

ĐẠI HỌC BÁCH KHOA HÀ NỘI  
TRƯỜNG CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG



**SOICT**

## PROJECT II

### BÁO CÁO CUỐI KỲ

# CHUẨN ĐOÁN BỆNH TIM MẠCH BẰNG SỬ DỤNG MÔ HÌNH SƠ ĐỒ TRI THỨC MỜ DẠNG CẶP FKG-PAIRS

Giảng viên hướng dẫn: PGS.TS Phạm Văn Hải

Sinh viên thực hiện: Đào Thành Mạnh 20211014

Hà Nội, Ngày 30 tháng 6 năm 2024

# MỤC LỤC

<b>MỤC LỤC</b> .....	<b>2</b>
<b>1. Giới thiệu tổng quan bài toán</b> .....	<b>4</b>
<b>2. Mục tiêu nghiên cứu</b> .....	<b>4</b>
2.1. Mục tiêu chung .....	4
2.2. Mục tiêu cụ thể .....	4
<b>3. Dữ liệu đầu vào</b> .....	<b>5</b>
3.1. Tập dữ liệu.....	5
3.2. Xử lý dữ liệu .....	7
<b>4. Kiến thức liên quan</b> .....	<b>8</b>
4.1. Tập mờ và logic mờ.....	8
4.2. Hệ suy diễn mờ.....	8
4.3. Đồ thị tri thức mờ .....	9
4.4. Suy luận xấp xỉ. ....	9
<b>5. Mô hình đề xuất</b> .....	<b>10</b>
5.1. Phát biểu bài toán .....	10
5.2. Mô hình đề xuất .....	10
5.2.1. Mô hình bài toán.....	10
5.2.2. Chức năng hệ thống.....	12
5.2.3. Phương pháp thực hiện .....	13
5.3. Thuật toán cài đặt FKG-Pairs .....	14
5.4. Lựa chọn công nghệ.....	15
<b>6. Phương pháp toán học</b> .....	<b>16</b>
6.1. Chuẩn đoán bệnh tim bệnh nhân dựa trên mô hình M-CFIS-FKG sử dụng FKG-Pairs-1 .....	16
6.2. Chuẩn đoán bệnh tim bệnh nhân dựa trên mô hình M-CFIS-FKG sử dụng FKG-Pairs-2.....	20
<b>7. Kết quả đánh giá bằng phương pháp toán học</b> .....	<b>27</b>
<b>8. Kết quả cài đặt</b> .....	<b>28</b>

8.1.	Phương pháp đánh giá .....	28
8.1.1.	Accuracy (Độ chính xác) .....	28
8.1.2.	Time Calculating (Thời gian tính toán) .....	28
8.1.3.	Precision (Độ chính xác tiên đoán).....	28
8.1.4.	Recall .....	28
8.1.5.	F1-Score .....	29
8.2.	Kết quả thực nghiệm.....	29
8.2.1.	Accuracy và Caculating time .....	29
8.2.2.	Precision, Recall và F1-Score .....	30
8.3.	Đánh giá kết quả thực nghiệm.....	33
8.3.1.	Mô hình FKG-Pairs 2.....	33
8.3.2.	Mô hình FKG-Pairs 3.....	33
8.4.	Giao diện Ứng dụng chẩn đoán bệnh tim bằng FKG-Pairs.....	33
8.4.1.	Màn hình chính .....	33
8.4.2.	Chức năng chẩn đoán bệnh .....	33
8.4.3.	Màn hình hướng dẫn .....	33
9.	<b>Kết luận .....</b>	<b>40</b>
10.	<b>Tài liệu tham khảo .....</b>	<b>41</b>

## 1. Giới thiệu tổng quan bài toán

Trong lĩnh vực chăm sóc sức khỏe, đặc biệt là trong việc đưa ra quyết định và chuẩn đoán về bệnh tim mạch, việc có sự hiểu biết sâu sắc về thông tin y tế và khả năng áp dụng các phương pháp và kiến thức y tế hiện đại là vô cùng quan trọng. Tuy nhiên, mặc dù đã có sự tiến bộ trong việc thu thập và tổ chức dữ liệu y tế, những vấn đề vẫn tồn tại do tính phức tạp và đa dạng của thông tin này, đặc biệt là khi đối mặt với các bệnh lý như bệnh tim mạch.

Sơ đồ tri thức (KG) truyền thống không thể hoàn toàn giải quyết các thách thức này. Mặc dù KG có thể tổ chức dữ liệu y tế và xây dựng các mối quan hệ giữa các yếu tố y tế khác nhau, nhưng nó thường không thể xử lý được tính mờ và không chắc chắn của thông tin y tế, cũng như không đủ linh hoạt để áp dụng các quy tắc và quan hệ phức tạp trong việc đưa ra quyết định và chuẩn đoán về bệnh tim mạch.

Trong khi đó, Sơ đồ Tri thức Mờ dạng cặp (FKG) mở ra một tiềm năng mới để cải thiện quá trình ra quyết định và chuẩn đoán về bệnh tim mạch. FKG không chỉ giúp biểu diễn thông tin y tế một cách linh hoạt hơn và xử lý tính mờ và không chắc chắn của dữ liệu, mà còn giải quyết vấn đề hiệu suất thấp của KG thông thường bằng cách sử dụng các cặp thông tin thay vì chỉ dùng một cặp đơn lẻ. Điều này tăng cường khả năng suy luận và đưa ra quyết định chính xác hơn trong việc điều trị và quản lý bệnh tim mạch, giúp nâng cao chất lượng chăm sóc sức khỏe và kết quả điều trị cho bệnh nhân.

## 2. Mục tiêu nghiên cứu

### 2.1. Mục tiêu chung

Phát triển một hệ thống/ứng dụng hỗ trợ ra quyết định và chuẩn đoán bệnh tim mạch dựa trên mô hình Sơ Đồ Tri Thức Mờ Dạng Cặp FKG-Pairs. Dự án tập trung vào xây dựng một cơ sở dữ liệu y tế đa dạng và chất lượng cao, sử dụng FKG-Pairs để biểu diễn thông tin y tế một cách linh hoạt và hiệu quả. Mục tiêu là phát triển một mô hình dự đoán bệnh tim mạch đáng tin cậy và chính xác, kết hợp cả khía cạnh dự đoán và phân tích giải thích. Cuối cùng, dự án sẽ đánh giá và tối ưu hóa hiệu suất của hệ thống trước khi triển khai vào thực tế y tế, nhằm cung cấp một công cụ hữu ích giúp nâng cao chất lượng chuẩn đoán và điều trị bệnh tim mạch, đồng thời tăng cường sự hiểu biết và tin cậy từ phía người sử dụng.

### 2.2. Mục tiêu cụ thể

Xây dựng các mô hình Sơ Đồ Tri Thức Mờ Dạng Cặp FKG-Pairs với các cặp FKG pairs  $k$  ( $k = 1, 2, 3$ ), phát triển thuật toán và phương pháp để xây dựng mô hình FKG-Pairs với độ chính xác và linh hoạt cao, tiến hành đánh giá hiệu năng của các mô hình FKG-Pairs dựa trên các tiêu chí như độ chính xác, độ phân loại và độ tin cậy, so sánh và phân tích kết quả của các mô hình FKG-Pairs để xác định mô hình tối ưu nhất cho việc chuẩn đoán bệnh tim mạch, tối ưu hóa và cải thiện

mô hình tối ưu nhất để đảm bảo hiệu suất và độ chính xác cao nhất trước khi triển khai vào thực tế y tế, cung cấp hướng dẫn và báo cáo chi tiết về quy trình và kết quả của dự án để chia sẻ kinh nghiệm và kiến thức cho cộng đồng y tế và nghiên cứu.

### 3. Dữ liệu đầu vào

#### 3.1. Tập dữ liệu

Dữ liệu thực nghiệm được thu thập trực tiếp từ trang chính của Kaggle: <https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset/data>

Bộ dữ liệu bao gồm 1025 dữ liệu bệnh nhân, 13 thuộc tính được lưu dưới dạng file **cardio\_train.csv**:

1. age
  2. sex
  3. chest pain type (4 values)
  4. resting blood pressure
  5. serum cholestoral in mg/dl
  6. fasting blood sugar > 120 mg/dl
  7. resting electrocardiographic results (values 0,1,2)
  8. maximum heart rate achieved
  9. exercise induced angina
  10. oldpeak = ST depression induced by exercise relative to rest
  11. the slope of the peak exercise ST segment
  12. number of major vessels (0-3) colored by flourosopy
  13. thal: 0 = normal; 1 = fixed defect; 2 = reversable defect
- The names and social security numbers of the patients were recently removed from the database, replaced with dummy values.

