Journal of Theoretical and Applied Information Technology

15th September 2023. Vol.101. No 17

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

FEATURE CLASSIFICATION BASED ON HETEROGENOUS DATA USING HYBRID MACHINE LEARNING: A REVIEW

AGUS NURSIKUWAGUS[1] , HERI PURWANTO[2] , DESHINTA ARROVA DEWI[3]

[1]Senior Lecturer. Universitas Komputer Indonesia, Informatics Management, Faculty of Engineering and Computer Science, Bandung, Indonesia and Research Fellow, INTI International University, Nilai, Malaysia

[2]Senior Lecturer, Department of Information Systems, Universitas Sangga Buana, Bandung, Indonesia

[3]Senior Lecturer, INTI International University, Faculty of Data Science and Information Technology, Nilai, Malaysia

E-mail: [1]agusnursikuwagus@email.unikom.ac.id, [2]heri.purwanto@usbypkp.ac.id, [3]deshinta.ad@newinti.edu.my

ABSTRACT

Heterogeneous data is a dataset with various types including data type and data source. Classification of heterogeneous data is still becoming a discussion in research in the field of intelligence artificial especially in learning classification. Based on data and machine development classification, then study this still relevant done. Machine classification that is still trending now is a hybrid engine known as the collaboration technique such as a fuzzy technique and neural network. The aim of this review paper is to find opportunity research on hybrid machine learning that perform classification on heterogeneous data with multi-class targets. There are several challenges on heterogeneous data such as 1)determining algorithm normalization and text processing as Step beginning from the input layer, 2) function formation variable linguistics for every case allow existence opportunity study for linguistic processes, 3) A membership function algorithm that can adapt of the dataset used can as opportunity research, 4) finding method shaper fuzzy rule as machine inference from a neural network, 5) Process structure of every task, 6) performance like efficiency memory for processing (management memory), complexity (process time), and validation architecture (accuracy, precision, recall, f-measure, specification, true prediction, false prediction). The result of the research obtained is the existence opportunity for improving or developing a hybrid classification machine that can handle heterogeneous data with multi-class targets.

Keywords: *Feature, Classification, Heterogenous Data, Fuzzy, Neural Network*

1. INTRODUCTION

Following the studies about heterogeneous data, many concepts about fuzzy are proposed to handle the data. The data if looking at the variance, show many distributions and structures. In the knowledge concepts, the data can separate into three kinds such as unstructured data, structured data, and hybrid data. If looking at the property of data, data consists of the field and the value of the field. The value of field data is the definition of the attribute to know the range and boundary of data. If the field or attribute consisted of various types of data, it can call as heterogenous or multimodal data.

Heterogenous data, if reviewed from progress until 2018 has been showing definition toward multi-source. Heterogenous this can be said as multi-type data or multi-source data. Heterogenous data is termed as variation data types and formats. Heterogenous data is often found in ambiguity, missing value, data redundancy, and not in accordance with fact. This is very difficult when want to integrate to Fulfill your desired goal. Miscellaneous heterogenous according to draft diversity data usage is as follows) heterogeneity by concept (conceptual heterogeneity) or more known term heterogeneity semantics or logical mismatch, log Thing difference in modeling but the same for the interest domain; 2) heterogeneity by terminology, heterogeneity occurs because variation naming when refers to a same entity but originated from different data sources ; 3) heterogeneity by semiotic (semiotic heterogeneity) or more known with heterogeneity pragmatic, this occur because interpretation something entity from someone [1].

The development of the classification process on heterogeneous data for algorithm classifier development Becomes toward machine classification intelligent hybrids. Machine classification intelligence is the machine that delivers appropriate dataset classification with its class label. Development of machine classification this already started since the existence of machines is helpful learning for analyzing a lot of data. Any assisted classification process with machine learning always involves datasets. Diverse machine learning is used for getting high performance on results dataset analysis. The data types in the dataset make source trigger Becomes found machine learning new. Development of machine classifiers like SVM, random forest, naive Bayes classifier, neural network multi-layer perceptron, decision tree, and machine other push analysis prediction against the dataset becomes more accurate with actual data. Machine classification intelligent hybrid started introduced in 1992 by PK Simpson with the use of merging technique fuzzy and neural network for complete problem classification with the technique fuzzy hyperbox. Hyperbox alone is a form-defined n-dimensional box with a score maximum and minimum for score function membership [2].

Using machine hybrid as a classification has two excellences. First, excellence is the use fuzzy rule-based system as an assistant for handling uncertainty and problems non-linear. Second, is the use of neural networks as part of the learning process of system rule fuzzy. Neural networks improve the fuzzy rule learning process from the data so that no need for knowledge external from experts during the learning process. The hybrid model, a combination of intelligence artificial and logic fuzzy, is a successful model for handling various of type problems [3].

The use of hybrid models has been tested on prediction increase price shares in Germany. Hybrid models can predict price share as well as predict the rise and fall of index stock on the day next. Hybrid models analyze chart development price sell buy stock. The results obtained are accurate with a number of determined factors like two, three, five, and ten factors. Hybrid models with the name hybrid neural fuzzy inference system (HyFis) have to get accuracy for whole indicator = 28%, two-factor indicator = 74.03 %, three factor = 76.24%, five factors = 73.48%, and 10 factors = 72.37% [4].

Machine evolving classification _ from the fuzzy set, which is based on from theory logic firm. The proposed fuzzy set Lotfi A Zadeh gives room to study control and linguistics [5]. Approach solving classification helped with the definition of linguistics from the problem studied. All machine classification that has been submitted ends up getting something result with high accuracy from the problem studied. Study how many the classification formed with the type-2 TSK fuzzy logic system. Classification involves _ results like a lot of attributes formed, number class, as well comparison technique classification with FCM, Subtractive, SVM, KNN, and Naive Bayes. Every amount classification formed, then as much that is also the fuzzy rule that is formed. Algorithm performance hybrid proposed learning gives excellent and stable performance with results in a lot of classes formed on datasets such as IRIS datasets with three classifications, wine dataset with three classifications, and breast cancer dataset with two classification type breast cancer. The accuracy obtained for the classification class and many attributes reached 88.33% [6].

Study about neural network classifier that can adapt classification text and compare with kind of classifier. This thing becomes very important in how to do exploitation and extraction of the data that has been collected. more again if data is in form text how to make classification text with topics certain but with ways efficient? Experimental results obtained _ with existing dataset simulation obtained 94% accuracy for the Reuters-21578 dataset and 95% for 20 newsgroup datasets [2].

Study machine classification for heterogenous data on a neural network (ANN) with mixing algorithm genetics so that is called heterogeneous neural network (HNN). The results obtained are calculated for every variable from observed documents to create a model of neurons that can accept heterogeneous input with make function similarity between the input and the weights. This model could process score mixture continuous, and discrete. HNN used is an algorithm developed for genetics. Results obtained classify missing values from each dataset used to experiment. The diabetic Pima dataset with amount attribute eight and target class is two class produce missing value by 10.6% [7].

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A study about hybrid data fusion with the use of Dempster–Shafer and ANFIS (DSANFI), obtained results accuracy of 91.50% compared fuzzy technique of 88.3% for problem classification content restricted information. Ascension score accuracy on _ every machine based on existing changes or modifications from the existing machine. Merger techniques fuzzy and neural networks, techniques Dempster-Shafer and ANFIS, and form others allow can get high accuracy [8].

Research that explains the use of ANFIS for analysis subjectivity. Destination research conducted is repair performance from algorithm hybrid for problem classification. Samir uses ANFIS for making variable linguistics. Linguistics variable from Fuzzy Becomes the input block as a score vector for processing on a Multilayer Artificial Neural Network (MANN). The result gives an accuracy of 92.16% for the supervised model use feature free in five languages in classification expression objective and subjective [8].

Study brain tumors to do tumor classification by utilizing machine hybrid structure descriptors and fuzzy logic based on radial basis function (RBF) with kernel support vector machine (SVM). The classification used _ like Meningiomas, Metastasis, Gliomas grade II, Gliomas grade III. Classification results with instances totaling 357 then Accuracy obtained were, class 1 (Meningioma) tumor type = 98.6%, class 2 (Metastasis) was 99.29%, class 3 (Gliomas grade II) was 97.87%, and class 4 (Gliomas grade III) was 98.6% [9].

Study big data and heterogenous data for accidents then traffic and mobility humans in Japan, who have successfully collected. Research this researched about how mobility man impact on accidents then cross. Study this conducted with the use of architecture deep for extracting features from mobility data from humans and carrying out the training process thoroughly to simulate risk accidents and then cross in scale large and real-time data. The result issued from the simulation is commemorated since early accidents then crosses with give advice for the choice of a safe path. Measurement results in this use MAE, MRE, and RMSE with score mean absolute error (MAE) = 0.96, mean relative error (MRE) = 0.39, and root-mean-square error (RMSE) = 1.00 [10].

