x), i = 1, 2$$
$$O_{1,i} = \mu N_i(y), i = 3, 4 \tag{10}$$

$$where: x \; and \; y \; are \; the \; inputs \; to \; node \; i$$
$$M_i \; and \; N_i \; are \; linguistic \; labels \; of \; high \; or \; low \; related \; to \; the \; node \; function \mu M_i(x) \; and \; \mu N_i \; 2(y) \; can \; adopt \; any \; fuzzy \; membership \; function$$

For instance, if the bell-shaped membership function is employed, $\mu M_i(x)$ is given by (11).

$$\boldsymbol{\mu M_i(x)} = \frac{\mathbf{1}}{\mathbf{1} + \left[\left(\frac{\boldsymbol{x - p_i}}{\boldsymbol{m_i}}\right)^{\mathbf{2}}\right] \boldsymbol{n_i}}, \boldsymbol{i = 1, 2} \tag{11}$$

Or the Gaussian membership function as in (12):

$$\boldsymbol{\mu M_i(x) = exp}\left[-\left(\frac{\boldsymbol{x - p_i}}{\boldsymbol{m_i}}\right)^{\mathbf{2}}\right] \tag{12}$$

$$where \; m_i, n_i \;, and \; p_i \; are \; the \; parameters \; of \; the \; membership \; function.$$

Layer II: In layer II, the nodes are fixed. This layer includes fuzzy operators, specifically utilizes the AND operator to fuzzifier required inputs. Hence, they are denoted with π, signifying simple performance of the multiplier. The output of Layer-II can be donated as in expression (13).

$$\boldsymbol{O_{2,i} = w_i = \mu M_i(x) * \mu N_i(y), i = 1, 2} \tag{13}$$

Basically, the equation is called the firing strength of the rules.

Layer III: Again, layer III has fixed nodes labeled by N to represent their capabilities to normalize the role to the firing strengths from layer II. The output of this layer can be denoted as (14).

$$\boldsymbol{O_{3,i} = \overline{w_i} = \frac{w_i}{w_1 + w_2}, i = 1, 2} \tag{14}$$

The output of this layer is generally called the normalized firing strength.

Layer IV: In layer IV, the nodes are adaptive. Meaning the output of each node in this layer is the product of the normalized firing strength and a first-order polynomial as represented in the first-order Sugeno model. The output of this layer is represented in (15).

$$\boldsymbol{O_{4,i} = \overline{w_i} f_i = \overline{w_i}(p_i x + q_i y + r_i), i = 1, 2} \tag{15}$$

Where $\overline{w_i}$ is the output of layer III, and $p_i, q_i, and \; r_i$ are the parameters settings.

Layer V: In layer V, there is only one single fixed node labeled with $\sum$. This node computes the summation of all incoming signals. The entire output of this model is represented by the function (16).

$$\boldsymbol{O_{5,i} = \sum_i \overline{w} f_i = \frac{\sum_i w_i f_i}{\sum_i w_i}} \tag{16}$$

5.1 ANFIS Hybrid Learning Algorithm

The ANFIS learning algorithm is called the hybrid algorithm because it is a combination of gradient descent and least square techniques. To determine the forward pass of the hybrid learning algorithm, node outputs go forward until layer IV and the subsequent parameters are adequately updated using the gradient descendent. According to Jang [48] the hybrid learning technique congregates much quicker by decreasing the search space dimensions of the initial backpropagation technique. Generally, the output can be represented by the following functionalities formed by functions in (17), (18), and (19) respectively.

$$\boldsymbol{f = \left[\frac{w_1}{w_1 + w_2}\right] * f_1 + \left[\frac{w_2}{w_1 + w_2}\right] * f_2} \tag{17}$$

$$\boldsymbol{f = \overline{w}(p_1 x + q_1 y + r_1) + \overline{w}(p_2 x + q_2 y + r_2)} \tag{18}$$

$$f = p_1(\overline{w_1}x) + q_1(\overline{w_1}y) + r_1(\overline{w_1}) + p_2(\overline{w_2}x) + q_2(\overline{w_2}y) + r_2(\overline{w_2}) \tag{19}$$

$where \; p_1, q_1, r_1, p_2, q_2, \; and \; r_2 \; are \; the \; linear \; subsequent \; paramters$

The optimal values of these parameters can be determined by the least squares' technique. Hence if the proposition parameters are not resolved, the search area becomes a challenge and larger enough with the convergence of the training becomes very slow. Notably, the ANFIS hybrid scheme combines two techniques, namely the least squares technique and the gradient decent technique, to enhance and resolve the problem of search space. The hybrid scheme comprises of a forward pass and a backward pass respectively.

The forward pass is the least squares technique which utilizes the optimization of the subsequent parameters while the backward pass is the gradient descent technique and utilizes the optimization of the proposition parameters. The output of the ANFIS is computed in the forward pass. And on the other hand, the output error is used to adapt the proposition parameters by a required technique of standard backpropagation algorithm. According Jang [48] , they show that the hybrid algorithm is significantly efficient in training the ANFIS system requirements. This makes the ANFIS system very robust for both classifications and nonlinear model using machine learning techniques.

International Journal of Scientific & Engineering Research, Volume 11, Issue 8, August-2020
ISSN 2229-5518

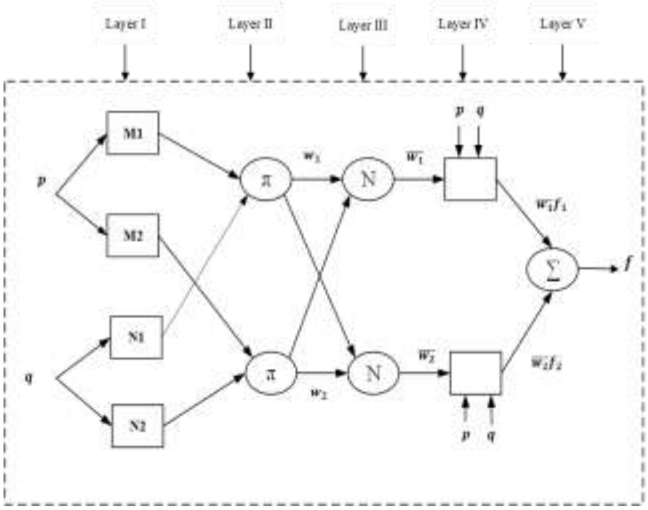


Fig. 4. Jang [19] ANFIS Architecture

6 ANFIS ARCHITECTURE FOR ADAPTIVE ONLINE LEARNING

ANFIS has played a significant role in modeling and controlling imprecise-defined and uncertain systems due to the intelligence of the Neuro-Fuzzy technique. Under considerations, ANFIS is based on the input and output datasets pairs of the system. In order to solve the problem of continuous changes in online learning settings by facilitating the delivery of adapted learning framework, ANFIS is selected. Therefore, the proposed ANFIS model can be used for modeling the learner framework. Essentially, the stages required to apply an ANFIS to learner modeling may include: define input and output values; define fuzzy sets for input the required values; define fuzzy rules, and create and train the Neural Network.

