

# A Novel Fuzzy Expert System Design for Predicting Heart Diseases

Arash Khormehr

MSc student of Software Engineering  
Department of Computer Engineering  
Sanandaj Branch, Islamic Azad University,  
Sanandaj, Kurdistan, Iran

Vafa Maihami

Faculty member  
Department of Computer Engineering  
Sanandaj Branch, Islamic Azad University,  
Sanandaj, Iran

## ABSTRACT

The volume of data generated by rapid technological progress also on the rise is too fast, use, the selection of useful data and its analysis of the issues that have been of interest to researchers, obtain conclusive results due to the uncertainty of this information to resolve these issues also are research priorities. Forecast diseases, including the risk factors in the selection of important and complex issues that concern is to get the correct result is that Heart disease is no exception. In this paper, using a fuzzy system a model is designed to predict heart disease that using design rules based on medical science works. By a physician, a series of rules designed, with this rules and fuzzy systems a good model with more efficient for predicting heart disease is presented. The proposed algorithm is based on data obtained from several cardiac patients and healthy individuals were tested in Tohid Hospital in Sanandaj city, the proposed algorithm's accuracy from 95% in people prone to heart disease to be identified with precision.

## Keywords

Predict heart disease, Fuzzy systems, Fuzzy inference engine, Risk factor.

## 1. INTRODUCTION

Expert systems and data mining a science that According to given the large volume of scientific data in various fields, including in the field of medical treatment and are widely used. Fuzzy logic, artificial neural networks and genetic algorithms are sub-Soft Computing, Soft Computing to a set of new ways of computing in computer science, artificial intelligence, machine learning and many other application areas means.

The theory of fuzzy sets was proposed in 1965 by Dr. Zadeh and found numerous applications in various sectors [20-25]. Each element has a degree of membership in binary logic zero and one, while membership in the fuzzy logic element is between zero and one. Soft computing and Fuzzy Systems for various sectors led performance in recent years of the logic used to predict heart disease[1-7].

S. Muthukaruppan "A hybrid particle swarm optimization based fuzzy expert system for the diagnosis of coronary artery disease" have offered [17]. They classify the factors of heart disease and used Mamdani fuzzy inference system and have used the triangular membership functions, their proposed classification accuracy was 93.27 percent. vahid Khatibi and colleagues study entitled "fuzzy inference engine designed to assess the risk of coronary heart disease" did [11]. In their study, the fuzzy algorithm is used, Risk factors for is classified set by the doctors and fuzzy systems that have been designed to have 89% accuracy.

V Krishnaiah and his colleagues in their study entitled "Diagnosis of heart disease using fuzzy classification techniques" used a combination of fuzzy techniques and k-

nearest neighbor [13], they proposed algorithm accuracy was 83.7%, those of mathematics zadeh used, to remove the extra data.

In this paper, to collect data, first search in the hospital information software (HIS) of Tohid Hospital in Sanandaj and were record numbers of heart patients, Since this study was to assess the heart patients Kurdistan To collect the required data, records of patients in Tohid Hospital in Sanandaj extracted archive and Risk factors previously identified by the consultant physician also noted that, The problem was that patients were required to continue that risk factors have been chosen as test LDL (bad fat), HDL (good fat) and CRP (a protein that increases the heart's blood disorder) have are, But not heart disease, such as to create rules or membership functions require non-cardiac patients as well. Outpatients perhaps had the tests but did not have a Dossier at the hospital to go to for

Other risk factors, and this was a problem, And since there was ever such a thing in the hospital, then to fix the problem to the programmer HIS (hospital information software) was referred.

In this study, risk factors, blood pressure, that sufficient. The lower harmful fats is better, Unlike the harmful fat, good fat the higher is better, CRP factor if the value of the fat is below 6 negative and positive means is greater than 6 and can even go up to 200. The output of this system consists of two classes - yes (of heart disease) and No (of heart disease)- is. The aim of this study is that the data collected will be Tohid Hospital in Sanandaj predict and Different learning algorithms run on them and compare, and ultimately proposed an optimal algorithm to important risk factors for developing heart disease is found.

Fuzzy systems are briefly described in the second part. Then, in the Third part of the proposed algorithm is fully detailed, forms and charts described. In the fourth part with the data in different systems with different volumes of run and comparisons made. The conclusion of the study is given in Section V.

