Soft Computing (2019) 23:10931–10938
https://doi.org/10.1007/s00500-018-3643-6

METHODOLOGIES AND APPLICATION

Intelligent hepatitis diagnosis using adaptive neuro-fuzzy inference system and information gain method

Waheed Ahmad[1] · Ayaz Ahmad[2,3] · Amjad Iqbal[4] · Muhammad Hamayun[5] · Anwar Hussain[5] · Gauhar Rehman[6] · Salman Khan[7] · Ubaid Ullah Khan[8] · Dawar Khan[9] · Lican Huang[1]

Published online: 24 November 2018

Abstract

Hepatitis, a common liver inflammation, is one of the major public health issues around the world. Proper interpretation of clinical data for the diagnosis of hepatitis is an important problem that needs to be addressed. In this study, a hybrid intelligent approach, combining information gain method and adaptive neuro-fuzzy inference system (ANFIS), is proposed for the diagnosis of fatal hepatitis disorder. Initially, the hepatitis dataset obtained from the University of California Irvine machine learning repository is preprocessed to make it suitable for the mining process. After the preprocessing stage, information gain method is applied to condense the number of features in order to decrease computation time and classification complexity. Selected features are then fed into the ANFIS classifier system. The performance of the proposed approach was evaluated using statistical methods, and the highest results for the classification accuracy, specificity, and sensitivity analysis of the proposed system reached were 95.24%, 91.7%, and 96.17%, respectively. The obtained results show that the proposed intelligent system has a good diagnosis performance and can be applied as a promising tool for the diagnosis of hepatitis.

Keywords Information gain (IG) · The adaptive neuro-fuzzy inference system (ANFIS) · Hepatitis · Diagnosis · Machine learning

Communicated by V. Loia.

✉ Waheed Ahmad
ahdwaheed@hotmail.com; ahdwaheed@gmail.com

✉ Lican Huang
licanhuang@zstu.edu.cn

Ayaz Ahmad
ahdayazb5@awkum.edu.pk

Amjad Iqbal
amjadiqbal@awkum.edu.pk

Muhammad Hamayun
hamayun@awkum.edu.pk

Anwar Hussain
dranwar@awkum.edu.pk

Gauhar Rehman
gauhar@awkum.edu.pk

Salman Khan
salman@awkum.edu.pk

Ubaid Ullah Khan
ubaid.awkum@gmail.com

Dawar Khan
dawar@ia.ac.cn

[1] Department of Computer Science, School of Informatics, Zhejiang Sci-Tech University, Hangzhou, China

[2] Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China

[3] Department of Biotechnology, Abdul Wali Khan University Mardan, Mardan, Pakistan

[4] Department of Agriculture, Abdul Wali Khan University Mardan, Mardan, Pakistan

[5] Department of Botany, Abdul Wali Khan University Mardan, Mardan 23200, Pakistan

[6] Department of Zoology, Abdul Wali Khan University Mardan, Mardan 23200, Pakistan

[7] Department of Computer Sciences, Abdul Wali Khan University Mardan, Mardan, Pakistan

[8] Department of Computer Sciences, Comsats Institute of Information Technology Abbottabad, Islamabad, Pakistan

[9] Institute of Automation, Chinese Academy of Science, Beijing, China

1 Introduction

Hepatitis is a fatal inflammation of the liver and is a direct threat to millions of human lives. It has five main types: hepatitis A, B, C, D, and E. According to the recent World Health Organization (WHO) data, hepatitis B and C infected people are 2 billion and 360 million, respectively. Only these two types of hepatitis infections (B and C) cause 57% of liver cirrhosis cases and 78% of primary liver cancer (Alshamrani and Osman 2017; World Health Organization 2013).

Hepatitis A and E are usually caused by consumption of contaminated food or water, whereas hepatitis B, C, and D occur due to parenteral contact with diseased body fluids of the infected individuals (Avci 2016). Hepatitis infections cause liver inflammation which may lead to even severe conditions such as cirrhosis and cancer leading to the death of the infected person (Ayaz et al. 2014; Schweitzer et al. 2015).

Hepatitis B is one of the most serious types of viral hepatitis which is transmitted through semen, exposure to infected blood, and other body fluids (Bilal et al. 2014). It can be transmitted to the infant from infected mother at the time of childbirth (Fisicaro et al. 2017). Furthermore, contaminated blood transfusion, use of injection during medical procedures, injective drug use, and careless use of hair and nail cutting instruments are among the risk factors of hepatitis B. Recently, as estimated by WHO, worldwide 0.78 million people die each year due to hepatitis B (Lok et al. 2016; Salman et al. 2016).