The proposed ANFIS architecture tool needed to be developed and implemented. The researchers utilize two main tools developed in the MATLAB environment: Fuzzy Logic Toolbox from MathWorks and an ANFIS architecture developed from scratch with user define the interface. This tool provides an environment to build and evaluate systems using the graphical user interface. From the MathWorks perspective, it consists of FIS editor, the rule editor, a membership function editor, the fuzzy inference viewer, and the output surface viewer. While the ANFIS implemented perspective, it consists of all the MathWorks parameters but with a defined user interface that helps improve the functionalities of the ANFIS system.

6.1 Learning Framework Settings Explained

The main objective of this research is to construct a model that will be useful to learners in order to accomplish learning activities by conveying adaptive learning content based on the learner's present framework. The procedure of the learner modeling system involves their key steps: collecting data associated with the learner, developing the learner model, and optimizing and updating the learner model. Essentially for clarity and understanding the learner's model the researchers needs to model the learner's learning construct and modularizes these systems into perspective by analyzing the condition and tasks involved for designing the learner model:

I. *It is 8:30 am Wednesday morning and Dmitri is traveling by public transport (bus/car/train) from his house to attend his lecture at the university which starts at 9:00 am. Dmitri checks his computer accessories (4G or 5G network, 1mps, or more bandwidth) to revise his lecture notes.*

II. *Isaac is in a Zoo for 30 minutes and wants to revise his assessments using his smartphone with the 3G network, 70kps, or more bandwidth.*

III. *It is weekend and Einstein is at the beech for two hours and wants to revise his academic materials using his Samsung phone with probably 4G network, and 500kbs bandwidth.*

We can depict several contextual parameters that contained and described in the above circumstances. Different activities can be identified such as places, events, time, locations, challenges, and other situations surrounding the online learning theory. Hence, we finalized by labeling the above circumstances as "Learning Task". We represented these tasks into different paradigms or features as shown in Table 1.

Table 1: Circumstances of Learning Tasks

No.	Learning Task	Features
1	Personal information	Name, ID number, gender, language, occupation, learning method
2	Time	Weekdays, weekends
3	Learning resources	Assignments, lecture notes
4	Location	Home, class-level, public transport, private transport, roadside
5	Electronic device	PDA, laptop, smartphone
6	Environmental settings	Sunny, windy, noisy, raining
7	Network	3G, 4G, 5G, GSM, WIFI
8	Bandwidth	Low, high, moderate/middle
9	Customizability & Adaptability	Customize feature, request, change platform looks

6.2 Proposed Online Learning Framework Inputs Selection

It is usual to have tens of potential inputs that could be used for online learning systems. The input criteria are essential when determining or creating an online learning model. However, an extreme increase in the number of inputs will increase the computational time in modeling the online learning system using ANFIS. Therefore, is important to select all inputs and design the model appropriately. To develop an accurate model for the prediction, substantial inputs must be chosen. There are many techniques that have been proposed for input selection in the literature review. According to ARD [2], CART [3], and $\delta - test$ [2], that determines the dependencies within data in order to select relevant inputs. David JC MacKay et al. [81], listed some practical considerations for input selections: eliminate noise or unrelated inputs,

International Journal of Scientific & Engineering Research, Volume 11, Issue 8, August-2020
ISSN 2229-5518

eliminate inputs that depends on other inputs, create the underlying system more brief and evident, and minimize the time complexity for model construction.

The existing technique presented in [2], takes the benefit of the ANFIS structure. It is based on the proposition that the ANFIS system with the minimum Root Mean Square Error (RMSE) after a small number of epochs has a greater possibility of achieving a lower RMSE when more epochs of training are provided. For instance, if there are 12 candidate inputs and we need to determine the most significant inputs, we therefore, can create $C_3^{12} = 220$ ANFIS models and model with minimum RMSE will be chosen.

In this paper for full implementation of an online learning model, hundreds of possible inputs can be selected for the model. If there are eight inputs, the four most significant inputs could be determined. Each time the RMSE is computed for combinations of variables such as ($C_4^8 = 70$) and the smallest RMSE is recorded. Dynamically, the algorithm runs for one input at a time and then two inputs at a time, and so on until it arrives the nine inputs at a time.

7 PROPOSED ANFIS-AOLM MODEL METHODOLOGY

In this section, we explain the proposed block diagram of the ANFIS-AOLM training model and the and the proposed architecture of the ANFIS-AOLM model. The reason is it to be able to understand the process used through the training method in the implementation and the architecture of the ANFIS-AOLM model.

7.1 Proposed Block-Diagram of ANFIS-AOLM Training Model

The training method of the ANFIS-AOLM model is shown in **Fig. 5**. The process is initiated by obtaining a training data with input and output data pairs. The training data is a set of both input and output vectors. Namely, the input vector and output vector are used to train the ANFIS-AOLM model. The training dataset is used to determine the proposition parameters for the membership functions. To determine the between the actual and desired output, a threshold value for the error is defined.

The subsequent parameters are established by using the least-squares method. If this error is greater than the threshold value, then the preposition parameters are updated using the gradient descent method. The process is terminated until when the error becomes less than the threshold value. The checking dataset is then used to compare the model with the actual system [48].

As stated initially in the literature, ANFIS training learning procedures use hybrid learning technique, by combining the least squares and gradient descent technique. The objective of utilizing ANFIS for adaptive online learning is to achieve the best performance possible. ANFIS training starts by developing adequate training data in order to be able to train the Neuro-Fuzzy system. The process of obtaining data must involve as many online learning conditions such as devices, locations, networks, bandwidth, and so on as feasible. ANFIS training uses the anfis-function and evaluation of the model compared to the desired output is determines by using the evalfis-function.

The first step is to prepare the training data to work with ANFIS in MATLAB. The dataset used as the input to the anfis-function and represented in a matrix form, where the last column in the matrix is the output, and the matrix contains as many columns as needed to denote the inputs to the system. The rows represent all the existing data situations. Since the proposed ANFIS online learning model is coded using MATLAB programming language with full utilization of MathWorks libraries and functions. The ANFIS architecture tool is developed in MATLAB purposely to enhance the ANFIS system and its performance peculiarities. Upon running the application displays three main parameter functions of which allow the user to select which kind of analysis the user may require to perform. These include Grid Partitioning (genfis1), Subtractive Clustering (genfis2), and Fuzzy Clustering Means (genfis3).

