HANOI UNIVERSITY OF SCIENCE AND TECHNOLOGY

SCHOOL OF INFORMATION AND COMMUNICATION TECHNOLOGY

-----🙟 🕮 🙜-----

PROJECT II

MID-TERM REPORT

DIAGNOSING CARDIOVASCULAR DISEASE USING FKG-PAIRS Fuzzy Knowledge Graph Model

Instructors: Assoc. Prof. PhD. Pham Van Hai

Implemented: Dao Thanh Manh 20211014

Hanoi , April 20 , 2024

TABLE OF CONTENTS

1. Problem

In health care, especially in decision making and diagnosis of cardiovascular disease, having a deep understanding of medical information and the ability to apply medical methods and knowledge is essential. Modernity is extremely important. However, although there has been progress in collecting and organizing medical data, problems still exist due to the complexity and diversity of this information, especially when dealing with medical conditions. like cardiovascular disease.

Traditional knowledge graphs (KGs) cannot completely address these challenges. Although KG can organize medical data and build relationships between different medical elements, it often cannot handle the fuzziness and uncertainty of medical information, nor does it Flexible enough to apply complex rules and relationships in cardiovascular disease decision-making and diagnosis.

Meanwhile, the pairwise Fuzzy Knowledge Graph (FKG) opens up a new potential to improve the decision-making and diagnosis process of cardiovascular disease. FKG not only helps to represent medical information more flexibly and handle data fuzziness and uncertainty, but also solves the low performance problem of conventional KG by using information pairs instead because only a single pair is used. This enhances the ability to reason and make more accurate decisions in the treatment and management of cardiovascular disease, helping to improve the quality of healthcare and patient outcomes.

2. Objectives of the study

2.1. General objective

Develop a system/application to support decision making and cardiovascular disease diagnosis based on the FKG-Pairs Fuzzy Knowledge Graph model. The project focuses on building a diverse and high-quality medical database, using FKG-Pairs to represent medical information flexibly and efficiently. The goal is to develop a reliable and accurate cardiovascular disease prediction model that incorporates both predictive and explanatory analytical aspects. Finally, the project will evaluate and optimize the performance of the system before deploying into medical practice, to provide a useful tool to help improve the quality of diagnosis and treatment of cardiovascular diseases, and at the same time. while increasing understanding and trust from users.

2.2. Detail goal

Build FKG-Pairs Fuzzy Knowledge Graph models with k FKG pairs (k = 1, 2, 3), develop algorithms and methods to build FKG-Pairs models with accuracy and highly flexible, conduct performance evaluation of FKG-Pairs models based on criteria such as accuracy, classification and reliability, compare and analyze the results of FKG-Pairs models to determine the most optimal model for cardiovascular disease diagnosis, optimize and improve the most optimal model to ensure the highest performance and accuracy before deploying into medical practice, providing Provide detailed guidance and reports on project processes and results to share experiences and knowledge with the medical and research community.

3. Input data

3.1. Data set

Experimental data is collected directly from Kaggle main page: https://www.kaggle.com/datasets/sulianova/cardiovascular-disease-dataset/data

The dataset includes 70 000 patient data, 11 attributes saved as file **cardio_train.csv** :

```
[ ] df_cardio_train
```

	id	age	gender	height	weight	ap_hi	ap_lo	cholesterol	gluc	smoke	alco	active	cardio
0	0	50.0	2	168	62.0	110	80	1	1	0	0	1	0
1	1	55.0	1	156	85.0	140	90	3	1	0	0	1	1
2	2	52.0	1	165	64.0	130	70	3	1	0	0	0	1
3	3	48.0	2	169	82.0	150	100	1	1	0	0	1	1
4	4	48.0	1	156	56.0	100	60	1	1	0	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...
69995	99993	53.0	2	168	76.0	120	80	1	1	1	0	1	0
69996	99995	62.0	1	158	126.0	140	90	2	2	0	0	1	1
69997	99996	52.0	2	183	105.0	180	90	3	1	0	1	0	1
69998	99998	61.0	1	163	72.0	135	80	1	2	0	0	0	1
69999	99999	56.0	1	170	72.0	120	80	2	1	0	0	1	0

70000 rows × 13 columns

Data description

There are 3 types of input features:

- Objective: factual information;
- Examination: health examination results;
- Subjective: information provided by the patient.

Properties:

Age | Objective features | age | int (day)

Height | Objective features | height | int (cm) |

Weight | Objective features | weight | float (kg) |

Gender | Objective features | gender | classification code |

Systolic blood pressure | Test feature | ap_hi | int |

Diastolic blood pressure | Test feature | ap_lo | int |

Cholesterol | Test feature | cholesterol | 1: normal, 2: above normal, 3: above normal |

Road | Test feature | Gluc | 1: normal, 2: above normal, 3: above normal |

Smoking | Subjective features | smoke | binary |

Drinking alcohol | Subjective features | wine | binary |

Physical activity | Subjective features | activity | binary |

Presence or absence of cardiovascular disease | Target variable | cardiovascular | binary |

All dataset values were collected at the time of medical examination.

3.2. Data processing

Process the data set from kaggle using Google Colab then save the processed data to Drive:

Link: https://colab.research.google.com/drive/1m6IaNr2wEebYCINRZ1peqyE_DYsZrUT-?usp=sharing

After processing, the data set was split into two data sets: patients with heart disease and patients without heart disease, divided in a 1:1 ratio for ease of use in the process of writing fuzzy knowledge rules and training. FKG-Pairs-k models .

After processing, the data is returned as shown below:

File **positive_data_3000.csv** :

```
positive_data_3000
```

	id	age	gender	height	weight	ap_hi	ap_lo	cholesterol	gluc	smoke	alco	active	cardio
0	1	55.0	1	156	85.0	140	90	3	1	0	0	1	1
1	2	52.0	1	165	64.0	130	70	3	1	0	0	0	1
2	3	48.0	2	169	82.0	150	100	1	1	0	0	1	1
3	4	62.0	2	178	95.0	130	90	3	3	0	0	1	1
4	5	46.0	2	172	112.0	120	80	1	1	0	0	0	1
...	...	...	...	...	...	...	...	...	...	...	...	...	...
2995	2996	52.0	1	161	76.0	140	80	1	1	0	0	1	1
2996	2997	53.0	1	160	70.0	110	70	2	1	0	0	1	1
2997	2998	51.0	2	170	69.0	140	90	1	3	0	0	0	1
2998	2999	60.0	2	164	77.0	150	120	1	1	0	0	1	1
2999	3000	59.0	1	151	103.0	140	90	1	1	0	0	1	1

3000 rows × 13 columns

File negative_data_3000.csv:

```
negative_data_3000
```

	id	age	gender	height	weight	ap_hi	ap_lo	cholesterol	gluc	smoke	alco	active	cardio
0	0	50.0	2	168	62.0	110	80	1	1	0	0	1	0
4	4	48.0	1	156	56.0	100	60	1	1	0	0	0	0
5	8	60.0	1	151	67.0	120	80	2	2	0	0	0	0
6	9	61.0	1	157	93.0	130	80	3	1	0	0	1	0
8	13	48.0	1	158	71.0	110	70	1	1	0	0	1	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...
6014	8545	49.0	1	155	87.0	120	80	1	1	0	0	1	0
6018	8552	48.0	1	158	88.0	120	80	1	1	0	0	0	0
6020	8555	54.0	1	147	51.0	110	70	1	1	0	0	1	0
6021	8556	40.0	1	168	67.0	120	80	1	1	0	0	1	0
6023	8559	62.0	2	175	85.0	120	90	1	1	0	0	1	0