Exposure to study classification using fuzzy models and neural networks until 2019 still have the opportunity for various uses. Opportunity-related development machines, process addiction with preprocessing, as well as applications still can do. Emphasis on aspects paper review this is found opportunity research on problems development algorithm classification with based on a hybrid model. This thing can be noticed because a number of the opportunity as ingredient research that can be delivered as follows: 1) determine algorithm normalization and text processing as Step beginning from the input layer, 2) function formation variable linguistics for every case allow existence opportunity study for linguistic processes, 3) Algorithm membership function that can adapt of the dataset used can as opportunity research, 4) Opportunity study for find method shaper rule fuzzy as machine inference from something neural network, 5) Process structure of every task performed, both on fuzzy and neural networks can Become opportunity in research, 6) performance like efficiency room for processing (management memory), complexity (time process), and validation architecture (accuracy, precision, recall, f-measure specification, true prediction, false prediction). With thereby results presentations and paper reviews carried out still allow did study in the machine area hybrid classification.

2. RESEARCH QUESTIONS

Many studies recorded the heterogeneous Frequent problems investigated in heterogeneous data areas are the unification of the value of the data. So that need there is a proper representation for could be processed by classification. Data representation for problem heterogeneous data can described in four stages. The stages are as follows:

- What different data types and formats of different sources?
- What is the representation with the term “unified representation”? This category of heterogenous data must be put together more formerly by converting individual attributes becomes information using the terminology “what-when-where”.
- How to reorganize the data such as shown with spatial grid on attributes thematic. Data like this must be aggregated to provide something data visualization and provide facility for to do queries.
- How heterogeneous data can be served by the whole by using the right operators.

3. RELATED RESEARCH

3.1 Heterogenous Data

Heterogenous data is another term for multi-type or multi-source data. The introduction to this term is to easily mention gathering data types for example real, integer, string text, and boolean in a dataset. Along with the development study here is an expansion definition of heterogeneous. start from only gathering real, integer, and string text and Boolean becomes gathering data values like knowledge from an expert about certain domains, textual descriptions, images, and sounds so that not only raw data.[2]

The use of various types of algorithms for heterogeneous data can see in Table I.

Table 1: Hybrid Algorithms On Heterogenous Data.

Author	Data	Method
[1]	Heterogenous data, no complete, complex, and dynamic	*Deep learning*
[3]	*Coronary artery disease (CAD)*	*k-NN, SVM*
[4]	*Data fusion, Heterogenous information, Information fusion*	*Dempster – Shafer and adaptive neuro-fuzzy inference*
[5]	Heterogenous Data	*Fuzzy – Rough set*
[6]	Hetergogenous Data, (ordinal, nominal, integer, continuous, and fuzzy)	*ANN* and *Genetic Algorithm*
[7]	Information Heterogenous. *Deep Relational Features, Heterogenous Dependencies*	*Convolution-Based Deep Relational Feature Earning*
[8]	*Medical Records*	*Deep Learning Feature Selection (ANN and SVM)*

Mention another for heterogenous data is data fusion. In the processing area image and system intelligence, data mining and data fusion are subject important. data fusion used by large for availability source and context from various areas. data fusion is combined of sensor data, image data, and information. Term fusion alone is a classification process for sensor data, image data, and information. With thereby data fusion usage identical with term heterogenous data that is carried out by the classification process [4] .

Study in field supporter decision about use heterogeneous data with various type data source. Data presented is follow the research, that heterogeneous data is defined as multi-type data [5]. Combining ordinal, nominal, continuous, and integer for helps in making supporter decisions by a group like Fig.1. Processing with the use of fuzzy can conduct get results with problem uncertainty [6] . Several things that can made problem is sourced from different data, unification data representation, data aggregation from spatial data into data something attributes thematic, determining proper learning _ for results unification and aggregation, obtaining conclusion learning about classification from heterogeneous the data. Solution classification on multi-document with representation vector of features, title feature, sentence length, proper noun, thematic word, and term weight use ANFIS with classification high or not high-quality summary. An illustration of heterogenous format can see in Fig.1. Illustration other ie in Fig.2 taken from the UCI dataset, describing examples of heterogeneous data.

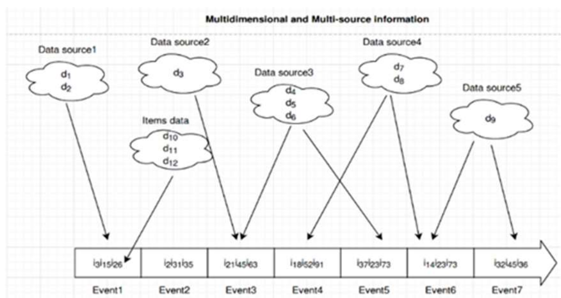

Figure 1: Illustration Heterogenous Data From Various Type Source that is Multi-Source and Multidimensional [10]

%X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15	class label
b	30.83	0	u	g	w	v	1.25	t	t	1	f	g	202	0	+
a	58.67	4.46	u	g	q	h	3.04	t	t	6	f	g	43	560	+
a	24.50	0.5	u	g	q	h	1.5	t	f	0	f	g	280	824	+
b	27.83	1.54	u	g	w	v	3.75	t	t	5	t	g	100	3	+
b	20.17	5.625	u	g	w	v	1.71	t	f	0	f	s	120	0	+
b	32.08	4	u	g	m	v	2.5	t	f	0	t	g	360	0	+
b	33.17	1.04	u	g	r	h	6.5	t	f	0	t	g	164	31285	+
a	22.92	11.585	u	g	cc	v	0.04	t	f	0	f	g	80	1349	+
b	54.42	0.5	y	p	k	h	3.96	t	f	0	f	g	180	314	+
b	42.50	4.915	y	p	w	v	3.165	t	f	0	t	g	52	1442	+
b	22.08	0.83	u	g	c	h	2.165	f	f	0	t	g	128	0	+
b	29.92	1.835	u	g	c	h	4.335	t	f	0	f	g	260	200	+
a	38.25	6	u	g	k	v	1	t	f	0	t	g	0	0	+

Figure 2: Illustration heterogenous dataset home-credit

3.2 Machine Learning for Data Classification

In 1993 it was proposed method hybrid this is known with name ANFIS (adaptive-network-based fuzzy inference system). ANFIS developed with use merging method fuzzy and method neural network [11]. Study about utilization machine hybrid (fuzzy and neural network) continues until repair or utilization method hybrid for problems other especially classification. Use method fuzzy min-max neural network for classification and clustering introduced by Bogdan Gabriels. Proposed method _ for solution irrigation system distribution that is hybrid clustering classification. Utilization of this method is a cluster and a classification without training data from early. So that result could have level high accuracy[12]. Model fixes for problem classification with method hybrid is repair neuromodel structure fuzzy with consider simplification rule fuzzy and constraint enhancement speed rule with inputs. Study repair performance from research conducted by PK Simpson [13] . Terminology classification is how classify data to class already determined. Formation class alone starts from the data already labeled. Various study the related classification with method hybrid this already can solved with good accuracy. Various research on object study like text summarization [9] , feature selection [14] - [15] , and electric production [16] leads to understanding about classification.

The development method proposed hybrid about analysis subjectivity that is repair structure hybrid Becomes combined fuzzy, ANFIS, and hidden Markov model. Repairs made is in trouble extraction feature with repair accuracy generated in research in 2013 [17] Explanation about architectural models neuro growing fuzzy. Model of Cooperative Neuro-Fuzzy Systems in fig.3, the technique used with customize and match technique inference fuzzy with use algorithm learning from neural network. Several methods performed algorithm namely: a) cooperative neuro-fuzzy systems for adjusting the membership function. Type this is the output (result) of processing Neural Network connected with system fuzzy, process on neural network is train data for get amount fuzzy value; b) cooperative neuro-fuzzy systems for extracting fuzzy rules from training data. Type this do the process the first time that is with learn defined rules, next _ is implement rule on the fuzzy system. Information previously given to fuzzy system to be processed; c) cooperative neuro-fuzzy systems for regularly updating the fuzzy system structure. Type this perform the matching process against the specified parameters that is with apply fuzzy system by straight away. Bait come back results from system fuzzy really needed on the structure this; d) cooperative neuro-fuzzy systems for identifying the weighted fuzzy rules. On processing type fourth this is method learning with use matching important rule _ in system fuzzy. Every generated rule system fuzzy given weight in accordance level interest rules.

Neural Networks Driven Fuzzy Reasoning on Fig. 4, architecture from method this divided on three part namely : divide inference rules , identify part from if (determination function membership), and identification part then (determination amount control from every rules) [18] .