Exposure to contaminated hepatitis type C virus (HCV)-infected blood and other infected blood products in the course of medical procedures cause hepatitis C. Approximately, 0.7 million people lose their lives because of this type of hepatitis annually. It is considered the main cause of chronic and acute liver infections. It has been reported recently that around 140 million people have chronic hepatitis infection globally (Ali et al. 2014).

Subjects infected with hepatitis type B are at high risk of type D infection also, and the result of double infection may have even worse outcomes. Hepatitis E virus enters the human body through the intestine and develops acute liver failure, also called fulminant hepatitis. Globally, 20 million infections and 0.56 million hepatitis E-related deaths are reported each year (Norton et al. 2017; World Health Organization 2013).

In the field of medicine, disease diagnosis is regarded as a complicated job (Ahmad et al. 2018; Alshamrani and Osman 2017). Identifying a disease from different features is an intricate and multilayer problem which may lead to false assumption and most probably accompanied by impulsive effects. Hence, the effort to utilize knowledge and experience of numerous medical practitioners and clinical data of the patients stored in databases to assist the diagnosis procedure is counted as a valuable option (Ahmad et al. 2017).

To address this issue, we propose a new technique to diagnose hepatitis disease. This technique hybridizes information gain method and adaptive neuro-fuzzy inference system in two stages. In the first stage, information gain reduces the number of features of the preprocessed dataset. It is used to select the most relevant attributes to the disease with respect to weight (using Entropy) which in turns reduces the computation time and increases classification accuracy. While, in the second stage, the selected attributes are fed to ANFIS to diagnose hepatitis disease. This approach is very promising, as compared to the applications reported previously, in terms of reduced computational time, classification complexity, and higher classification accuracy.

2 Literature review

Myriad attempts have been made earlier to diagnose hepatitis using different single and hybrid models. Some of the classification systems used for clinical diagnosis problems are discussed in this section, i.e., Bascil and Temurtas (2011) proposed a multilayer neural network structure and obtained 91.87% classification accuracy. Polat and Güneş (2006) introduced a hybrid approach combining two different techniques (FS-AIRS) to forecast hepatitis disorders and achieved classification accuracy results about 92.59%. Ster and Dobnikar (1996) made several attempts using different techniques such as ASI, MLP-BP, LDA, and FDA and succeeded with 86.4% classification accuracy. Land and Verheggen (2003) applied different methods such as AdaBoost and ANFIS and obtained 57.60% and 59.90% classification accuracy, respectively. Özyıldırım and Yıldırım (2003) used PLM and RBF and obtained 80% classification accuracy. Dogantekin et al. (2009) used LDA-ANFIS and obtained 94.24% classification accuracy. Recently, Prasad et al. (2016) introduced various hybrid methods (rough dataset and machine learning algorithms) for medical diagnosis and obtained 93% accuracy. Liu et al. (2017) presented RFRS a hybrid classification system and obtained 92.59% accuracy.

Performance of the system's classification should be as high as possible in order to have a real-life system (Ahmad et al. 2017; Cahan and Cimino 2017). Disease diagnosis is considered a significant yet intricate task that requires to be accomplished accurately and efficiently (Alickovic and Subasi 2016). Since hepatitis is a deadly disease and the previous diagnosis methods reported in the literature are not very accurate, we were intrigued to propose a hybrid decision support system using information gain method and ANFIS to provide more accurate classification. We combined two different techniques in which each one has its own significance. Information gain method reduces the dimensionality of the dataset, while ANFIS performs classification.

