THE PREDICTION AND ANALYSIS OF LUNG CANCER BY USING FIS AND ANN HYBRID TECHNIQUE

Meena Rawat[1], Pooja Pathak[2], Pooja Vats[3], Rahul Boadh[4*]

Abstract
Lung cancer is a global health concern, and early detection plays a crucial role in improving patient outcomes. In recent years, various computational approaches have been developed to aid in lung cancer prediction. This article presents a comprehensive comparative analysis of two popular machine learning techniques: Fuzzy Inference Systems (FIS) and Artificial Neural Networks (ANN). The aim is to evaluate their performance and suitability for lung cancer prediction. A dataset containing clinical and demographic features of lung cancer patients is used for experimentation. The analysis includes model training, evaluation, and interpretation, along with the incorporation of graphical figures for better understanding. The results shed light on the strengths and weaknesses of FIS and ANN in the context of lung cancer prediction, aiding researchers, and medical professionals in choosing the most appropriate technique for their applications.

Keywords: Lung Cancer, FIS, ANN, Prediction, Comparative Analysis, Healthcare

[1,3] Department of Mathematics, School of Basic and Applied Science, K. R. Mangalam University, Gurugram, Haryana, India-122103
[2] Department of Mathematics, GLA University, Mathura, Uttar Pradesh, India-281406
[4*]Department of Mathematics, Shyam Lal College, University of Delhi, G. T. Road, Shahdara, Delhi, India-110032

***Corresponding Author:** Dr. Rahul Boadh
***Department of Mathematics, Shyam Lal College, University of Delhi, G. T. Road, Shahdara, Delhi, India-110032, Email: rboadhmath@shyamlal.du.ac.in

DOI: - 10.48047/ecb/2023.12.si5a.0494

Table 1: Abbreviations

Fuzzy Inference Systems	FIS
Artificial Neural Networks	ANN
Lung Cancer	LC
Non-Small Cell Lung carcinoma	NSCLC
Small Cell Lung Cancer	SCLC
Neural Networks	NN
Membership Function	MF

1. Introduction

Lung cancer is a major cause of cancer-related deaths worldwide, making early prediction and diagnosis crucial for effective treatment. Computational techniques, such as FIS and ANN, have shown promise in this field. This article aims to compare and analyze the performance of FIS and ANN for LC prediction, considering their strengths, limitations, and overall suitability.

One of the cancer forms that causes death is lung cancer according to Kuruvilla and Gunavathi (2014). The lung diseases, which damage the lungs, are unpredictable medical ailments that can occur anywhere in the world, but especially in India. On a global scale, LC accounts for 12.8% of all cancer types, 17.8% of cancer-related fatalities, and it is growing by 0.5% annually. Males account for 38.6% of cancer cases, while females account for 5.2%. However, it has been reported that 15% of LC patients survive for five years or longer after receiving their diagnosis, along with this early diagnosis and the administration of medications when Krishnaiah et al. (2013) and Chauhan et al. (2016) done the analysing of these factors, it enables computer assisted diagnosis of lung cancer. Today, 42% of patients have the prospect of surviving for a year or more following diagnosis, compared to 37% in the late 1970s (Mahmoudi et al. 2010, Manikandan and Bharathi, 2011, Dominguez et al. 2011, Kumar et al. 2011)

The disease known as lung cancer is brought on by the development of abnormal tissues in the lung. These abnormal tissues then move to other regions of the body, a process known as metastasis. However, in some cases of cancer, the early lung growths are referred to be carcinomas because they originated from epithelial cells. If this unchecked development can be effectively diagnosed in its early phases, there is a lower risk of invasive surgery and a higher chance of survival. Asthma, Chronic Obstructive Pulmonary (COPD) disease, and other breathing issues such as pneumonia, influenza, and tuberculosis are all symptoms of lung illness, which affects different regions of the lungs by D'Cruz et al. (2019). The FIS and FIS-ANN used by the other authors to solve different problems in real life (Kaur et al. 2022, Kalra et al. 2022 and Bhagat et al. 2022)