## 2. FUZZY SYSTEMS

Fuzzy system is a system that is modeled phenomena in the form of fuzzy sets.

These systems are based on knowledge and sophisticated systems in place, to be classified. Fuzzy systems have been defined on the basis of which they can be fuzzy/defuzzy system noted. In this system that fuzzy rule base, After the fuzzyfication of the input data with the fuzzy inference system, Fuzzy Inference Engine using inference rules extended to apply appropriate rules of the input data is a set of fuzzy rules and Fuzzy result defuzzier returns it to the final output. The combination of rules and regulations for each of

the two approaches could be combined results from the union or the issue of benefit sharing.

Linguistic variables; in the final set of variables can take different amounts of numbers. Fuzzy sets can be varied in addition to different numbers, different values accept language. If the variable can accept the terms of a natural language as a variable called language. The advantages of the theory of fuzzy linguistic terms by fuzzy collections.

Approximate Reasoning; every argument involves a prayer that could be valid or invalid and the two types is comparative and Inductive. Rules of inference arguments are acclaimed for the

Propositions hold. Despite the uncertainty in the space of fuzzy sets has caused arguments are not definitive encounter with approximate reasoning.

Membership function; uncertainty in a fuzzy set can be solved by membership functions. The membership value of a member such as x in the functions between 0 and 1. This means that a function get an entry such as the x and it's the interval [0,1] defines. The following are two types of membership functions of triangular and trapezoidal noted that the most used, Triangular membership function is defined by Equation 2 and Equation 3 is defined as a trapezoidal membership function:

$$A[x] \begin{cases} 0 & \text{if } x \leq a \\ \frac{x-a}{b-a} & \text{if } x \in (a, b) \\ \frac{c-x}{c-m} & \text{if } x \in (b, c) \\ 0 & \text{if } x \geq c \end{cases}$$

Equation 2 - triangular membership function

$$SG[x] \begin{cases} 0 & \text{if } x \neq m \\ 1 & \text{if } x = m \end{cases}$$

Equation 3 - Singleton membership functions

In the proposed system of Mamdani fuzzy inference system is used, this system in 1975 by Mamdani for fuzzy control system for the steam engine was first introduced at the University of London. He has a set of rules used to express the human experience in his system. More modern fuzzy control system of a set of rules defined by Mamdani made. Also in this article through the center of gravity to Defuzzy the system is used.

### 3. THE NEW SYSTEM FUZZY PROVIDED TO PREDICT HEART DISEASE

After pre-processing of data collected Enter to fuzzy systems are designed and then the decision will be based on output, how preprocessing and the decision will be discussed later in this. (figure 1).

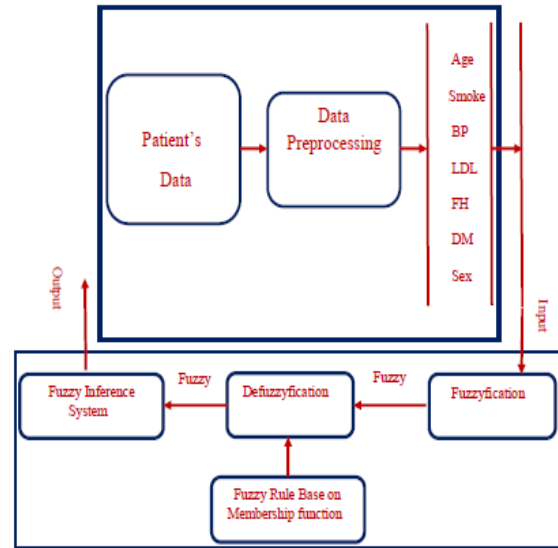


Figure 1 - The overall shape of the proposed algorithm

### 3.1 Identify and classify risk factors

After consultation with the medical team advisor was decided that the factors listed in Table 1 is extracted from the patient's dossier:

Table 1 - Categories of risk factors

Risk Factor	Rank	
Age	Young	15-30
	Med	30-50
	Med-Old	50-65
	Old	65 بالای
Sex	Male Fmale	
HDL	Normal	>50
	Low	30-50
	Very Low	<30
Smoking	yes no	
CRP	Negative	<6
	Positive	>6
Diabet	yes no	
Physical Inactivity(PI)	yes no	
Family History(FH)	yes no	
Blood Pressure (BP)	Low	<10
	Normal	10-14
	High	14-16
	Very High	>16
LDL	Normal	<100
	High	100-130
	Very High	130-160
	Danger	>160
Paste History (PH)	yes no	