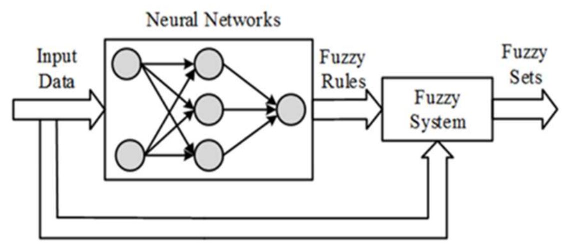

Figure 3: Architecture General Cooperative Neuro-Fuzzy System [18]

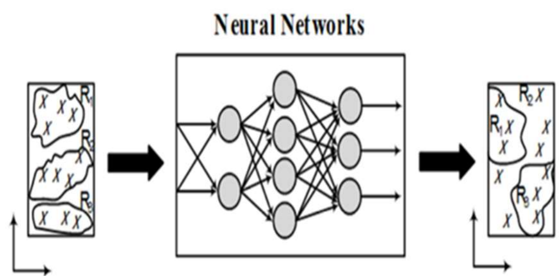

Figure 4: Architecture Neural Networks – Driven Fuzzy Reasoning [18]

Hybrid Neural Networks Based Systems on Fig. 5, Method this implement system hybrid - nya with form structure parallel. System fuzzy and neural network work by synchronous. Synchronization process can be seen from something results parallel from existing fuzzy system _ in structure NN. System Fuzzy can be a sub block of Layers that exist in neural network. Method this developed return Becomes two methods namely: Adaptive-Network-Based and Fuzzy Inference System (ANFIS).

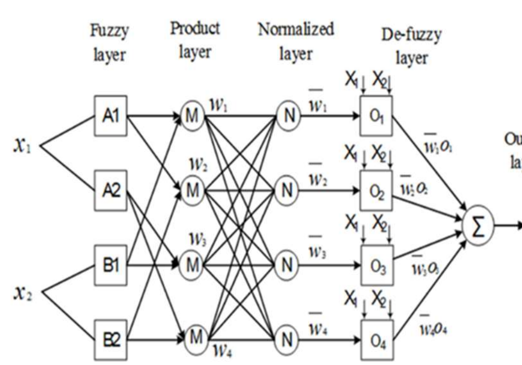

Figure 5: ANFIS Model with Five Layers [18]

4. SELECTED STUDY

4.1 Multilayer Adaptive Neural Network (MANN)

Research conducted on ANFIS combined with meta-heuristic techniques (Bee Algorithm) provides the dominant technique aimed at difficult problems in business areas. ANFIS (Adaptive Neuro-Fuzzy Inference System) is an efficient combination for research problems that are non-linear , complex, and dynamic systems . ANFIS The system in fig.6 can estimate the production of each factory with a number of defined rules. In ANFIS the number of rules defined and parameters matched results in the process increasing exponentially when there are too many input nodes. Referring to the case that occurred, the ANFIS modeling combined with the bee algorithm was cut to search for a classification match. The results shown in the study with a combination of ANFIS and the bee algorithm showed a smaller error than the ANFIS . model [19] .

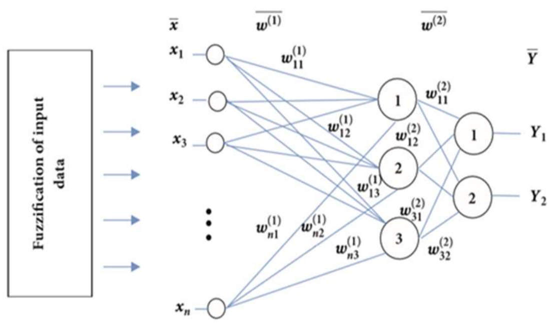

Figure 6: MANN Structure in ANFIS [20]

4.2 Multiclass Adaptive Neuro-Fuzzy Classifier (MC-NFC)

The fuzzy system has the ability to handle uncertain and imprecise information, but cannot update and fine-tune fuzzy parameters automatically. Research on image classification conducted by [21], proves that there is an increase in accuracy in image interpretation . This study uses the Sugeno . model Fuzzy Inference System (SFIS) where the consequence of the given result is constant. This SFIS is known as the "zero-order Sugeno type". A simple architecture depicting a neuro-fuzzy procedure in Fig. 7. The improvement offered in this research is the provision of new features for the input classifier . This study proposes a method, starting by using many features, and then reducing the number by using a variable dimension reduction technique. This alternative solution saves a lot of time for classifier development. The proposed method uses multiple error patterns, but can detect all possible real patterns for the error in question. Development of Multiclass Adaptive Neuro-Fuzzy Classifier (MC-NFC) with discrimination between five different errors in PVA . The Neuro-Fuzzy Architecture used in the MC-NFC research can be seen in Fig. 7 [21] .

Hybrid Structure Descriptor and Fuzzy Logic based SVM, details information with use property like level gray, texture, and color with use hybrid technique that is Hybrid structure descriptor and Fuzzy logic, and RBF kernel SVM explain that classification for picture can conducted for get details information. Study this offer method that can clarify that. Method the explained with Suite step as follows. The first process is the segmentation process. Segmentation process that is to do feature extraction from every region with hybrid structure descriptors. Stages end of this process is classification picture Becomes four class difference between tumor and training done by Fuzzy kernel based support vector machine [22] .

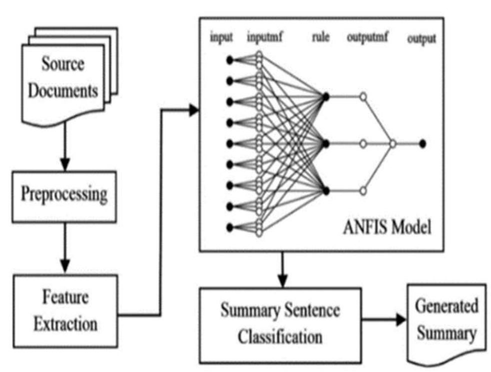

Figure 7: Multiclass Adaptive Neuro-Fuzzy Classifier (MC-NFC) architecture for image classification [21]

4.3 Self-Constructing Neural Fuzzy Inference Network (SONFIN)

Hybrid-Neuro-Fuzzy System and Adaboost -Classifier on Fig. 8, development technique hybrid classification has trigger existence repair in Thing class. Joseph explained algorithm adaboost Classifiers. adaboost classifier is a classifier combined from many weak classifiers (such as linear classifier). Every classifier works classification only on vector one dimensions for inputs. That thing known with weak classifiers. Structure this named self-constructing neural fuzzy inference network (SONFIN) with divided on six layers. Structure this made alone without follow structure fuzzy and neural existing networks. While the dataset used in

15th September 2023. Vol.101. No 17

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

the research this is IRIS dataset, WISCONSIN dataset, and CSMU dataset [23] .

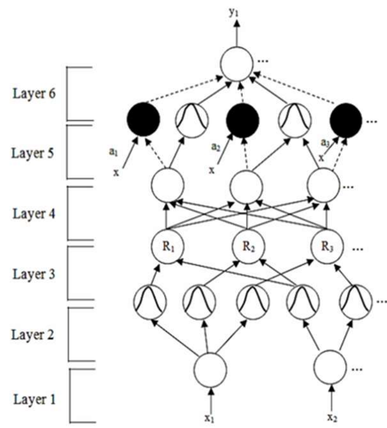

Figure 8: Hybrid-neuro-fuzzy system architecture and Adaboost - classifier SONFIN classifier [23].

4.4 Algorithm Axiomatic Fuzzy Set NN (AFSNN)

AFS can provide a more rational mechanism for generating membership functions. Therefore, it was inspired by the weaknesses and strengths of the previous method related to the method based on the theory of axiomatic fuzzy set (AFS) , neural network random weight (NNRW), and neuro-fuzzy . AFSNN Algorithm on Fig . 8 which integrates AFS theory with NNRW to reduce the complexity of the previous classification based AFS in selecting complex concepts and providing classification results with several improved interpretations. Improvements include the following aspects 1) A simple concept coherence membership function is inserted into a hidden node, and a hidden layer that maps the original data into a feature space that describes data with complex concepts in AFS theory. 2) The selection between attributes, simple concepts, and complex concepts is determined randomly at the beginning , assigning random input weights in NNRW . This practice significantly reduces the complexity of selecting complex concepts to describe the sample. 3) With the help of Moore-Penrose inverse generalization computation, it is possible to directly evaluate the feasibility of complex concepts for each class with output weights instead of optimizing the membership level of these complex concepts. 4) Benefiting from the characteristics of AFS theory in knowledge discovery, NNRW overcomes the black box phenomenon and makes classification results that can be interpreted logically semantics [24] .

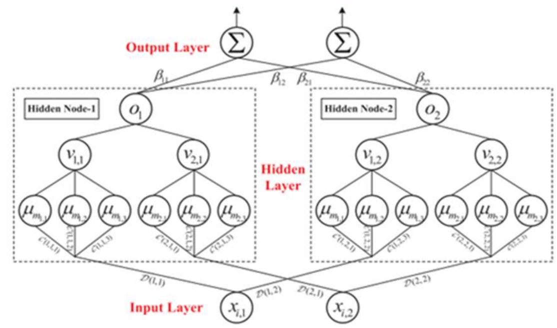

Figure 9: Structure of Axiomatic Fuzzy set NN (AFSNN) [24]

Research conducted on this AFSNN, using Nemenyi test. Nemenyi test describe fuzzy more conducive for find knowledge that has implications for the original dataset. The two parameters and L are used as consideration in classification. On experiment L value is more sensitive. So that probability on the AFSNN model is improvement in the determination of L. L is the size of room generation relation Among attribute , while is _ function control from score range wide from weight [24] . In Fig.10, Fig.11, and in Fig.12 it is algorithm AFSNN.