NSCLC and SCLC, sometimes known as oat cell carcinoma, are the two main forms of LC (Krishnaiah et al., 2013) and each is treated differently based on how it spreads and grows. Small cell/large cell cancer that exhibits symptoms from both cancer types is referred to as miscellaneous minor cell/bulky cubicle cancer. Compared to SCLC, non-small cell lung adenocarcinoma (NSCLA) is more prevalent and typically develops and spreads more slowly. Smoking characteristics are associated with SCLC, which develops more quickly by creating a sizable tumour that can disseminate throughout the body. Doctors can spot anomalies earlier and more quickly with the aid of computer-aided diagnosis (CAD) systems, and they can use them as a second opinion before recommending a biopsy test by Tiwari (2016). There are several procedures for diagnosing lung cancer, but they are extra luxurious, time-taking, and fewer effective at finding the disease. In order to predict lung cancer in its early stages, a new prediction method is necessary.

The current study uses an ANN with FIS to propose an integrated framework for predicting LC. NN use a variety of problem-solving techniques, and the neurons are trained and evaluated using a database (Tiwari, 2016). Based on specific features employed in the system, the LC features are retrieved in order to predict the cancer stage. The computation approach is given with less cancer cells to feature selection, which finds predicted subsets of cancer cells inside a database. It is possible to improve performance by eliminating some features.

2. Background

2.1 Lung Cancer and Early Detection

The largest cause of cancer-related fatalities worldwide is lung cancer. The key to enhancing patient outcomes and survival rates is early identification. Lung cancer in its early stages frequently exhibits mild symptoms, making diagnosis difficult. Low-dose computed tomography (LDCT), for example, is a medical imaging and screening tool that has advanced, allowing for the early diagnosis of LC. Lung cancer death rates have decreased thanks to LDCT screening, which has demonstrated promising outcomes. A timely diagnosis enables the start of treatment right away, which may improve the prognosis and raise the likelihood of a successful intervention.

2.2 Fuzzy Inference Systems

Fuzzy Inference Systems (FIS) are computer models designed to handle uncertainty and imprecision in data that are inspired by fuzzy logic. A fuzzy rule base, a fuzzy membership function, and an inference engine make up the three core parts of FIS. The fuzzy rule base uses linguistic rules to establish the relationship between the input and output variables. Based on its linguistic word, the membership function determines the degree of membership for each input variable. To provide output values, the inference engine combines the membership functions and rules. The visualise of FIS has been shown in Fig. 1.

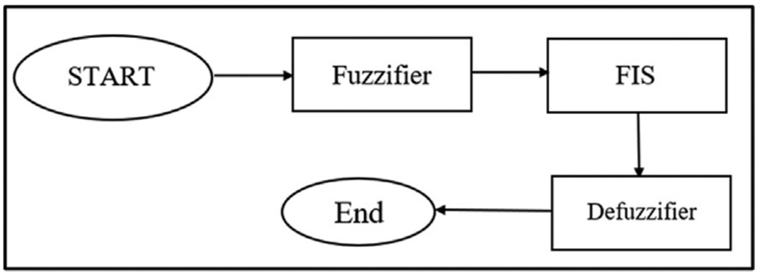

Fig1: Block diagram of Process of Fuzzy Inference System (Boadh et al. 2022)

For activities including diagnosis, prediction, and decision-making, FIS has been extensively employed in a variety of industries (Boadh et al. 2022), including healthcare (Kumar et al 2022). Since it can manage ambiguous and uncertain information, it is appropriate for medical applications where data may be missing or unclear. The Fig. 1 shows how FIS components are visualised, giving a precise depiction of the behaviour of the system, and assisting in the interpretation of outcomes.

2.3 Artificial Neural Networks

Computational models known as ANN are modelled after the structure and operation of the human brain. Neurons are networks of interconnected nodes that process and transfer information. An input layer, one or more hidden layers, and an output layer are the layers that make up an ANN. Each neuron in the network takes in input signals, performs a mathematical operation, and then sends the outcome to the layer below it. Weighted connections between neurons enable the network to learn relationships and patterns from training data. As seen in the graphic below, neural network diagrams can be used to visualise ANN models.