Grid Partitioning: Upon selection of grid partitioning allow users to define or enter genfis1 parameter functions such as *membership functions (MFs), input membership function type (IMF/Gaussian), and output membership function type (linear)*. Determining these parameters, the system allows and display the genfis3 which permit users to define the required values for a maximum number of *epochs, error goal, initial step size decrease rate, and step size increase rate*. We defined default values for the genfis3 which automatically compute and plot the resulting functions of the main genfis1 parameters which are the initial step of the membership functions. This function is responsible for producing the most accepted output.

Subtractive Clustering: The subtractive clustering (genfis2) allow users to define the influence radius of the genfis2 thereby permitting the model to compute display the next stage of the genfis3 with the maximum number of *epochs, error goal, initial step size decrease rate, and step size increase rate*. All of these essentially allow the genfis2 clustering to be enhance and thus have default values.

Fuzzy Clustering-Means: The FCM (genfis3) displays the corresponding parameter functions of genfis3 that include a *number of clusters, partition matrix exponent, maximum number of iterations and minimum improvement*. The FCM function further requires defining the maximum number of *epochs, error goal, initial step size decrease rate, and step size increase rate*, after defining the previous parameters of the fuzzy clustering means. This helps the system determining the robustness of the ANFIS online learning approach which is explained in the next section of this paper.

International Journal of Scientific & Engineering Research, Volume 11, Issue 8, August-2020
ISSN 2229-5518

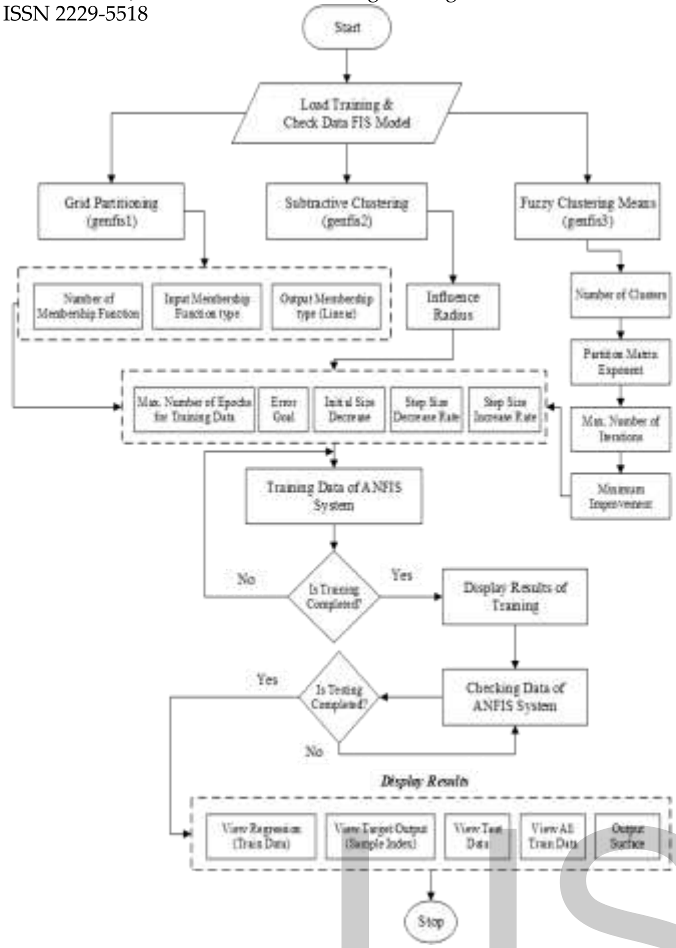


Fig. 5. Proposed Block Diagram of ANFIS-AOLM Training Model

7.2 Proposed Architecture of ANFIS-AOLM Model

Upon the completion of the training model, ANFIS offers a technique to study and evaluate the model performance by using the evalfis-function. The fuzzy system that was founded based on the completed training is used. By entering the input datasets into the fuzzy model, the process starts by studying and evaluating the robustness of the model. It is important to note that these datasets do not include output values. The evaluation of the evalfi-function depicts the response of the final output of the ANFIS model. This response output can be measured by means of correlations between the desired learner framework and the learning content format as the model result or input/output. Immediately the ANFIS is trained, we can experiment the system against different sets of data values to check the functionality and performance of the proposed model.

To determine the criteria set format, the condition was set by the researcher expert in determining the learning framework. Each of the six input conditions is denoted by the setting the following conditions: LL has four membership functions denoted as *home, campus, class outdoor*. NB has three membership functions denoted as *low, middle, and high*. NA has three membership functions denoted as *weak, moderate, and strong*. CPC has three membership functions denoted as *low, half, and full*. DSC has five membership functions denoted as *text, text+pdf, text+video, text+pdf+video, and text+audio*. CA has three membership functions denoted as *customize, requested, platform-looks*. The proposed ANFIS-AOLM model is represented in **Figure 6** for the adapted online learning model.

We represented the Sugenor-type fuzzy rules-based system in the following forms:

$$IF\ (LL\ =\ Outdoor)\ AND\ (NB\ =\ High)\ AND\ (CPC\ =\ Full)\ AND\ (NA\ =\ Strong)\ AND\ (CPC\ =\ Full)\ AND\ (DSC\ =\ Text + PDF + Video)\ AND\ (CA\ =\ Requested)\ THEN\ (Learning\ Framework\ Format\ is\ Video)$$

$$IF\ (LL\ =\ Class)\ AND\ (NB\ =\ Low)\ AND\ (CPC\ =\ Half)\ AND\ (NA\ =\ Weak)\ AND\ (CPC\ =\ Half)\ AND\ (DSC\ =\ Text + PDF + Video)\ AND\ (CA\ =\ Requested)\ THEN\ (Learning\ Framework\ Format\ is\ PDF)$$

In MATLAB, the genfis1 is the function used to create an initial single output of FIS matrix from the training data. By initiating our model with default values for membership functions “20 and 25” and of types such as “Gaussian curve” called the gaussmf-function, “Triangular-shaped” called the trimf-function and “Generalized bell-shaped” called the gbellmf-function. The defaults values provide membership functions on each of the six inputs making a total of 21 altogether. The generated fuzzy inference system contains 1620 fuzzy rules. The linear type by default has one output membership function generated by each of the rules of genfis1. The ANFIS is a multi-input with a single structure of each of the fuzzy rules. The main objective of the gradient vector is to compute the alteration of the parameters of the membership function hence indicating how satisfactory the ANFIS is modeled by providing various set of training data comprising of certain conditions.