4.5 Fuzzy Min-Max (FMM) Neural Network

Method fuzzy min max (FMM) neural network for classification, FMM neural network classification, create definition class with use approach supervised learning. Meanwhile, FMM itself for processing membership function which is degrees feature from hyperbox. Function this for find proximity hyperbox area with investigated instance. _ Steps offered FMM for classification are: 1) Expansion Criteria, namely the expansion process from the given input for determine class and grade maximum from membership functions. 2) Overlap Test, process for determine the overlap area between two classes with different hyperbox classes. Overlapping areas with class similar to hyperbox could allowed. 3) Contraction Test, process for telling you that limit different class on everlap process hyperbox no could defined with right. 4) Adaptive Expansion Algorithm, Algorithm this carry method new for expansion of hyperbox area on FMM neural network for pattern classification. Expansion made _ is with added pattern input on the same class in hyperbox. identifier for expansion use variable AREA_HB with 0 AREA_HB with 1. AREA_HB is defined as maximum area coverage from something hyperbox like Fig.13. Study this produce classification against the iris dataset. Approach method this is use score created limit in form hyperbox [25] .

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Require: The training dataset, X_{train}; the parameter for controlling weight function, σ; the number of hidden neurons, L; the number of simple concepts on every feature, N_{sc}.
Ensure: The output weight, β.
1: Randomly generate a two-dimensional matrix $\mathcal{D}$ and a three-dimensional matrix $\mathcal{C}$;
2: **for** each feature $f_j \in F$ **do**
3: Find the minimum and maximum value, $f_j^{\min}$ and $f_j^{\max}$;
4: Calculate $d_{\text{interval}} = \frac{f_j^{\max} - f_j^{\min}}{N_{sc} - 1}$;
5: Determine the parameters $P_j = \{f_j^{\min} + (k-1)\, d_{\text{interval}} | k = 1, 2, \ldots, N_{\text{sc}}\}$;
6: Construct weight function with parameter set P_j and predetermined parameter σ;
7: Define the corresponding set of simple concept $M_j = \{m_{j,1}, m_{j,2}, \ldots, m_{j,N_{sc}}\}$;
8: **for** each sample $x_i \in X_{\text{train}}$ **do**
9: Compute the weights $\rho_{m_{j,k}}(x_{ij}), k \in [1, N_{\text{sc}}]$;
10: **end for**
11: **end for**
12: **for** $x_p, x_q \in X_{\text{train}}$ **do**
13: Construct linearly ordered relation $\tau(x_p, x_q)$ by Definition 1.
14: **for** $m \in M$ **do**
15: Get the AFS structure $A^{\tau}(x_p)$, $A = \{m\}$.
16: **end for**
17: **end for**
18: Build hidden layer R with L nodes.
19: **for** $x_i \in X_{\text{train}}$ **do**
20: **for** $R_l \in R$ **do**
21: **for** each $f_j \in F$ **do**
22: Calculate the membership degree of $\mu_{\xi_j}(x_{ij})$, $\xi_j = m_{j1} + m_{j2} + \cdots + m_{jN_{sc}}$ with the use of (12), (15), $\mathcal{D}$ and $\mathcal{C}$.
23: **end for**
24: Obtain the firing strength o_{il} by (16).
25: **end for**
26: **end for**
27: Combine o_{il} as input matrix H with $|X_{\text{train}}|$ rows and $|R|$ columns.
28: Create the target matrix T according to the class labels of samples in X_{train}.
29: Solve the output weights by $\beta = (H^T H)^{-1} H^T$
30: **return** β.

Figure 10: AFSNN Algorithm [24]

Algorithm 2: Forming a fuzzy description for each class.
Require: The training dataset, X_{train}; the set of class labels, C.
Ensure: The fuzzy description set, ξ;
1: Combine all complex concepts as a set ψ.
2: **for** each complex concept $\psi_i \in \psi$ **do**
3: determine the class label of ψ_i by (17).
4: **end for**
5: Divide ψ into $\psi_{C_1}, \psi_{C_2}, \ldots, \psi_{C_{|C|}}$ according to obtained class label in Step 3.
6: **for** each $\psi_{C_i} \subseteq \psi$ **do**
7: Get the overall optimized complex concept by (19).
8: Obtain the local optimized complex concept by (20).
9: Generate the fuzzy description ξ_{C_i} by (18)
10: **end for**
11: **return** ξ.

Figure 11: Fuzzy Algorithm for describe every class [24]

Algorithm 3: Classifying sample with AFSNN classifier.
Require: The testing dataset, X_{test}.
Ensure: The class label set, C_{test};
1: **for** each sample $x_i \in X_{\text{test}}$ **do**
2: Embed x_i into X_{train} and generate a new dataset X_i.
3: Execute Step 2–Step 2 (Algorithm 1) on X_i to built weight function and calculate the weights.
4: Carry out Step 12–Step 17 (Algorithm 1) on X_i to construct AFS structure which is applied in the AFS membership function.
5: Run Step 20–Step 25 (Algorithm 1) to obtain the firing strength o_{il}.
6: Assemble o_{il} as an output weight vector of hidden layer, H'.
7: Solve the weight of x_i belonging to every class, T'.
8: Determine the class label of x_i, $C_{\text{test}}^i = \arg\max_{1 \leqslant k \leqslant N_C} T_k{}'$
9: **end for**
10: **return** C_{test}^i.

Figure 12: Algorithm Classification sample with AFSNN Classifier

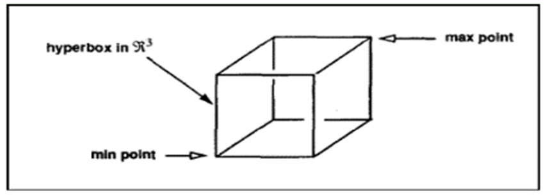

Figure 13: Hyperbox format [25]

Expansion models for FMM this is when new input pattern entered, with specified class, then the class must enter in hyperbox which has class and grade the same membership. Expansion this no could exceed size hyperbox. Rule expansion this shown in formula $\sum_{j=1}^{n} \max(w_{ji}, a_{hi}) - \min(v_{ji}, a_{hi}) \leq n\theta$, and changes the min and max values with $v_{ji}^{new} = \min(v_{ji}^{old}, a_{hj})$and $w = \min(w_{ji}^{old}, a_{hj})$. Advantages FMM created on this research compared FMM first andrepair problem accuracy for reduce amount error at time classification. Then reduce complex structure at minimum value in hyperbox area with sample minimum training. The result of study allow existence repair for problem minimum coverage on hyperbox. Adapt this model taken from study [12] , then conducted manipulation algorithm by [25] . The deficiency of this model is a process that is carried out not yet use text area problem classification, so that proof accuracy that reaches 100% still doubtful in Thing result. As comparison use method hybrid, some a hybrid method on the problem heterogenous can show in table II.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Require: The training dataset, X_{train}; the parameter for controlling weight function, σ; the number of hidden neurons, L; the number of simple concepts on every feature, N_{sc}.
Ensure: The output weight, β.
1: Randomly generate a two-dimensional matrix $\mathcal{D}$ and a three-dimensional matrix $\mathcal{C}$;
2: **for** each feature $f_j \in F$ **do**
3: Find the minimum and maximum value, $f_j^{\min}$ and $f_j^{\max}$;
4: Calculate $d_{\text{interval}} = \frac{f_j^{\max} - f_j^{\min}}{N_{sc}-1}$;
5: Determine the parameters $P_j = \{f_j^{\min} + (k-1) d_{\text{interval}} | k = 1, 2, \ldots, N_{\text{sc}}\}$;
6: Construct weight function with parameter set P_j and predetermined parameter σ;
7: Define the corresponding set of simple concept $M_j = \{m_{j,1}, m_{j,2}, \ldots, m_{j,N_{sc}}\}$;
8: **for** each sample $x_i \in X_{\text{train}}$ **do**
9: Compute the weights $\rho_{m_{j,k}}(x_{ij}), k \in [1, N_{\text{sc}}]$;
10: **end for**
11: **end for**
12: **for** $x_p, x_q \in X_{\text{train}}$ **do**
13: Construct linearly ordered relation $\tau(x_p, x_q)$ by Definition 1.
14: **for** $m \in M$ **do**
15: Get the AFS structure $A^{\tau}(x_p), A = \{m\}$.
16: **end for**
17: **end for**
18: Build hidden layer R with L nodes.
19: **for** $x_i \in X_{\text{train}}$ **do**
20: **for** $R_l \in R$ **do**
21: **for** each $f_j \in F$ **do**
22: Calculate the membership degree of $\mu_{\xi_j}(x_{ij})$, $\xi_j = m_{j1} + m_{j2} + \cdots + m_{jN_{sc}}$ with the use of (12), (15), $\mathcal{D}$ and $\mathcal{C}$.
23: **end for**
24: Obtain the firing strength o_{il} by (16).
25: **end for**
26: **end for**
27: Combine o_{il} as input matrix H with $|X_{\text{train}}|$ rows and $|R|$ columns.
28: Create the target matrix T according to the class labels of samples in X_{train}.
29: Solve the output weights by $\beta = (H^T H)^{-1} H^T$
30: **return** β.