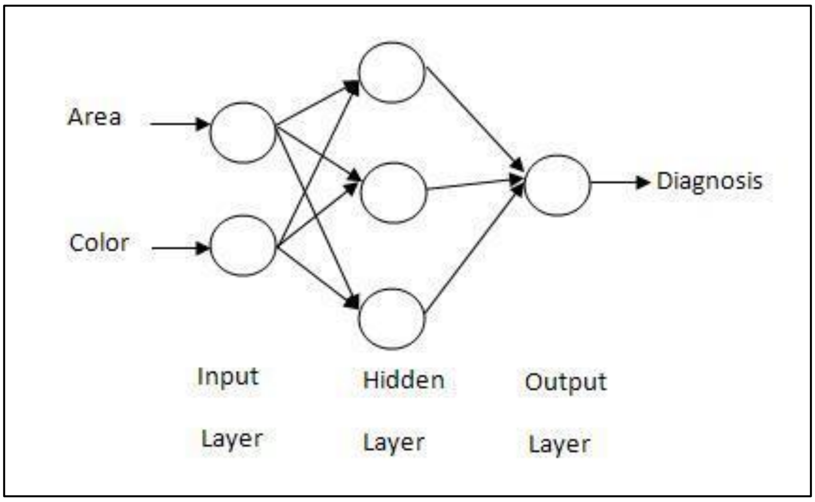

Fig. 2: Three layers of ANN graphical representation (Taken from Google)

For many different tasks, including classification, regression, and pattern recognition, ANNs are frequently utilised. They are very good at handling non-linear, complex relationships in data. The picture provides a visual depiction of the modelling system by showing the structure of an ANN and its associated neurons and layers. When used for tasks like disease prediction, image analysis, and drug discovery, ANN models have demonstrated extraordinary performance in a variety of industries, including healthcare.

3. Methodology

The pre-processing step of the procedure involves making the image acceptable for subsequent processing by boosting contrast and removing noise. In this investigation, the noise was removed using linear filtering. Following pre-processing, picture segmentation was completed. The region expanding algorithm also contributes to the segmentation of the lungs. Next, decisions were made by categorising tumours as benign, malignant, or progressed using FIS. The FIS system's output is then tested, and the model's correctness is determined, using an artificial neural network. The MRI centre in Hamedan has used this technique on a range of given pictures, some of which are shown in Fig. 3 (Hashemi et al., 2013).

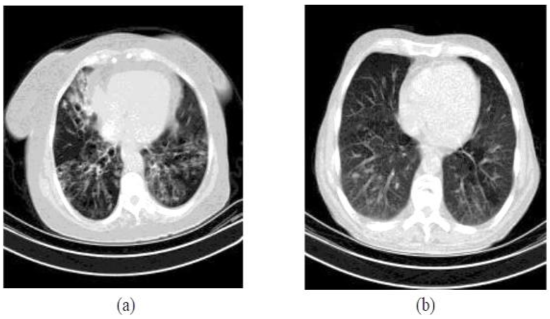

Fig. 3: Mass is lungs (MRI image)

3.1 The Proposed Algorithm

The suggested algorithms consist of seven phases:
Phase 1: The initial seed points and threshold value must be established in the region increasing segmentation.
Phase 2: The size of the image is determined in order to assess the loop's condition.
Phase 3: Use the threshold value to compare the first initial seed point to nearby pixels.
Phase 4: If the neighbouring pixels are comparable to the seed point, we can add them to the region. For instance, if (p-q)<T for some threshold T, we can add pixel q to the area of pixel p.
Phase 5: When no more pixels are discovered that match the requirements for inclusion in that region. Then, a different seed point is chosen that is not a part of any other region.
Phases 6: This procedure is repeated up until every pixel (equivalent to the image's size) is contained within a single section.
Phase 7: The area of the regions is finally computed.
It is true that segmentation is a procedure that divides a picture, say B, into 'n' subregions, say B1, B2, and B3,..., Bn, so that:

$$U_{i=1}^{n} B_i = B$$
$$B_i \cap B_j = \emptyset, for\ all\ i\ and\ j, i \neq j$$

When each pixel is a part of a region, the first criterion signifies that segmentation is finished. The second requirement specifies that the area must be disconnected. Four separate lung pictures are subjected to the described region-growing segmentation algorithm, and the consequences remain shown in Fig 4.