Furthermore, the ANFIS training method is initiated by shaping the fuzzy sets and number of each set input variable and determining of their membership function. The training data is permitted to pass through the Neural Network. The process therefore, adjust the input parameters to determine the associations between input/output, and to mitigate the errors. RMSE, Error Mean and Error standard deviation are computed and used to monitor the training errors which are defined in (20):

$$\boldsymbol{RMSE} = \sqrt{\frac{\mathbf{1}}{\boldsymbol{N_p}}\sum_{\boldsymbol{j=1}}^{\boldsymbol{N_p}}(\boldsymbol{y_i} - \boldsymbol{\bar{y}_j})^{\mathbf{2}}} \tag{20}$$

The mean absolute error (MAE) is used and defines as in (21):

$$\boldsymbol{MAE} = \frac{\mathbf{1}}{\boldsymbol{N_p}}\sum_{\boldsymbol{j=1}}^{\boldsymbol{N_p}}\left|\boldsymbol{y_i} - \boldsymbol{\bar{y}_j}\right| \tag{21}$$

The standard deviation error is used and defined as in (22):

$$\boldsymbol{SDE} = \frac{\mathbf{1}}{\sqrt{\boldsymbol{N_p}}}\sqrt{\frac{\mathbf{1}}{\boldsymbol{N_p}}\sum_{\boldsymbol{j=1}}^{\boldsymbol{N_p}}(\boldsymbol{y_i} - \boldsymbol{\bar{y}_j})^{\mathbf{2}}} \tag{22}$$

$where\ N_p\ is\ the\ total\ number\ of\ predictions, \bar{y}_j\ is\ the\ prediction\ time\ series, and\ y_i\ is\ the\ original\ series$

To circumvent the overfitting challenge, we test our model by

International Journal of Scientific & Engineering Research, Volume 11, Issue 8, August-2020
ISSN 2229-5518

setting training epoch with different values. Consequently, allow the training data to optimal epoch number with the smallest RMSE. It then generates and plots an outputs surface map for the system presented by 3D surface model.

In our research, an ANFIS model based on both Neural Network and Fuzzy Logic has been developed and coded for non-linear model to adapt learning framework formats for online learning system. The simulation exercises were divided into six main ANFIS organizations to show learning tasks in various settings. The simulation through experimentation is demonstrated in MATLAB were statistical validation indexed were used to determine the robustness of the best proposed model within each framework. Following the validation of both inputs and output have been identified, which is appropriate to authenticate the quality of each model results after training the data. Upon validation of the model hence authenticate the accuracy and validity whether the model can be easily understood and interpreted.

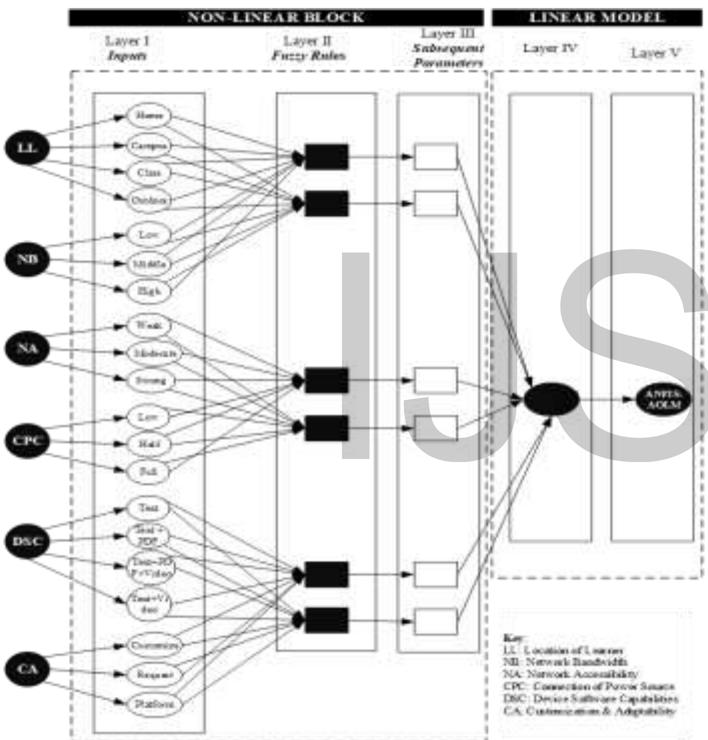


Fig. 7. Proposed ANFIS-AOLM Architecture with six input and one output

The robustness of the ANFIS models for both training and examining data were evaluated and the best training data through the dataset was selected with reference value of RMSE. By comparing the difference of predicted and measure values, the prediction accuracy was determined and computed appropriately. In order to obtain better result different parameter values of the dataset items namely: number of epochs, membership function type and number, number of both increase and decrease step size to achieve finest performance. Analysis and results for the prediction for both training and testing scenarios are presented in next section of this paper.

8 Experimental Result And Discussions

Different experimental simulations were carried out and the sizes of training and checking data pairs were determined by taking the prediction accuracies and classifications into consideration. The dataset was divided into two columns: inputs and target data. The training data was used to train the ANFIS, while the checking data (target data) was used to validate the accuracy and performance of the trained ANFIS model for the adaptation of the ANFIS-AOLM context. We needed to investigate the number of factors that play a significant role in addressing the challenge of the online learner model was determined

Essentially, to determine the optimal ANFIS-AOLM model framework the smallest RMSE will be based on the comparison of RMSE value for different epoch numbers and the number of membership functions assigned to each ANFIS structure. There are four main considerations for the ANFIS training system: overfitting, number of membership functions, type of membership functions, and training options (training data, test data, and epoch maximum number).

ANFIS-I: The first system was streamlined by selecting three inputs and feeding them into the network. Three main inputs and membership functions (9). The selection criteria were based on the output of the training ANFIS system for the online learner application. The three-parameter controlling the inputs were involved namely: Location of Learner, Network Bandwidth, and Device Software Capabilities. Where LL has three-parameter functions denoted as Home, Campus, and Outdoor. NB has three membership functions denoted as Low, Middle, and High. DSC has their membership functions denoted as *Text, Text+PDF, and Text+PDF+Video*.

Our first simulation was to test the training data with values for membership functions (9, 18, 20, and 25) with membership function type Gaussian. These default membership functions enhance the membership functions on each of the three inputs making a total of nine (9) altogether. The number of fuzzy rules generated by the ANFIS-I structure for 9, 18, 20 and 25 (see Table 1). The membership functions perform efficiently with a minimum value of RMSE (17.257) and maximum epoch number of 200 as indicated in Table 1.

ANFIS-I was tested again for four inputs and membership functions number (18). Secondly, the simulation involves four inputs: LL has three membership functions denoted as *Home, Class, and Outdoor*. NB has their membership functions denoted as Low, Middle, and High. DSC has three membership functions denoted as *Text, Text+PDF, and Text+PDF+Video*. CPC has their membership functions denoted as *Low, Half, and Full*.

