A Type-2 Fuzzy Expert System for Diagnosis of Leukemia

Ali Akbar Sadat Asl(✉) and Mohammad Hossein Fazel Zarandi

Department of Industrial Engineering,
Amirkabir University of Technology, Tehran, Iran
{a.sadatasl, zarandi}@aut.ac.ir

Abstract. Medical field, especially in the diagnosis and treatment, is facing with inherent uncertainty. Causes of leukemia can be different factors that determining of them is with uncertainty. Owing to the high potential of the fuzzy expert systems for managing uncertainty associated to the medical diagnosis, in this paper, we propose a type-2 fuzzy expert system for Leukemia diagnosis. In this system, we use Mamdani-style inference that has high interpretability to clarify the results of system to experts. The classification accuracy of the type-2 fuzzy system for Leukemia diagnosis has obtained about 94% which demonstrate its capability for helping experts to early diagnosis of the disease.

Keywords: Leukemia · Type-2 fuzzy · Expert system

1 Introduction

1.1 Leukemia

Leukemia is a cancer that affects the blood and bone marrow where blood cells are made. Usually, Leukemia involves the production of abnormal white blood cells. The cells are responsible for fighting infection. However, the abnormal cells in Leukemia do not function in the same way as normal white blood cells. The Leukemia cells continue to grow and divide, eventually crowding out the normal blood cells. The end result is that it becomes difficult for the body to fight infections, control bleeding, and transport oxygen. Leukemia is a general term for four types of malignant disease of the blood and bone marrow [1].

Leukemia can be described as fast-growing (acute) or slow growing (chronic). The different types of Leukemia have varied outlooks and treatment options. There are two main types of acute Leukemia containing: acute myeloid Leukemia (AML) and acute lymphoblastic Leukemia (ALL). Also, there are three main types of chronic Leukemia containing: chronic myeloid Leukemia (CML), chronic lymphocytic Leukemia (CLL) and hairy cell Leukemia (HCL).

Chronic Leukemias are generally slow-developing, long-term conditions. Hairy cell Leukemia is a very rare type of chronic Leukemia. The most commonly diagnosed Leukemia in adults is CLL and AML [2].

P. Melin et al. (eds.), *Fuzzy Logic in Intelligent System Design*,
Advances in Intelligent Systems and Computing 648, DOI 10.1007/978-3-319-67137-6_6

1.2 Expert System

Expert systems are programs for reconstructing the expertise and reasoning capabilities of qualified specialists within limited domains. Expert systems require detailed information about a special domain and the strategies for applying the information to problem solving. To construct an expert system, the knowledge should be formalized, represented in the computer and manipulated according to some problem-solving methods [3].

Any expert system consists of a knowledge base, a database and an inference engine. These three units, together with some interface for communicating with the user, form the minimal configuration that may still be called an expert system. The knowledge base contains general knowledge related to the problem domain. The purpose of the database is to store data for each specific task of the expert system. The inference engine of an expert system operates on a series of production rules and makes inferences [4] (Fig. 1).

Klir and Yuan [4] considered below architecture for an expert system:

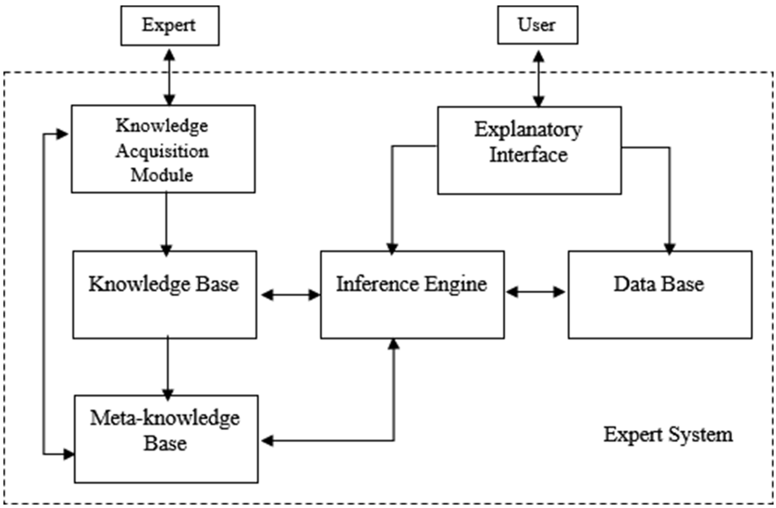

Fig. 1. Architecture of an expert system [4]