Figure 10: AFSNN Algorithm [24]

Algorithm 2: Forming a fuzzy description for each class.
Require: The training dataset, X_{train}; the set of class labels, C.
Ensure: The fuzzy description set, ξ;
1: Combine all complex concepts as a set ψ.
2: **for** each complex concept $\psi_i \in \psi$ **do**
3: determine the class label of ψ_i by (17).
4: **end for**
5: Divide ψ into $\psi_{C_1}, \psi_{C_2}, \ldots, \psi_{C_{|C|}}$ according to obtained class label in Step 3.
6: **for** each $\psi_{C_i} \subseteq \psi$ **do**
7: Get the overall optimized complex concept by (19).
8: Obtain the local optimized complex concept by (20).
9: Generate the fuzzy description ξ_{C_i} by (18)
10: **end for**
11: **return** ξ.

Figure 11: Fuzzy Algorithm for describe every class [24]

Algorithm 3: Classifying sample with AFSNN classifier.
Require: The testing dataset, X_{test}.
Ensure: The class label set, C_{test};
1: **for** each sample $x_i \in X_{\text{test}}$ **do**
2: Embed x_i into X_{train} and generate a new dataset X_i.
3: Execute Step 2–Step 2 (Algorithm 1) on X_i to built weight function and calculate the weights.
4: Carry out Step 12–Step 17 (Algorithm 1) on X_i to construct AFS structure which is applied in the AFS membership function.
5: Run Step 20–Step 25 (Algorithm 1) to obtain the firing strength o_{il}.
6: Assemble o_{il} as an output weight vector of hidden layer, H'.
7: Solve the weight of x_i belonging to every class, T'.
8: Determine the class label of x_i, $C_{\text{test}}^i = \arg\max_{1 \leqslant k \leqslant N_C} T_k{}'$
9: **end for**
10: **return** C_{test}^i.

Figure 12: Algorithm Classification sample with AFSNN Classifier

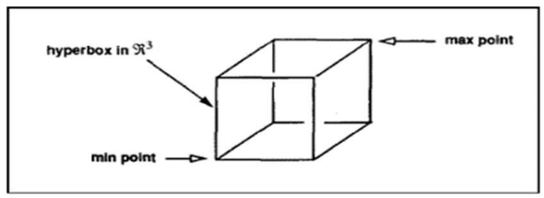

Figure 13: Hyperbox Format [25]

Expansion models for FMM this is when new input pattern entered, with specified class, then the class must enter in hyperbox which has class and grade the same membership. Expansion this no could exceed size hyperbox. Rule expansion this shown in formula $\sum_{j=1}^{n} \max(w_{ji}, a_{hi}) - \min(v_{ji}, a_{hi}) \leq n\theta$, and changes the min and max values with $v_{ji}^{new} = \min(v_{ji}^{old}, a_{hj})$and $w = \min(w_{ji}^{old}, a_{hj})$. Advantages FMM created on this research compared FMM first andrepair problem accuracy for reduce amount error at time classification. Then reduce complex structure at minimum value in hyperbox area with sample minimum training. The result of study allow existence repair for problem minimum coverage on hyperbox. Adapt this model taken from study [12] , then conducted manipulation algorithm by [25] . The deficiency of this model is a process that is carried out not yet use text area problem classification, so that proof accuracy that reaches 100% still doubtful in Thing result. As comparison use method hybrid, some a hybrid method on the problem heterogenous can show in table II. Table II is placed by the end of the paper.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The algorithm for making conversion score assertive to score fuzzy. Stages this done after receive input in the form of features that have been found at the input layer stage. In stages this be delivered is method for defining specified membership function from analysis features at the input layer stage. Stages this also finds linguistic variables for made score membership function on input layer and fuzzy layer [18] [37].

Function model Gaussian for finding score memberships. The proposed equation is as following:

$$\mu_{ij} = \begin{cases} 0 & if\ \sigma_{ij} = 0\ and\ x_i \neq \bar{x}_{ij} \\ e^{-\left(\frac{1}{2}\frac{(x_i-\bar{x}_{ij})^2}{\sigma_{ij}^2}\right)}, & if\ \sigma_{ij} \neq 0 \\ 1 & if\ \sigma_{ij} = 0\ and\ x_i = \bar{x}_{ij} \end{cases} \quad (1)$$

Index i state many formed feature and j is membership variable linguistics from feature input. The statement x_i is the input feature , $\bar{x}_{ij}$is the average, and $\sigma_{ij}^2$is the standard deviation of the cluster j on the feature i [37]. Proposed function _ for determine membership function can also use curve bell , with equality as following [18] :

$$\mu_{A_i^m}(x_i) = \frac{1}{1+\left[\left(\frac{x_i-c_p}{a_p}\right)^2\right]^{b_p}} \quad (2)$$

$A_i^m(x_i)$ is the linguistic label on the function node, x is the node for the input layer.

Identify the hidden layer model for product layer and normalization Layers. At stage part product layer which is the place keep fuzzy rules. In stages this for determine rule can use function proposed by [15] . That every feature input supposes where A = {x1, x2, …, x3} and a set of class C = {c1, c2, c3, …, cm} which is the label of feature, then can served in something vector feature values. Let $x = x_1\, x x_2\, x x_3\, x\, \ldots\, x\, x_n$ so can say as something room multidimensional from feature Formation fuzzy rules can follow every existing value _ the feature value vector containing can value 0 or 1. So that can write as following:

$$R_j: \text{IF}\ s_1 \wedge x_1 = A_{1j}\ \text{AND}\ s_2 \wedge x_2 = A_{2j}\ \text{AND} \ldots \text{AND}\ s_n \wedge x_n = A_{nj}\ \text{THEN class} = j = 1, \ldots, R,$$

Rule this can recognized a lot with the letter R, and j is amount fuzzy rules that are formed. A_{ij} is the resulting fuzzy term from value of $s_1 \wedge x_1$. Value in s_k can state existence feature or no with value 0 or 1 [15] .

In part layer normalization, is count weights on each normalized node. Formulation this can written as following:

$$w_m = \mu_{A_i^m}(x_1) \otimes \mu_{A_i^m}(x_2),$$
$$i = 1,2;\ j = 1,2;\ m = 1,\ldots,4. \quad (3)$$

$$\bar{w}_m = \frac{w_m}{\sum_{p=1}^{4} w_m} \quad (4)$$

Wm is fire rules _ with the input features presented, while $\overline{w_m}$ is the average value of the weights obtained from Wm.

Determine *defuzzification* from results inference *fuzzy rule,* then next determine which rules to be _ suitable from results calculation *feature vectors.* After getting score from *nodes* before, then the process continues for find algorithm proper *defuzzification.* Result from the *defuzzification* process this so will form every instance be in *class* which *label.*

Determine the evaluator, each machine *learning* that has been arranged, then must made method evaluation for see how much accurate machine proposed *learning* _ for problem *heterogeneous* data. In a paper that has done , at [3][8][10][17][44] not yet showing how method for evaluate machine *learning* with *fuzzy* and *NN* . The part will be proposal alone to get a suitable evaluator with machine *learning* used.

5. RESEARCH OPPORTUNITIES

Research that has been performed by various researcher has given description about classification with use various method hybrid [7][10][22][25]–[35] . Opportunity for get study renewable can reviewed from research that has been done. When judging by back to research before , then a number of possibility study update on issues classification for heterogeneous data still can conducted [1] . Merger method fuzzy and neural network give description still existence opportunity to purposeful research gets more hybrid models accurate.

In table III it is shown step every algorithm on research about the PDA process heterogeneous data. There is a process similarity in formation structure in layer shape. Displayed column is a process that is carried out with collaboration Among fuzzy and neural networks. Several studies have seen how to get a solution in his research. As the essence of table III, that if the data used already in feature vector so no preprocessing is done. But on the contrary if the dataset contains gathering text so will preprocessing for _ get something feature vector. The process will work directly using techniques from each stage of fuzzy and neural network. Referring to every research, novelty offered repair or add something algorithm new to use complete flaws in existing algorithms there is.

Determination object scoring with heterogeneous data using neural network and fuzzy, has get score small error l, for _ MLP the error only 3.6% in class 1, 0.4% for class 2, and 0.4% for class general. This thing already proves that very reliable neuro-fuzzy for complete problem classification.

Opportunity research that can be taken from research that has been conducted that is about what just must do task done until get suitable result with destination research. A number of the opportunity for according to table III for get novelty is as following: 1) determine algorithm normalization and text processing as Step beginning from the input layer [5], 2) function formation variable linguistics for every case allow existence opportunity study for linguistic processes [15] ; 3) Algorithm membership function that can adapt of the dataset used can as opportunity study [29][37] ; 4) Opportunity study for find method shaper rule fuzzy as meisn inference from something neural network [30] [27] ; 5) Process structure of every task performed , both on fuzzy and neural networks can Becomes opportunity in study [12] [18] [24][39]–[41] ; 6) ; performance like efficiency room for processing management memory), complexity (process time), and validation architecture (accuracy, precision, recall, f-measure, specification, True prediction, False Prediction) [47] .