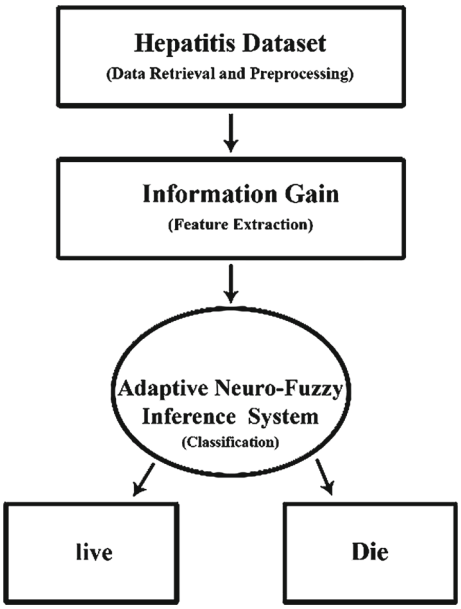

Fig. 1 Block diagram of the proposed system

We used hepatitis dataset to examine the effectiveness of our approach and achieved 95.24% classification accuracy, which proves that our proposed hybrid technique is more successful than the previously applied models and can be utilized as a promising tool for other diseases also.

3 Materials and methods

3.1 Proposed system

In the proposed system, information gain and ANFIS techniques are combined to diagnose hepatitis disorders. Information gain method is defined in the field of machine learning as the term's goodness criterion (Uğuz 2011). Its purpose in this study is to select quality attributes while deselecting unimportant and redundant attributes from the dataset retrieved from the UCI machine learning repository. In the output, it provides a set of attributes whose ranking values are higher than the rest, which is then provided as input to ANFIS to train and test the recommended system. The first stage of the proposed system includes data retrieval and feature extraction. The second stage is called classification stage as shown in the block diagram (Fig. 1).

3.2 Information gain and WEKA

In datasets, every attribute has a specific rank and importance, based on which machine can learn about a certain problem (Witten et al. 2016). Information gain filter "*InfoGainAttributeVal*" attribute evaluator was applied with a searching method "*Ranker-T-1*" on hepatitis dataset to achieve the result of information gain method. This measure is associated with the reduction in entropy in the training set when the value of the feature is known. The algorithms calculate the worth of attribute by computing the information gain according to classification. To approximate the quality of each attribute, the information gain method is implied that uses entropy by means of estimating the difference between prior entropy and post entropy (Ashraf et al. 2010).

The prior entropy of X is given in Eq. (1), where X and Y are considered discrete variables:

$$-\sum_{X} P(X)\log_2 P(X). \tag{1}$$

Here $P(X)$ represents the probability function of X. The conditional entropy of X given by post entropy Y will be as in Eqs. (2) and (3).

$$\text{Entropy}(X|Y) = -\sum_{V} P(Y)\text{Entropy}(X|Y), \tag{2}$$

$$= -\sum_{Y} P(Y)\text{Entropy}(X|Y)\log_2 P(X|Y). \tag{3}$$

The information gain defined in Eqs. (4) and (5) and represented with InfoGain(X;Y):

$$\text{InfoGain}\,(X;Y) = \text{Entropy}\,(X) - \text{Entropy}\,(X|Y)\,, \tag{4}$$

$$\text{InfoGain}\,(X;Y) = -\sum_{X} P\,(X)\log_2 P\,(X) - \sum_{Y}(-P\,(Y) \times \sum_{X} P\,(X|Y)\log_2 P(X|Y)). \tag{5}$$

In this study, we used the Waikato Environment for Knowledge Analysis (WEKA) to compute information gain. WEKA is written in Java language and is an open-source machine learning software that provides an environment for the calculation of information gain (Frank et al. 2004; Hall et al. 2009). It comprises of various machine learning and data-mining methods for data processing, visualization, association, classification, clustering, and regression. We propose that combining information gain and adaptive neuro-fuzzy inference system for the diagnosis of hepatitis can be a potential and efficient alternative for the physician to decide about the infection.

3.3 Adaptive neuro-fuzzy inference system

ANFIS (adaptive neuro-fuzzy inference system) is the combination of neural network and fuzzy inference system. To train a hybrid learning algorithm, ANFIS uses least square