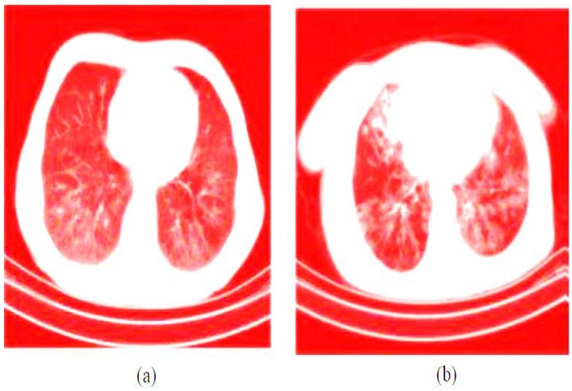

Fig. 4: Region increasing segmentation algorithm's output

4. Results and Discussion

4.1 Rule base by FIS

A FIS is a technique for mapping an input space to an output space using fuzzy logic and IF-THEN rules. Different architectures can be used while designing FIS to improve system performance.
Finding the ideal number of directions, choosing the proper MFs, and modification both must all be considered by Al-Daoud (2010). The architecture of the FIS we created is depicted in Fig. 5. This fuzzy system has a diagnosis output and two inputs of "color" and "area".

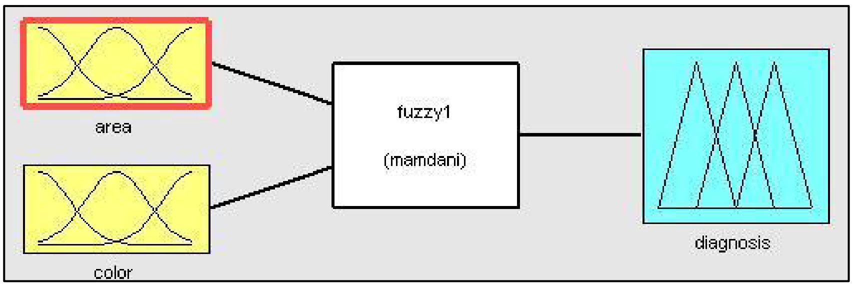

Fig. 5: The design of a FIS

In Fig. 5, three levels of the fuzzy system are depicted. The input variables are shown in the top layer. The third shows the output variables, whereas the second shows fuzzy rules. In order to identify the nodule using fuzzy MFs as output and input, a FIS is used. This fuzzy system has a rule foundation made up of four fuzzy IF-THEN rules, two inputs (colour and area) and one output (diagnostic).