1.3 Type-2 Fuzzy

It is known that type-2 fuzzy sets let us model and minimize the effects of uncertainties in rule-based fuzzy logic systems (FLSs) [5]. There are at least four sources of uncertainties in type-1 FLSs: (1) the meanings of the words that are used in the antecedents and consequents of rules can be uncertain. (2) Consequents may have a histogram of values associated with them, especially when knowledge is extracted from a group of experts who do not all agree. (3) Measurements that activate a type-1 FLS may be noisy and therefore uncertain. (4) The data that are used to tune the parameters of a type-1 FLS may also be noisy. All of these uncertainties translate into uncertainties about fuzzy set membership functions. Type-1 fuzzy sets are not able to directly model such uncertainties because their membership functions are totally crisp. On the other hand, type-2 fuzzy sets are able to model such uncertainties because their membership functions are themselves fuzzy [6] (Fig. 2).

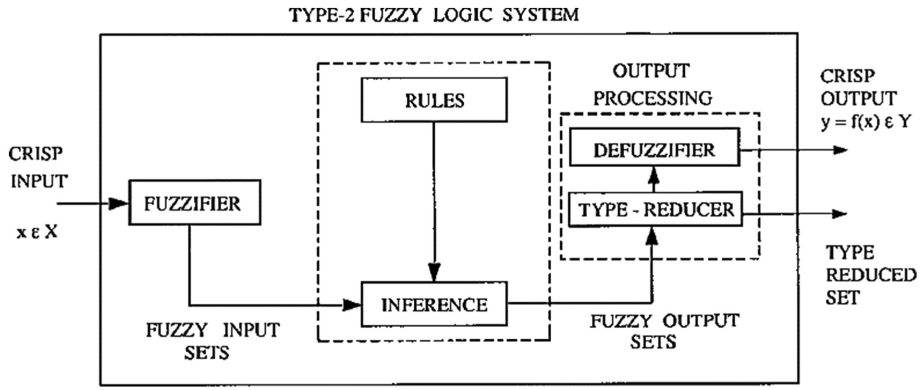

Fig. 2. Components of type-2 fuzzy logic system [7]

Medical field, especially in the diagnosis and treatment, is facing with inherent uncertainty. Causes of Leukemia can be different factors that determining the correspondence between Leukemia and its causes are with uncertainty. In other words, doctor diagnosis is with the uncertainty which it can affect diagnosis results and all treatment process and if a mistake is made, it can result in irreparable damage to the patient. Therefore, in this paper, we present a type-2 fuzzy intelligent system that is capable of handling uncertainties in the diagnosis process of Leukemia.

2 Literature Review

Medical issues such as the diagnosis are always associated with uncertainty. Using the expert systems by different logics can assist the experts in any time. There are numerous expert systems in medical fields, for example, CREAM systems in the field of cardiology, DIAS in the field of diabetes and MYCIN to diagnose bacterial infections. Because of the capability of fuzzy logic in uncertainty modeling we focus on the papers which fuzzy logic has been used to inference.

Polat and Güneş detected on diabetes disease using principal component analysis (PCA) and adaptive neuro-fuzzy inference system (ANFIS). The aim of their study is to improve the diagnostic accuracy of diabetes disease combining PCA and ANFIS [8]. Muthukaruppan and Er presented a particle swarm optimization (PSO)-based fuzzy expert system for the diagnosis of coronary artery disease (CAD) [9]. Keleş et al. developed an expert system for diagnosis of breast cancer. In their system, the fuzzy rules which used in inference engine were found by using neuro-fuzzy method [10]. Hayashi proposed a fuzzy neural network and the learning method using fuzzy teaching input. As an application, a fuzzy neural expert system (FNES) for diagnosing hepatobiliary disorders has been developed [11]. Biyouki et al. presented a fuzzy rule-based expert system for diagnosis thyroid's disease. This proposed system includes three steps: pre-processing (feature selection), neuro-fuzzy classification and system evaluating [12]. Maftouni et al. designed a type-2 fuzzy rule-based expert system for ankylosing spondylitis diagnosis. In this system, the medical expertise and evidences are used for simulating the expert's manner in diagnosis [13]. Zarandi et al. developed a type-II fuzzy expert system for brain tumor imageprocessing. The main contributions in this paper were the aggregation of the available image pre-processing methods, development of a Type-II fuzzy cluster analysis for segmentation, and presenting a Type-II fuzzy expert system for approximate reasoning [14].