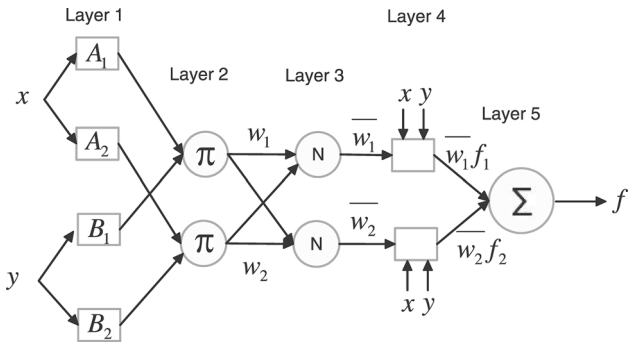

Fig. 2 ANFIS architecture

estimate (LSE) with gradient descent method (Adeli et al. 2013; Jang 1993, 1996). One cycle of hybrid learning algorithm consists of two passes: forward and backward pass. LSE method is used to identify consequent parameters, while a signal travels until layer 4 in the forward pass. Gradient descent then performs updating of premise parameters, whereas error propagates backward. In order to achieve the lowest possible error, the same process is repeated again and again (Kalaiselvi and Nasira 2014). To understand ANFIS architecture, suppose a fuzzy system with two input x and y. The fuzzy system comprises of two Sugeno fuzzy rules:

$$\text{IF } x \text{ is } A_1 \text{ AND } y \text{ is } B_1, \text{ THEN } f_1 = p_1 x + q_1 y + r_1, \quad (6)$$

$$\text{IF } x \text{ is } A_2 \text{ AND } y \text{ is } B_2, \text{ THEN } f_2 = p_2 x + q_2 y + r_2. \quad (7)$$

In above rules 06 and 07, given inputs are indicated by x and y, fuzzy sets by A_i and B_i, output specified by fuzzy rules indicated by f_i and design parameters are denoted by p_i, q_i, and r_i. ANFIS structure is shown in Fig. 2. There are five layers, and each contains a different number of nodes. Nodes in the same layer have the same functions.

Layer [1]: Each square node i represented by A_i in layer[1] has a node function as described in Eq. (8).

$$o_i^1 = \mu_{A_i}(x). \quad (8)$$

In Eq. (8), x represents input to node i. Here o_i^1 denotes A_i membership function. It specifies the degree to which x satisfies the A_i quantifier. $\mu_{A_i}(x)$ is a generalized bell-shaped function as defined in Eq. (9). Here a_i and c_i are premise parameters.

$$\mu_{A_i}(x) = \exp\left[-\left(\frac{x - c_i}{a_i}\right)^2\right] \quad (9)$$

Layer[2]: Nodes are represented by circles and labeled with "π" in Fig. 2. It takes layer[1]'s outputs as input and multiplies them to produce weight. The output of layer[2] indicates the firing strength of the rules.

$$\bar{w} = \mu_{A_i}(x) \times \mu_{B_i}(x), \quad i = 1, 2 \quad (10)$$

Layer[3]: It consists of circled nodes labeled with "N" compute implication of each output member function by normalizing weight of a certain node comparing with the weights of other nodes.

$$\bar{w} = \frac{w_i}{\sum_j w_j} \quad i = 1, 2. \quad j = 2 \quad (11)$$

Layer[4]: This layer is represented by square nodes. Equation (12) describes how to write a linear format for the output of a rule based on Sugeno inference system. In this equation, r_i represents bias and p_i and q_i indicates consequent parameters.

$$o_i^4 = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i) \quad (12)$$

Layer[5]: This is the aggregation layer. It computes the summation of rules and produces a single output.

$$o_i^5 = \text{output} = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (13)$$

3.4 Data retrieval and preprocessing

Hepatitis dataset provided by Gail Gong (collected from Carnegie-Mellon University) was acquired from the UCI machine learning repository (Alshamrani and Osman 2017). The reason behind using this dataset was to compare our experimental results with previously used techniques for hepatitis diagnosis. The dataset comprises 155 records of hepatitis patients with 19 attributes for each patient. Table 1 shows attributes of hepatitis dataset and its values provided by the repository. There are two types of records in the class attributes of this dataset based on the condition of the patient, i.e., live class (represented with 1) and die class (represented as 0). The dataset contains 123 (79.4%) records of "live" class and 32 (20.6%) records of "die" class as shown in Table 2.

Hepatitis dataset contains a matrix of 155 by 19, where (155) rows denote the number of patients (records) and columns (19) denote the number of experiments. In this work, experiments are called as attributes. The dataset was preprocessed, and attribute values were normalized to make it suitable for the mining process.