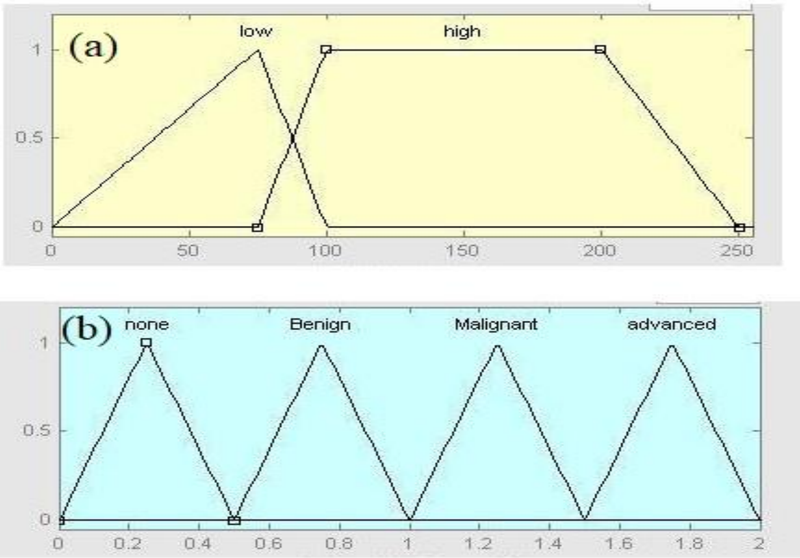

Fig. 6: (a) Input parameter colour (b) Output in diagnosis in FIS

The scale of "color" is occupied between1 to 250 and is divided into two sections, "low" and "high," which are represented by triangles and trapezoids, respectively, as illustrated in Fig. 6a. As a result, The scale of "area" is considered between 1 and 5300, which are likewise divided hooked on "high" and "low" parts in the trapezius shape. The output "diagnosis" scale is used between 0 to 2, which are further divided between four different sections: "Advanced," "Malignant," "Benign," and "None," as shown in Fig. 6b.
How to create MFs is one of the key issues in fuzzy clustering methods. One can find various membership functions (IF-THEN rules) that remain utilised to find LC, per research, in Table 2 and the scale viewer (Fig. 7).

Table 2: IF- THEN Rules

S.N.	Rule
1	If area is high & color is low then diagnosis is Malignant
2	If area is high & color is high then diagnosis is advanced
3	If area is low & color is high then diagnosis is Benign
4	If area is low & color is low then diagnosis is none

Because the result is 1.85, which is within the range of 1.4 to 2, the scale observer demonstrates that the specified malignant form of CT scan picture. As a result, the production stands of the malignant type. The area parament of input in the FIS is 3500 (high), and grey value used as input for color is 120.8 (low), it is clearly shown in Fig.7.

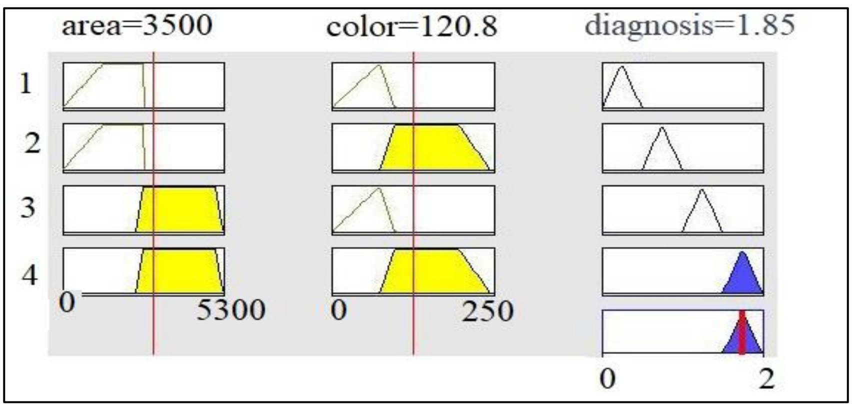

Fig. 7: FIS rule visualization

4.2ANN to test FIS output

In this section, the artificial neural network (described in section 2.3) is used to obtain the fuzzy system's accuracy. In this approach, a neural network is trained using the fuzzy system's input and output. System error remains measured after the trying stage. It is important to note that this portion of the work was completed with the aid of a specialist.

An ANN with three layers and a 2-3-1 nodal network topology makes up this model. This system, as depicted, has a single output called "diagnosis" and two inputs (features) called "colour" and "area." For the two clinical neurons, characteristics found on the CT were present in the input film. There were three tansig-functioning neurons in the buried layer. As is typical in ANN applications by Chen et al. (2012), the number of neurons in the buried layer was determined experientially. Lastly, the output film contained a single neuron through the purlin occupation, providing an output charge between 0 and 2.