The input parameter values (LL, NB, CBC, and DSC) are the default model values of membership function type (20 and 25) Gaussian, generalized bell-shaped, and Triangular-shaped are utilized. These four inputs combined altogether making twelve (12) membership functions. ANFIS-I structure generated fuzzy rules with tested membership functions numbers and types are represented in Table 1. The number of fuzzy rules generated by the ANFIS-I structure was for membership function of 9, 18, 20, 25. The membership functions perform efficiently with a minimum value of RMSE (**2.46**) and maximum epoch number of 200 with generated fuzzy rules of 81, 324, 400, and 625 respectively (see Table 1 and Figure 1).

International Journal of Scientific & Engineering Research, Volume 11, Issue 8, August-2020
ISSN 2229-5518

Table 2: The ANFIS-I, II, and III Architecture Information

Description	Result obtained			
ANFIS (I) Grid Partitioning				
Number of membership functions (NMF)	9	18	20	25
Number of nodes	203	725	885	1355
Number of linear parameters	243	972	1200	1875
Number of non-linear parameters	36	72	80	100
Total number of parameters	279	1044	1280	1975
Number of training data pairs	1019			
Number of Inputs	3, 4, 5			
Input Membership Function Type	Gaussian Curve			
Number of combinations	LL, NB, and CBC		LL, NB, CBC, and DSC	
Number of fuzzy rules	81	324	400	625
Designated epoch number reached ANFIS training completed	200			
Minimal RMSE	17.257	3.939	4.448	3.153
ANFIS (II) Subtractive Clustering			**ANFIS (III) Fuzzy Clustering Means**	
Influence radius	0.3			
Number of clusters			50	100
Number of nodes	131		305	605
Number of linear parameters	63		150	300
Number of non-linear parameters	84		200	400
Total number of parameters	147		350	700
Number of fuzzy rules	21		50	100
Number of training data pairs	1019			
Input Membership Function Type	Triangular Shape		Generalized bell-Shape	
Designated epoch number reached ANFIS training completed	200			
Minimal RMSE	19.791		16.416	

ANFIS-II (Subtractive Clustering): The number of input membership functions were evaluated with 9 and 18 that represent four inputs. LL has three membership functions as "Home, Class, and Outdoor". NB has three membership functions represented as "Low, Middle, and High". DSC has three membership functions represented as Text, Text+PDF+Video". LL, NB, CBC, and DSC are the inputs, with values for membership function (20 and 25); membership function type is Triangular shape was used (see Fig. 8). The generated fuzzy inference system structure contains 21. The result shows a minimal RMSE value of 19.791.

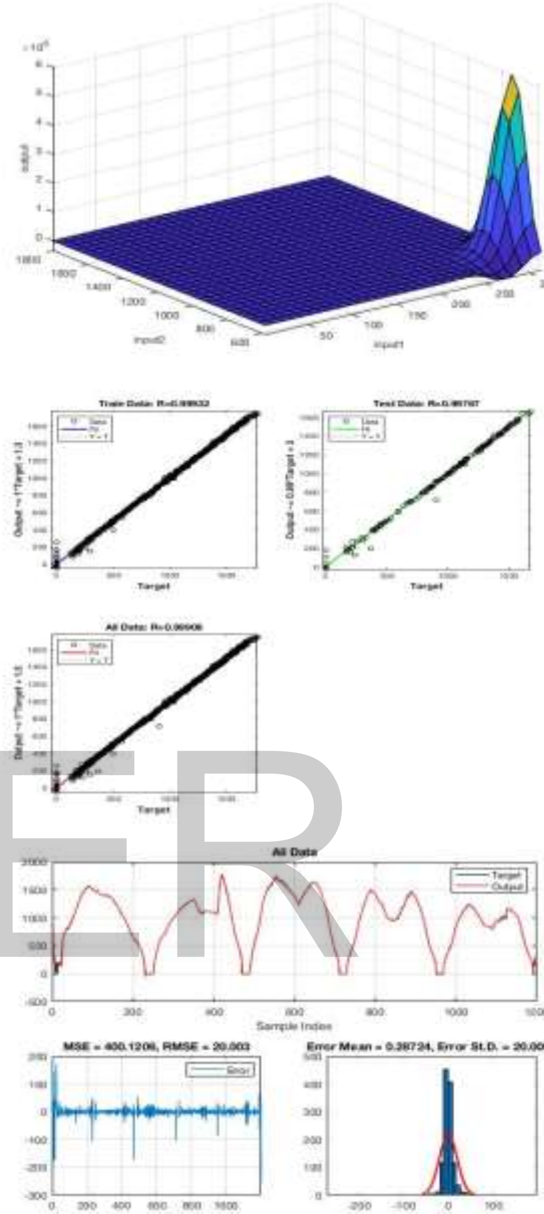


Fig. 8. Results obtained after ANFIS-I Analysis with 9 Membership functions

ANFIS-III (Fuzzy Clustering Means): The number of clusters at 50 and 100 has experimented. The results show that at cluster 50, the output node, linear parameter, non-linear parameter, and fuzzy rules values double the result obtained when 50 and 100 number of clusters was inserted into the Neural Network. The minimal value of RMSE at clusters number 50 and 100 remain the same (16.416). The values of the RMSE in ANFIS-I with membership function at 9, 18, 20, and 25 are lesser than ANFIS-II and ANFIS-III except for a membership function of 9. Therefore, ANFIS-I and ANFIS-III was selected as the two predictions of the ANFIS-AOLM model to conduct the online learning model construct.

ANFIS-I, in terms of the number of inputs, shows a variation in the number of membership functions and the training data and testing data sample. According to the results ANFIS-I, the predic-