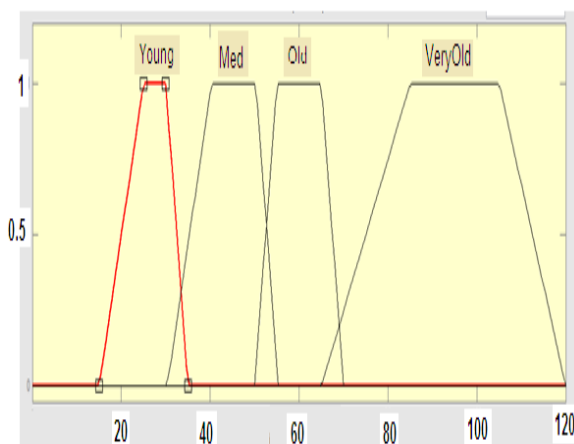
### 3.2 Description of the algorithm

The proposed algorithm consists of 4 steps that will be explained below. The data collected in Step 1 is read. Duplicate and out-of-range data are corrected and Empty fields are filled and if the record is a lot of empty fields that record should be removed, In phase 2 data from the pre-stage be given to fuzzy system is to be processed. In this paper, fuzzy systems, including 7 inputs and one output. In phase 3, fuzzy output compared with a threshold value and the final result is obtained. in step 4, the threshold obtained from the third step is compared with the fuzzy output and Based on the patient's class is determined and it is determined that the patient is a heart patient or not. Risk factors are a lower priority should have been removed to a more important risk factors apply fuzzy systems. Risk factors are listed in Table 2.

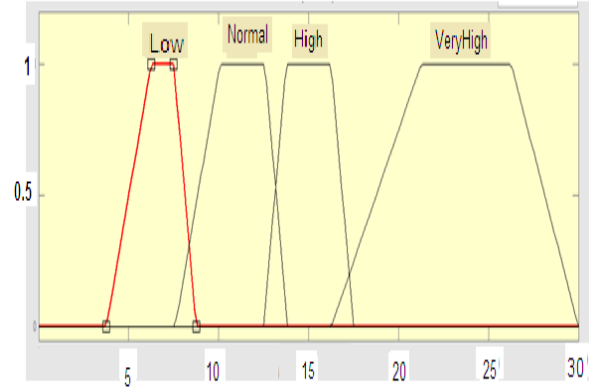
**Table 2 - Risk factors finalized and ready for input to be proposed fuzzy system**

Age			
Young	15-30	Med	30-50
Med-Old	>65	Old	>65
B-P			
Low	<10	10-14	Normal
High	14-16	>16	Very-High
LDL			
Normal	<100	High	100-130
Very-High	130-160		
Smoke			
No		Yes	
F-H			
No	0-5	Yes	5-10
Sex			
Female	0-5	Male	5-10
OutPut			
No	0-0.5	Yes	0.5-1

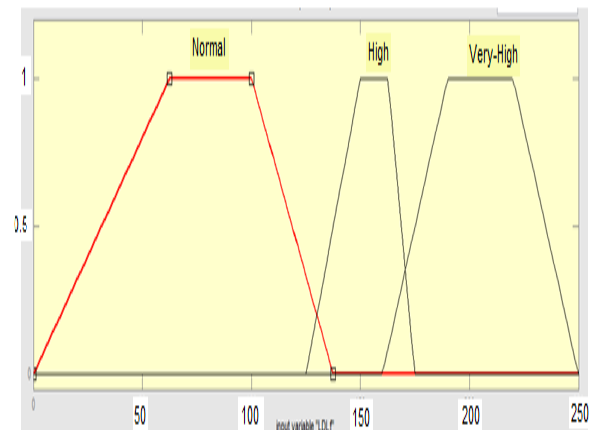
Membership functions of the system are shown in Figures 2 to 6. System According to the inputs and existing rules in the rule base gives the output is required.



**Figure 2 - fuzzy definition of age**



**Figure 3 - The definition phase blood pressure**



**Figure 4 - The definition of harmful fat phase**

Age, consists of four levels and is defined as a trapezoidal membership function. Blood pressure consists of four levels and is defined as a trapezoidal membership function. Harmful fats, the three levels and a trapezoidal membership function is used to define it. History of diabetes, the two levels of membership function sigmf been used to define it. Output, the output of two members of the "yes" and "no" and the membership function is defined sigmf.