3000 rows × 13 columns

4. Related knowledge

4.1. Fuzzy sets and fuzzy logic.

Fuzzy sets were first introduced by Zadeh in 1965, introduced as a new mathematical tool for solving problems with ambiguous, uncertain information. Unlike normal sets, which evaluate the membership relationship of a set according to binary logic "an element belongs or does not belong to the set", fuzzy logic evaluates the membership relationship of an element through a function. membership $\mu \to [0,1]$, represents the membership of an element to a set

4.2. Fuzzy inference system.

Fuzzy inference is the process of finding conclusions for a set of input values, based on a synthesized fuzzy rule system. Fuzzy inference methods are often referred to as Mamdani fuzzy inference, Takagi-Sugeno fuzzy inference, etc. The above inference systems are also known as classical inference methods, which have been widely used in automatic control systems. Fuzzy knowledge graph is known as a new, effective, and more accurate inference method than previous inference methods. The general rule for applying the fuzzy inference system is shown in three steps:

- Fuzzification: In this step, we need to determine the value scale and corresponding level terms of each input attribute of the data set, followed by the conversion process from explicit values. of the input data set into fuzzy values, based on the value scale combined with the previously built membership function, finally combining the fuzzy values of each input data sample using operators. fuzzy (AND, OR, NOT) to provide representation rules in the form of IF-THEN clauses and put them into the fuzzy rule base system
- Fuzzy inference: Use fuzzy inference method to find output results based on the fuzzy rule system built in step 1
- Defuzzification: Converts the fuzzy output values found in step 2 into clear values, giving the results of the problem.

4.3. Fuzzy knowledge graph

The term fuzzy knowledge graph was first introduced and integrated in the M-CFIS-FKG model with the initial purpose of expanding the M-CFIS-R model to make the inference process of this model in the testing becomes faster. Inheriting the characteristics of knowledge graphs, formally, a fuzzy knowledge graph includes vertices representing the linguistic labels of attributes and output labels of rules, the corresponding edges are arcs. connection between vertices.

The way to calculate the weight values of edges of fuzzy knowledge graphs has been presented in detail in [7], and is briefly summarized as follows:

- For edges connecting two attribute vertices, for each pair of values $\left(X_i, X_j\right), 1 \leq i \leq j \leq m$, in rule t $R_t, t = \overline{1,k}$, the weight $A_{ij}^t$of this edge is calculated according to the formula:

$$A_{ij}^t = \frac{|X_i \text{ quan hệ với } X_j \text{ trong luật thứ } t|}{|R|}$$

- For edges connecting attribute vertices and output label vertices, for each pair $(X_i, l), 1 \leq i \leq j \leq m,\ l = \overline{1, C}$, in rule t $R_t, t = \overline{1, k}$, the weight $B_{ij}^t$of this edge is calculated according to the formula:

$$B_{il}^t = \left(\sum A_{ij}^t\right) \times \frac{|X_i\ quan\ hệ\ với\ nhãn\ l\ trong\ luật\ thứ\ t|}{|R|}$$

The results of the two sets of weights are stored in an adjacency matrix, representing the constructed fuzzy knowledge graph.

4.4. Approximate inference.

Approximate reasoning is defined as a tool for reasoning from propositions whose meaning is not clearly defined through fuzzy logic. Normally, the approximate reasoning method's accuracy of results is not as high as conventional reasoning techniques for clear data, however, the advantage of approximate reasoning is that it can perform argumentative reasoning. With linguistic variables, or natural language is data with ambiguous and unclear meanings

5. Models and algorithms

5.1. State the problem

Input: To build the fuzzy knowledge graph used in this problem, we need to set patient samples that have been diagnosed by doctors and experts based on the given attributes. This sample data set goes through preprocessing (fuzzy) and is saved into a fuzzy rule base system as shown in table 1. This fuzzy rule base system includes n rules $R_1, R_2, \dots, R_n$representing patient samples, m $S_1, S_2, \dots, S_m$representative attributes. represents the symptoms of the disease, and C output labels 1, 2, 3, …, C represents the doctor's diagnostic conclusion

In addition, there is one new patient outside the above rule system, shown as follows

IF $S_1$is “Low” and $S_2$is “Low” and $S_3$is “High” and $S_4$is “Very high” and … and $S_{m-1}$is “High” and $S_m$is “Low” THEN Conclusion = ?

	S_1	S_2	…	S_{m-1}	S_m	Conclude
R_1	High	High	…	Very high	High	first
R_2	Medium	Medium	…	Medium	Low	2

…	…	…	…	…	…	…
R_{n-1}	Medium	Medium	…	Medium	Medium	2
R_n	Low	Medium	…	Low	Low	3
Patient's symptoms	High Medium Low	High Medium	…	Very high High Medium Low	High Medium High	0,1,…, C

6. *Table 1: Fuzzy rule base system*

Output: Results of the new patient output diagnostic system based on existing fuzzy rules.

5.2. Proposed model

5.2.1. Problem model.

The system includes 2 stages that need to be processed:

- Data processing phase: Data collected from doctors and experts is raw data, this data goes into the data processing phase, the pre-processing step includes scale division, design corresponding level terms, then combined with the rule generation tool to provide a fuzzy rule base system of the problem, as a basis for the phase of building a fuzzy knowledge graph and diagnosis.
- Graph construction and diagnosis phase: based on the fuzzy rule base system formed in the above step, proceed to build the fuzzy knowledge graph FKG by calculating edge weight sets and saving it as an adjacency matrix. , thereby using the data of that FKG fuzzy knowledge graph to diagnose new cases

The model for the above two stages of the problem has been described in [8], and is redrawn below.