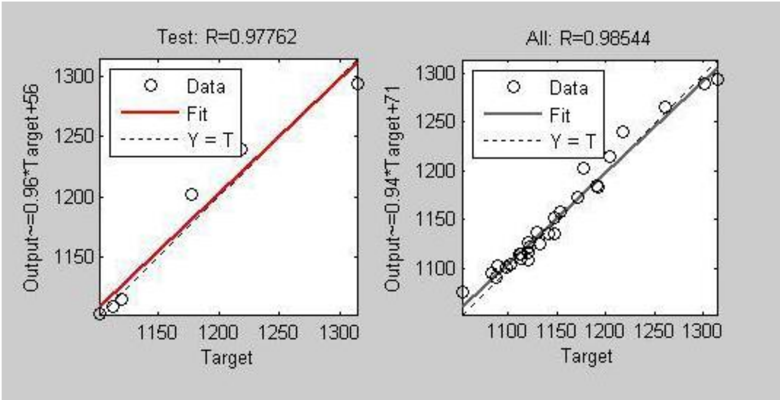

Fig. 8: Regression Analysis of ANN

Fig. 8 displays the system's outcome. The system is doing well because the rotary themes are near towards the color positions. The technology has a 97% accuracy rate as of today. In fact, neural networks' capacity for generalisation is one of their key features. As a result, a trained net may classify data that it has never seen before that belongs to the same class as the learning data. Developers of real-world applications typically only have access to a small portion of all potential patterns for creating neural networks. The dataset should be divided into three sections based on Ahmed at el. (2022) in order to achieve the best generalisation:

a. A neural network is trained using the training set. This dataset's error is reduced during training.

b. A neural network's performance on patterns that aren't trained during learning is evaluated using the validation set.

c. A test set for assessing a neural network's overall performance.

5. Conclusion

Fuzzy Inference Systems and Artificial Neural Networks for Lung Cancer Prediction: A Comparative Analysis provides Important Insights into Performance and Interpretability. It might be challenging to distinguish amongst kind and malicious abnormalities. Present aim of current work is to reduce hominoid error in detecting besides interpreting facts that the radiologist might have missed. For this, a FIS-based decision-making system was employed. In this step, features including area and colour (grey values) were extracted and sent as input to the algorithm. To determine the kind of abnormality, the systems use membership functions that can be described as if-then rules. After that, this paper compares the FIS system's diagnostic performance using ANNs. This approach is used on 500 CT slices from 60 patients with lung tumours, each of which included 60 tumours. When using both methods in the context of lung cancer prediction, it might be helpful for researchers and medical practitioners to have a clear understanding of the characteristics of both approaches.

Acknowledgement

The first author is thankful to K. R. Mangalam University for providing the research facility to conduct the research. All the authors are very grateful to the anonymous reviewers for their valuable suggestions for improving the quality of this research.

References

1. Chauhan, D., Solan, H.P. and Jaiswal, V., 2016. Development of Computational Tool for Lung Cancer Prediction Using Data Mining.
2. Krishnaiah, V., Narsimha, D.G. and Chandra, D.N.S., 2013. Diagnosis of lung cancer prediction system using data mining classification techniques. International Journal of Computer Science and Information Technologies, 4(1), pp.39-45.
3. Kuruvilla, J. and Gunavathi, K., 2014. Lung cancer classification using neural networks for CT images. Computer methods and programs in biomedicine, 113(1), pp.202-209.
4. D'Cruz, M.J., Jadhav, M.A., Dighe, M.A., Chavan, M.V. and Chaudhari, J.L., 2019.