**Hình 3.1: Dữ liệu 13 thuộc tính của bộ dữ liệu bệnh tim**

Mô tả dữ liệu:

STT	Thuộc tính	Thang đo	Biến ngôn ngữ	Ký hiệu biến	Ghi chú
1	age	<40	Low	L1	Tuổi của bệnh nhân (tính theo năm)
		40-60	Medium	M1	
		61-80	High	H1	
2	sex	0	Female	0	Giới tính của bệnh nhân (0: nữ, 1: nam)
		1	Male	1	
3	cp	0	Typical angina	0	Loại đau ngực (0: Đau thắt ngực điển hình, 1: Đau thắt ngực không điển hình, 2: Đau không do thắt ngực, 3: Không triệu chứng)
		1	Atypical angina	1	
		2	Non-anginal pain	2	
		3	Asymptomatic	3	
4	trestbps	<90	Low	L4	Huyết áp khi nghỉ ngơi (tính bằng mm Hg) của bệnh nhân khi nhập viện
		90-120	Medium	M4	
		>120	High	H4	
5	chol	<200	Medium	M5	Mức cholesterol trong huyết thanh (tính bằng mg/dl)
		200-240	High	H5	
		>240	Very High	VH5	
6	fbs	0	Medium	M6	Mức đường huyết lúc đói (trên 120 mg/dl được xem là 1 và <= 120 mg/dl được xem là 0)
		1	High	H6	
7	restecg	0	Medium	M7	Kết quả điện tâm đồ khi nghỉ ngơi. Được phân loại thành ba loại: 0: Bình thường 1: Có bất thường sóng ST-T (đảo ngược sóng T và/hoặc ST tăng hoặc giảm > 0.05 mV) 2: Cho thấy có khả năng hoặc chắc chắn phì đại thất trái theo tiêu chí của Estes
		1	High	H7	
		2	Very High	VH7	
8	thalach	<100	Medium	M8	Nhịp tim tối đa đạt được
		>100	High	H8	
9	exang	0	No	0	Đau thắt ngực do gắng sức (1 = có, 0 = không)
		1	Yes	1	
10	oldpeak	<2	Low	L9	Suy giảm ST do gắng sức so với khi nghỉ ngơi (đơn vị -> suy giảm)
		>=2	High	H9	
11	slope	0	Upsloping	0	Độ dốc của đoạn ST khi gắng sức. Được phân loại thành ba loại: 0: Độ dốc tăng 1: Bằng phẳng 2: Độ dốc giảm
		1	Flat	1	
		2	Downsloping	2	
12	ca	0	Normal	0	Số lượng mạch máu chính (0-3) được nhuộm màu bằng cách chụp cắt lớp thallium: Kết quả kiểm tra cân bằng thallium (0: Bình thường, 1: Khuyết tật cố định, 2: Khuyết tật có thể đảo ngược, 3: Không được mô tả)
		1	Fixed defect	1	
		2	Reversible defect	2	
		3	Not described	3	
		4		4	
13	thal	0	Normal	0	Bệnh máu gọi là thalassemia. Được phân loại thành ba loại: 1: Bình thường 2: Khuyết tật cố định 3: Khuyết tật có thể đảo ngược
		1	Fixed	1	
		2	Reversible	2	
		3		3	

**Hình 3.2: Mô tả dữ liệu của 13 thuộc tính bệnh tim**

Tất cả các giá trị tập dữ liệu được thu thập tại thời điểm kiểm tra y tế.

File dữ liệu sau khi tải về được tổng hợp vào sheet **heart** trong file excel **Data.xlsx**:

age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
52	1	0	125	212	0	1	168	0	1	2	2	3	0
53	1	0	140	203	1	0	155	1	3,1	0	0	3	0
70	1	0	145	174	0	1	125	1	2,6	0	0	3	0
61	1	0	148	203	0	1	161	0	0	2	1	3	0
62	0	0	138	294	1	1	106	0	1,9	1	3	2	0
58	0	0	100	248	0	0	122	0	1	1	0	2	1
58	1	0	114	318	0	2	140	0	4,4	0	3	1	0
55	1	0	160	289	0	0	145	1	0,8	1	1	3	0
46	1	0	120	249	0	0	144	0	0,8	2	0	3	0
54	1	0	122	286	0	0	116	1	3,2	1	2	2	0
71	0	0	112	149	0	1	125	0	1,6	1	0	2	1
43	0	0	132	341	1	0	136	1	3	1	0	3	0
34	0	1	118	210	0	1	192	0	0,7	2	0	2	1
51	1	0	140	298	0	1	122	1	4,2	1	3	3	0
52	1	0	128	204	1	1	156	1	1	1	0	0	0
34	0	1	118	210	0	1	192	0	0,7	2	0	2	1
51	0	2	140	308	0	0	142	0	1,5	2	1	2	1
54	1	0	124	266	0	0	109	1	2,2	1	1	3	0
50	0	1	120	244	0	1	162	0	1,1	2	0	2	1
58	1	2	140	211	1	0	165	0	0	2	0	2	1
60	1	2	140	185	0	0	155	0	3	1	0	2	0
67	0	0	106	223	0	1	142	0	0,3	2	2	2	1
45	1	0	104	208	0	0	148	1	3	1	0	2	1
63	0	2	135	252	0	0	172	0	0	2	0	2	1
42	0	2	120	209	0	1	173	0	0	1	0	2	1
61	0	0	145	307	0	0	146	1	1	1	0	3	0
44	1	2	130	233	0	1	179	1	0,4	2	0	2	1
58	0	1	136	319	1	0	152	0	0	2	2	2	0
56	1	2	130	256	1	0	142	1	0,6	1	1	1	0
55	0	0	180	327	0	2	117	1	3,4	1	0	2	0
44	1	0	120	169	0	1	144	1	2,8	0	0	1	0
50	0	1	120	244	0	1	162	0	1,1	2	0	2	1
57	1	0	130	131	0	1	115	1	1,2	1	1	3	0
70	1	2	160	269	0	1	112	1	2,9	1	1	3	0
50	1	2	129	196	0	1	163	0	0	2	0	2	1
46	1	2	150	231	0	1	147	0	3,6	1	0	2	0

Hình 3.3: Dữ liệu bệnh tim trong file excel trước khi xử lý

### 3.2. Xử lý dữ liệu

Xử lý tập dữ liệu lấy từ kaggle bằng Visual Studio Code sau đó lưu dữ liệu đã xử lý vào các sheets của file excel gốc **Data.xlsx**

age	sex	cp	trestbps	chol	fb	restecg	thalach	exang	oldpeak	slope	ca	thal	target
Medium	1	0	High	High	Medium	High	High	0	Low	2	2	3	0
Medium	1	0	High	High	High	Medium	High	1	High	0	0	3	0
High	1	0	High	Medium	Medium	High	High	1	High	0	0	3	0
High	1	0	High	High	Medium	High	High	0	Low	2	1	3	0
High	0	0	High	Very High	High	High	High	0	Low	1	3	2	0
Medium	0	0	Medium	Very High	Medium	Medium	High	0	Low	1	0	2	1
Medium	1	0	Medium	Very High	Medium	Very High	High	0	High	0	3	1	0
Medium	1	0	High	Very High	Medium	Medium	High	1	Low	1	1	3	0
Medium	1	0	Medium	Very High	Medium	Medium	High	0	Low	2	0	3	0
Medium	1	0	High	Very High	Medium	Medium	High	1	High	1	2	2	0
High	0	0	Medium	Medium	Medium	High	High	0	Low	1	0	2	1
Medium	0	0	High	Very High	High	Medium	High	1	High	1	0	3	0
Low	0	1	Medium	High	Medium	High	High	0	Low	2	0	2	1
Medium	1	0	High	Very High	Medium	High	High	1	High	1	3	3	0
Medium	1	0	High	High	High	High	High	1	Low	1	0	0	0
Low	0	1	Medium	High	Medium	High	High	0	Low	2	0	2	1
Medium	0	2	High	Very High	Medium	Medium	High	0	Low	2	1	2	1
Medium	1	0	High	Very High	Medium	Medium	High	1	High	1	1	3	0
Medium	0	1	Medium	Very High	Medium	High	High	0	Low	2	0	2	1
Medium	1	2	High	High	High	Medium	High	0	Low	2	0	2	1
Medium	1	2	High	Medium	Medium	Medium	High	0	High	1	0	2	0
High	0	0	Medium	High	Medium	High	High	0	Low	2	2	2	1
Medium	1	0	Medium	High	Medium	Medium	High	1	High	1	0	2	1
High	0	2	High	Very High	Medium	Medium	High	0	Low	2	0	2	1
Medium	0	2	Medium	High	Medium	High	High	0	Low	1	0	2	1
High	0	0	High	Very High	Medium	Medium	High	1	Low	1	0	3	0
Medium	1	2	High	High	Medium	High	High	1	Low	2	0	2	1
Medium	0	1	High	Very High	High	Medium	High	0	Low	2	2	2	0
Medium	1	2	High	Very High	High	Medium	High	1	Low	1	1	1	0
Medium	0	0	High	Very High	Medium	Very High	High	1	High	1	0	2	0
Medium	1	0	Medium	Medium	Medium	High	High	1	High	0	0	1	0
Medium	0	1	Medium	Very High	Medium	High	High	0	Low	2	0	2	1
Medium	1	0	High	Medium	Medium	High	High	1	Low	1	1	3	0
High	1	2	High	Very High	Medium	High	High	1	High	1	1	3	0
Medium	1	2	High	Medium	Medium	High	High	0	Low	2	0	2	1
Medium	1	2	High	High	Medium	High	High	0	High	1	0	2	0