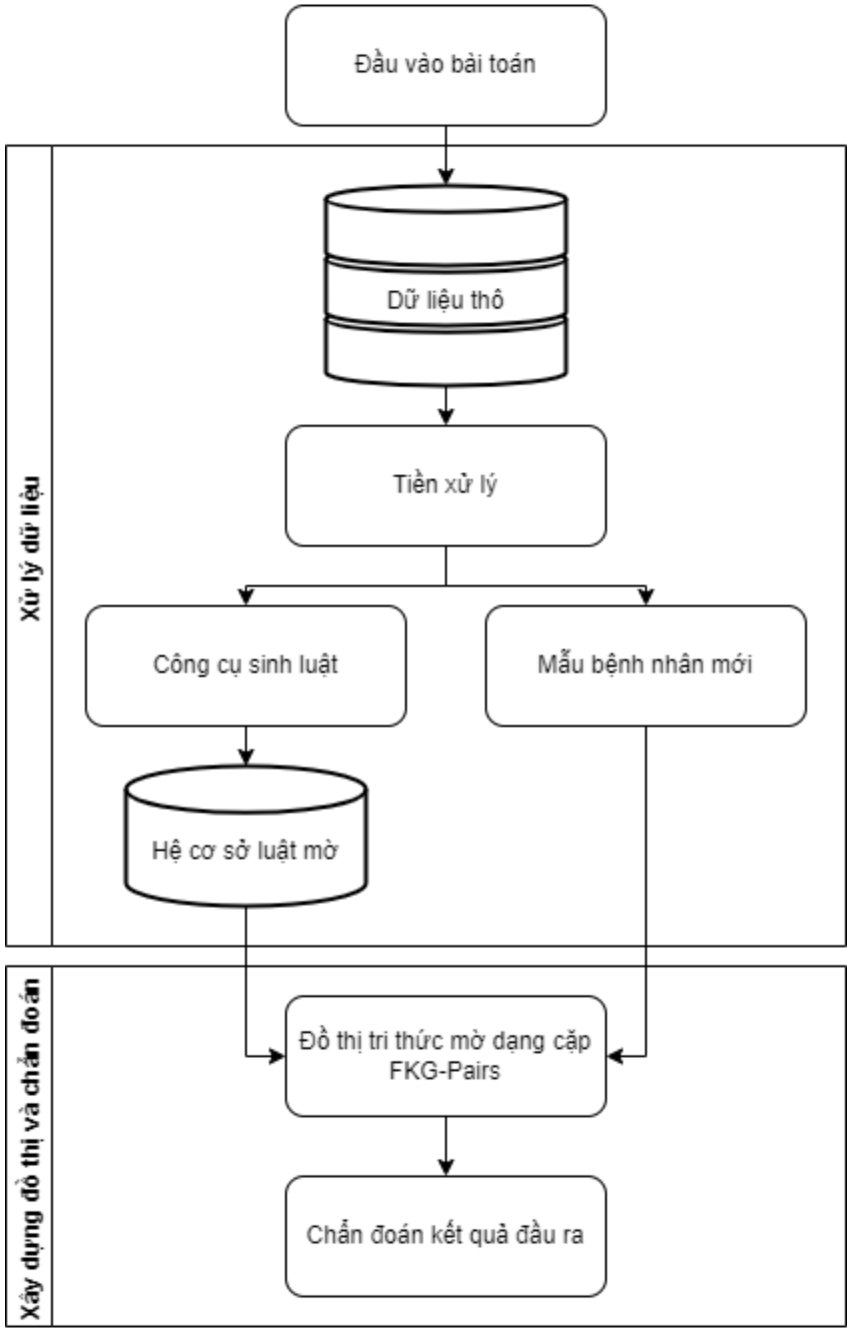

Figure 1: Problem model for diagnosing heart disease using FKG-Pairs

5.2.2. System functions

The system needs to have functions that can perform the two stages mentioned above

- For the data preparation phase: the raw data collected by the system designer as input to the problem needs to be analyzed, noise and errors removed, and a complete input data set produced. This data set needs to be blurred based on the scale and level terminology agreed upon with medical experts in the above field, through the rule generation tool to synthesize a fuzzy rule base system for use in medical research. subsequent functionality of the system

- For the performance and diagnosis phase: After obtaining the fuzzy rule base system, the system will have to build a fuzzy knowledge graph FKG and store this graph for use in the diagnosis step. At the diagnosis step, the system will receive input, combined with the fuzzy knowledge graph FKG to conduct diagnosis and draw conclusions.

5.2.3. Methods of implementation

As described in the system functions section, the problem will be installed according to the following steps:

- Step 1: Validate and process data

  After building the scale, build functions and procedures to control automatic data processing with conditions for each attribute based on the scale, combined with a rule generation tool, to give input. The output is the fuzzy rule base system as shown in table 1.

- Step 2: Calculate the weight sets$\tilde{A}, \tilde{B}$

  The set of weights $\tilde{A}$of the pairwise fuzzy knowledge graph is the weight of the edge connecting the linguistic labels of the attributes in rules t (R_t). These weights are calculated by the following formula:

$$\tilde{A}^t_{ij\dots k} = \frac{\left|S_i \to S_j \to \cdots \to S_{k+1}\ trong\ luật\ thứ\ t\ \right|}{|R|} \quad (1)$$

  In there$t = \overline{1, n}, 1 \le i \le j < \cdots < k < m - 1.$

  The set of weights $\tilde{B}$of the pairwise fuzzy knowledge graph are the weights of the edges connecting the labels of the attribute pairs with the output label in the t-rule. (R_t).These weights are calculated by the following formula:

$$\tilde{B}^t_{ij\dots kl} = \left(\sum \tilde{A}^t_{ij\dots k+1}\right) \times \min\left(\frac{|S_i \to l\ trong\ luật\ t|}{|R|}, \frac{\left|S_j \to l\ trong\ luật\ t\right|}{|R|}, \dots, \frac{|S_k \to l\ trong\ luật\ t|}{|R|}\right) \quad (2)$$

  In there$t = \overline{1, n}, 1 \le i \le j < \cdots < k < m - 1, l = \overline{1, C}.$

- Step 3: Store the pairwise fuzzy knowledge graph

The weight sets after calculation need to be stored as adjacency matrices for convenience in calculating the next steps. In this design, the adjacency matrices of the weight sets are stored in 2 sheets. of an excel file belonging to the file system.

- Step 4: Apply the approximate inference process and provide disease diagnosis results.

In [4], author Luong Thi Hong Lan used the FISA algorithm to approximate the output of the problem of determining the output of a new law based on the fuzzy knowledge graph, then author Cu Kim Long improved the above algorithm in his article [8] to be suitable for determining the output of new rules based on pairwise fuzzy knowledge graphs. The algorithm is described through the steps below:

First, it is necessary to calculate the weighted sum of the edges $(\tilde{C})$from the super vertices to the output label using the following formula:

$$\tilde{C}_{ij\ldots kl} = \sum_{t} \tilde{B}_{ij\ldots kl}^{t} \quad (3)$$

In there$t = \overline{1,n}, 1 \leq i \leq j < \cdots < k < m-1, l = \overline{1,C}$.

Then, apply Max-Min operations to calculate the values $(\widetilde{D})$, these values help approximate the new attribute pairs in the Testing set with the corresponding attribute pairs in the pairwise fuzzy knowledge graph, to find out the level of influence on the output labels. The value of $(\widetilde{D})$is calculated based on the Max-Min operators according to the following formula:

$$\widetilde{D}_l = Max_{1\leq i\leq j<\cdots\leq k}\left(\tilde{C}_{ij\ldots kl}\right) + Min_{1\leq i\leq j<\cdots\leq k}\left(\tilde{C}_{ij\ldots kl}\right) \quad (4)$$

In there$t = \overline{1,n}, 1 \leq i \leq j < \cdots < k < m-1, l = \overline{1,C}$

Finally, the output label diagnosis result of the new rule is concluded using the Max operation as follows:

$$Label = p \; If \; \widetilde{D}_p = Max_{l=\overline{1,C}}\left(\widetilde{D}_l\right) \quad (5)$$

5.3. FKG-Pairs installation algorithm

	Paired FKG installation algorithm
1	**Input data** : Test data set, m: Number of attributes of each rule, n: Number of samples of the data set, C: Number of labels of each attribute.
2	**Output data** : Label of new sample.
3	**Begin**

4 Enter values ;
5 Get test data set;
6 Conduct fuzzification of the test data set;
7 $\boldsymbol{for\ i = 1\ to\ m\ do}$
8 $\boldsymbol{for\ t = 1\ to\ n\ do}$
9 $\boldsymbol{for\ l = 1\ to\ C\ do}$
10 $\boldsymbol{while\ 1 \leq i \leq j < \cdots \leq k\ do}$
11 $\boldsymbol{Tính\ \widetilde{C}^{t}_{ij\ldots kl} = \sum_t \widetilde{B}^{t}_{ij\ldots kl}}$
12 $\boldsymbol{Tính\ \widetilde{D}_l = Max_{1\leq i\leq j<\cdots\leq k}\left(\widetilde{C}_{ij\ldots kl}\right) + Min_{1\leq i\leq j<\cdots\leq k}\left(\widetilde{C}_{ij\ldots kl}\right)}$
13 $\boldsymbol{end}$
14 Determine the label of sample t:$\boldsymbol{Label = p\ If\ \widetilde{D}_p = Max_{l=\overline{1,C}}\left(\widetilde{D}_l\right)}$
15 $\boldsymbol{end}$
16 Get the label of sample t and repeat steps 5 to 12 to find the labels of other samples until the end
17 $\boldsymbol{end}$
18 $\boldsymbol{end}$
19 $\boldsymbol{end}$