### 3.3 Defined system rules

Number 56 rules defined in the system, the uncertainty was modeled in assessing the risk of heart disease that some of these rules are shown in Table 3.

**Table 3 - Some of the proposed rules**

- 1 If( Age is Young)and(Smoke is yes)and(BP is High)and(LDL is Very-High )and(F-H is yes)and(DM is yes)and(Sex is F)Then (Heart is Yes)
- 2 If(Age is Young)and(Smoke is yes)and(BP is High)and(LDL is Very-High )and(F-H is yes)and(DM is yes)and(Sex is M)Then (Heart is Yes)
- 3 If(Age is Med)and(Smoke is No)and(BP is Normal)and(LDL is High)and(F-H is yes)and(DM is yes)and(Sex is M)Then (Heart is No)

The fuzzy system to compare and enhance the relationship between the input and the output can be used from the Surface Chart, Figures 1 and 2 show some diagrams Surface of the system:

In Figure 8, the output based on input blood pressure and harmful fats have been shown, As can be seen by increasing the amount of each of the two input increases the likelihood of heart disease, For example, high levels of harmful fat around 130 and the blood pressure values of the top 16 around the output value is closer to 1, In amounts less than the probability of disease is said to be close to zero. In Figure 8, we see that blood pressure and heart disease are likely to have a direct relationship with age, As can be seen with increasing input values increases the likelihood of heart disease, About variable blood pressure at higher doses for age 12 and older than 50 years about the value of the output value is closer to 1, In less than said amount is zero probability of disease.

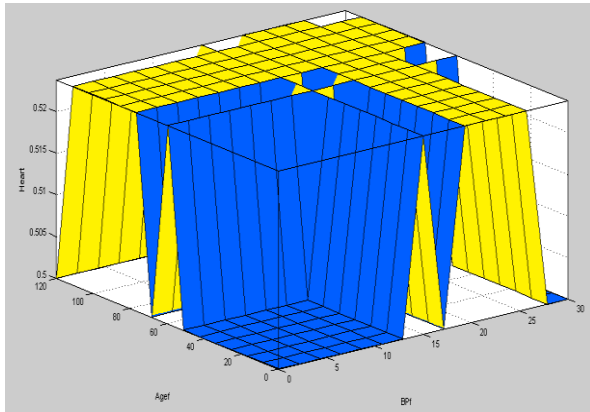


Figure 7 - charts the relationship between blood pressure and harmful fats

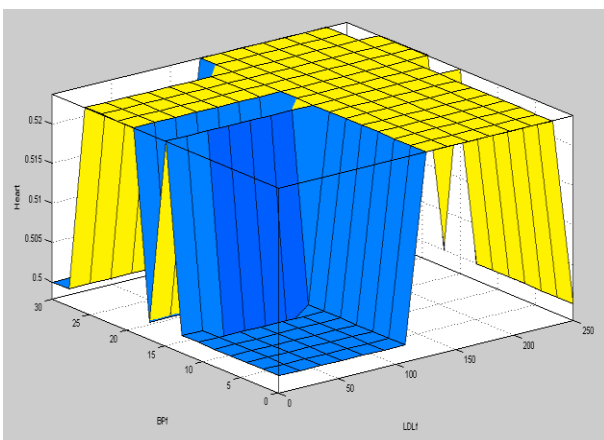


Figure 8 - charts the relationship between blood pressure and age

### 3.4 Read input files and apply fuzzy system

Nexus 4 Gives seven input to the fuzzy files, this input in the fuzzy system are evaluation, And using membership functions and limits defined for which the rules are applied on They and The result puts x in output, the result is a number between zero and one. Based on the output fuzzy should classes be determined, Records in the class "yes" are that fuzzy result of them is greater than the threshold and records in the class "no" are that fuzzy result of them

is less than the threshold.