Hình 3.4: Dữ liệu bệnh tim trong file excel sau khi xử lý bằng ngưỡng

## 4. Kiến thức liên quan

### 4.1. Tập mờ và logic mờ.

Tập mờ lần đầu được giới thiệu bởi Zadeh vào năm 1965 , được giới thiệu như một công cụ toán học mới để giải quyết các vấn đề với thông tin mơ hồ, không chắc chắn. Khác với các tập hợp thông thường, đánh giá quan hệ thành viên của tập hợp theo logic nhị phân “một phần tử là thuộc hoặc không thuộc tập hợp”, logic mờ đánh giá quan hệ thành viên của một phần tử thông qua một hàm liên thuộc  $\mu \rightarrow [0,1]$ , thể hiện độ thuộc của một phần tử đối với một tập hợp

### 4.2. Hệ suy diễn mờ.

Suy diễn mờ là quá trình tìm ra kết luận cho tập giá trị đầu vào, dựa trên hệ luật mờ đã được tổng hợp. Các phương pháp suy diễn mờ thường được nhắc đến như suy diễn mờ Mamdani, suy diễn mờ Takagi-Sugeno, ... các hệ suy diễn kể trên còn được biết đến là các phương pháp suy diễn cổ điển, đã được sử dụng rộng rãi trong các hệ điều khiển tự động. Đồ thị tri thức mờ được biết đến như một phương pháp suy diễn mới, hiệu quả, và



có độ chính xác cao hơn so với các phương pháp suy diễn trước đó. Quy tắc chung cho quá trình áp dụng hệ suy diễn mờ được thể hiện qua ba bước:

- Mờ hóa (fuzzification) : Trong bước này, chúng ta cần xác định được thang đo giá trị và thuật ngữ mức độ tương ứng của từng thuộc tính đầu vào của bộ dữ liệu, sau đó là quá trình chuyển đổi từ các giá trị rõ của bộ dữ liệu đầu vào thành các giá trị mờ, dựa trên thang đo giá trị kết hợp với hàm thuộc đã được xây dựng trước đó, cuối cùng là kết hợp các giá trị mờ của từng mẫu dữ liệu đầu vào bằng các toán tử mờ ( AND, OR, NOT ) để đưa ra các luật biểu diễn dưới dạng các mệnh đề IF-THEN và đưa vào hệ cơ sở luật mờ
- Suy diễn mờ ( fuzzy inference ) : Sử dụng phương pháp suy diễn mờ để tìm ra kết quả đầu ra dựa trên hệ cơ sở luật mờ đã xây dựng ở bước 1
- Giải mờ ( deuzification ) : Chuyển đổi các giá trị đầu ra mờ được tìm ra ở bước 2 thành các giá trị rõ, đưa ra kết quả của bài toán.

#### 4.3. Đồ thị tri thức mờ

Thuật ngữ đồ thị tri thức mờ được giới thiệu lần đầu được tích hợp trong mô hình M-CFIS-FKG với mục đích ban đầu là mở rộng mô hình M-CFIS-R làm cho quá trình suy luận của mô hình này trong pha testing trở nên nhanh hơn. Kế thừa đặc điểm của đồ thị tri thức, nên về mặt hình thức, một đồ thị tri thức mờ bao gồm các đỉnh đại diện cho các nhân ngôn ngữ của thuộc tính và nhãn đầu ra của các luật, các cạnh tương ứng là cung nối giữa các đỉnh.

Cách tính giá trị trọng số của các cạnh của đồ thị tri thức mờ đã được trình bày chi tiết trong [7], và được tóm tắt gọn lại như sau:

- Đối với các cạnh nối giữa hai đỉnh thuộc tính, với mỗi cặp giá trị  $(X_i, X_j)$ ,  $1 \leq i \leq j \leq m$ , trong luật thứ  $t$   $R_t$ ,  $t = \overline{1, k}$ , trọng số  $A_{ij}^t$  của cạnh này được tính theo công thức:

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

- Đối với các cạnh nối giữa đỉnh thuộc tính và đỉnh nhãn đầu ra, với mỗi cặp  $(X_i, l)$ ,  $1 \leq i \leq j \leq m$ ,  $l = \overline{1, C}$ , trong luật thứ  $t$   $R_t$ ,  $t = \overline{1, k}$ , trọng số  $B_{il}^t$  của cạnh này được tính theo công thức:

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

Kết quả của hai bộ trọng số được lưu trữ trong một ma trận kề, đại diện cho đồ thị tri thức mờ đã được xây dựng.

#### 4.4. Suy luận xấp xỉ.

Suy luận xấp xỉ được định nghĩa như một công cụ để suy luận từ các mệnh đề có ý nghĩa không cụ thể rõ ràng thông qua logic mờ . Thông thường phương pháp suy luận xấp xỉ có độ chính xác của kết quả không cao bằng các kỹ thuật suy luận thông thường đối với các

dữ liệu rõ, tuy nhiên ưu thế của suy luận xấp xỉ là nó có thể thực hiện lập luận đối với các biến ngôn ngữ, hay ngôn ngữ tự nhiên là dữ liệu có nghĩa nhập nhằng, không rõ ràng

## 5. Mô hình đề xuất

### 5.1. Phát biểu bài toán

Đầu vào: Để xây dựng đồ thị tri thức mờ sử dụng trong bài toán này, ta cần tập các mẫu bệnh nhân đã được chẩn đoán bởi các bác sĩ và chuyên gia dựa trên các thuộc tính được đặt ra. Tập dữ liệu mẫu này thông qua quá trình tiền xử lý (mờ hóa) và lưu vào hệ cơ sở luật mờ như bảng 1, hệ cơ sở luật mờ này bao gồm  $n$  luật  $R_1, R_2, \dots, R_n$  đại diện cho các mẫu bệnh nhân,  $m$  thuộc tính  $S_1, S_2, \dots, S_m$  đại diện cho các triệu chứng của bệnh, và  $C$  nhãn đầu ra  $1, 2, 3, \dots, C$  đại diện cho kết luận chẩn đoán của bác sĩ

Ngoài ra, có thêm một bệnh nhân mới nằm ngoài hệ luật trên, được biểu diễn như sau

IF  $S_1$  là “Low” và  $S_2$  là “Low” và  $S_3$  là “High” và  $S_4$  là “Very high” và ... và  $S_{m-1}$  là “High” và  $S_m$  là “Low” THEN Kết luận = ?

	$S_1$	$S_2$	...	$S_{m-1}$	$S_m$	Kết luận
$R_1$	High	High	...	Very high	High	1
$R_2$	Medium	Medium	...	Medium	Low	2
...	...	...	...	...	...	...
$R_{n-1}$	Medium	Medium	...	Medium	Medium	2
$R_n$	Low	Medium	...	Low	Low	3
Triệu chứng của bệnh nhân	High Medium Low	High Medium	...	Very high High Medium Low	High Medium High	0,1,..., C

**Bảng 5.1: Hệ cơ sở luật mờ**

Đầu ra: Kết quả của hệ thống chẩn đoán đầu ra của bệnh nhân mới dựa trên cơ sở luật mờ đã có.