Algorithm for setting up pairwise FKG problem

7. Real-life example of FKG-Pairs-1 and FKG-Pairs-2 diagnosing heart disease

7.1. Diagnosis of patient's heart disease based on the M-CFIS-FKG model using FKG-Pairs-1

Input: Suppose the input of the problem is a list of 6 patients $\{R_1, R_2, R_3, R_4, R_5, R_6\}$, each patient has test results expressed through attributes $\{S_1, S_2, S_3, S_4, S_5, S_6\}$. The above patient cases have been examined and diagnosed based on test results by doctors, the diagnostic conclusions "Normal", "Heart" and "Severe heart" are shown respectively. labels 0, 1, 2. After going through the "Data Processing" stage, a fuzzy rule base system is obtained as shown in Table 1.

	S_1	S_2	S_3	S_4	S_5	S_6	Conclude
R_1	High	Medium	High	Medium	High	High	2
R_2	High	Medium	Medium	Medium	High	High	2
R_3	Medium	Medium	Medium	Medium	Medium	Medium	0
R_4	Medium	High	Medium	High	Medium	High	first
R_5	High	High	Medium	High	Low	High	first

R_6	High	High	High	Medium	High	Medium	2

Table 1: The fuzzy rule base system assumes the medical examination results of six patients who have been diagnosed by a doctor.

Besides, Input has an additional new patient case shown as follows:
IF *S_1= "High", S_2= "Medium", S_3= "Medium", S_4= "Medium", S_5= "High", S_6= "Medium"* ***THEN*** *Conclusion = ?*
Output: Provide diagnostic conclusions for the above patient, based on the fuzzy rule base system given by Input.

The steps for the above problem are performed sequentially as follows:

- **Step 1:** Calculate the weight sets $\tilde{A}$, $\tilde{B}$
  The weight set $\tilde{A}$includes edges connecting the linguistic labels of the patient's attributes, calculated according to formula (1)
  For example, in case {1}, the weights $\tilde{A}$will be calculated as follows:

$$\tilde{A}_{12}^{1} = \frac{|High \to Medium|}{R} = \frac{1}{3}$$
$$\tilde{A}_{13}^{1} = \frac{|High \to High|}{R} = \frac{1}{3}$$
$$\tilde{A}_{14}^{1} = \frac{|High \to Medium|}{R} = \frac{1}{2}$$

  The weight set $\tilde{B}$of the edge connecting the labels of the attribute pairs with the output labels is calculated according to the formula:

$$\tilde{B}_{il}^{t} = \left(\sum \tilde{A}_{i+1}^{t}\right) \times \min\left(\frac{|S_i \to l\ in\ rule\ t|}{|R|}, \frac{|S_j \to l\ in\ rule\ t|}{|R|}, \dots, \frac{|S_k \to l\ in\ rule\ t|}{|R|}\right)$$

$$\tilde{B}_{1l}^{1} = (\tilde{A}_{12}^{1} + \tilde{A}_{13}^{1} + \tilde{A}_{14}^{1} + \tilde{A}_{15}^{1} + \tilde{A}_{16}^{1} + \tilde{A}_{23}^{1} + \tilde{A}_{24}^{1} + \tilde{A}_{25}^{1} + \tilde{A}_{26}^{1} + \tilde{A}_{34}^{1} + \tilde{A}_{35}^{1} + \tilde{A}_{36}^{1} + \tilde{A}_{45}^{1} + \tilde{A}_{46}^{1} + \tilde{A}_{56}^{1}) \times \frac{|High \to 2|}{R} = \frac{11}{4}$$

  The calculation results of the entire weight matrix are shown in Tables 2 and 3.

	R_1	R_2	R_3	R_4	R_5	R_6

$\tilde{A}_{12}^{t}$	1/3	1/3	1/6	1/6	1/3	1/3
$\tilde{A}_{13}^{t}$	1/3	1/3	1/3	1/3	1/3	1/3
$\tilde{A}_{14}^{t}$	1/2	1/2	1/6	1/6	1/6	1/2
$\tilde{A}_{15}^{t}$	1/2	1/2	1/3	1/3	1/6	1/2
$\tilde{A}_{16}^{t}$	1/2	1/2	1/6	1/6	1/2	1/6
$\tilde{A}_{23}^{t}$	1/6	1/3	1/3	1/3	1/3	1/6
$\tilde{A}_{24}^{t}$	1/2	1/2	1/2	1/3	1/3	1/6
$\tilde{A}_{25}^{t}$	1/3	1/3	1/6	1/6	1/6	1/6
$\tilde{A}_{26}^{t}$	1/3	1/3	1/6	1/3	1/3	1/6
$\tilde{A}_{34}^{t}$	1/3	1/3	1/3	1/3	1/3	1/3
$\tilde{A}_{35}^{t}$	1/3	1/6	1/3	1/3	1/6	1/3
$\tilde{A}_{36}^{t}$	1/6	1/2	1/6	1/2	1/2	1/6
$\tilde{A}_{45}^{t}$	1/2	1/2	1/6	1/6	1/6	1/2
$\tilde{A}_{46}^{t}$	1/3	1/3	1/3	1/3	1/3	1/3
$\tilde{A}_{56}^{t}$	1/3	1/3	1/6	1/6	1/6	1/6

Table 2: Results of calculating the weight matrix$\tilde{A}$

	R_1	R_2	R_3	R_4	R_5	R_6
$\tilde{B}_{1l}^{t}$	April 11	Decembe r 35	23/36	25/36	13/18	June 13
$\tilde{B}_{2l}^{t}$	June 11	35/18	23/36	25/18	Septembe r 13	13/18
$\tilde{B}_{3l}^{t}$	June 11	35/36	23/36	25/18	Septembe r 13	Septembe r 13
$\tilde{B}_{4l}^{t}$	April 11	Decembe r 35	23/36	25/18	Septembe r 13	June 13
$\tilde{B}_{5l}^{t}$	April 11	Decembe r 35	23/36	25/36	13/18	June 13

$\tilde{B}_{6l}^{t}$	June 11	35/18	23/36	25/18	Septembe r 13	13/18

Table 3: Results of calculating the weight matrix$\tilde{B}$

⇨ The weight sets will be combined with the fuzzy rule base system to represent the fuzzy knowledge graph.

- **Step 2:** Apply approximate inference method to provide disease diagnosis results.
After obtaining the fuzzy knowledge graph (represented based on the weight sets and fuzzy rule base system in Step 1), continue to diagnose the new patient's results using the approximate inference method. First, it is necessary to calculate the total weight of the edges $(\tilde{C})$from the super vertices to the output label based on the formula:

$$\tilde{C}_{ij\ldots kl} = \sum_{t} \tilde{B}_{ij\ldots kl}^{t}$$

$$\tilde{C}_{11} = \sum_{t} \tilde{B}_{11}^{t} = \tilde{B}_{11}(Rule\ 4) + \tilde{B}_{11}(Rule\ 5) = \frac{25}{36} + \frac{13}{18} = \frac{17}{12}$$

In there$t = \overline{1,n}, 1 \leq i \leq j < \cdots < k < m-1, l = \overline{1,C}$.