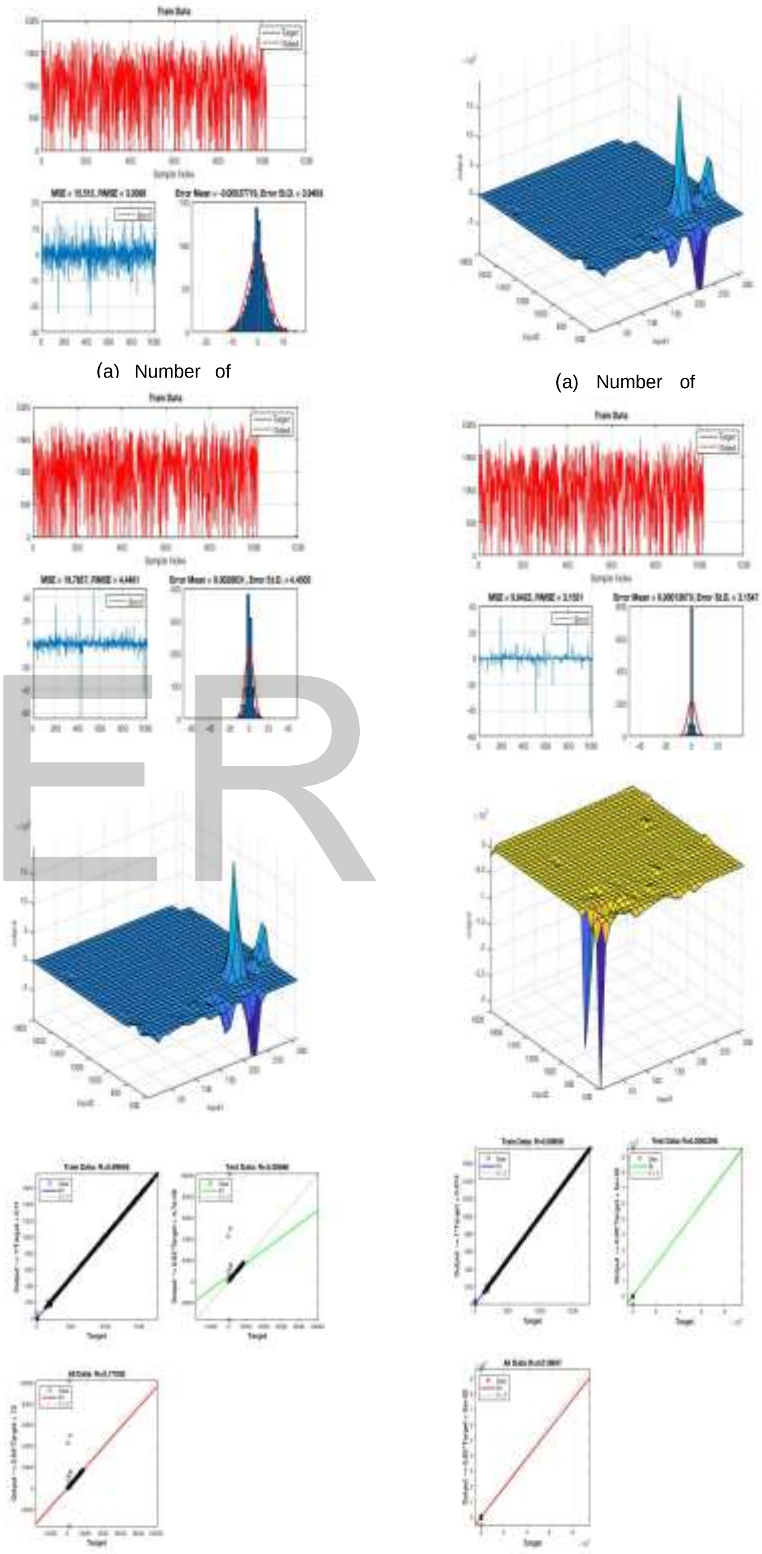
International Journal of Scientific & Engineering Research, Volume 11, Issue 8, August-2020
ISSN 2229-5518

tions mechanism is better than ANFIS-II and ANFIS-III. Conversely, ANFIS-I consumed a lot of time to compute when the MF is inserted into the network, especially at NMF of 20 and 25. Therefore, the higher the MF, the more time is required to compute the parameters and the hence better result obtained (in terms of RMSE value).

For each ANFIS I, II, and III using different epoch number and using the same epoch number do not necessarily change the performance of the network or prediction value. However, it improves and resolves the challenges of overfitting. Several tests were conducted in Table 1 and Fig. shows that the proposed ANFIS-AOLM performance overall, accurate, and suitable where some rules are determined by a human expert and ANFIS system to determine and discovered the missing rules using the training data.

Furthermore, a 3D surface plot for ANFIS I, II, and III to see how a response variable relates to two predictor variables (train data and test data) was explored. A 3D surface shows a plot of a three-dimensional graph that is useful for investigating desirable response values and operating conditions. The peaks and valleys correspond with combinations of input1 and input2 (see **Fig. 9**) that produces maxima or minima. The plot shows the relationship between the training data and testing data settings used to produce the quality of the proposed ANFIS-AOLM model for learner's framework. A smaller dataset results in a low performance of the ANFIS (output). However, providing enough data for the training data combined with the highest performance of the ANFIS and also result in a low value of RMSE. The peak on the plot corresponds with the highest output scores. **Fig. 9** and **10** illustrate the output generated by the ANFIS adaptive learning context for the proposed model. The ANFIS-AOLM was trained and tested with different membership functions. In ANFIS-I, the regression predicted was 99.7% accuracy (testing data) and 99.9% accuracy (training data). At MF of 18, the regression prediction was 99.9% accuracy (training data). At MF of 20, the prediction accuracy was 77.0%. This further shows that ANFIS-I performs better considerably than ANFIS-II and III respectively. However, at lower MF the prediction accuracy increases and decreases when MF was increased for NAFIS-I. Conversely, for a lower value of RMSE ANFIS-I performed better whilst at a lower value of MF ANFIS-I perform worst.

In this research work, the applicability and fitness of the ANFIS technique are evaluated through the exploration of the dataset, membership function, and type of membership function. The main contribution of this study is to learning theory to ANFIS in the Neural Network context, the opportunity of the learner to learn and to make decisions based on the availability of management of online learning techniques. Furthermore, to support the learner through an exploration of various learning techniques and to better provide a mechanism that improves an online learning method. By adaptation of the principles such as learner's context and construct a learner's model which eventually changes with time. The utilization of ANFIS for a reasoning mechanism would update the learner context and make accurate predictions (adaptation) in the decision-making process through machine learning theories and techniques.



(a) Number of

(a) Number of

(b) Number of

(c) Number of

Fig. 9. Results obtained after ANFIS-I Analy-

International Journal of Scientific & Engineering Research, Volume 11, Issue 8, August-2020
ISSN 2229-5518