### 5.2. Mô hình đề xuất

#### 5.2.1. Mô hình bài toán.

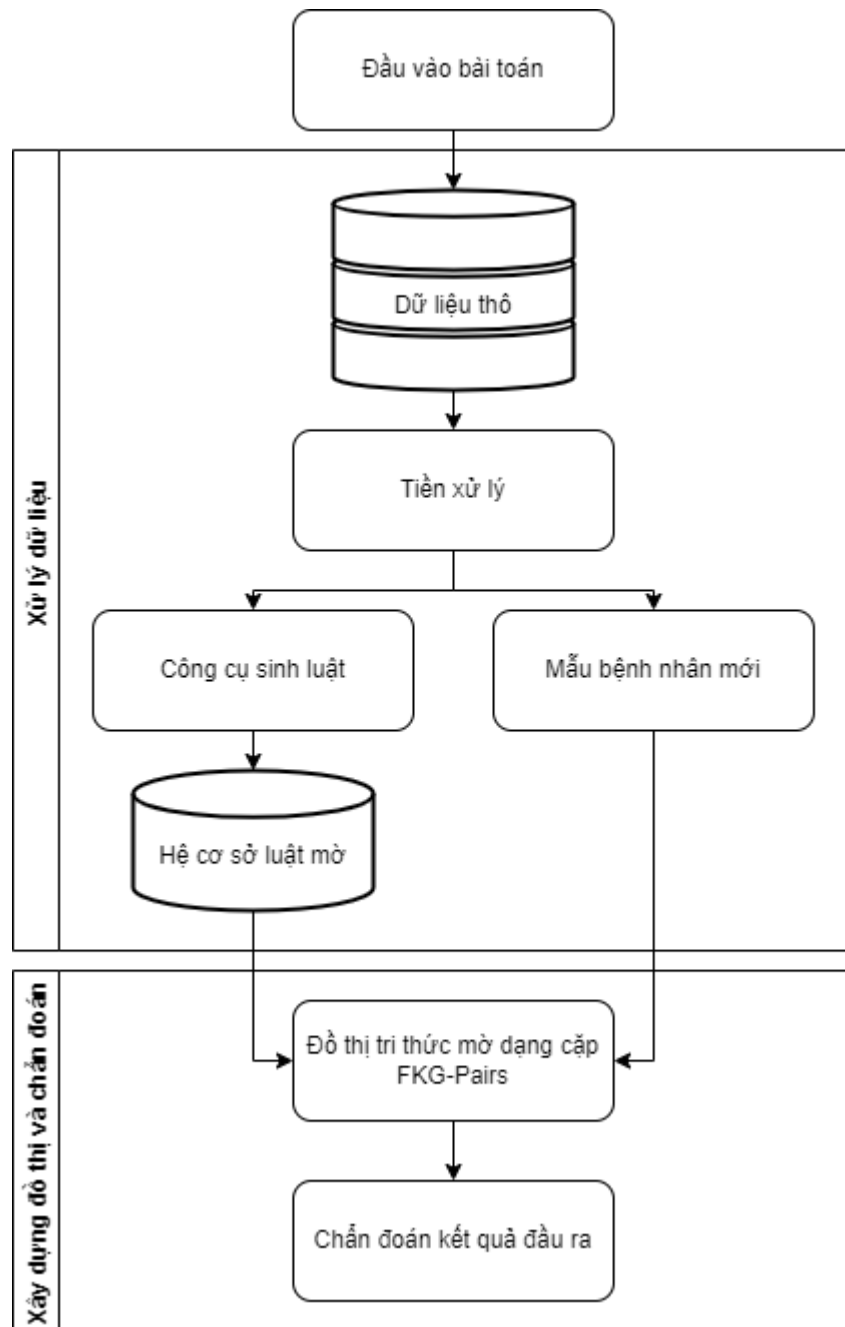
Hệ thống gồm 2 giai đoạn cần xử lý:

- Giai đoạn xử lý dữ liệu: Dữ liệu thu thập được từ các bác sĩ và chuyên gia là dữ liệu thô, dữ liệu này đi vào pha xử lý dữ liệu, bước tiền xử lý

bao gồm công đoạn phân chia thang đo, thiết kế thuật ngữ mức độ tương ứng, sau đó kết hợp với công cụ sinh luật để đưa ra hệ cơ sở luật mờ của bài toán, làm cơ sở cho giai đoạn xây dựng đồ thị tri thức mờ và chẩn đoán

- Giai đoạn xây dựng đồ thị và chẩn đoán : dựa trên hệ cơ sở luật mờ đã được hình thành ở bước trên, tiến hành xây dựng đồ thị tri thức mờ FKG bằng cách tính toán các bộ trọng số cạnh và lưu thành ma trận kề, từ đó sử dụng dữ liệu của đồ thị tri thức mờ FKG đó để chẩn đoán ca bệnh mới

Mô hình cho hai giai đoạn trên của bài toán đã được vẽ lại dưới đây:



**Hình 5.1: Mô hình bài toán chẩn đoán bệnh tim bằng FKG-Pairs**

### 5.2.2. Chức năng hệ thống

Hệ thống cần có các chức năng thực hiện được 2 giai đoạn kể trên

- Đối với giai đoạn chuẩn bị dữ liệu : dữ liệu thô được thu thập bởi người thiết kế hệ thống làm đầu vào của bài toán cần được phân tích, loại bỏ nhiễu, lỗi, và đưa ra bộ dữ liệu đầu vào hoàn chỉnh. Bộ dữ liệu này cần được làm mờ dựa trên thang đo và thuật ngữ mức độ được thống nhất

với các chuyên gia y tế trong lĩnh vực trên, thông qua công cụ sinh luật để tổng hợp thành hệ cơ sở luật mờ sử dụng cho các chức năng sau đó của hệ thống

- Đối với giai đoạn biểu diễn và chẩn đoán : Sau khi có được hệ cơ sở luật mờ, hệ thống sẽ phải xây dựng được đồ thị tri thức mờ FKG và lưu trữ đồ thị này sử dụng cho bước chẩn đoán. Tại bước chẩn đoán, hệ thống sẽ nhận đầu vào, kết hợp với đồ thị tri thức mờ FKG để tiến hành chẩn đoán và đưa ra kết luận.

### 5.2.3. Phương pháp thực hiện

Theo như mô tả trong phần chức năng hệ thống, bài toán sẽ được cài đặt theo các bước sau:

- Bước 1: Xác thực và xử lý dữ liệu  
Sau khi đã xây dựng được thang đo, tiến hành xây dựng các hàm và thủ tục để điều khiển xử lý tự động dữ liệu với các điều kiện cho từng thuộc tính dựa trên thang đo, kết hợp với công cụ sinh luật, cho đầu ra là hệ cơ sở luật mờ như trong bảng 1.
- Bước 2: Tính các bộ trọng số  $\tilde{A}, \tilde{B}$   
Bộ trọng số  $\tilde{A}$  của đồ thị tri thức mờ dạng cặp là trọng số của cạnh nối giữa các nhãn ngôn ngữ của các thuộc tính trong các luật  $t (R_t)$ . Các trọng số này được tính bởi công thức sau:

$$\tilde{A}_{ij...k}^t = \frac{|S_i \rightarrow S_j \rightarrow \dots \rightarrow S_{k+1} \text{ trong luật thứ } t|}{|R|} \quad (1)$$

Trong đó  $t = \overline{1, n}, 1 \leq i \leq j < \dots < k < m - 1$ .

Bộ trọng số  $\tilde{B}$  của đồ thị tri thức mờ dạng cặp là các trọng số của các cạnh nối giữa các nhãn của các cặp thuộc tính với nhãn đầu ra trong luật  $t (R_t)$ . Các trọng số này được tính bởi công thức sau :

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

Trong đó  $t = \overline{1, n}, 1 \leq i \leq j < \dots < k < m - 1, l = \overline{1, C}$ .

- Bước 3: Lưu trữ đồ thị tri thức mờ dạng cặp  
Các bộ trọng số sau khi tính toán cần được lưu trữ dưới dạng các ma trận kề để tiện cho việc sử dụng cho tính toán các bước tiếp theo, trong

thiết kế này, ma trận kê của các bộ trọng số được lưu trong 2 sheet của một file excel thuộc tệp hệ thống.

- Bước 4: Áp dụng quá trình suy luận xấp xỉ và đưa ra kết quả chẩn đoán bệnh.

Trong [4], tác giả Lương Thị Hồng Lan đã sử dụng thuật toán FISA để suy luận xấp xỉ đầu ra của bài toán xác định đầu ra của luật mới dựa trên đồ thị tri thức mờ, sau đó, tác giả Cù Kim Long đã cải tiến thuật toán trên trong bài báo [8] của mình cho phù hợp với việc xác định đầu ra của luật mới dựa trên đồ thị tri thức mờ dạng cặp, thuật toán được mô tả qua các bước dưới đây:

Trước tiên, cần tính toán tổng trọng số của các cạnh ( $\tilde{C}$ ) từ các siêu đỉnh đến nhãn đầu ra bằng cách sử dụng công thức sau:

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

Trong đó  $t = \overline{1, n}, 1 \leq i \leq j < \dots < k < m - 1, l = \overline{1, C}$ .