In the field of Leukemia diagnosis some papers by using fuzzy expert systems have been presented which we provide an overview of these papers. Obi and Imianvan presented a hybrid neuro fuzzy expert system to help in diagnosis of Leukemia using a set of symptoms. The designed system is an interactive system that tells the patient his current condition as regards Leukemia [15]. Azar and Alizadeh proposed an expert system to diagnose and recommend treatment method for Leukemia [16]. Latifi et al. introduced a fuzzy inference system (FIS) for diagnosing of acute lymphocytic Leukemia in children. The fuzzy expert system applies Mamdani reasoning model that has high interpretability to explain system results to experts in a high level. The system has been designed based on the specialist physician's knowledge [17].

3 Methodology

3.1 Leukemia Dataset

The procedure of diagnosing a patient suffering from Leukemia is synonymous to the general approach to medical diagnosis. The physician may carry out a precise diagnosis, which requires a complete physical evaluation to determine whether the patient have Leukemia. The examining physician accounts for possibilities of having Leukemia through an interview, physical examination and laboratory test. Many primary health care physicians may require tools for Leukemia evaluation [15].

In this study, the Leukemia dataset obtained according to the Obi and Imianvan [15]. The purpose of the dataset is to predict the presence or suspicion of presence or absence of the Leukemia disease given the results of various medical tests carried out on a patient. If the patient is having five or more of the symptoms, he is having severe Leukemia and should go for treatment urgently. If it is approximately four of the symptoms he is having, he might be suffering from Leukemia and hence should see a physician right away, but if it is three or lesser of the symptoms, he may not be having Leukemia. This dataset contains 14 attributes. The dataset contains 500 samples belonging to three different classes (274 "with Leukemia" cases, 100 "Might be Leukemia" cases, 126 "Not Leukemia" cases).

To design our type-2 fuzzy system for diagnosis of Leukemia, we designed a system which consists of a set of symptoms needed for the diagnosis. The Leukemia symptoms have been shown in Table 1.

Table 1. Leukemia symptoms [14]

No.	Symptom	No.	Symptom
1	Paleness	8	Thrombocytopenia
2	Shortness of breath	9	Granulocytopenia
3	Nose bleeding	10	Asthenia
4	Frequent infection	11	Palpitations
5	Anemia	12	Digestive bleeding
6	Epistaxis	13	Enlargedspleen
7	Bone pain	14	Fatigue

3.2 Determining the Number of Rules

We should use a cluster validity index to determine the most suitable number of clusters. In this work, the validity index proposed by Zarandi et al. is applied. This validity index V_{ECAS} (an Exponential compactness and separation index) can find the number of clusters as the maximum of its function with respect to c. This index is defined as [18]:

$$\boldsymbol{V_{ECAS} = ECAS(c) = \frac{EC_{comp}(c)}{max_c\left(EC_{comp}(c)\right)} - \frac{ES_{sep}(c)}{max_c\left(ES_{sep}(c)\right)}}, \tag{1}$$

where $EC_{comp}(c)$ and $ES_{sep}(c)$ are Exponential compactness and Exponential separation measures, respectively and are defined as follows [18]:

$$\boldsymbol{EC_{comp}(c) = \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^{m} \exp[-(\frac{\left\|x_i - v_j\right\|^2}{\beta_{comp}} + \frac{1}{c+1})]}, \tag{2}$$

$$ES_{sep}(c) = \sum_{i=1}^{c} \exp\left[-min_{i\neq k}\left(\frac{(c-1)\|v_i - v_k\|^2}{\beta_{sep}}\right)\right]. \tag{3}$$

β_{comp} is defined as the sample covariance for cluster i, i.e. [18]:

$$\boldsymbol{\beta_{comp} = \frac{\sum_{k=1}^{n}\left\|x_k - \bar{v}\right\|^2}{n(i)}}, \tag{4}$$

where $n(i)$ is the number of data in cluster i.

β_{sep} is defined as the total average distance measure for all clusters, i.e. [18]:

$$\beta_{sep} = \frac{\sum_{l=1}^{n} \|v_l - \bar{v}\|^2}{c}, \tag{5}$$

with $\bar{v} = \frac{\sum_{j=1}^{n} x_j}{n}$.

We apply this cluster validity index to determine the most suitable number of clusters or rules. The best number of clusters based on this cluster validity index is obtained in three clusters.

3.3 The Proposed Type-2 Fuzzy Model

For many application problems, classifiers can be used to support a decision-making process. In some areas like medical, it is not preferable to use black box approaches. The user should be able to understand the classifier and to evaluate its results. Fuzzy rule-based classifiers are especially suitable because they consist of simple linguistically interpretable rules and do not have some drawbacks of symbolic or crisp rule-based classifiers. Classifiers must often be created from data by a learning process because there is not enough expert knowledge to determine their parameters completely [19].