Things that already disclosed in opportunity research that can done, then for focus on manufacture machine classification with method hybrid, several work that can conducted for find opportunity study special talk about hybrid engine. First, a rs i texture NN and Fuzzy, fuzzy process Becomes one unity with NN. Every node that is formed will represent a fuzzy process. NN they will share above: a) first layer is the input node consisting of from matrix vector results feature extraction documents, images, and signals; b) second layer is operationalization linguistics k variable. Worked on this node only there is three nodes namely: no linguistic for linguistics variable e-documents, node linguistics for variable linguistics images, node linguistics for linguistics variable signal; c) third layer is membership function with score linguistics from every variable linguistics. Every score linguistics will translated Becomes number of nodes; d) fourth layer is every node will Becomes machine inference containing rule base. Node this will receive input from the third layer. Then will conduct inference for get appropriate rule base; e) fifth layer is the defuzzification process, which accepts input from the 4th layer. The defuzzification process conducted with proper technique is center of gravity y (COG) or technique other; f) last process is for classification with function activation specified. Second, an architecture NN and Fuzzy conducted by sequential. How to do it is as following: a) process NN conducted more formerly until get output, Fuzzy process done after get output value of NN, as the activation process function from NN, b) own fuzzy process follow algorithm of the fuzzy process, which can be written as following: Determine score form that obtain from calculation NN, determine linguistics variable for handle score assertive as input value, specify degree membership function, with linguistics value already obtained from every linguistics variable, creates rule if-then, from results linguistics existing values for look for connectedness, do inference to firing rules with if-then rule , get use technique Sugeno, Mamdani, Takagi Sugeno Kang (TSK) , or other , get results inference, get score assertive return with defuzzification, c) classification obtained from results defuzzification.

Table 3 by the end of the paper.

The combination of architecture fuzzy and neural network has produced significant accuracy to problem faced. However, there is still results accuracy yet significant, if involve processing text. Thereby the algorithm already discovered by various studies still can repaired or reconstructed so that could produce high accuracy. When propped up with research by Lidong Wang [1] , then combination techniques and completion targets heterogenous still allow conducted research . This thing is based on

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

results studies that include heterogeneous data for _ problem classification. more again on paper [1] , already give opportunity study in field heterogeneous related data with heterogenous syntax, heterogenous semantic, heterogeneous terminology, and heterogenous semiotics.

Confirm return from opportunity research that has been collected, with thereby can concluded steps processing for something machine hybrid classification. Steps this arranged based on every solution the problem you want solved and can Becomes novelty in hybrid method. Every defined machine so tasks in study for machine classification , among others: a) trial test hybrid, pair technique Fuzzy and NN ; b) m making variable linguistics from matrix obtained vector; c) determination membership function , many chart function like trapezoid , triangle , bell , and S curves ; d) determine weight for each NN node; e) determine function for each node on each NN layer; f) determine function summarization for output node on NN; g) determine function Activation for output nodes; h) determine class to be generated by the hybrid model Fuzzy and NN. Besides talks about hybrid engine, then need existence i identify the dataset used related heterogenous data, identification imbalance class and method completion, identification solution for normalization of data and features to be used as variable linguistics on fuzzification, determination matching graphics for determination membership function degree, identify amount neurons with limitation feature selection for hidden layer , identify defuzzification for the value of the data obtained , make algorithm for function activation , and make plot evaluation for model accuracy.

On identification function data normalization is to do normalization of existing data in the dataset. Normalization this conducted to data value in the form of text. So that every dataset is already in form multidimensional. this step for find feature choices presented _ in form data dimensions [9][10][22][42] . Destination from feature extraction and feature selection is for get really feature _ relevant and not redundant. From other to get data normalization via feature extraction and feature selection are use algorithm genetics , hill climbing , and simulated annealing [1] .

6. CONCLUSION

Based on exposure about possibility opportunity research, then can concluded that study for classification with method hybrid still open. Things that have been served about opportunity study has shown still existence gap that is collaboration machine with various algorithm processing text still not yet defined with right. So, the hybrid process used sometimes does not involve processing text and the NN algorithm. With thereby opportunity study for this hybrid method can give opportunity with : 1) determine algorithm normalization and text processing as Step beginning from the input layer, 2) function formation variable linguistics for every case allow existence opportunity study for the linguistic process, 2) Algorithm membership function that can adapt of the dataset used can as opportunity research , 3) Opportunity study for find method shaper rule fuzzy as machine inference from something neural network, 4) Process structure of every task performed , both on fuzzy and neural networks can Becomes opportunity in research , 5) ; performance _ like efficiency room for processing (management memory), complexity (process time), and validation architecture (accuracy , precision , recall, f-measure, specification, True prediction, False Prediction).

Linkages with machine the classification to be defined, then still can determine: **first,** NN and Fuzzy architecture, the fuzzy process becomes one unity with NN. Every node that is formed will represent a fuzzy process. the NN will shared above: a) first layer is the input node; b) second layer is operationalization of linguistic variable; c) third layer is a membership function with linguistic value of every linguistic variable; d) fourth layer is each node will Becomes machine inference containing the rule base; e) fifth layer is the defuzzification process, which receives input from the 4th layer; f) last process is to do classification with function specified activation as well. **Second,** NN and Fuzzy architecture are carried out sequentially. How to do it is as the following: a) the NN process is carried out more formerly until get the output, fuzzy process is done after getting the output value of NN, as the activation process function from NN, b) fuzzy process itself follow algorithm of the fuzzy process.

REFERENCES:

[1] L. Wang, “Heterogeneous Data and Big Data Analytics”, Vol. 3, No. 1, 2017, pp. 8-15.

[2] C. Borgelt, J. Gebhardt, and R. Kruse, “Chapter F1.2: Inference Methods”, *Handbook of Fuzzy Computation n. Institute of Physics Publishing Ltd.*, 1998.

15th September 2023. Vol.101. No 17

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[3] Y. Motai, NA Siddique, and H. Yoshida, “Heterogenous data analysis: Online learning for medical-image-based diagnosis”, *Pattern Recognit.*, Vol. 63, No. May 2016, 2017, pp. 612-624.

[4] E. Brumancia, S. Justin Samuel, LM Gladence, and K. Rathan, “Hybrid data fusion model for restricted information using Dempster-Shafer and adaptive neuro-fuzzy inference (DSANFI) system”, *Soft Comput.*, Vol. 1, 2019.

[5] B. Sun, W. Ma, X. Chen, and X. Li, “Computers & Industrial Engineering Heterogeneous multigranulation fuzzy rough set-based multiple attribute group decision making with heterogeneous preference information”, *Comput. Eng. Eng.*, Vol. 122, No. September 2017, 2018, pp. 24-38.

[6] B. Mu, S. Inform, U. Polit, and C. Barcelona, “Modeling Heterogenous Data Sets with Neural”, *Innov. Hybrid Intel. Sys.*, Vol. ASC 44, 2007, pp. 96-103.

[7] Y. Zhang, J. Mi, Y. Xiong, X. Kong, S. Li, and Y. Zhu, “Deep Collective Classification in Heterogeneous Information Networks”, 2018, pp. 399-408.

[8] Deep Collective Classification in Heterogeneous Information Networks, “Learning Deep Representations from Heterogeneous Patient Data for Predictive Diagnosis”, 2017, pp. 115-213.

[9] M. Azhari and Y. Jaya Kumar, “Improving text summarization using neuro-fuzzy approach”, *J. Inf. Telecommun.*, Vol. 1, No. 4, 2017, pp. 367-379.

[10] JB Daher et al., “A Review on Heterogeneous , Multi-source and Multi-dimensional data mining To cite this version : HAL Id : p-01811232 A Review on Heterogeneous, Multi-source and Multi-dimensional data mining”, 2018.

[11] N. Walia, “ANFIS: Adaptive Neuro-Fuzzy Inference System- A Survey”, Vol. 123, No. 13, 2015, pp. 32-38.

[12] B. Gabrys and A. Bargiela, “General fuzzy min-max neural network for clustering and classification”, *IEEE Trans. Neural Networks*, Vol. 11, No. 3, 2000, pp. 769-783.

[13] R. Hata, MAH Akhand, MM Islam, and K. Murase, “Simplified Real-, Complex-, and Quaternion-Valued Neuro-Fuzzy Learning Algorithms”, *Int. J. Intell. syst. app*, Vol. 10, No. 5, 2018, pp. 1-13.

[14] S. Zobeidi, M. Naderan, and SE Alavi, “Effective text classification using multi-level fuzzy neural network”, *in 5th Iranian Joint Congress on Fuzzy and Intelligent Systems - 16th Conference on Fuzzy Systems and 14th Conference on Intelligent Systems , CFIS 2017*, 2017, pp. 91-96.

[15] M. Bardamova et al., “A Fuzzy Classifier with Feature Selection Based on the Gravitational Search Algorithm”, *Symmetry (Basel).*, Vol. 10, No. 11, 2018, pp. 609.