Sau đó, áp dụng các phép toán Max-Min để tính các giá trị ( $\tilde{D}$ ), các giá trị này giúp ước tính gần đúng các cặp thuộc tính mới trong tập Testing với các cặp thuộc tính tương ứng trong đồ thị tri thức mờ dạng cặp, nhằm tìm ra mức độ ảnh hưởng đến các nhãn đầu ra. Giá trị của ( $\tilde{D}$ ) được tính dựa trên các toán tử Max-Min theo công thức sau:

$$\tilde{D}_l = \text{Max}_{1 \leq i \leq j < \dots < k} (\tilde{C}_{ij\dots kl}) + \text{Min}_{1 \leq i \leq j < \dots < k} (\tilde{C}_{ij\dots kl}) \quad (4)$$

Trong đó  $t = \overline{1, n}, 1 \leq i \leq j < \dots < k < m - 1, l = \overline{1, C}$

Cuối cùng, kết quả chẩn đoán nhãn đầu ra của luật mới được kết luận bằng cách sử dụng phép toán Max như sau:

$$\text{Label} = p \text{ If } \tilde{D}_p = \text{Max}_{l=\overline{1, C}} (\tilde{D}_l) \quad (5)$$

### 5.3. Thuật toán cài đặt FKG-Pairs

#### Thuật toán cài đặt FKG dạng cặp

- 1 **Dữ liệu vào:** Tập dữ liệu kiểm thử, m: Số thuộc tính của mỗi luật, n: Số mẫu của tập dữ liệu, C: Số nhãn của mỗi thuộc tính.
- 2 **Dữ liệu ra:** Nhãn của mẫu mới.
- 3 **Begin**
- 4   Nhập các giá trị ;
- 5   Nhận tập dữ liệu kiểm thử;
- 6   Tiến hành mờ hóa tập dữ liệu kiểm thử;
- 7   **for**  $i = 1$  **to**  $m$  **do**

---

```

8      for  $t = 1$  to  $n$  do
9          for  $l = 1$  to  $C$  do
10             while  $1 \leq i \leq j < \dots \leq k$  do
11                 Tính  $\tilde{C}_{ij\dots kl}^t = \sum_t \tilde{B}_{ij\dots kl}^t$ 
12                 Tính  $\tilde{D}_l = \text{Max}_{1 \leq i \leq j < \dots \leq k}(\tilde{C}_{ij\dots kl}) + \text{Min}_{1 \leq i \leq j < \dots \leq k}(\tilde{C}_{ij\dots kl})$ 
13             end
14             Xác định nhãn của mẫu  $t$ :  $\text{Label} = p$  If  $\tilde{D}_p = \text{Max}_{l=1, \dots, C}(\tilde{D}_l)$ 
15         end
16         Nhận nhãn của mẫu  $t$  và lặp lại các bước từ 5 đến 12 để tìm nhãn của
        các mẫu khác cho đến khi kết thúc
17     end
18 end
19 end

```

---

#### *Thuật toán cài đặt bài toán FKG dạng cặp*

#### 5.4. Lựa chọn công nghệ

Sau khi phân tích bài toán, cũng như các bước tính toán và lập trình, để phù hợp cho việc xây dựng hệ thống một cách hiệu quả nhất, các công nghệ được lựa chọn cho thiết kế hệ thống cụ thể như sau:

a. Lựa chọn ngôn ngữ lập trình.

Ngôn ngữ lập trình chính được lựa chọn dùng cho thiết kế hệ thống trên là Python cụ thể là phiên bản Python 3.10 vì các ưu điểm dưới đây:

- Python là một ngôn ngữ lập trình dễ học, dễ sử dụng với các cấu trúc câu lệnh đơn giản, dễ hiểu và thực thi. Đồng thời, Python cũng thể hiện tính linh hoạt, hiệu quả, độ tin cậy và tốc độ cao, có thể sử dụng được trong nhiều môi trường khác nhau
- Cộng đồng hỗ trợ mạnh mẽ, do Python hiện là một ngôn ngữ ngày càng trở nên phổ biến và mạnh mẽ, nên cộng đồng người sử dụng cũng ngày càng lớn hơn. Khi gặp các vấn đề trong quá trình thiết kế thì việc tìm và gỡ lỗi sẽ dễ dàng hơn.
- Được hỗ trợ hàng trăm thư viện và Framework Python, do nhận được nhiều sự quan tâm từ người dùng cũng như là các doanh nghiệp nổi tiếng, nên Python được xây dựng rất nhiều thư viện hỗ trợ mạnh mẽ, ngoài ra còn có nhiều dịch vụ truyền thông đám mây cung cấp hỗ trợ đa nền tảng thông qua các công cụ giống như thư viện.
- Python được sử dụng như ngôn ngữ lập trình cốt lõi trong học thuật do vô số ứng dụng của nó trong Trí tuệ nhân tạo, Học sâu, Khoa học dữ liệu
- Tính tự động hóa, do được hỗ trợ rất nhiều công cụ và module có sẵn, nên Python cũng là công cụ tăng cường hiệu suất tốt nhất trong quá trình tự động hóa, kiểm thử phần mềm.

- IDE Pycharm hỗ trợ soạn thảo code Python nhanh, với nhiều extension hỗ trợ soạn thảo code nhanh hơn và miễn phí hầu hết các tính năng với tài khoản edu

b. Lựa chọn công cụ thiết kế giao diện ứng dụng.

Công cụ đi kèm với Python để thiết kế giao diện được lựa chọn là Kivy, đây là một công cụ mạnh mẽ, dùng cho nhiều nền tảng, được hỗ trợ bởi lượng người dùng lớn trong cộng đồng lập trình.

## 6. Phương pháp toán học

### 6.1. Chuẩn đoán bệnh tim bệnh nhân dựa trên mô hình M-CFIS-FKG sử dụng FKG-Pairs-1

Input: Giả sử đầu vào của bài toán là một danh sách gồm 6 bệnh nhân  $\{R_1, R_2, R_3, R_4, R_5, R_6\}$ , mỗi bệnh nhân có các kết quả xét nghiệm được thể hiện qua các thuộc tính  $\{S_1, S_2, S_3, S_4, S_5, S_6\}$ . Những trường hợp bệnh nhân nói trên đã được kiểm tra và chẩn đoán dựa trên các kết quả xét nghiệm bởi bác sĩ, các kết luận chẩn đoán “Bình thường”, “Tim ” và “Tim nặng” được thể hiện tương ứng với các nhãn 0, 1, 2. Sau khi qua giai đoạn “Xử lý dữ liệu” thu được một hệ cơ sở luật mờ như trong Bảng 1.

	$S_1$	$S_2$	$S_3$	$S_4$	$S_5$	$S_6$	Kết luận
$R_1$	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	1
$R_5$	High	High	Medium	High	Low	High	1
$R_6$	High	High	High	Medium	High	Medium	2

**Bảng 1: Hệ cơ sở luật mờ giả định kết quả khám bệnh của sáu bệnh nhân đã được kết luận chẩn đoán bởi bác sĩ**

Bên cạnh đó, Input có thêm một trường hợp bệnh nhân mới được biểu diễn như sau:

**IF**  $S_1 = \text{“High”}$ ,  $S_2 = \text{“Medium”}$ ,  $S_3 = \text{“Medium”}$ ,  $S_4 = \text{“Medium”}$ ,  $S_5 = \text{“High”}$ ,  $S_6 = \text{“Medium”}$  **THEN** Kết luận = ?

Output: Đưa ra kết luận chẩn đoán cho bệnh nhân trên, dựa trên hệ cơ sở luật mờ đã cho bởi Input.

Các bước thực hiện cho bài toán trên được thực hiện lần lượt như sau:

- **Bước 1:** Tính toán các bộ trọng số  $\tilde{A}$ ,  $\tilde{B}$



Bộ trọng số  $\tilde{A}$  gồm cạnh nối giữa các nhãn ngôn ngữ của các thuộc tính của bệnh nhân, được tính toán theo công thức (1)

Ví dụ, trong trường hợp  $\{1\}$ , trọng số  $\tilde{A}$  sẽ được tính toán như sau:

$$\begin{aligned}\tilde{A}_{12}^1 &= \frac{|High \rightarrow Medium|}{R} = \frac{1}{3} \\ \tilde{A}_{13}^1 &= \frac{|High \rightarrow High|}{R} = \frac{1}{3} \\ \tilde{A}_{14}^1 &= \frac{|High \rightarrow Medium|}{R} = \frac{1}{2}\end{aligned}$$

Bộ trọng số  $\tilde{B}$  của cạnh nối giữa nhãn của các cặp thuộc tính với các nhãn kết quả đầu ra được tính theo công thức:

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

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

Kết quả tính toán của toàn bộ ma trận trọng số được thể hiện qua Bảng 2 và Bảng 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

**Bảng 2: Kết quả tính ma trận trọng số  $\tilde{A}$**

	$R_1$	$R_2$	$R_3$	$R_4$	$R_5$	$R_6$
$\tilde{B}_{1l}^t$	11/4	35/12	23/36	25/36	13/18	13/6
$\tilde{B}_{2l}^t$	11/6	35/18	23/36	25/18	13/9	13/18
$\tilde{B}_{3l}^t$	11/6	35/36	23/36	25/18	13/9	13/9
$\tilde{B}_{4l}^t$	11/4	35/12	23/36	25/18	13/9	13/6
$\tilde{B}_{5l}^t$	11/4	35/12	23/36	25/36	13/18	13/6
$\tilde{B}_{6l}^t$	11/6	35/18	23/36	25/18	13/9	13/18

**Bảng 3: Kết quả tính ma trận trọng số  $\tilde{B}$**

⇒ Các bộ trọng số sẽ được kết hợp với hệ cơ sở luật mờ để biểu diễn đồ thị tri thức mờ.