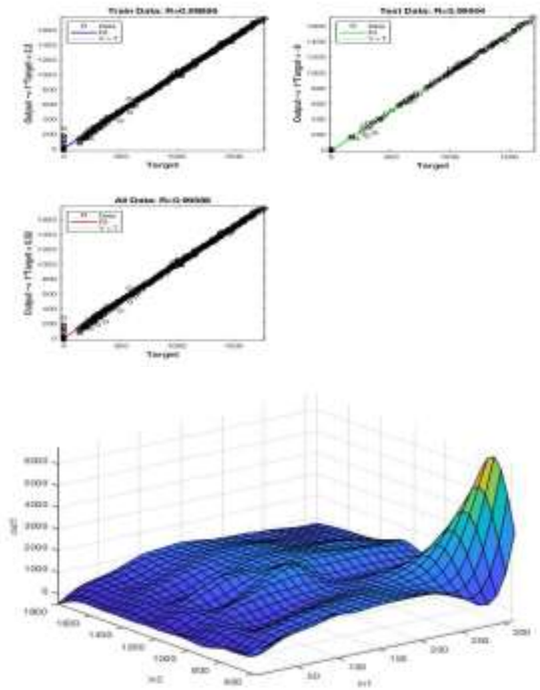


Fig. 10. Results obtained after ANFIS-II Analysis with an Influence Radius = 0.3

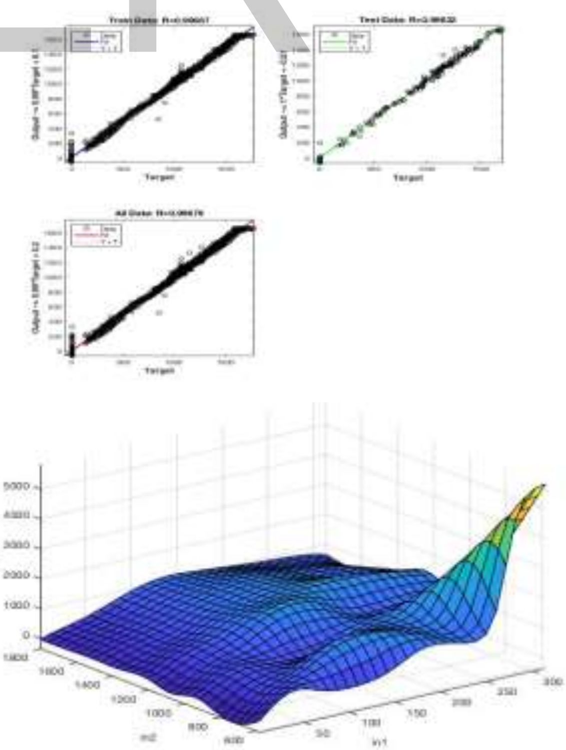


Fig. 11. Results obtained after ANFIS-III Analysis with Number of Cluster = 50, 100

International Journal of Scientific & Engineering Research, Volume 11, Issue 8, August-2020
ISSN 2229-5518

9 Conclusion

The proposed ANFIS-AOLM has being designed, constructed, and evaluated for an online learning framework to each learner's context and update the learner's model with newly derived facts. An overview of recently developed techniques of ANFIS with regression, classification, and multiple-kernel learning approaches.

The paper also explained a theoretical concept to utilize the proposed ANFIS-AOLM system and to a general framework, first-order, second-order, and online learning expert advice. The performance of the proposed ANFIS-AOLM was evaluated using standard error measurements and predictability accuracy (see **Fig 9, 10, and 11**). Because human thinking skills cannot provide adequate rules, the number of fuzzy rules obtained was considered a challenge. Therefore, ANFIS provides a hybrid approach with a combined fuzzy rules inference system and the Neural Network determines a complete learning technique was also explored with the prediction accuracy of 99.7% (ANFIS-I). For future work, we intend to use other machine learning techniques models to see whether there are possible models that would mitigate the challenges of response time when the number of the membership function is increased considerably.

References

[1] Ralph G Andrzejak et al. "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state". In: Physical Review E 64.6 (2001), p. 061907.
[2] Peter C Austin et al. "Using methods from the data-mining and machine-learning literature for disease classification and prediction: a case study exam-ining classification of heart failure subtypes". In: Journal of clinical epidemi-ology 66.4 (2013), pp. 398–407.
[3] William G Baxt. "Use of an artificial neural network for data analysis in clin-ical decision-making: the diagnosis of acute coronary occlusion". In: Neural computation 2.4 (1990), pp. 480–489.
[4] Suliman Yousef Belal et al. "Automatic detection of distorted plethysmogram pulses in neonates and paediatric patients using an adaptive-network-based fuzzy inference system". In: Artificial intelligence in medicine 24.2 (2002), pp. 149–165.
[5] Leo Breiman et al. "Classification and regression trees. Wadsworth & Brooks". In: Cole Statistics/Probability Series (1984).
[6] Emre Comak, ¸ Ahmet Arslan, and Ibrahim T¨urko˘glu. "A decision support sys-tem based on support vector machines for diagnosis of the heart valve diseases". In: Computers in biology and Medicine 37.1 (2007), pp. 21–27.
[7] Koby Crammer et al. "Online passive-aggressive algorithms". In: Journal of Machine Learning Research 7. Mar (2006), pp. 551–585.
[8] Abhinandan S Das et al. "Google news personalization: scalable online collab-orative filtering". In: Proceedings of the 16th international conference on World Wide Web. 2007, pp. 271–280.
[9] Norman R Draper and Harry Smith. "Applied Regression Analysis, John Wiley and Sons". In: New York 407 (1981).
[10] Mark Dredze, Koby Crammer, and Fernando Pereira. "Confidence-weighted linear classification". In: Proceedings of the 25th international conference on Machine learning. 2008, pp. 264–271.
[11] Didier Dubois and Henri Prade. "An introduction to fuzzy systems". In: Clinica Chimica Acta 270.1 (1998), pp. 3–29.
[12] Richard M Felder, Linda K Silverman, et al. "Learning and teaching styles in engineering education". In: Engineering education 78.7 (1988), pp. 674–681.
[13] Yoav Freund and Robert E Schapire. "A desicion-theoretic generalization of on-line learning and an application to boosting". In: European conference on computational learning theory. Springer. 1995, pp. 23–37.
[14] Hong Jiang and Yufen Tian. "Fuzzy image fusion based on modified Self-Generating Neural Network". In: Expert Systems with Applications 38.7 (2011), pp. 8515–8523.
[15] Enrique Frias-Martinez et al. "Modeling human behavior in user-adaptive sys-tems: Recent advances using soft computing techniques". In: Expert Systems with Applications 29.2 (2005), pp. 320–329.
[16] Robert Full´er. "Neural fuzzy systems". In: (1995).
[17] Patricio Garcıa et al. "Evaluating Bayesian networks' precision for detecting students' learning styles". In: Computers & Education 49.3 (2007), pp. 794– 808.
[18] Hubert Gijzen. "Big data for a sustainable future". In: Nature 502.7469 (2013), pp. 38–38.
[19] Inan Guler and Elif Derya Ubeyli. "Application of adaptive neuro-fuzzy inference system for detection of electrocardiographic changes in patients with 27.3 (2004), pp. 323–330
[20] Inan Guler and Elif Derya Ubeyli. "Detection of ophthalmic artery stenosis byleast-mean squares backpropagation neural network". In: Computers in Biology and Medicine 33.4 (2003), pp. 333–343.
[21] Stephanie E Hampton et al. "Big data and the future of ecology". In: Frontiers in Ecology and the Environment 11.3 (2013), pp. 156–162.
[22] Ahmed Al-Hmouz, Jun Shen, and Jun Yan. "A machine learning based framework for adaptive mobile learning". In: International Conference on Web-Based Learning. Springer. 2009, pp. 34–43.
[23] Ahmed Al-Hmouz et al. "Enhanced learner model for adaptive mobile learning". In: Proceedings of the 12th international conference on information integration and web-based applications & services. 2010, pp. 783–786.
[24] Anthony Jameson. "Numerical uncertainty management in user and student modeling: An overview of systems and issues". In: User Modeling and User-Adapted Interaction 5.3-4 (1995), pp. 193–251.
[25] J-SR Jang. "ANFIS: adaptive-network-based fuzzy inference system". In: IEEE transactions on systems, man, and cybernetics 23.3 (1993), pp. 665–685.
[26] J-SR Jang and Chuen-Tsai Sun. "Neuro-fuzzy modeling and control". In: Proceedings of the IEEE 83.3 (1995), pp. 378–406.
[27] Jyh-Shing Roger Jang, Chuen-Tsai Sun, and Eiji Mizutani. "Neuro-fuzzy and soft computing-a computational approach to learning and machine intelligence [Book Review]". In: IEEE Transactions on automatic control 42.10 (1997), pp. 1482–1484.
[28] S Mo Jang and P Sol Hart. "Polarized frames on "climate change" and "global warming" across countries and states: Evidence from Twitter big data". In: Global Environmental Change 32 (2015), pp. 11–17.
[29] M Dursun Kaya et al. "A new approach to estimate anthropometric measure-ments by adaptive neuro-fuzzy inference system". In: International Journal of Industrial Ergonomics 32.2 (2003), pp. 105–114.
[30] Alfred Kobsa. "Generic user modeling systems". In: User modeling and user-adapted interaction 11.1-2 (2001), pp. 49–63.
[31] Priyan Malarvizhi Kumar and Usha Devi Gandhi. "A novel three-tier Internet of Things architecture with machine learning algorithm for early detection of heart diseases". In: Computers & Electrical Engineering 65 (2018), pp. 222– 235.
[32] Ludmila I Kuncheva and Friedrich Steimann. Fuzzy diagnosis. 1999.
[33] B Li, D Sahoo, and SCH Hoi. "Olps: A toolbox for online portfolio selection". In: Journal of Machine Learning Research (JMLR) (2015).
[34] Bin Li and Steven CH Hoi. "Online portfolio selection: A survey". In: ACM Computing Surveys (CSUR) 46.3 (2014), pp. 1–36.
[35] Nick Littlestone, Manfred K Warmuth, et al. The weighted majority algorithm. University of California, Santa Cruz, Computer Research Laboratory, 1989.