Table 1 Attributes of hepatitis disease dataset

S. no.	Attribute	Values of attribute
1	Age	10, 20, 30, 40, 50, 60, 70, 80
2	Sex	Male, female
3	Steroid	Yes, no
4	Antivirals	Yes, no
5	Fatigue	Yes, no
6	Malaise	Yes, no
7	Anorexia	Yes, no
8	Liver big	Yes, no
9	Liver firm	Yes, no
10	Spleen palpable	Yes, no
11	Spiders	Yes, no
12	Ascites	Yes, no
13	Varices	Yes, no
14	Bilirubin	0.39, 0.80, 1.20, 2.00, 3.00, 400
15	Alk phosphate	33, 80, 120, 160, 200, 250
16	SGOT	13, 100, 200, 300, 400, 500
17	ALBUMIN	2.1, 3.0, 3.8, 4.5, 5.0, 6.0
18	PROTIME	10, 20, 30, 40, 50, 60, 70, 80, 90
19	HISTOLOGY	Yes, no

Table 2 Class distribution

Class values	Number of instances
1 (live)	123
0 (die)	32

4 Experimental results

Experiments were performed on hepatitis dataset to examine the effectiveness of our proposed technique. The dataset was acquired from the UCI machine learning repository, and attribute values of the dataset were normalized for the next phases of the approach as discussed in Sect. 3.4. After preprocessing the dataset went through two phases. In the first phase of the proposed technique, the dimension of the hepatitis dataset was reduced using information gain method. Information gain filter "*InfoGainAttributeVal*" was applied with a searching method *Ranker-T-1* on hepatitis dataset to achieve the results (details Sect. 3.2).

Information gain is one of the popular approaches for reducing the number of features (Uğuz 2011). It aims at ranking the subset of features based on high information gain entropy in decreasing order. In this phase, the effects of individual features ranking operation by the IG method on classifier performance are examined. Waikato Environment for Knowledge Analysis (WEKA) application version 3.8 was used for the said attribute selection of hepatitis dataset. Out of 19 attributes, eight attributes were selected which showed the highest ranking and were recommended to be applied as an input for the ANFIS system. Sample attributes with ranking values are shown in Fig. 3. The dataset with selected attributes was divided into two sets for the final phase, i.e., training and testing set. Training set shared 35% and testing set shared 65% of the actual dataset's sample size.

In the second part, we applied the Sugeno fuzzy inference system to construct a fuzzy inference system (FIS). As shown in Fig. 4, FIS maps attributes to attributes membership functions (MF), attributes MP maps to rules, rules maps to output, output maps to output MF, and the output MP to output; which is a single-valued output.

ANFIS is a hybrid of two intelligent system models. It combines the low-level computational power of a neural network with the high-level reasoning capability of a fuzzy inference system. This means that it is combining the best aspects of the two technologies while limiting the drawbacks (Nazmy et al. 2010).

The specific advantages of ANFIS hybrid system are:

- ANFIS uses the neural network's ability to classify data and find patterns.
- It then develops a fuzzy expert system that is more transparent to the user and also less likely to produce memorization errors than a neural network.
- Furthermore, ANFIS keeps the advantages of a fuzzy expert system, while removing (or at least reducing) the need for an expert.

In the proposed experimental study, eight inputs (y_1, y_2, $y_3, y_4, y_5, y_6, y_7, y_8$) were provided to ANFIS which produced one output (z). The expression given below is a rule set for first-order Sugeno fuzzy model with base fuzzy if–then rules.

If $y_1 H_1$ and $y_2 G_1$ and $y_3 F_1$ and $y_4 E_1$ and $y_5 D_1$ and $y_6 C_1$ and $y_7 B_1$ and $y_8 A_1$, then

$$f_1 = ky_1 + kky_2 + my_3 + mmy_4 + ny_5 + nny_6 + py_7 + ppy_8 + u \quad (14)$$

Linear output parameters are represented by k, kk, m, mm, n, nn, p, pp, u in expression (14). Structure of the proposed approach is given in Fig. 5. Figure 6 represents MATLAB ANFIS structure.

In order to evaluate the classification accuracy and verify the validity of our proposed technique, we used k-fold cross-validation approach. The value of k in this study was set to 10. The classification accuracy for the information gain-ANFIS was found to be 95.24% for hepatitis disease.