- **Bước 2:** Áp dụng phương pháp suy luận xấp xỉ để đưa ra kết quả chẩn đoán bệnh.

Sau khi có được đồ thị tri thức mờ (được biểu diễn dựa trên các bộ trọng số và hệ cơ sở luật mờ ở Bước 1), tiếp tục tiến hành chẩn đoán kết quả của bệnh nhân mới bằng phương pháp suy luận xấp xỉ. Đầu tiên, cần tính toán tổng trọng số của các cạnh ( $\tilde{C}$ ) từ các siêu đỉnh đến nhãn đầu ra dựa theo công thức:

$$\tilde{C}_{ij...kl} = \sum_t \tilde{B}_{ij...kl}^t$$

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

Trong đó  $t = \overline{1, n}, 1 \leq i \leq j < \dots < k < m - 1, l = \overline{1, C}$ .

Kết quả tính toán cụ thể được tổng hợp lại trong Bảng 4.

	<i>Label 0</i>	<i>Label 1</i>	<i>Label 2</i>
$\tilde{C}_{1l}$	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

**Bảng 4: Kết quả tính ma trận trọng số  $\tilde{C}$**

Dựa vào bảng trên, tiếp tục tính các giá trị ( $\tilde{D}$ ) theo công thức:

$$\tilde{D}_l = \max_{1 \leq i \leq j < \dots \leq k} (\tilde{C}_{ij \dots kl}) + \min_{1 \leq i \leq j < \dots \leq k} (\tilde{C}_{ij \dots kl})$$

Trong đó  $t = \overline{1, n}, 1 \leq i \leq j < \dots < k < m - 1, l = \overline{1, C}$

Input có thêm một trường hợp bệnh nhân mới được biểu diễn như sau:

**IF**  $S_1 = \text{"High"}$ ,  $S_2 = \text{"Medium"}$ ,  $S_3 = \text{"Medium"}$ ,  $S_4 = \text{"Medium"}$ ,  $S_5 = \text{"High"}$ ,  $S_6 = \text{"Medium"}$  **THEN** Kết luận = ?

Label 0:

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

$$\begin{aligned} \tilde{D}_0 &= \max(\tilde{C}_{High1,0}, \tilde{C}_{Medium2,0}, \tilde{C}_{Medium3,0}, \tilde{C}_{Medium4,0}, \tilde{C}_{High5,0}, \tilde{C}_{Medium6,0}) \\ &+ \min(\tilde{C}_{High1,0}, \tilde{C}_{Medium2,0}, \tilde{C}_{Medium3,0}, \tilde{C}_{Medium4,0}, \tilde{C}_{High5,0}, \tilde{C}_{Medium6,0}) \\ &= 23/36 \end{aligned}$$

Label 1:

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

$$\begin{aligned} \tilde{D}_1 &= \max(\tilde{C}_{High1,1}, \tilde{C}_{Medium2,1}, \tilde{C}_{Medium3,1}, \tilde{C}_{Medium4,1}, \tilde{C}_{High5,1}, \tilde{C}_{Medium6,1}) \\ &+ \min(\tilde{C}_{High1,1}, \tilde{C}_{Medium2,1}, \tilde{C}_{Medium3,1}, \tilde{C}_{Medium4,1}, \tilde{C}_{High5,1}, \tilde{C}_{Medium6,1}) \\ &= 17/6 \end{aligned}$$

Label 2:

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

$$\begin{aligned} \tilde{D}_2 &= \max(\tilde{C}_{High1,2}, \tilde{C}_{Medium2,2}, \tilde{C}_{Medium3,2}, \tilde{C}_{Medium4,2}, \tilde{C}_{High5,2}, \tilde{C}_{Medium6,2}) \\ &+ \min(\tilde{C}_{High1,2}, \tilde{C}_{Medium2,2}, \tilde{C}_{Medium3,2}, \tilde{C}_{Medium4,2}, \tilde{C}_{High5,2}, \tilde{C}_{Medium6,2}) \\ &= 47/6 + 13/18 \\ &= 77/9 \end{aligned}$$

Từ đó ta có:

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

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

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

Theo công thức:

$$Label = p \text{ If } \tilde{D}_p = \text{Max}_{l=\overline{1,C}}(\tilde{D}_l)$$

$$Ta \text{ có: } \text{Max}_{l=\overline{0,2}}(\tilde{D}_l) = D_2 = \frac{77}{9}$$

$$Do \text{ đó } Label = 2$$

Sử dụng phép toán Max ta thu được nhãn đầu ra của bệnh nhân mới là  $2 \left( D_2 = \frac{77}{9} \right)$ , từ đó có thể kết luận, bệnh nhân mới có dấu hiệu mắc bệnh Tim nặng.

## 6.2. Chuẩn đoán bệnh tim bệnh nhân dựa trên mô hình M-CFIS-FKG sử dụng FKG-Pairs-2

Input: Giả sử đầu vào của bài toán là một danh sách gồm 6 bệnh nhân  $\{R_1, R_2, R_3, R_4, R_5, R_6\}$ , mỗi bệnh nhân có các kết quả xét nghiệm được thể hiện qua các thuộc tính  $\{S_1, S_2, S_3, S_4, S_5, S_6\}$ . Những trường hợp bệnh nhân nói trên đã được kiểm tra và chẩn đoán dựa trên các kết quả xét nghiệm bởi bác sĩ, các kết luận chẩn đoán “Bình thường”, “Tim” và “Tim nặng” được thể hiện tương ứng với các nhãn 0, 1, 2. Sau khi qua giai đoạn “Xử lý dữ liệu” thu được một hệ cơ sở luật mờ như trong Bảng 1.

	$S_1$	$S_2$	$S_3$	$S_4$	$S_5$	$S_6$	Kết luận
$R_1$	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	1
$R_5$	High	High	Medium	High	Low	High	1

$R_6$	High	High	High	Medium	High	Medium	2
-------	------	------	------	--------	------	--------	---

**Bảng 1: Hệ cơ sở luật mờ giả định kết quả khám bệnh của sáu bệnh nhân đã được kết luận chẩn đoán bởi bác sĩ**

Bên cạnh đó, Input có thêm một trường hợp bệnh nhân mới được biểu diễn như sau:

**IF**  $S_1 = \text{"High"}$ ,  $S_2 = \text{"Medium"}$ ,  $S_3 = \text{"Medium"}$ ,  $S_4 = \text{"Medium"}$ ,  $S_5 = \text{"High"}$ ,  $S_6 = \text{"Medium"}$  **THEN** Kết luận = ?

Output: Đưa ra kết luận chẩn đoán cho bệnh nhân trên, dựa trên hệ cơ sở luật mờ đã cho bởi Input.

Các bước thực hiện cho bài toán trên được thực hiện lần lượt như sau:

- **Bước 1:** Tính toán các bộ trọng số  $\tilde{A}$ ,  $\tilde{B}$

Bộ trọng số  $\tilde{A}$  gồm cạnh nối giữa các nhãn ngôn ngữ của các thuộc tính của bệnh nhân, được tính toán theo công thức (1)

Ví dụ, trong trường hợp  $\{1\}$ , trọng số  $\tilde{A}$  sẽ được tính toán như sau:

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

Bộ trọng số  $\tilde{B}$  của cạnh nối giữa nhãn của các cặp thuộc tính với các nhãn kết quả đầu ra được tính theo công thức:

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

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

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

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

Kết quả tính toán của toàn bộ ma trận trọng số được thể hiện qua Bảng 2 và Bảng 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

**Bảng 2: Kết quả tính ma trận trọng số  $\tilde{A}$**

	$R_1$	$R_2$	$R_3$	$R_4$	$R_5$	$R_6$
$\tilde{B}_{12l}^t$	17/9	11/6	11/18	25/36	25/36	25/36
$\tilde{B}_{13l}^t$	17/9	11/12	11/18	25/36	25/36	25/18
$\tilde{B}_{14l}^t$	17/6	11/4	11/18	25/36	25/36	25/12
$\tilde{B}_{15l}^t$	17/6	11/4	11/18	25/36	25/36	25/12
$\tilde{B}_{16l}^t$	17/9	11/6	11/18	25/36	25/36	25/36
$\tilde{B}_{23l}^t$	17/18	11/12	11/18	25/18	25/18	25/36
$\tilde{B}_{24l}^t$	17/9	11/6	11/18	25/18	25/18	25/36
$\tilde{B}_{25l}^t$	17/9	11/6	11/18	25/18	25/18	25/36
$\tilde{B}_{26l}^t$	17/9	11/6	11/18	25/18	25/18	25/36

$\tilde{B}_{34l}^t$	17/9	11/12	11/18	25/18	25/18	25/18
$\tilde{B}_{35l}^t$	17/9	11/6	11/18	25/36	25/36	25/18
$\tilde{B}_{36l}^t$	17/18	11/12	11/18	25/18	25/18	25/36
$\tilde{B}_{45l}^t$	17/6	11/4	11/18	25/36	25/36	25/12
$\tilde{B}_{46l}^t$	17/9	11/6	11/18	25/18	25/18	25/36
$\tilde{B}_{56l}^t$	17/9	11/6	11/18	25/36	25/36	25/36

**Bảng 3: Kết quả tính ma trận trọng số  $\tilde{B}$**

⇒ Các bộ trọng số sẽ được kết hợp với hệ cơ sở luật mờ để biểu diễn đồ thị tri thức mờ.