In the Type-2 fuzzy model, we obtain the model with three rules, fourteen inputs, and one output. The inputs are Leukemia symptoms which presented in Table 1. A universal set of symptoms of Leukemia disease is set up for diagnosis where the patient is expected to pick from the set of symptoms fed into the system. We use Mamdani-style inference, min–max operators and centroid defuzzification methods. In the proposed model, Gaussian membership function was used for fuzzy sets description. The rule-based of the proposed system consists of three general rules. The rules of the proposed system are as follows:

If (PALENESS is in1cluster c) and (SHORTNESS OF BREATH is in2cluster c) and (NOSE BLEEDING is in3cluster c) and (FREQUENT INFECTION is in4cluster c) and (ANEMIA is in5cluster c) and (EPISTAXIS is in6cluster c) and (BONE PAIN is in7cluster c) and (THROMBOCYTOPENIA is in8cluster c) and (GRANULOCYTOPENIA is in9cluster c) and (ASTHENIA is in10cluster c) and (PALPITATIONS is in11cluster c) and (DIGESTIVE BLEEDING is in12cluster c) and (ENLARGE SPLEEN is in13cluster c) and (FATIGUE is in14cluster c) then (output is out1cluster c), where c = $\{1, 2, 3\}$.

Figure 3 represents the fuzzy rules of the proposed system.

3.4 Performance Evaluation

For performance evaluation of the proposed system, the dataset divided into two sets containing: The training set and the test set which include 400 and 100 samples, respectively. These samples are applied to demonstrate the performance of the proposed system. The classification accuracy of the type-2 fuzzy system for Leukemia diagnosis has obtained about 94%.