Detection of Lung Cancer Using Backpropagation Neural Networks and Genetic Algorithm.
5. Tiwari, A.K., 2016. Prediction of lung cancer using image processing techniques: a review". Advanced Computational Intelligence: An International Journal (ACII), 3(1).
6. Boadh, R., S. N. Yadav, Agnivesh Tiwari, Yogendra Kr. Rajoria, Jitendra Singh, 2022, Application of fuzzy inference system (FIS) for assessment and predication of compressive asset of concrete containing fly ash, Materials Today: Proceedings. https://doi.org/10.1016/j.matpr.2022.08.160.
7. Satish Kumar, Gunjan Kalra, Hari Kishan Bhardwaj, Yogendra Kumar Rajoria, Deepak Kumar, Rahul Boadh, (2022). Internet of Medical Thing and FIS Evaluation for Selecting and Delivering the best Health Insurance Coverage. Journal of Pharmaceutical Negative Results, Vol. 13 (08), pp. 3438–3446. DOI: 10.47750/pnr.2022.13.S08.422.
8. Atiyeh Hashemi, A., Pilevar, A. H. and Rafeh, R. 2013, Mass Detection in Lung CT Images Using Region Growing Segmentation and Decision Making Based on Fuzzy Inference System and Artificial Neural Network, I.J. *Image, Graphics and Signal Processing,* 2013, 6, 16-24. DOI: 10.5815/ijigsp.2013.06.03
9. E. Al-Daoud, "Cancer Diagnosis Using Modified Fuzzy Network," presented at the Universal Journal of Computer Science and Engineering Technology, 2010.
10. H. Chen, J. Zhang, Y. Xu, B. Chen, and K. Zhang, "Performance comparison of artificial neural network and logistic regression model for differentiating lung nodules on CT scans," Expert Systems with Applications, vol. 30, pp. 11503–11509, 2012.
11. S. E. Mahmoudi, A. Akhondi-Asl, R. Rahmani, S. Faghih-Roohi, V. Taimouri, A. Sabouri, et al., "Web-based interactive 2D/3D medical image processing and visualization software," computer methods and programs in biomedicine, vol. 9 8, pp. 172–182, 2010.
12. Kaur, K., Hari Kishan Bhardwaj, Yogendra Kumar Rajoria, Anil Yadav, Dinesh Kumar Maurya, Rahul Boadh, (2022). Fuzzy Expert System (FES) Integration for Oral Cancer Risk (OCR) Diagnosis. Journal of Pharmaceutical Negative Results, Vol. 13(08), pp. 3433–3437. DOI:https://doi.org/10.47750/pnr.2022.13.S08.421.
13. Kalra, G., Kamal Kishore, Avneesh Kumar, Yogendra Kumar Rajoria, Anil Yadav, Rahul Boadh, (2022). FIS-Based Prediction and Estimation of Health Insurance for Workers in the Manufacturing Sector. Journal of Pharmaceutical Negative Results, Vol. 13(08), pp. 2636–2645.
14. DOI:https://doi.org/10.47750/pnr.2022.13.S08.331.
15. Bhagat, R. K., Anil Yadav, Yogendra Kr. Rajoria, Sandeep Raj and Rahul Boadh, 2022. Study of Fuzzy and Artificial Neural Network (ANN) based Techniques to Diagnose Heart Disease, Journal of Pharmaceutical Negative Results, Vol. 6 (S05), pp. 1023-1029. DOI: 10.47750/pnr.2022.13. S05.161.
16. T. Manikandan and N. Bharathi, "Lung Cancer Diagnosis from CT Images Using Fuzzy Inference System," Communications in Computer and Information Science vol. 250, pp. 642-647, 2011.
17. J. Quintanilla-Dominguez, B. Ojeda-Magaña, M. G. Cortina-Januchs, R. Ruelas, A. Vega-Corona, and D. Andina, "Image segmentation by fuzzy and possibilistic clustering algorithms for the identification of microcalcifications," Sharif University of Technology Scientia Iranica, vol. 18, pp. 580–589, Received 21 July 2010; revised 26 October 2010; accepted 8 February 2011 2011.
18. S. A. Kumar, Dr.J.Ramesh, Dr.P.T.Vanathi, and Dr.K.Gunavathi, "ROBUST AND AUTOMATED LUNG NODULE DIAGNOSIS FROM CT IMAGES BASED ON FUZZY SYSTEMS," IEEE, pp. 1-6, 2011.
19. Ahmed M. Montaser, Korany R. Mahmoud. "Design of Intelligence Reflector Metasurface using Deep Learning Neural Network for 6G Adaptive Beamforming", IEEE Access, 2022