- **Bước 2:** Áp dụng phương pháp suy luận xấp xỉ để đưa ra kết quả chẩn đoán bệnh.

Sau khi có được đồ thị tri thức mờ (được biểu diễn dựa trên các bộ trọng số và hệ cơ sở luật mờ ở Bước 1), tiếp tục tiến hành chẩn đoán kết quả của bệnh nhân mới bằng phương pháp suy luận xấp xỉ. Đầu tiên, cần tính toán tổng trọng số của các cạnh ( $\tilde{C}$ ) từ các siêu đỉnh đến nhân đầu ra dựa theo công thức:

$$\tilde{C}_{ij...kl} = \sum_t \tilde{B}_{ij...kl}^t$$

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

Trong đó  $t = \overline{1, n}, 1 \leq i \leq j < \dots < k < m - 1, l = \overline{1, C}$ .

Kết quả tính toán cụ thể được tổng hợp lại trong Bảng 4.

	<i>Label 0</i>	<i>Label 1</i>	<i>Label 2</i>
$\tilde{C}_{12l}$	11/18	25/18	53/12
$\tilde{C}_{13l}$	11/18	25/18	79/9
$\tilde{C}_{14l}$	11/18	25/18	23/3
$\tilde{C}_{15l}$	11/18	25/18	23/3
$\tilde{C}_{16l}$	11/18	25/18	53/2
$\tilde{C}_{23l}$	11/18	25/9	23/9
$\tilde{C}_{24l}$	11/18	25/9	53/12
$\tilde{C}_{25l}$	11/18	25/9	53/12
$\tilde{C}_{26l}$	11/18	25/9	53/12
$\tilde{C}_{34l}$	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	53/12
$\tilde{C}_{56l}$	11/18	25/18	53/12

**Bảng 4: Kết quả tính ma trận trọng số  $\tilde{C}$**

Dựa vào bảng trên, tiếp tục tính các giá trị ( $\tilde{D}$ ) theo công thức:

$$\tilde{D}_l = \text{Max}_{1 \leq i \leq j < \dots \leq k} (\tilde{C}_{ij \dots kl}) + \text{Min}_{1 \leq i \leq j < \dots \leq k} (\tilde{C}_{ij \dots kl})$$

Trong đó  $t = \overline{1, n}, 1 \leq i \leq j < \dots < k < m - 1, l = \overline{1, C}$

Input có thêm một trường hợp bệnh nhân mới được biểu diễn như sau:

**IF**  $S_1 = \text{"High"}$ ,  $S_2 = \text{"Medium"}$ ,  $S_3 = \text{"Medium"}$ ,  $S_4 = \text{"Medium"}$ ,  $S_5 = \text{"High"}$ ,  $S_6 = \text{"Medium"}$  **THEN** Kết luận = ?

Label 0:

$$\begin{aligned} \tilde{C}_{\text{High}1 \rightarrow \text{Medium}2,0} &= 0, \\ \tilde{C}_{\text{High}1 \rightarrow \text{Medium}3,0} &= 0, \\ \tilde{C}_{\text{High}1 \rightarrow \text{Medium}4,0} &= 0, \\ \tilde{C}_{\text{High}1 \rightarrow \text{High}5,0} &= 0, \\ \tilde{C}_{\text{High}1 \rightarrow \text{Medium}6,0} &= 0, \\ \tilde{C}_{\text{Medium}2 \rightarrow \text{Medium}3,0} &= 11/18, \\ \tilde{C}_{\text{Medium}2 \rightarrow \text{Medium}4,0} &= 11/18, \\ \tilde{C}_{\text{Medium}2 \rightarrow \text{High}5,0} &= 0, \\ \tilde{C}_{\text{Medium}2 \rightarrow \text{Medium}6,0} &= 11/18, \\ \tilde{C}_{\text{Medium}3 \rightarrow \text{Medium}4,0} &= 11/18, \\ \tilde{C}_{\text{Medium}3 \rightarrow \text{High}5,0} &= 0, \\ \tilde{C}_{\text{Medium}3 \rightarrow \text{Medium}6,0} &= 11/18, \\ \tilde{C}_{\text{Medium}4 \rightarrow \text{High}5,0} &= 0, \\ \tilde{C}_{\text{Medium}4 \rightarrow \text{Medium}6,0} &= 11/18, \\ \tilde{C}_{\text{High}5 \rightarrow \text{Medium}6,0} &= 0 \end{aligned}$$

$$\tilde{D}_0 =$$

$\max($

$$\tilde{C}_{\text{High}1 \rightarrow \text{Medium}2,0}, \tilde{C}_{\text{High}1 \rightarrow \text{Medium}3,0}, \tilde{C}_{\text{High}1 \rightarrow \text{Medium}4,0}, \tilde{C}_{\text{High}1 \rightarrow \text{High}5,0}, \tilde{C}_{\text{High}1 \rightarrow \text{Medium}6,0}, \\ \tilde{C}_{\text{Medium}2 \rightarrow \text{Medium}3,0}, \tilde{C}_{\text{Medium}2 \rightarrow \text{Medium}4,0}, \tilde{C}_{\text{Medium}2 \rightarrow \text{High}5,0},$$

$$\tilde{C}_{\text{Medium}2 \rightarrow \text{Medium}6,0}, \tilde{C}_{\text{Medium}3 \rightarrow \text{Medium}4,0}, \tilde{C}_{\text{Medium}3 \rightarrow \text{High}5,0}, \tilde{C}_{\text{Medium}3 \rightarrow \text{Medium}6,0}, \\ \tilde{C}_{\text{Medium}4 \rightarrow \text{High}5,0}, \tilde{C}_{\text{Medium}4 \rightarrow \text{Medium}6,0}, \tilde{C}_{\text{High}5 \rightarrow \text{Medium}6,0} ) +$$

$\min($



$$\begin{aligned}
& \tilde{C}_{High1 \rightarrow Medium2,0}, \tilde{C}_{High1 \rightarrow Medium3,0}, \tilde{C}_{High1 \rightarrow Medium4,0}, \tilde{C}_{High1 \rightarrow High5,0}, \tilde{C}_{High1 \rightarrow Medium6,0}, \\
& \tilde{C}_{Medium2 \rightarrow Medium3,0}, \tilde{C}_{Medium2 \rightarrow Medium4,0}, \tilde{C}_{Medium2 \rightarrow High5,0}, \\
& \tilde{C}_{Medium2 \rightarrow Medium6,0}, \tilde{C}_{Medium3 \rightarrow Medium4,0}, \tilde{C}_{Medium3 \rightarrow High5,0}, \tilde{C}_{Medium3 \rightarrow Medium6,0}, \\
& \tilde{C}_{Medium4 \rightarrow High5,0}, \tilde{C}_{Medium4 \rightarrow Medium6,0}, \tilde{C}_{High5 \rightarrow Medium6,0} ) \\
& = 11/18
\end{aligned}$$

Label 1:

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

$$\begin{aligned}
& \tilde{D}_1 = \\
& \max( \\
& \tilde{C}_{High1 \rightarrow Medium2,1}, \tilde{C}_{High1 \rightarrow Medium3,1}, \tilde{C}_{High1 \rightarrow Medium4,1}, \tilde{C}_{High1 \rightarrow High5,1}, \tilde{C}_{High1 \rightarrow Medium6,1}, \\
& \tilde{C}_{Medium2 \rightarrow Medium3,1}, \tilde{C}_{Medium2 \rightarrow Medium4,1}, \tilde{C}_{Medium2 \rightarrow High5,1}, \\
& \tilde{C}_{Medium2 \rightarrow Medium6,1}, \tilde{C}_{Medium3 \rightarrow Medium4,1}, \tilde{C}_{Medium3 \rightarrow High5,1}, \tilde{C}_{Medium3 \rightarrow Medium6,1}, \\
& \tilde{C}_{Medium4 \rightarrow High5,1}, \tilde{C}_{Medium4 \rightarrow Medium6,1}, \tilde{C}_{High5 \rightarrow Medium6,1} ) + \\
& \min( \\
& \tilde{C}_{High1 \rightarrow Medium2,1}, \tilde{C}_{High1 \rightarrow Medium3,1}, \tilde{C}_{High1 \rightarrow Medium4,1}, \tilde{C}_{High1 \rightarrow High5,1}, \tilde{C}_{High1 \rightarrow Medium6,1}, \\
& \tilde{C}_{Medium2 \rightarrow Medium3,1}, \tilde{C}_{Medium2 \rightarrow Medium4,1}, \tilde{C}_{Medium2 \rightarrow High5,1}, \\
& \tilde{C}_{Medium2 \rightarrow Medium6,1}, \tilde{C}_{Medium3 \rightarrow Medium4,1}, \tilde{C}_{Medium3 \rightarrow High5,1}, \tilde{C}_{Medium3 \rightarrow Medium6,1}, \\
& \tilde{C}_{Medium4 \rightarrow High5,1}, \tilde{C}_{Medium4 \rightarrow Medium6,1}, \tilde{C}_{High5 \rightarrow Medium6,1} )
\end{aligned}$$