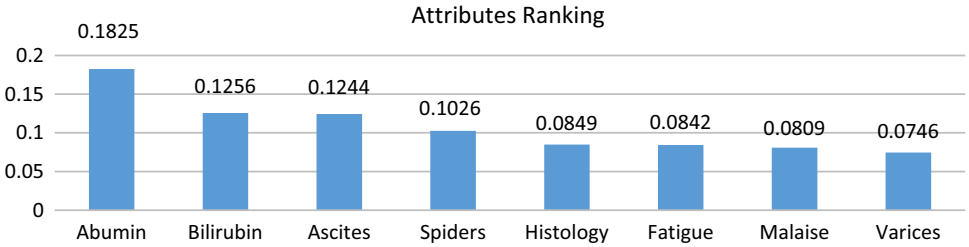

Fig. 3 Information gain ranking of selected attributes

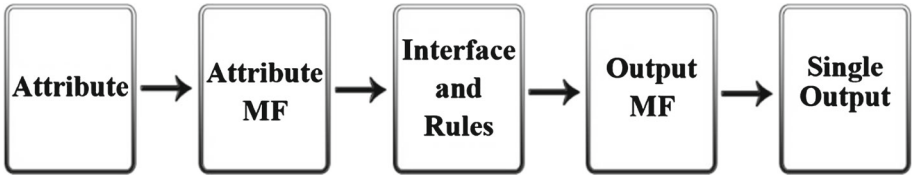

Fig. 4 Fuzzy inference system structure

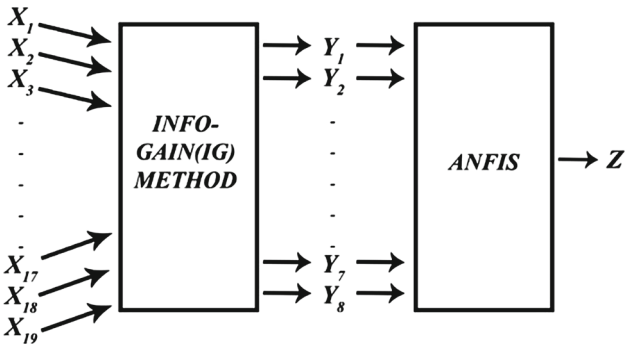

Fig. 5 Structure of the proposed approach

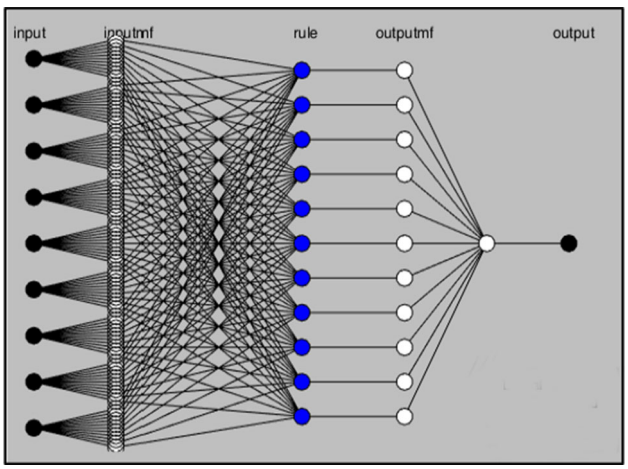

Fig. 6 ANFIS structure on MATLAB

4.1 Specificity and sensitivity analysis

Specificity and sensitivity of information gain-ANFIS were calculated for the exact values of hepatitis diagnosis. Equations (15) and (16) are used to calculate the sensitivity and specificity (Ashraf et al. 2010; Dogantekin et al. 2009).

$$\text{Sensitivity} = \frac{\text{TP(true positive)}}{\text{TP(true positive)} + \text{FN(false negative)}}(\%) \tag{15}$$

$$\text{Specificity} = \frac{\text{TN(true negative)}}{\text{FP(false postive)} + \text{TN(true negative)}}(\%) \tag{16}$$

In Eqs. (15) and (16), TP (true positive) indicates that according to the automatic clinician's optic nerve diagnosis, the input was considered "die". While true negative (TN) represents that input was considered "live" and was also labeled as "live" by the automatic clinicians. False positive (PF) indicates that input was labeled as "live" but was considered "die" by the automatic clinicians, and false negative (FN) shows that according to automatic clinicians' input was considered as "live" with an optic nerve diagnosis.

The values obtained using specificity and sensitivity analysis for information gain (IG)-ANFIS diagnosis system were 91.7% and 96.17%, respectively.