International Journal of Scientific & Engineering Research, Volume 11, Issue 8, August-2020
ISSN 2229-5518

[36] D Lopez and G Sekaran. “Climate change and disease dynamics-a big data per-spective”. In: International Journal of Infectious Diseases 45 (2016), pp. 23– 24.

[37] Daphne Lopez and Jagan Mohan J Gunasekaran Manogaran. “Modelling the H1N1 influenza using mathematical and neural network approaches”. In: (2017).

[38] Daphne Lopez and M Gunasekaran. “Assessment of vaccination strategies us-ing fuzzy multi-criteria decision making”. In: Proceedings of the fifth interna-tional conference on fuzzy and neuro computing (FANCCO-2015). Springer. 2015, pp. 195–208.

[39] Daphne Lopez and Gunasekaran Manogaran. “Big data architecture for cli-mate change and disease dynamics”. In: The human element of big data: issues, analytics, and performance (2016), pp. 301–331.

[40] Daphne Lopez and Gunasekaran Manogaran. “Parametric model to predict H1N1 influenza in Vellore District, Tamil Nadu, India”. In: Handbook of statis-tics. Vol. 37. Elsevier, 2017, pp. 301–316.

[41] Daphne Lopez et al. “Spatial big data analytics of influenza epidemic in Vel-lore, India”. In: 2014 IEEE international conference on big data (Big Data). IEEE. 2014, pp. 19–24.

[42] Jake Luo et al. “Big data application in biomedical research and health care: a literature review”. In: Biomedical informatics insights 8 (2016), BII–S31559.

[43] Zhi-hui Luo, Min Wu, and Yi-qing Zhao. “Big data applications in biomedical informatics”. In: Journal of Medical Informatics 36.5 (2015), pp. 2–9.

[44] David JC MacKay et al. “Bayesian nonlinear modeling for the prediction com-petition”. In: ASHRAE transactions 100.2 (1994), pp. 1053–1062.

[45] Jihen Malek, Mona Laroussi, and Alain Derycke. “A multi-layer ubiquitous middleware for bijective adaptation between context and activity in a mobile and collaborative learning”. In: 2006 International Conference on Systems and Networks Communications (ICSNC’06). IEEE. 2006, pp. 39–39.

[46] Gunasekaran Manogaran and Daphne Lopez. “A Gaussian process based big data processing framework in cluster computing environment”. In: Cluster Computing 21.1 (2018), pp. 189–204.

[47] Gunasekaran Manogaran and Daphne Lopez. “A survey of big data architec-tures and machine learning algorithms in healthcare”. In: International Journal of Biomedical Engineering and Technology 25.2-4 (2017), pp. 182–211.

[48] Gunasekaran Manogaran and Daphne Lopez. “Disease surveillance system for big climate data processing and dengue transmission”. In: Climate Change and Environmental Concerns: Breakthroughs in Research and Practice. IGI Global, 2018, pp. 427–446.

[49] Gunasekaran Manogaran and Daphne Lopez. “Spatial cumulative sum algo-rithm with big data analytics for climate change detection”. In: Computers & Electrical Engineering 65 (2018), pp. 207–221.

[50] Gunasekaran Manogaran, Chandu Thota, and Daphne Lopez. “Human-computer interaction with big data analytics”. In: HCI challenges and privacy preserva-tion in big data security. IGI Global, 2018, pp. 1–22.

[51] Gunasekaran Manogaran, R Varatharajan, and MK Priyan. “Hybrid recom-mendation system for heart disease diagnosis based on multiple kernel learning with adaptive neuro-fuzzy inference system”. In: Multimedia tools and appli-cations 77.4 (2018), pp. 4379–4399.

[52] Gunasekaran Manogaran et al. “Big data knowledge system in healthcare”. In: Internet of things and big data technologies for next generation healthcare. Springer, 2017, pp. 133–157.

[53] AS Miller, BH Blott, et al. “Review of neural network applications in medical imaging and signal processing”. In: Medical and Biological Engineering and Computing 30.5 (1992), pp. 449–464.