Specific calculation results are summarized in Table 4.

	$Label\ 0$	$Label\ 1$	$Label\ 2$
$\tilde{C}_{1l}$	23/36	17/12	47/2
$\tilde{C}_{2l}$	23/36	17/3	149/18
$\tilde{C}_{3l}$	23/36	17/3	271/36
$\tilde{C}_{4l}$	23/36	17/3	47/2
$\tilde{C}_{5l}$	23/36	17/12	47/2
$\tilde{C}_{6l}$	23/36	17/3	149/18

Table 4: Results of calculating the weight matrix$\tilde{C}$

Based on the table above, continue to calculate the values $(\widetilde{D})$according to the formula:

$$\widetilde{D}_{l} = Max_{1 \leq i \leq j < \cdots \leq k}\left(\tilde{C}_{ij\ldots kl}\right) + Min_{1 \leq i \leq j < \cdots \leq k}\left(\tilde{C}_{ij\ldots kl}\right)$$

In there$t = \overline{1,n}, 1 \leq i \leq j < \cdots < k < m-1, l = \overline{1,C}$

Input has an additional new patient case represented as follows:
IF *S_1= "High", S_2= "Medium", S_3= "Medium", S_4= "Medium", S_5= "High", S_6= "Medium"* ***THEN*** *Conclusion = ?*

Label 0:

$\tilde{C}_{High1,0}$= 0, $\tilde{C}_{Medium2,0}$= 23/36, $\tilde{C}_{Medium3,0}$= 23/36,
$\tilde{C}_{Medium4,0}$= 23/36, $\tilde{C}_{High5,0}$= 0, $\tilde{C}_{Medium6,0}$= 23/36

$\widetilde{D}_0$= max($\tilde{C}_{High1,0}$, $\tilde{C}_{Medium2,0}$, $\tilde{C}_{Medium3,0}$, $\tilde{C}_{Medium4,0}$, $\tilde{C}_{High5,0}$, $\tilde{C}_{Medium6,0}$)
+ min($\tilde{C}_{High1,0}$, $\tilde{C}_{Medium2,0}$, $\tilde{C}_{Medium3,0}$, $\tilde{C}_{Medium4,0}$, $\tilde{C}_{High5,0}$, $\tilde{C}_{Medium6,0}$)

= 23/36

Label 1:

$\tilde{C}_{High1,1}$= 13/18, $\tilde{C}_{Medium2,1}$= 0, $\tilde{C}_{Medium3,1}$= 17/6,
$\tilde{C}_{Medium4,1}$= 0, $\tilde{C}_{High5,1}$= 0, $\tilde{C}_{Medium6,1}$= 0

$\widetilde{D}_1$= max($\tilde{C}_{High1,1}$, $\tilde{C}_{Medium2,1}$, $\tilde{C}_{Medium3,1}$, $\tilde{C}_{Medium4,1}$, $\tilde{C}_{High5,1}$, $\tilde{C}_{Medium6,1}$)
+ min($\tilde{C}_{High1,1}$, $\tilde{C}_{Medium2,1}$, $\tilde{C}_{Medium3,1}$, $\tilde{C}_{Medium4,1}$, $\tilde{C}_{High5,1}$, $\tilde{C}_{Medium6,1}$)

= June 17

Label 2:

$\tilde{C}_{High1,2}$= 47/6, $\tilde{C}_{Medium2,2}$= 34/9, $\tilde{C}_{Medium3,2}$= 35/36,
$\tilde{C}_{Medium4,2}$= 47/6, $\tilde{C}_{High5,2}$= 47/6, $\tilde{C}_{Medium6,2}$= 13/18

$\widetilde{D}_2$= max($\tilde{C}_{High1,2}$, $\tilde{C}_{Medium2,2}$, $\tilde{C}_{Medium3,2}$, $\tilde{C}_{Medium4,2}$, $\tilde{C}_{High5,2}$, $\tilde{C}_{Medium6,2}$)
+ min($\tilde{C}_{High1,2}$, $\tilde{C}_{Medium2,2}$, $\tilde{C}_{Medium3,2}$, $\tilde{C}_{Medium4,2}$, $\tilde{C}_{High5,2}$, $\tilde{C}_{Medium6,2}$)

= 47/6 + 13/18

= 77/9

From there we have:

$$D_0 = \frac{23}{36}$$

$$D_1 = \frac{17}{6}$$

$$D_2 = \frac{77}{9}$$

According to the formula:

$$Label = p \ If \ \widetilde{D}_p = Max_{l=\overline{1,C}}(\widetilde{D}_l)$$

We have: $Max_{l=\overline{0,2}}(\widetilde{D}_l) = D_2 = \frac{77}{9}$

Therefore Label = 2

Using the Max operation, we obtain the output label of the new patient as 2 $\left(D_2 = \frac{77}{9}\right)$, from which we can conclude that the new patient shows signs of severe heart disease.

7.2. Diagnosis of patient's heart disease based on the M-CFIS-FKG model using FKG-Pairs-2

Input: Suppose the input of the problem is a list of 6 patients $\{R_1, R_2, R_3, R_4, R_5, R_6\}$, each patient has test results expressed through attributes $\{S_1, S_2, S_3, S_4, S_5, S_6\}$. The above patient cases have been examined and diagnosed based on test results by doctors, the diagnostic conclusions "Normal", "Heart" and "Severe heart" are shown respectively. labels 0, 1, 2. After going through the "Data Processing" stage, a fuzzy rule base system is obtained as shown in Table 1.

	S_1	S_2	S_3	S_4	S_5	S_6	Conclude
R_1	High	Medium	High	Medium	High	High	2
R_2	High	Medium	Medium	Medium	High	High	2
R_3	Medium	Medium	Medium	Medium	Medium	Medium	0
R_4	Medium	High	Medium	High	Medium	High	first
R_5	High	High	Medium	High	Low	High	first

R_6	High	High	High	Medium	High	Medium	2

Table 1: The fuzzy rule base system assumes the medical examination results of six patients who have been diagnosed by a doctor.

Besides, Input has an additional new patient case shown as follows:
IF *S_1= "High", S_2= "Medium", S_3= "Medium", S_4= "Medium", S_5= "High", S_6= "Medium"* ***THEN*** *Conclusion = ?*
Output: Provide diagnostic conclusions for the above patient, based on the fuzzy rule base system given by Input.