4.2 Classification accuracy

Classification accuracy for the proposed approach was measured with the help of Eq. (17).

$$\text{diag-accuracy}(A) = \frac{\sum_{i-1}^{|A|} \text{assess}(a_i)}{|A|}, \quad a_i \in A \tag{17}$$

$$\text{assess}(a) = \begin{cases} 1, & \text{if classify}(a) = a \cdot c \\ 0, & \text{otherwise,} \end{cases}$$

In Eq. (17), "A" represents test set of data items to be classified, $a_i \epsilon A$; $a \cdot c$ indicates item "a" class and classification

Table 3 Classification accuracies obtained using hepatitis diagnosis systems

Method used	Author of the article	Classification accuracy (%)
AdaBoost	Land et al.	57.60
ANFIS	Land et al.	59.90
MLP	Özyıldırım and Yıldırım	74.37
MLP + BP (Tooldiag)	Adamczak	77.4
ASI	Stern and Dobnikar	82.0
MLP with BP	Stern and Dobnikar	82.1
RBF	Özyıldırım and Yıldırım	83.75
LDA	Stern and Dobnikar	86.4
PCA-AIRS	Polat and Gunes	94.12
LDA-ANFIS	Dogantekin et al.	94.16
RFRS	Liu et al.	92.59
Rough dataset	Prasad et al.	93
IG-ANFIS	Proposed approach	95.24

of a_i by ANFIS classifier returns class (a_i). The classification accuracy value obtained using information gain-ANFIS method was 92.24%.

The results of information gain-ANFIS for hepatitis are compared with earlier methods. Classification accuracies of information gain-ANFIS system and other previous hepatitis diagnosis systems are given in Table 3.

5 Discussion and conclusion

In this paper, we proposed a novel hybrid medical diagnosis system which combines information gain and ANFIS method to diagnose hepatitis disease. Attributes were reduced using information gain to an optimal number and then fed the novel set of attributes to the adaptive neuro-fuzzy inference system (ANFIS) to diagnose hepatitis disorders. We got the promising result of classification accuracy, specificity, and sensitivity through the proposed approach which was 95.24%, 91.7%, and 96.17%, respectively. Comparison with earlier diagnosing methods is shown in Table 3 which proves that the proposed system is better than the previous diagnosis methods reported in the literature.

In the future, we will further improve this method to reduce the time of computation, increase the accuracy of the system and apply it to more datasets. In addition, a desktop application can be made for easy and efficient hepatitis diagnosis.

Compliance with ethical standards

Conflict of interest All authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

Adeli M, Bigdeli N, Afshar K (2013) New hybrid hepatitis diagnosis system based on genetic algorithm and adaptive network fuzzy inference system. In: 2013 21st Iranian conference on electrical engineering (ICEE), 2013. IEEE, pp 1–6

Ahmad W, Huang L, Ahmad A, Shah F, Iqbal A (2017) Thyroid diseases forecasting using a hybrid decision support system based on ANFIS, k-NN and information gain method. J Appl Environ Biol Sci 7:78–85

Ahmad W, Ahmad A, Lu C, Khoso BA, Huang L (2018) A novel hybrid decision support system for thyroid disease forecasting. Soft Comput 22:1–7

Ali S, Ahmad A, Khan RS, Khan S, Hamayun M, Khan SA, Iqbal A, Khan AA, Wadood A, Ur Rahman T, Baig AH (2014) Genotyping of HCV RNA reveals that 3a is the most prevalent genotype in mardan, pakistan. Adv Virol. https://doi.org/10.1155/2014/606201

Alickovic E, Subasi A (2016) Medical decision support system for diagnosis of heart arrhythmia using DWT and random forests classifier. J Med Syst 40:108

Alshamrani BS, Osman AH (2017) Investigation of hepatitis disease diagnosis using different types of neural network algorithms. Int J Comput Sci Netw Secur (IJCSNS) 17:242

Ashraf M, Le K, Huang X (2010) Information gain and adaptive neuro-fuzzy inference system for breast cancer diagnoses. In: 2010 5th international conference on computer sciences and convergence information technology (ICCIT), 2010. IEEE, pp 911–915

Avci D (2016) An automatic diagnosis system for hepatitis diseases based on genetic wavelet kernel extreme learning machine. J Electr Eng Technol 11:993–1002

Ayaz A et al (2014) Computational analysis reveals three micro-RNAs in hepatitis A virus genome. J Appl Environ Biol Sci 4:34–39

Bascil MS, Temurtas F (2011) A study on hepatitis disease diagnosis using multilayer neural network with levenberg marquardt training algorithm. J Med Syst 35:433–436