[16] M. Dhimish, V. Holmes, B. Mehrdadi, M. Dales, and P. Mather, “Photovoltaic fault detection algorithm based on theoretical curves modeling and fuzzy classification system”, *Energy*, Vol. 140, 2017, pp. 276-290.

[17] S. Rustamov, E. Mustafayev, and MA Clements, “Context analysis of customer requests using a hybrid adaptive neuro fuzzy inference system and hidden markov models in the natural language call routing problem”, *Open Eng.*, Vol. 8, No. 1, 2018, pp. 61-68.

[18] I. Dawy, “The Most General Intelligent Architectures of the Hybrid Neuro-Fuzzy Models”, *Am. J. Artif. Intell.*, Vol. 2, No. 1, 2018, pp. 1.

[19] TN Mohd Najib, Mohd Salleh and KH Talpur, “A Modified Neuro-Fuzzy System Using Metaheuristic Approaches for Data Classification”, *Intech Open*, Vol. 2, 2018, pp. 64.

[20] S. Rustamov, “A Hybrid System for Subjectivity Analysis”, *Adv. Fuzzy System*, Vol. 2018, 2018, pp. 64.

[21] A. Belaout, F. Krim, A. Mellit, B. Talbi, and A. Arabi, “Multiclass adaptive neuro-fuzzy classifier and feature selection techniques for photovoltaic array fault detection and classification”, *Renew. Energy*, Vol. 127, No. May, 2018, pp. 548-558.

[22] A. Jayachandran and R. Dhanasekaran, “Multi Class Brain Tumor Classification of Mri Images Using Hybrid Structure Descriptor and Fuzzy Logic Based Rbf Kernel Svm”, *Iran. J. Fuzzy Syst.*, Vol. 14, No. 3, 2017, pp. 41-54.

[23] JH Leung, Y. Kuo, T. Weng, and C. Chin, “Hybrid-Neuro-Fuzzy System and Adaboost-Classifier for Classifying Breast Calcification”, *J. Comput.*, Vol. 26, No. 5, 2017, pp. 29-42.

[24] . Duan, Y. Wang, W. Pedrycz, X. Liu, C. Wang, and Z. Li, “AFSNN: A Classification Algorithm Using Axiomatic Fuzzy Sets and Neural Networks”, *IEEE Trans. Fuzzy System.*, Vol. 26, No. 5, 2018, pp. 3151-3163.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[25] VA Arvindrao, "Adaptive Expansion Algorithm for Fuzzy Min-Max Neural Networks in Pattern Classification", No. 1, 2017, pp. 741-744.

[26] A. Bilski, "A review of artificial intelligence algorithms in document classification", *Int. J. Electron. Telecommun.*, Vol. 57, No. 3, 2011, pp. 263-270.

[27] A. Caliskan and ME Yuksel, "Classification of coronary artery disease data sets by using a deep neural network", *EuroBiotech J.*, Vol. 1, No. 4, 2017, pp. 271-277.

[28] S. Chaki, A. Routray, WK Mohanty, and M. Jenamani, "Development of a hybrid learning system based on SVM, ANFIS and domain knowledge: DKFIS", *2th IEEE Int. conf. electrons. Energy, Environ. comm. Comput. Control (E3-C3), INDICON 2015*, 2016.

[29] Q. Chen, X. Song, H. Yamada, and R. Shibasaki, "Learning Deep Representation from Big and Heterogeneous Data for Traffic Accident Inference", *Aaai*, 2016, pp. 338-344.

[30] F. García, F. Guijarro, J. Oliver, and R. Tamošiūnienė, "Hybrid Fuzzy Neural Network To Predict Price Direction in the German Dax-30 Index", *Technol. econ. Dev. Econ.*, Vol. 24, No. 6, 2018, pp. 2161-2178.

[31] S. Gorbachev and V. Syryamkin, "High-Performance Adaptive Neurofuzzy Classifier with a Parametric Tuning", *MATEC Web of Conferences*, Vol. 155, 2018.

[32] B. Haznedar, MT Arslan, and A. Kalinli, "Using Adaptive Neuro-Fuzzy Inference System for Classification of Microarray Gene Expression Cancer Profiles", Vol. 2018, No. 29, 2018, pp. 1-13.

[33] BR Hoard, "Application of adaptive-network-based fuzzy inference systems to the parameter optimization of a biochemical rule-based model", 2018.

[34] OJ Kravets and S. Alekseevich, "Processes of Multidimensional Classification of Scoring Objects with Heterogeneous Features Based on the Neural Networks Modeling 2 Neural Network Approach to Solving the Multidimensional Classification Problem 4 Features of Learning a Neural", Vol. 119, No. 7, 2018, pp. 875-880.

[35] A. Omotosho, AE Oluwatobi, OR Oluwaseun, AE Chukwuka, and A. Adekanmi, "A Neuro-Fuzzy Based System for the Classification of Cells as Cancerous or Non-Cancerous", *Int. J. Med. Res. Heal. Science.*, Vol. 7, No. 5, 2018, pp. 155-166.

[36] A. Puzanov and K. Cohen, "Deep Reinforcement One-Shot Learning for Artificially Intelligent Classification Systems", 2018, pp. 1-27.

[37] HR Singh and SK Biswas, "Transparent Neuro-fuzzy model for Linguistic variables selection and rule-based classification", Vol. 118, No. 19, 2018, pp. 85-100.

[38] P. Ashtari et al., "Supervised Fuzzy Partitioning", *CoRR*, Vol. 2018, No. 4, 2018, pp. 1-13.

[39] J. Cai, J. Luo, S. Wang, and S. Yang, "Feature selection in machine learning: A new perspective", *Neurocomputing*, Vol. 300, 2018, pp. 70-79.

[40] D. Caragea, J. Pathak, and VG Honavar, "Learning Classifiers from Semantically Heterogeneous Data", 2011, pp. 963-980.

[41] F. Chang, C. Lin, and C. Chen, "A Hybrid Method for Multiclass Classification and Its Application to Handwritten Character Recognition", pp. 1-23.

[42] P. D'Urso and M. . Gil, "Fuzzy data analysis and classification", *Adv. Anal Data. Classif.*, Vol. 11, No. 4, 2017, pp. 645-657.

[43] DEN Frossard, IO Nunes, and RA Krohling, "An approach to dealing with missing values in heterogeneous data using k-nearest neighbors", 2016.

[44] VS Kodogiannis, I. Petrounias, and JN Lygouras, "Intelligent classification using adaptive fuzzy logic systems", *2008 4th Int. IEEE Conf. Intell. syst. IS 2008*, Vol. 1, No. 4, 2008, pp. 98-913.

[45] HW Lim, YW Hau, CW Lim, and MA Othman, "Artificial intelligence classification methods of atrial fibrillation with implementation technology", *Comput. Assists. Heaven.*, Vol. 21, 2016, pp. 155-162.

[46] M. Pratama, W. Pedrycz, and E. Lughofer, "Evolving Ensemble Fuzzy Classifier", *IEEE Trans. Fuzzy System.*, Vol. 26, No. 5, 2018, pp. 2552-2567.

[47] MA Sanchez, JR Castro, V. Ocegueda-Miramontes, and L. Cervantes, "Hybrid learning for general type-2 TSK fuzzy logic systems," Algorithms", *Algorithms*, Vol. 10, No. 3, 2017.

[48] H. Wang, Z. Xu, and W. Pedrycz, "An overview on the roles of fuzzy set techniques in big data processing: Trends, challenges and opportunities", *Knowledge-Based System.*, Vol. 118, 2017, pp. 118.

15th September 2023. Vol.101. No 17

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[49] X. Zhang and Z. Xu, "The Extended TOPSIS Method for Multi-criteria Decision Making Based on Hesitant Heterogeneous Information", No. November, 2014.

[50] C. Zhou, "Linguistic and Numerical Heterogeneous Data Integration with Reinforcement Learning for Humanoid Robots", *Control*.

[51] R. Sheikhpour, MA Sarram, S. Gharaghani, and MAZ Chahooki, "A Survey on semi-supervised feature selection methods", *Pattern Recognit.*, Vol. 64, No. November 2016, 2017, pp. 141-158.

15th September 2023. Vol.101. No 17

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 2: Hybrid Algorithm On Heterogenous Data.