The steps for the above problem are performed sequentially as follows:

- **Step 1:** Calculate the weight sets $\tilde{A}$, $\tilde{B}$
  The weight set $\tilde{A}$includes edges connecting the linguistic labels of the patient's attributes, calculated according to formula (1)
  For example, in case {1}, the weights $\tilde{A}$will be calculated as follows:

$$\tilde{A}_{123}^{1} = \frac{|High \rightarrow Medium \rightarrow\ High|}{R} = \frac{1}{6}$$
$$\tilde{A}_{124}^{1} = \frac{|High \rightarrow High \rightarrow\ Medium|}{R} = \frac{1}{3}$$
$$\tilde{A}_{125}^{1} = \frac{|High \rightarrow Medium \rightarrow High|}{R} = \frac{1}{2}$$

  The weight set $\tilde{B}$of the edge connecting the labels of the attribute pairs with the output labels is calculated according to the formula:

$$\tilde{B}_{il}^{t} = \left(\sum \tilde{A}_{ij}^{t}\right) \times \min\left(\frac{|S_i \rightarrow l\ in\ rule\ t|}{|R|}, \frac{|S_j \rightarrow l\ in\ rule\ t|}{|R|}, \dots, \frac{|S_k \rightarrow l\ in\ rule\ t|}{|R|}\right)$$

$$\tilde{B}_{12l}^{1} = (\tilde{A}_{123}^{1} + \tilde{A}_{124}^{1} + \tilde{A}_{125}^{1} + \tilde{A}_{126}^{1} + \tilde{A}_{134}^{1} + \tilde{A}_{135}^{1} + \tilde{A}_{136}^{1} + \tilde{A}_{145}^{1} + \tilde{A}_{146}^{1} + \tilde{A}_{156}^{1} + \tilde{A}_{234}^{1} + \tilde{A}_{235}^{1} + \tilde{A}_{236}^{1} + \tilde{A}_{245}^{1} + \tilde{A}_{246}^{1} + \tilde{A}_{256}^{1} + \tilde{A}_{345}^{1} + \tilde{A}_{346}^{1} + \tilde{A}_{356}^{1} + \tilde{A}_{456}^{1}) \times \frac{|High \rightarrow Medium \rightarrow 2|}{R}$$

$$= \frac{17}{3} \times \frac{1}{3}$$

$$= \frac{17}{9}$$

The calculation results of the entire weight matrix are shown in Tables 2 and 3.

	R_1	R_2	R_3	R_4	R_5	R_6
$\tilde{A}_{123}^{t}$	1/6	1/6	1/6	1/6	1/6	1/6
$\tilde{A}_{124}^{t}$	1/3	1/3	1/6	1/6	1/6	1/6
$\tilde{A}_{125}^{t}$	1/3	1/3	1/6	1/6	1/6	1/6
$\tilde{A}_{126}^{t}$	1/3	1/3	1/6	1/6	1/6	1/6
$\tilde{A}_{134}^{t}$	1/3	1/6	1/6	1/6	1/6	1/3
$\tilde{A}_{135}^{t}$	1/3	1/6	1/3	1/3	1/6	1/3
$\tilde{A}_{136}^{t}$	1/6	1/3	1/6	1/6	1/3	1/6
$\tilde{A}_{145}^{t}$	1/2	1/2	1/6	1/6	1/6	1/2
$\tilde{A}_{146}^{t}$	1/3	1/3	1/6	1/6	1/6	1/6
$\tilde{A}_{156}^{t}$	1/3	1/3	1/6	1/6	1/6	1/6
$\tilde{A}_{234}^{t}$	1/6	1/3	1/3	1/3	1/3	1/6
$\tilde{A}_{235}^{t}$	1/6	1/6	1/6	1/6	1/6	1/6
$\tilde{A}_{236}^{t}$	1/6	1/6	1/6	1/3	1/3	1/6
$\tilde{A}_{245}^{t}$	1/3	1/3	1/6	1/6	1/6	1/6
$\tilde{A}_{246}^{t}$	1/3	1/3	1/6	1/3	1/3	1/6
$\tilde{A}_{256}^{t}$	1/3	1/3	1/6	1/6	1/6	1/6
$\tilde{A}_{345}^{t}$	1/3	1/6	1/6	1/6	1/6	1/3
$\tilde{A}_{346}^{t}$	1/6	1/6	1/6	1/3	1/3	1/6
$\tilde{A}_{356}^{t}$	1/6	1/6	1/6	1/6	1/6	1/6
$\tilde{A}_{456}^{t}$	1/3	1/3	1/6	1/6	1/6	1/6

Table 2: Results of calculating the weight matrix$\tilde{A}$

	R_1	R_2	R_3	R_4	R_5	R_6
$\tilde{B}^t_{12l}$	September 17	June 11	11/18	25/36	25/36	25/36
$\tilde{B}^t_{13l}$	September 17	December 11	11/18	25/36	25/36	25/18
$\tilde{B}^t_{14l}$	June 17	April 11	11/18	25/36	25/36	December 25
$\tilde{B}^t_{15l}$	June 17	April 11	11/18	25/36	25/36	December 25
$\tilde{B}^t_{16l}$	September 17	June 11	11/18	25/36	25/36	25/36
$\tilde{B}^t_{23l}$	17/18	December 11	11/18	25/18	25/18	25/36
$\tilde{B}^t_{24l}$	September 17	June 11	11/18	25/18	25/18	25/36
$\tilde{B}^t_{25l}$	September 17	June 11	11/18	25/18	25/18	25/36
$\tilde{B}^t_{26l}$	September 17	June 11	11/18	25/18	25/18	25/36
$\tilde{B}^t_{34l}$	September 17	December 11	11/18	25/18	25/18	25/18
$\tilde{B}^t_{35l}$	September 17	June 11	11/18	25/36	25/36	25/18
$\tilde{B}^t_{36l}$	17/18	December 11	11/18	25/18	25/18	25/36
$\tilde{B}^t_{45l}$	June 17	April 11	11/18	25/36	25/36	December 25
$\tilde{B}^t_{46l}$	September 17	June 11	11/18	25/18	25/18	25/36
$\tilde{B}^t_{56l}$	September 17	June 11	11/18	25/36	25/36	25/36

Table 3: Results of calculating the weight matrix $\tilde{B}$

⇨ The weight sets will be combined with the fuzzy rule base system to represent the fuzzy knowledge graph.

- **Step 2:** Apply approximate inference method to provide disease diagnosis results.
  After obtaining the fuzzy knowledge graph (represented based on the weight sets and fuzzy rule base system in Step 1), continue to diagnose the new patient's results using the approximate inference method. First, it is necessary to calculate the total weight of the edges $(\tilde{C})$from the super vertices to the output label based on the formula:

$$\tilde{C}_{ij\dots kl} = \sum_t \tilde{B}^t_{ij\dots kl}$$

$$\tilde{C}_{121} = \sum_t \tilde{B}^t_{121} = \tilde{B}^t_{121}(Rule\ 4) + \tilde{B}^t_{121}(Rule\ 5) = \frac{25}{36} + \frac{25}{36} = \frac{25}{18}$$

In there$t = \overline{1, n}, 1 \le i \le j < \dots < k < m - 1, l = \overline{1, C}$.

Specific calculation results are summarized in Table 4.