Bilal M et al (2014) Computational prediction of micro-RNAs in hepatitis B virus genome. J Appl Environ Biol Sci 4:106–113

Cahan A, Cimino JJ (2017) A learning health care system using computer-aided diagnosis. J Med Internet Res 19:11282–11286

Dogantekin E, Dogantekin A, Avci D (2009) Automatic hepatitis diagnosis system based on linear discriminant analysis and adaptive network based on fuzzy inference system. Expert Syst Appl 36:11282–11286

Fisicaro P et al (2017) PS-053-Proteasome dysfunction as a reversible defect underlying virus-specific CD8 cell exhaustion in chronic hepatitis B. J Hepatol 66:S30

Frank E, Hall M, Trigg L, Holmes G, Witten IH (2004) Data mining in bioinformatics using Weka. Bioinformatics 20:2479–2481

Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH (2009) The WEKA data mining software: an update. ACM SIGKDD Explor Newsl 11:10–18

Jang J-S (1993) ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans Syst Man Cybern 23:665–685

Jang J-S (1996) Input selection for ANFIS learning. In: Proceedings of the fifth IEEE international conference on fuzzy systems, 1996. IEEE, pp 1493–1499

Kalaiselvi C, Nasira G (2014) A new approach for diagnosis of diabetes and prediction of cancer using ANFIS. In: 2014 World congress on computing and communication technologies (WCCCT), 2014. IEEE, pp 188–190

Land W, Verheggen E (2003) Experiments using an evolutionary programmed neural network with adaptive boosting for computer aided diagnosis of breast cancer. In: Proceedings of the 2003 IEEE international workshop on soft computing in industrial applications, 2003. SMCia/03, 2003. IEEE, pp 167–172

Liu X, Wang X, Su Q, Zhang M, Zhu Y, Wang Q, Wang Q (2017) A hybrid classification system for heart disease diagnosis based on the RFRS method. Comput Math Methods Med. https://doi.org/10.1155/2017/8272091

Lok AS et al (2016) Antiviral therapy for chronic hepatitis B viral infection in adults: a systematic review and meta-analysis. Hepatology 63:284–306

Nazmy T, El-Messiry H, Al-Bokhity B (2010) Adaptive neuro-fuzzy inference system for classification of ECG signals. In: 2010 The 7th international conference on informatics and systems (INFOS), 2010. IEEE, pp 1–6

Norton B, McMurry C, Gover M, Cunningham C, Litwin A (2017) THU-238-feasibility and acceptability of a group medical visit intervention to improve hepatitis C virus treatment uptake among persons who inject drugs (PWID) in a primary care setting. J Hepatol 66:S294–S295

Ozyilmaz L, Yildirim T (2003) Artificial neural networks for diagnosis of hepatitis disease. In: Proceedings of the international joint conference on neural networks, 2003. IEEE, pp 586–589

Polat K, Güneş S (2006) A hybrid medical decision making system based on principles component analysis, k-NN based weighted pre-processing and adaptive neuro-fuzzy inference system. Digit Signal Proc 16:913–921

Prasad V, Rao TS, Babu MSP (2016) Thyroid disease diagnosis via hybrid architecture composing rough data sets theory and machine learning algorithms. Soft Comput 20:1179–1189

Salman, Shahzad, Riaz A, Waheed A (2016) The mobile apps and literature review on the major causes of deaths according to WHO (World Health Organization). J Appl Environ Biol Sci 6:16–24

Schweitzer A, Horn J, Mikolajczyk RT, Krause G, Ott JJ (2015) Estimations of worldwide prevalence of chronic hepatitis B virus infection: a systematic review of data published between 1965 and 2013. Lancet 386:1546–1555

Šter B, Dobnikar (1996) A neural networks in medical diagnosis: comparison with other methods. In: International conference on engineering applications of neural networks, 1996. pp 427–430

Uğuz H (2011) A two-stage feature selection method for text categorization by using information gain, principal component analysis and genetic algorithm. Knowl-Based Syst 24:1024–1032

Witten IH, Frank E, Hall MA, Pal CJ (2016) Data Mining: practical machine learning tools and techniques. Morgan Kaufmann, Burlington

World Health Organization (2013) Hepatitis. World Health Organization. http://www.who.int/immunization/topics/hepatitis/en/. Accessed 9 Feb 2017