For i<-1:n

FR(i) = evalfis([Inputs],fismat);

End

Nexus 4 - input fuzzy system

---

If FR(i)>= Bound & IC=2

Then Compare(i) <- 2

End

---

If FR(i)<Bound & IC=1

Then Compare(i) <- 1

End

If Compare(i)= IC

Then Result <- Result+1

End

---

Nexus 4 - specify the number of classes foreseen Field

In Nexus 5, IC variable indicates the patient is sick or healthy. Results and tests, algorithm with Six data series by fuzzy algorithm by comparing the was intelligent was implemented, Then apply the algorithms and see the results. Compare intelligent algorithm was tested for 6-series data, The result is shown in Diagram 1, Precision algorithm for 10 patients 70 percent, to 100 patients 80 percent, 82 percent for the 200 patients, for 400 patients percent, 89 percent to 500 sick and 95 percent for 1000 patient that's mean The accuracy of the proposed algorithm with the number of cases rises, From this it can be concluded that the number of cases rises learning capability of the proposed algorithm is also on the rise. In each column with a colored graph and in the beginning we see that in the number of patients percent gradient algorithm increases accuracy but the dip in the number of patients is low.

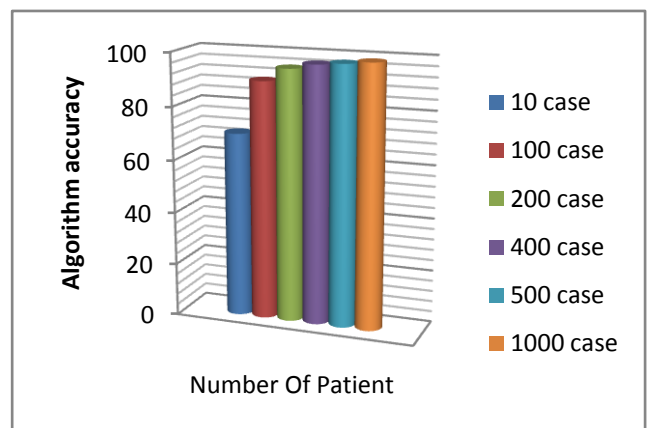


Diagram 1 - fuzzy algorithm test results

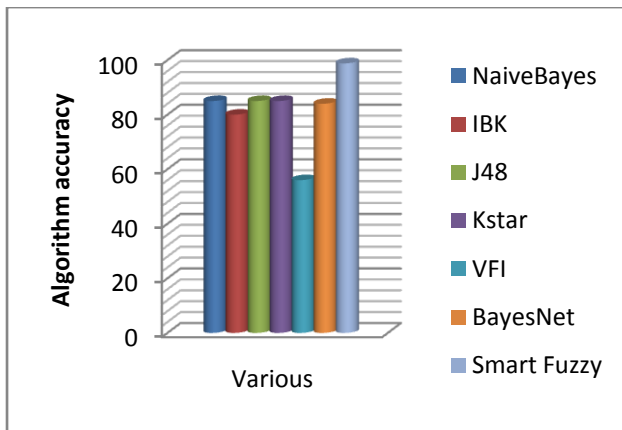


Diagram 2 – compare proposed algorithm with other algorithms with this dataset

Compare the proposed algorithm with other algorithms, This data set with a series of other algorithms was tested using the software Weka and The results were compared with the output of the proposed algorithm, The result is shown in diagram 2, The accuracy of prediction algorithms have been shown in this graph, Above all, they last cylinder, which refers to the proposed algorithm is 95 percent, Placed second on an algorithm that is nave bayes, And other algorithms in the next orders placed, Of course, this does not mean that other algorithms are not suitable but for this job haven't good accuracy. The proposed algorithm has been compared with previous works, Compared with previous work done on the proposed algorithm is shown in Table 4.

Table 4 - Comparison algorithm with Previous works

Author	Method	Accuracy - the percentage
Sellappan Palaniappan[19]	Method1	86.53
	Method2	94.93
Jesmin Nahar[18]	Method1	83.83
	Method2	84.89
	Method3	76.90
	Method4	83.50
R. Chitra[3]	His proposed article	81
Vahid Khatibi[11]	His proposed article	89
V.Krishnaiah[13]	His proposed article	83.7
This paper	The proposed approach	95

#### 4. CONCLUSION

In this paper, we propose a new algorithm with the idea of fuzzy systems for Predict heart disease patients in Tohid Hospital in Sanandaj In order to increase forecast accuracy and in order to study the patients were in the hospital was presented. How it was that information was extracted from the patient's Dossier and sorted so it can be into the fuzzy system. The fuzzy algorithm system has seven inputs and one output, inputs the number of risk factors involved in heart disease and Output has two values: "Yes" and "No", input Was given to fuzzy system and if the fuzzy result

is greater than the threshold record in the class "yes" and If it is less than the threshold class "No" is placed.