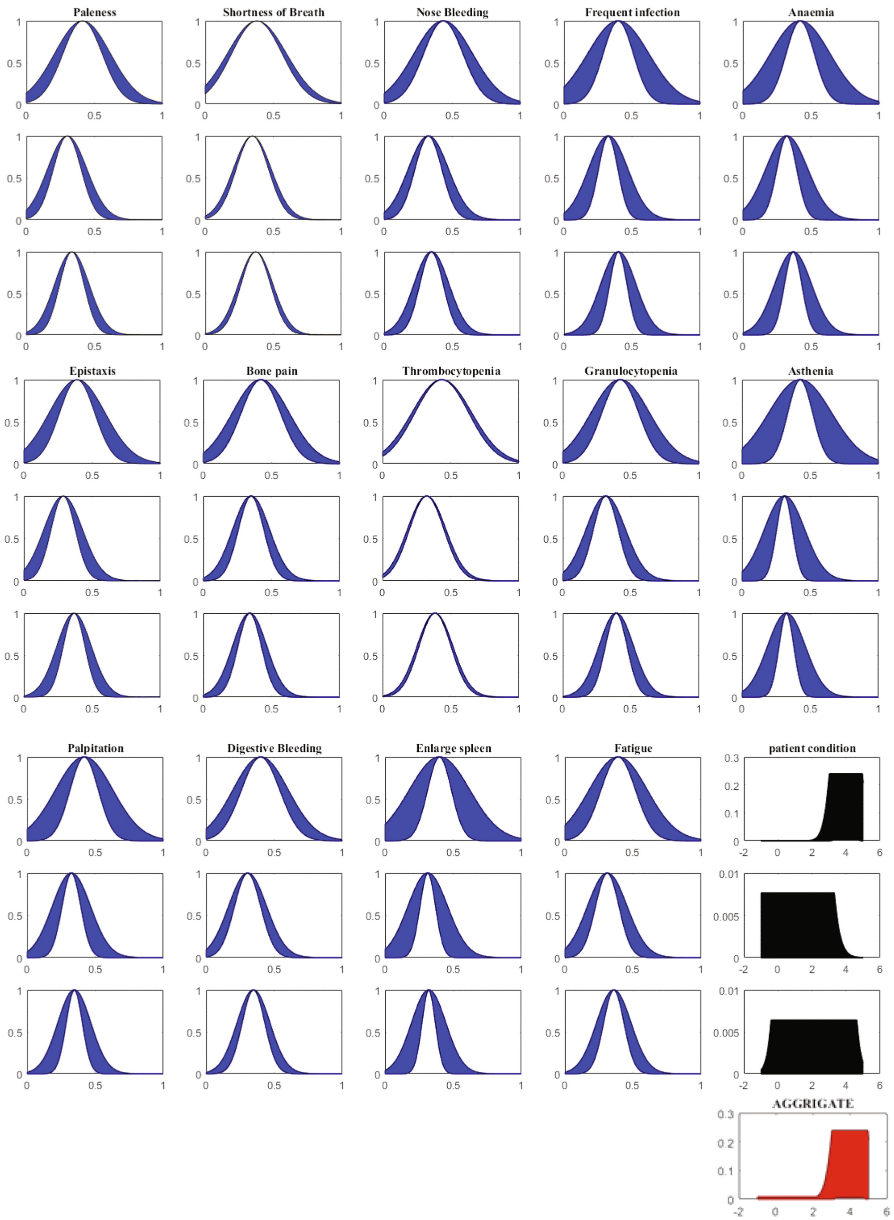

Fig. 3. Type-2 fuzzy rule based

4 Conclusion

In this paper, we proposed a type-2 fuzzy expert system to Leukemia diagnosis. In this system, for simulating the expert's manner in diagnosis, the medical expertise and evidences are used. Because of the structure and semantic of Leukemia diagnosis, which is with uncertainty, we used the type-2 fuzzy for uncertainty modeling. By relying on the results, the type-2 fuzzy expert system can diagnose Leukemia with the average accuracy of about 94%.

References

1. MedicineNet. Leukemia (2011). http://www.medicinenet.com/
2. Cancertutor Website. https://www.cancertutor.com/types-of-leukemia/
3. Puppe, F.: Systematic Introduction to Expert Systems, Knowledge Representations and Problem-Solving Methods. Springer, Heidelberg (1993)
4. Klir, G.J., Yuan, B.: Fuzzy Sets and Fuzzy Logic: Theory and Application. Prentice Hall PTR, Upper Saddle River (1995)
5. Mendel, J.M.: Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions. Prentice Hall PTR, Upper Saddle River (2001)
6. Mendel, J.M., John, R.B.: Type-2 fuzzy sets made simple. IEEE Trans. Fuzzy Syst. **10**(2), 117–127 (2002)
7. Karnik, N.N., Mendel, J.M., Liang, Q.: Type-2 fuzzy logic systems. IEEE Trans. Fuzzy Syst. **7**(6), 643–658 (1999)
8. Polat, K., Güneş, S.: An expert system approach based on principal component analysis and adaptive neuro-fuzzy inference system to diagnosis of diabetes disease. Digit. Sig. Proc. **17**(4), 702–710 (2007)
9. Muthukaruppan, S., Er, M.J.: A hybrid particle swarm optimization based fuzzy expert system for the diagnosis of coronary artery disease. Expert Syst. Appl. **39**(14), 11657–11665 (2012)
10. Keleş, A., Keleş, A., Yavuz, U.: Expert system based on neuro-fuzzy rules for diagnosis breast cancer. Expert Syst. Appl. **38**(5), 5719–5726 (2011)
11. Hayashi, Y.: Neural expert system using fuzzy teaching input and its application to medical diagnosis. Inf. Sci.-Appl. **1**(1), 47–58 (1994)
12. Biyouki, S.A., Turksen, I.B., Zarandi, M.F.: Fuzzy rule-based expert system for diagnosis of thyroid disease. In: 2015 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB). IEEE (2015)
13. Maftouni, M., et al.: Type-2 fuzzy rule-based expert system for ankylosing spondylitis diagnosis. In: 2015 Annual Conference of the North American Fuzzy Information Processing Society (NAFIPS) held Jointly with 2015 5th World Conference on Soft Computing (WConSC). IEEE (2015)
14. Zarandi, M.F., Zarinbal, M., Izadi, M.: Systematic image processing for diagnosing brain tumors: a Type-II fuzzy expert system approach. Appl. Soft Comput. **11**(1), 285–294 (2011)
15. Obi, J.C., Imianvan, A.A.: Interactive neuro-fuzzy expert system for diagnosis of leukemia. Glob. J. Comput. Sci. Technol. **11**(12), 73–80 (2011)
16. Azar, A.G., Alizadeh, Z.M.: Designing an Expert System to Diagnose and Propose about Therapy of Leukemia. Int. J. Comput. Inf. Technol. (IJOCIT) (2013). ISSN 2345-3877

17. Latifi, F., Hosseini, R., Mazinai, M.: A fuzzy expert system for diagnosis of acute lymphocytic leukemia in children. Int. J. Inf. Secur. Syst. Manag. **4**(2), 424–429 (2015)
18. Zarandi, M.F., Faraji, M.R., Karbasian, M.: An exponential cluster validity index for fuzzy clustering with crisp and fuzzy data. Sci. Iranica. Trans. E Ind. Eng. **17**(2), 95 (2010)
19. Nauck, D., Kruse, R.: Obtaining interpretable fuzzy classification rules from medical data. Artif. Intell. Med. **16**(2), 149–169 (1999)