	$Label$ 0	$Label$ 1	$Label$ 2
$\tilde{C}_{12l}$	11/18	25/18	12/53
$\tilde{C}_{13l}$	11/18	25/18	79/9
$\tilde{C}_{14l}$	11/18	25/18	March 23
$\tilde{C}_{15l}$	11/18	25/18	March 23
$\tilde{C}_{16l}$	11/18	25/18	53/2
$\tilde{C}_{23l}$	11/18	25/9	September 23
$\tilde{C}_{24l}$	11/18	25/9	12/53
$\tilde{C}_{25l}$	11/18	25/9	12/53
$\tilde{C}_{26l}$	11/18	25/9	12/53
$\tilde{C}_{34l}$	11/18	25/9	151/36
$\tilde{C}_{35l}$	11/18	25/18	46/9
$\tilde{C}_{36l}$	11/18	25/9	23/9
$\tilde{C}_{45l}$	11/18	25/18	23/3
$\tilde{C}_{46l}$	11/18	25/9	12/53
$\tilde{C}_{56l}$	11/18	25/18	12/53

Table 4: Results of calculating the weight matrix$\tilde{C}$

Based on the table above, continue to calculate the values $(\widetilde{D})$according to the formula:

$$\widetilde{D}_l = Max_{1\le i\le j<\cdots\le k}(\tilde{C}_{ij\ldots kl}) + Min_{1\le i\le j<\cdots\le k}(\tilde{C}_{ij\ldots kl})$$

In there$t = \overline{1,n}, 1 \le i \le j < \cdots < k < m-1, l = \overline{1,C}$

Input has an additional new patient case represented as follows:

IF *S_1= "High", S_2= "Medium", S_3= "Medium", S_4= "Medium", S_5= "High", S_6= "Medium"* ***THEN*** *Conclusion = ?*

Label 0:

$\tilde{C}_{High1\to Medium2,0}$= *0,*

$\tilde{C}_{High1\to Medium3,0}$= *0,*

$\tilde{C}_{High1\to Medium4,0}$= *0,*

$\tilde{C}_{High1\to High5,0}$= *0,*

$\tilde{C}_{High1\to Medium6,0}$= *0,*

$\tilde{C}_{Medium2\to Medium3,0}$= *11/18,*

$\tilde{C}_{Medium2\to Medium4,0}$= *11/18,*

$\tilde{C}_{Medium2\to High5,0}$= *0,*

$\tilde{C}_{Medium2\to Medium6,0}$= *11/18,*

$\tilde{C}_{Medium3\to Medium4,0}$= *11/18,*

$\tilde{C}_{Medium3\to High5,0}$= *0,*

$\tilde{C}_{Medium3\to Medium6,0}$= *11/18,*

$\tilde{C}_{Medium4\to High5,0}$= *0,*

$\tilde{C}_{Medium4\to Medium6,0}$= *11/18,*

$\tilde{C}_{High5\to Medium6,0}$= *0*

$\widetilde{D}_0$=
max(
$\tilde{C}_{High1\to Medium2,0}, \tilde{C}_{High1\to Medium3,0}, \tilde{C}_{High1\to Medium4,0}, \tilde{C}_{High1\to High5,0}, \tilde{C}_{High1\to Medium6,0},$
$\tilde{C}_{Medium2\to Medium3,0}, \tilde{C}_{Medium2\to Medium4,0}, \tilde{C}_{Medium2\to High5,0},$
$\tilde{C}_{Medium2\to Medium6,0}, \tilde{C}_{Medium3\to Medium4,0}, \tilde{C}_{Medium3\to High5,0}, \tilde{C}_{Medium3\to Medium6,0},$

$\tilde{C}_{Medium4\rightarrow High5,0}, \tilde{C}_{Medium4\rightarrow Medium6,0}, \tilde{C}_{High5\rightarrow Medium6,0}) +$
$min($
$\tilde{C}_{High1\rightarrow Medium2,0}, \tilde{C}_{High1\rightarrow Medium3,0}, \tilde{C}_{High1\rightarrow Medium4,0}, \tilde{C}_{High1\rightarrow High5,0}, \tilde{C}_{High1\rightarrow Medium6,0},$
$\tilde{C}_{Medium2\rightarrow Medium3,0}, \tilde{C}_{Medium2\rightarrow Medium4,0}, \tilde{C}_{Medium2\rightarrow High5,0},$
$\tilde{C}_{Medium2\rightarrow Medium6,0}, \tilde{C}_{Medium3\rightarrow Medium4,0}, \tilde{C}_{Medium3\rightarrow High5,0}, \tilde{C}_{Medium3\rightarrow Medium6,0},$
$\tilde{C}_{Medium4\rightarrow High5,0}, \tilde{C}_{Medium4\rightarrow Medium6,0}, \tilde{C}_{High5\rightarrow Medium6,0})$

$= 11/18$

Label 1:

$\tilde{C}_{High1\rightarrow Medium2,1} = 0,$
$\tilde{C}_{High1\rightarrow Medium3,1} = 25/36,$
$\tilde{C}_{High1\rightarrow Medium4,1} = 0,$
$\tilde{C}_{High1\rightarrow High5,1} = 0,$
$\tilde{C}_{High1\rightarrow Medium6,1} = 0,$
$\tilde{C}_{Medium2\rightarrow Medium3,1} = 0,$
$\tilde{C}_{Medium2\rightarrow Medium4,1} = 0,$
$\tilde{C}_{Medium2\rightarrow High5,1} = 0,$
$\tilde{C}_{Medium2\rightarrow Medium6,1} = 0,$
$\tilde{C}_{Medium3\rightarrow Medium4,1} = 0,$
$\tilde{C}_{Medium3\rightarrow High5,1} = 0,$
$\tilde{C}_{Medium3\rightarrow Medium6,1} = 0,$
$\tilde{C}_{Medium4\rightarrow High5,1} = 0,$
$\tilde{C}_{Medium4\rightarrow Medium6,1} = 0,$
$\tilde{C}_{High5\rightarrow Medium6,1} = 0$

$\widetilde{D}_1 =$
$max($
$\tilde{C}_{High1\rightarrow Medium2,1}, \tilde{C}_{High1\rightarrow Medium3,1}, \tilde{C}_{High1\rightarrow Medium4,1}, \tilde{C}_{High1\rightarrow High5,1}, \tilde{C}_{High1\rightarrow Medium6,1},$
$\tilde{C}_{Medium2\rightarrow Medium3,1}, \tilde{C}_{Medium2\rightarrow Medium4,1}, \tilde{C}_{Medium2\rightarrow High5,1},$
$\tilde{C}_{Medium2\rightarrow Medium6,1}, \tilde{C}_{Medium3\rightarrow Medium4,1}, \tilde{C}_{Medium3\rightarrow High5,1}, \tilde{C}_{Medium3\rightarrow Medium6,1},$

$\tilde{C}_{Medium4\to High5,1}, \tilde{C}_{Medium4\to Medium6,1}, \tilde{C}_{High5\to Medium6,1}$) +
min(
$\tilde{C}_{High1\to Medium2,1}, \tilde{C}_{High1\to Medium3,1}, \tilde{C}_{High1\to Medium4,1}, \tilde{C}_{High1\to High5,1}, \tilde{C}_{High1\to Medium6,1},$
$\tilde{C}_{Medium2\to Medium3,1}, \tilde{C}_{Medium2\to Medium4,1}, \tilde{C}_{Medium2\to High5,1},$
$\tilde{C}_{Medium2\to Medium6,1}, \tilde{C}_{Medium3\to Medium4,1}, \tilde{C}_{Medium3\to High5,1}, \tilde{C}_{Medium3\to Medium6,1},$
$\tilde{C}_{Medium4\to High5,1}, \tilde{C}_{Medium4\to Medium6,1}, \tilde{C}_{High5\to Medium6,1}$)
= 25/36

Label 2:

$\tilde{C}_{High1\to Medium2,2}$ = 67/18,
$\tilde{C}_{High1\to Medium3,2}$ = 11/12,
$\tilde{C}_{High1\to Medium4,2}$ = March 23,
$\tilde{C}_{High1\to High5,2}$ = March 23,
$\tilde{C}_{High1\to Medium6,2}$ = 25/36,
$\tilde{C}_{Medium2\to Medium3,2}$ = 11/12,
$\tilde{C}_{Medium2\to Medium4,2}$ = 67/18,
$\tilde{C}_{Medium2\to High5,2}$ = 67/18,
$\tilde{C}_{Medium2\to Medium6,2}$ = 0,
$\tilde{C}_{Medium3\to Medium4,2}$ = 11/12,
$\tilde{C}_{Medium3\to High5,2}$ = June 11,
$\tilde{C}_{Medium3\to Medium6,2}$ = 0,
$\tilde{C}_{Medium4\to High5,2}$ = March 23,
$\tilde{C}_{Medium4\to Medium6,2}$ = 25/36,
$\tilde{C}_{High5\to Medium6,2}$ = 25/36