#### 5. REFERENCES

- [1] Anooj, P. (2012). Clinical decision support system: Risk level prediction. Journal of King Saud University – Computer and Information Sciences, 14.
- [2] Austin, P. C. (2013). Using methods from the data-mining and machine-learning literature for disease classification and prediction: a case study examining classification of heart failure subtypes. Elsevier, 10.
- [3] Chitra, R. (2013). Heart Disease Prediction System Using Supervised Learning Classifier. Bonfring International Journal of Software Engineering and Soft Computing, 7.
- [4] Czabanski, R. (2013). Fetal state assessment using fuzzy analysis of fetal heart rate signals—Agreement with the neonatal outcome. Elsevier.
- [5] Dai, W. (2014). Prediction of hospitalization due to heart diseases by supervised learning methods. Elsevier, 9.
- [6] Das, R. (2010). Effective diagnosis of heart disease through neural networks ensembles. Elsevier, 6.
- [7] Imam, T. (2013). Association rule mining to detect factors which contribute to heart disease. Elsevier, 8.
- [8] jabbar, M. (2013). Classification of Heart Disease Using K- Nearest Neighbor and Genetic Algorithm. elsevier, 10.
- [9] Jee, S. H. (2015). A coronary heart disease prediction model: the Korean Heart Study. dx.doi.org, 10.
- [10] K.Rajeswari. (2012). Feature Selection in Ischemic Heart Disease Identification using Feed Forward Neural Networks. elsevier, 6.
- [11] Khatibi, V. (2010). A fuzzy-evidential hybrid inference engine for coronary heart disease. elsevier, 7.
- [12] Kolus, A. (2015). Adaptive neuro-fuzzy inference systems with k-fold cross-validation for energy expenditure predictions based on heart rate. Elsevier, 11.
- [13] Krishnaiah, V. (2013). Diagnosis of Heart Disease Patients Using Fuzzy Classification Technique.
- [14] Kumar, Y. (2014). Research Aspects of Expert System. International Journal of Computing & Business Research, 11.
- [15] Laura Sabiani, M. (2015). Intra- and interobserver agreement among obstetric experts in court regarding the review of abnormal fetal heart rate tracings and obstetrical management. Am J Obstet Gynecol 2015;213:856.e1-8., p. 8.
- [16] M, I. (2015). Efficient Data Mining Method to Predict the Risk of Heart Diseases. 4th International Conference on Eco-friendly Computing and Communication Systems, ICECCS 2015 (p. 7). Procedia.
- [17] Muthukaruppan, S. (2012). A hybrid particle swarm optimization based fuzzy expert system for the diagnosis of coronary artery disease. elsevier, 9.
- [18] Nahar, J. (2013). Computational intelligence for heart disease diagnosis: A medical knowledge driven approach. elsevier, 9.
- [19] Palaniappan, S. (2008). Intelligent Heart Disease Prediction System Using Data Mining Techniques.

IJCSNS International Journal of Computer Science and Network Security, 8.

- [20] Petkovic, D. (2013). Adaptive neuro fuzzy selection of heart rate variability parameters affected by autonomic nervous system. Elsevier, 6.
- [21] Setoa, E. (2012). Developing healthcare rule-based expert systems: Case study of a heart failure telemonitoring system. *ijmijournal*, 10.
- [22] Sharma, A. (2014). Emerging Applications of Data Mining for Healthcare Management - A Critical Review Management - A Critical Review. 6.
- [23] Tay, D. (2015). A novel neural-inspired learning algorithm with application to clinical risk prediction. *Journal of Biomedical Informatics*, 10.
- [24] Fathi, M., Maihami, V., & Moradi, P. (2013). Reinforcement Learning for Multiple Access Control in Wireless Sensor Networks: Review, Model, and Open Issues. *Wireless personal communications*, 72(1), 535-547.
- [25] Mandeh, A., Khamforoosh, K., & Maihami, V. (2015). Data Fusion in Wireless Sensor Networks using Fuzzy Systems. *International Journal of Computer Applications*, 125(12).