Author	Dataset	Pre-processing	Method	µ	Fuzzy Rules	Defuzz	NN	Evaluator	Advantages	Opportunity
Azhari , M., & Jaya Kumar, Y. (2017) [9]	DUC 2000	Feature Extraction	ANFIS	Gaussian MF	FIS	CoG	LSE, Back-propagation	Recall Oriented Understudy for Gisting Evaluation (ROUGE)	Have High precision and recall	Tuning variable in membership function is necessary upgraded
Belaout , A., Krim , F., Mellit, A., Talbi , B., & Arabi, A. (2018) [21]	5790 IV curves	feature space dimensionality reduction techniques	Multiclass Adaptive Neuro-Fuzzy Classifier (MC-NFC	Trapezoi -dal	Sugeno FIS	CoG	LSE, Back-propagation	RMS, R2	Have accuracy tall	Need existence combination class error for more test accuracy
Bilski , A. (2011) [26]	Multi-document	vector space model	Artificial Neural Networks (ANN)	Trapezoi -dal	Sugeno FIS	CoG	Back-propagation	RMSE	reduce amount dimension and increase quality classification	Not showing yet combination combined with fuzzy
Brumancia , E., Justin Samuel, S., Gladence , LM, & Rathan, K. (2019) [4]	MADELOON	Dempster – Shafer	ANFIS	Trapezoi -dal	Sugeno FIS	CoG	LSE, Back-propagation	RMSE	Succeed get fusion info from	Trial _ to hybrid models still is partial in each method
Caliskan , A., & Yuksel , ME (2017) [27]	coronary artery disease (CAD	Feature Extraction	deep neural network (DNN)	Trapezoi -dal	Sugeno FIS	CoG	Broyden – Fletcher– Goldfarb– Shanno	Accuracy	Accuracy tall for four languages	Fuzzy method yet embedded _ _
Chaki , S., Routray , A., Mohanty , WK, & Jenamani , M. (2016) [28]	Indian onshore	z-score normalization	ANFIS-SVM	Figellmf	Sugeno FIS	CoG	LSE, Back-propagation	Accuracy	Accuracy for make bi-class already tall	At stage classification still added fuzzy process
Chen, Q., Song, X., Yamada, H., & Shibasaki , R.	Traffic accident data		Deep Learning				SDAE Model	MAE, MRE, RMSE	Small error rate compared method DT, SVM	still can appointed a new evaluator

(2016) [29]										
Dhimish , M., Holmes, V., Mehrdadi , B., Dales, M., & Mather, P. (2017) [16]	Curves		Fuzzy	Gaussian and Triangle	Mamdani			Efficiency , Detection Accuracy	Reach accuracy tall for multiple fault detection	Developed algorithm _ still depending on the instrument
Gabrys , B., & Bargiela , A. (2000) [12]	IRIS, wine, ionosphere		Fuzzy Min-Max Neural Network	Simpson Hyperbox Method	Sugeno FIS		Three Layer (Input, Htperbox, Class)	Precision, Recall, Accuracy	Definition hyperbox can be adaptable	Identification proper hyperbox _ if amount instance increase, need to be added text processing elements
García, F., Guijarro , F., Oliver, J., & Tamošiūnienė , R. (2018) [30]	German DAX-30	Analysis Factor	HyFIS	Gaussian membership functions	Mamdani model	CoG	Self-Construct	MAE, MRE, RMSE	Accuracy already up to 74%	Election variable and dimensional reduction
Gorbachev, S., & Syryamkin , V. (2018) [31]	multicriteria problems	Multi-dimensional features objects	neural-fuzzy networks	triangular fuzzy numbers	AND-neuron	degree of rules	back propagation	accuracy	have ability for interpretation accumulation knowledge	Not seen yet accuracy because new proposal NN and Fuzzy structure
Haznedar , B., Arslan, MT, & Kalinli , A. (2018) [32]	DNA structure	Clustering	ANFIS	Figellmf	first-order Sugeno fuzzy model		genfis2	RMSE	Get 100% accuracy	Definition multidimensional variables need considered, so that need a process that regulates right variable _
Hoard, BR (2018) [33]	biochemical	cutoff distance	ANFIS	Gaussian MF	biological rule-based model		LSE and Backpropagation	MAE, MRE, RMSE	get optimal cluster with small RMSE	Still need to test with fuzzy and nn by independent for knowing cluster results
Kravets , OJ, & Alekseevich , S. (2018) [34]	Article	fuzzy features	MLP and Fuzzy	Gaussian MF	Sugeno FIS	CoG	forward	RMSE	Get less mis classification and	on preprocessing still there is repair featuring

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

									small RMSE	technique so that you can reduce amount feature
Leung, JH, Kuo , Y., Weng, T., & Chin, C. (2017) [23]	WISCONSIN	NLP	self-constructing neural fuzzy inference network (SONFIN	Gaussian MF	Sugeno FIS	CoG	forward	Accuracy	Already produce accuracy tall	Allow for copy SONFIN work with make structure new
Mohd Najib, Mohd Salleh, TN, & Talpur , KH (2018	Iris, Breast Cancer		ANFIS	Gaussian	Sugeno FIS	CoG	gradient descent (GD) and least squares estimator (LSE	RMSE	Have small MSE level	Make the right membership function for ANFIS standard and ANFIS modification for handle time series and clustering data in classification
Mu, B., Inform, S., Polit, U., & Barcelona, C. (2007) [6]	Pima Diabetes	vector space model, normalization	ANN	Trapezoi -dal	Sugeno FIS	CoG	LSE and Backpropagation	RMSE	Succeed identify missing values	The use of the fuzzy process has not seen clear
Omotosho , A., Oluwatobi , AE, Oluwaseun , OR, Chukwuka , AE, & Adekanmi , A. (2018) [35]	Topography	Feature Extraction	Fuzzy, ANN	Trapezoi -dal	Sugeno FIS	CoG	back-propagation	Feature extraction success	Success in To do import CT scan from normal image becomes not normal	The multiclassification process is still not yet conducted research and on real time data
Puzanov , A., & Cohen, K. (2018) [36]	Amazon	NLP	Fuzzy, ANN	Trapezoi -dal	Sugeno FIS	CoG	genfis2	Accuracy	Have accuracy tall	Repair system for more review problems from one person
Singh, HR, & Biswas, SK (2018) [37]	IRIS	Supervise Dynamic Clustering Algorithm (SDCA .)		Trapezoi -dal	Sugeno FIS	CoG	genfis2	Accuracy	Succeed get amount right features –	Need existence mix Among fuzy and NN so that the process gets feature more good again.

15th September 2023. Vol.101. No 17

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 3: Research Opportunities On Hybrid Methods

Author	Layer I (Input Node)		Layer II	Layer III	Layer IV	Layer V	Output Layer
	Normalization	Feature Extraction	Linguistic	Membership Function	Inference	Defuzzification	Function Activation
Azhari, M., & Jaya Kumar, Y. (2017)	fuzzy rules scoring	fuzzy rules scoring	fuzzy rules scoring	sentence features	Sugeno - type system	CoG	multi-document text summarization
Belaout, A., Krim, F., Mellit, A., Talbi, B., & Arabi, A. (2018		feature space dimensionality reduction techniques	inverse method	sigmoidal	zero-order Sugeno type	CoG	Multi-fault
Bilski, A. (2011	Rocchio algorithm	Vector space model	local feature sets				Expectation-Maximization
Brumancia, E., Justin Samuel, S., Gladence, LM, & Rathan, K. (2019		fuzzy equivalent model			nine rules, FIS	CoG	fused information
Caliskan, A., & Yuksel, ME (2017		Feature Extraction					softmax classifier
Chaki, S., Routray, A., Mohanty, WK, & Jenamani, M. (2016	Z-score and min-max	SVM	Category	Figellmf		modified membership grades	Qualitatively Constrained Functions (QCF)
Chen, Q., Song, X., Yamada, H., & Shibasaki, R. (2016		feature representation				gradient-base optimization technique	Stack denoise Autoencoder (SdAE)
Gabrys, B., & Bargiela, A. (2000	n - dimensional space	fuzzy hyperboxes		hyperbox expansion constraint proposed			adaptive maximum size of a hyperbox scheme
García, F., Guijarro, F., Oliver, J., & Tamošiūnienė, R. (2018		linguistic variables	linguistic variables	Gaussian membership functions	Mamdani model	CoG	forecast

15th September 2023. Vol.101. No 17

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Gorbachev, S., & Syryamkin, V. (2018	numerical values	multidimensional feature objects	numerical values	fuzzy decision tree function	Relative degree (weight) of Fuzzy Rule	sum of the degrees of truth	forecast class
Haznedar, B., Arslan, MT, & Kalinli, A. (2018	fuzzification process	Clustering / fuzzification process	fuzzification process	gaussmf	first-order Sugeno fuzzy model	gradient descent	Breast Cancer Accuracy
Hoard, BR (2018		Rule clusters using fuzzy c-means	number of clusters			Monte Carlo Model Prediction	FIS-predicted rule rate and cutoff distance
Kravets, OJ, & Alekseevich, S. (2018	frequency of the occurrence of a term	signal conversion unit (SCU)	number of signal features	fuzzy variable membership function	SCU / MLP Algorithm		neuron activation function
Leung, JH, Kuo, Y., Weng, T., & Chin, C. (2017	data sampling techniques	Adaboost algorithm Input:	feature combination	Gaussian	Mamdani model		classification accuracy
Mohd Najib, Mohd Salleh, TN, & Talpur, KH (2018		Numeric values		gaussian	metaheuristics optimization algorithm artificial bee colony	gradient based learning	
Mu, B., Inform, S., Polit, U., & Barcelona, C. (2007	similarity function	similarity value	regression		breeder genetic algorithm	RBF	
Singh, HR, & Biswas, SK (2018)		Supervise Dynamic Clustering Algorithm (SDCA)	linguistic variable selection	Gaussian function			sigmoid activation function