$\widetilde{D}_2$ =
max(
$\tilde{C}_{High1\to Medium2,2}, \tilde{C}_{High1\to Medium3,2}, \tilde{C}_{High1\to Medium4,2}, \tilde{C}_{High1\to High5,2}, \tilde{C}_{High1\to Medium6,2},$
$\tilde{C}_{Medium2\to Medium3,2}, \tilde{C}_{Medium2\to Medium4,2}, \tilde{C}_{Medium2\to High5,2},$
$\tilde{C}_{Medium2\to Medium6,2}, \tilde{C}_{Medium3\to Medium4,2}, \tilde{C}_{Medium3\to High5,2}, \tilde{C}_{Medium3\to Medium6,2},$
$\tilde{C}_{Medium4\to High5,2}, \tilde{C}_{Medium4\to Medium6,2}, \tilde{C}_{High5\to Medium6,2}$) +
min(

$\tilde{C}_{High1\rightarrow Medium2,2}, \tilde{C}_{High1\rightarrow Medium3,2}, \tilde{C}_{High1\rightarrow Medium4,2}, \tilde{C}_{High1\rightarrow High5,2}, \tilde{C}_{High1\rightarrow Medium6,2},$
$\tilde{C}_{Medium2\rightarrow Medium3,2}, \tilde{C}_{Medium2\rightarrow Medium4,2}, \tilde{C}_{Medium2\rightarrow High5,2},$
$\tilde{C}_{Medium2\rightarrow Medium6,2}, \tilde{C}_{Medium3\rightarrow Medium4,2}, \tilde{C}_{Medium3\rightarrow High5,2}, \tilde{C}_{Medium3\rightarrow Medium6,2},$
$\tilde{C}_{Medium4\rightarrow High5,2}, \tilde{C}_{Medium4\rightarrow Medium6,2}, \tilde{C}_{High5\rightarrow Medium6,2}\,)$
= March 23

From there we have:

$$D_0 = \frac{11}{18}$$

$$D_1 = \frac{25}{36}$$

$$D_2 = \frac{23}{3}$$

According to the formula:

$$Label = p\ If\ \widetilde{D}_p = Max_{l=\overline{1,C}}\left(\widetilde{D}_l\right)$$

We have: $Max_{l=\overline{0,2}}\left(\widetilde{D}_l\right) = D_2 = \frac{23}{3}$
Therefore Label = 2

Using the Max operation, we obtain the output label of the new patient as 2 $\left(D_2 = \frac{23}{3}\right)$, from which we can conclude that the new patient shows signs of severe heart disease.

8. Results of evaluating two models FKG-Pairs-1 and FKG-Pairs-2 using mathematical methods

Model	FKG-Pairs-1	FKG-Pairs-2
D_0	23/36	11/18
D_1	17/6	25/26
D_2	77/9	23/3

Based on the results of Pairs-1 and Pairs-2 and evaluating the difference between the values D_0, D_1, D_2, we can make some comments as follows:

1. Performance of the Pair 2 model: Pair 2 has a smaller difference between values than Pair 1. This suggests that the Pair 2 model has the ability to clearly and clearly classify and predict risks. D_0, D_1, D_2 more accurate than Pair 1. Consistency in prediction across labels also enhances model reliability.
2. Pair 1 model accuracy: Pair 1 has a larger difference between values D_0, D_1, D_2, indicating that the model may have difficulty clearly classifying risk groups. Large discrepancies may suggest that the model needs to be adjusted or improved to increase the accuracy and reliability of predictions.
3. Flexibility and applicability: Both models use the M-CFIS-FKG model with FKG-Pairs-k to predict the risk of severe or mild heart disease. However, their performance can depend on how the model is built and how the data is processed. For this particular application, Pair 2 appears to be more flexible and more effective in predicting risk.
    - ⇨ Continued research and improvement are needed: Model evaluation is only part of the process. Continued research and development of new methods are needed to improve model performance in predicting heart disease risk. Continued monitoring and evaluation of models is important to ensure the accuracy and reliability of predictions in real-world applications.

9. Expected progress of the next work

- Consult with doctors and look up thresholds to determine the thresholds of typical attributes such as height, weight, systolic blood pressure, diastolic blood pressure, etc. from there write fuzzy knowledge rules
    - ⇨ Build a set of fuzzy knowledge base rules for the data set to build a complete data set for training the heart disease diagnosis model.
- Find more different heart disease diagnosis data sets to serve the process of training and testing the heart disease diagnosis model.
- Continue to develop a complete disease diagnosis system/application using Python.

10. References

[1] Abdel-Basset M, Gamal A, Manogaran G, Son LH, Long HV (2019) A novel group decision making model based on neutrosophic sets for heart disease diagnosis. Multimed Tools Appl 79:9977–10002. https://doi.org/10.1007/s11042-019-07742-7

[2] Hai V.Pham , Cu Kim Long, Phan Hung Khanh and Ha Quoc Trung (2023) A Fuzzy Knowledge Graph Pairs-Based Application for Classification in Decision Making: Case Study of Preeclampsia Signs. https://www.mdpi.com/2078-2489/14/2/104

[3] Alves MA et al (2021) Explaining machine learning based diagnosis of COVID-19 from routine blood tests with decision trees and criteria graphs. Comput Biol Med 132. https://doi.org/10.1016/j.compbiomed.2021.104335

[4] Bai W, Ding J, Zhang C (2020) Dual hesitant fuzzy graphs with applications to multi-attribute decision making. Int J Cogn Comput Eng 1:18–26. https://doi.org/10.1016/j.ijcce.2020.09.002

[5] Bakhshipour A et al (2020) Application of decision trees and fuzzy inference system for quality classification and modeling of black and green tea based on visual features. In: Proc. Food Meas. Characterization, pp. 1–15

[6] FKG-Group (2021). Datasets and source codes of this paper are available at the following: https://github.com/CodePaper/FKG-Group

[7] Luong Thi Hong Lan, Tran Manh Tuan and others (2020) A New Complex Fuzzy Inference System With Fuzzy Knowledge Graph and Extensions in Decision Making. https://ieeexplore.ieee.org/abstract/document/9184876

[8] *Pham Minh Chuan , Tran Manh Tuan , Cu Kim Long , Nguyen Hong Tan (2023) APPLICATION OF FuzzY KNOWLEDGE GRAPH IN SUPPORTING DIAGNOSIS FOR PATIENTS WITH DIABETES.* https://jst.tnu.edu.vn/jst/article/view/9132

[9] Xiaonan Li, Kai Zhang, Guanyu Li & Bin Zhu (2021) A Chinese Knowledge Graph for Cardiovascular Disease. https://link.springer.com/chapter/10.1007/978-981-15-8411-4_239

[10] Boya Cheng, Yuan Zhang, Cai D ejun , Wan Qiu & 1 other (2018) Construction of traditional Chinese medicine Knowledge Graph using Data Mining and Expert Knowledge. https://typeset.io/papers/construction-of-traditional-chinese-medicine-knowledge-graph-crk2w3ri1f