$$= 25/36$$

Label 2:

$$\begin{aligned}\tilde{C}_{High1 \rightarrow Medium2,2} &= 67/18, \\ \tilde{C}_{High1 \rightarrow Medium3,2} &= 11/12, \\ \tilde{C}_{High1 \rightarrow Medium4,2} &= 23/3, \\ \tilde{C}_{High1 \rightarrow High5,2} &= 23/3, \\ \tilde{C}_{High1 \rightarrow Medium6,2} &= 25/36, \\ \tilde{C}_{Medium2 \rightarrow Medium3,2} &= 11/12, \\ \tilde{C}_{Medium2 \rightarrow Medium4,2} &= 67/18, \\ \tilde{C}_{Medium2 \rightarrow High5,2} &= 67/18, \\ \tilde{C}_{Medium2 \rightarrow Medium6,2} &= 0, \\ \tilde{C}_{Medium3 \rightarrow Medium4,2} &= 11/12, \\ \tilde{C}_{Medium3 \rightarrow High5,2} &= 11/6, \\ \tilde{C}_{Medium3 \rightarrow Medium6,2} &= 0, \\ \tilde{C}_{Medium4 \rightarrow High5,2} &= 23/3, \\ \tilde{C}_{Medium4 \rightarrow Medium6,2} &= 25/36, \\ \tilde{C}_{High5 \rightarrow Medium6,2} &= 25/36\end{aligned}$$

$$\tilde{D}_2 =$$

$\max($

$$\begin{aligned}&\tilde{C}_{High1 \rightarrow Medium2,2}, \tilde{C}_{High1 \rightarrow Medium3,2}, \tilde{C}_{High1 \rightarrow Medium4,2}, \tilde{C}_{High1 \rightarrow High5,2}, \tilde{C}_{High1 \rightarrow Medium6,2}, \\ &\tilde{C}_{Medium2 \rightarrow Medium3,2}, \tilde{C}_{Medium2 \rightarrow Medium4,2}, \tilde{C}_{Medium2 \rightarrow High5,2}, \\ &\tilde{C}_{Medium2 \rightarrow Medium6,2}, \tilde{C}_{Medium3 \rightarrow Medium4,2}, \tilde{C}_{Medium3 \rightarrow High5,2}, \tilde{C}_{Medium3 \rightarrow Medium6,2}, \\ &\tilde{C}_{Medium4 \rightarrow High5,2}, \tilde{C}_{Medium4 \rightarrow Medium6,2}, \tilde{C}_{High5 \rightarrow Medium6,2} ) +\end{aligned}$$

$\min($

$$\begin{aligned}&\tilde{C}_{High1 \rightarrow Medium2,2}, \tilde{C}_{High1 \rightarrow Medium3,2}, \tilde{C}_{High1 \rightarrow Medium4,2}, \tilde{C}_{High1 \rightarrow High5,2}, \tilde{C}_{High1 \rightarrow Medium6,2}, \\ &\tilde{C}_{Medium2 \rightarrow Medium3,2}, \tilde{C}_{Medium2 \rightarrow Medium4,2}, \tilde{C}_{Medium2 \rightarrow High5,2}, \\ &\tilde{C}_{Medium2 \rightarrow Medium6,2}, \tilde{C}_{Medium3 \rightarrow Medium4,2}, \tilde{C}_{Medium3 \rightarrow High5,2}, \tilde{C}_{Medium3 \rightarrow Medium6,2}, \\ &\tilde{C}_{Medium4 \rightarrow High5,2}, \tilde{C}_{Medium4 \rightarrow Medium6,2}, \tilde{C}_{High5 \rightarrow Medium6,2} ) \\ &= 23/3\end{aligned}$$

Từ đó ta có:

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

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

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

Theo công thức:

$$Label = p \text{ If } \tilde{D}_p = \text{Max}_{l=\overline{1,C}}(\tilde{D}_l)$$

$$\text{Ta có: } \text{Max}_{l=\overline{0,2}}(\tilde{D}_l) = D_2 = \frac{23}{3}$$

Do đó  $Label = 2$

Sử dụng phép toán Max ta thu được nhãn đầu ra của bệnh nhân mới là 2 ( $D_2 = \frac{23}{3}$ ), từ đó có thể kết luận, bệnh nhân mới có dấu hiệu mắc bệnh Tim nặng.

## 7. Kết quả đánh giá bằng phương pháp toán học

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

**Bảng 7.1: Kết quả đánh giá 2 mô hình FKG-Pairs-1 và FKG-Pairs-2 bằng phương pháp toán học**

Dựa trên kết quả của hai cặp Pairs-1 và Pairs-2 và đánh giá chênh lệch giữa các giá trị  $D_0, D_1, D_2$ , ta có thể đưa ra một số nhận xét như sau:

- Hiệu suất của mô hình Pairs 2: Pairs 2 có độ chênh lệch giữa các giá trị  $D_0, D_1, D_2$  nhỏ hơn so với Pair 1. Điều này gợi ý rằng mô hình Pair 2 có khả năng phân loại và dự đoán nguy cơ một cách rõ ràng và chính xác hơn so với Pair 1. Sự nhất quán trong việc dự đoán giữa các nhãn cũng tăng cường tính đáng tin cậy của mô hình.
- Độ chính xác của mô hình Pairs 1: Pairs 1 có độ chênh lệch lớn hơn giữa các giá trị  $D_0, D_1, D_2$ , cho thấy mô hình có thể gặp khó khăn trong việc phân loại rõ ràng giữa các nhóm nguy cơ. Sự chênh lệch lớn có thể gợi ý rằng mô hình cần được điều chỉnh hoặc cải thiện để tăng tính chính xác và đáng tin cậy của dự đoán.
- Tính linh hoạt và tính ứng dụng: Cả hai mô hình đều sử dụng mô hình M-CFIS-FKG với FKG-Pairs-k để dự đoán nguy cơ mắc bệnh tim nặng hay nhẹ. Tuy nhiên, hiệu suất của chúng có thể phụ thuộc vào cách thức xây dựng mô hình và cách xử lý dữ liệu. Đối với ứng dụng cụ thể này, Pair 2 có vẻ linh hoạt hơn và hiệu quả hơn trong việc dự đoán nguy cơ.

## 8. Kết quả cài đặt

### 8.1. Phương pháp đánh giá

#### 8.1.1. Accuracy (Độ chính xác)

Accuracy là tỷ lệ các dự đoán đúng trên tổng số dự đoán. Công thức tính như sau:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

- TP (True Positive): Số lượng dự đoán đúng và thực tế cũng đúng.
- TN (True Negative): Số lượng dự đoán sai và thực tế cũng sai.
- FP (False Positive): Số lượng dự đoán đúng nhưng thực tế sai.
- FN (False Negative): Số lượng dự đoán sai nhưng thực tế đúng.

#### 8.1.2. Time Calculating (Thời gian tính toán)

Thời gian tính toán là thời gian mà mô hình cần để hoàn thành quá trình huấn luyện hoặc dự đoán. Thời gian này được tính bằng giây, phút hoặc giờ tùy thuộc vào độ phức tạp của mô hình và kích thước của dữ liệu.

#### 8.1.3. Precision (Độ chính xác tiên đoán)

Precision là tỷ lệ các dự đoán đúng (True Positive) trên tổng số các dự đoán được xác định là đúng (True Positive + False Positive):

$$\text{Precision} = \frac{TP}{TP+FP}$$

Precision cao có nghĩa là trong các trường hợp được dự đoán là dương tính, phần lớn là chính xác.

#### 8.1.4. Recall

Recall là tỷ lệ các dự đoán đúng (True Positive) trên tổng số các trường hợp thực tế là đúng (True Positive + False Negative):

$$\text{Recall} = \frac{TP}{TP+FN}$$

Recall cao có nghĩa là mô hình có thể phát hiện được hầu hết các trường hợp dương tính.

### 8.1.5. F1-Score

F1-Score là trung bình điều hòa của Precision và Recall, cung cấp một thước đo cân bằng giữa Precision và Recall:

$$\text{Precision} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

F1-Score cao cho thấy mô hình có cả Precision và Recall cao.

## 8.2. Kết quả thực nghiệm

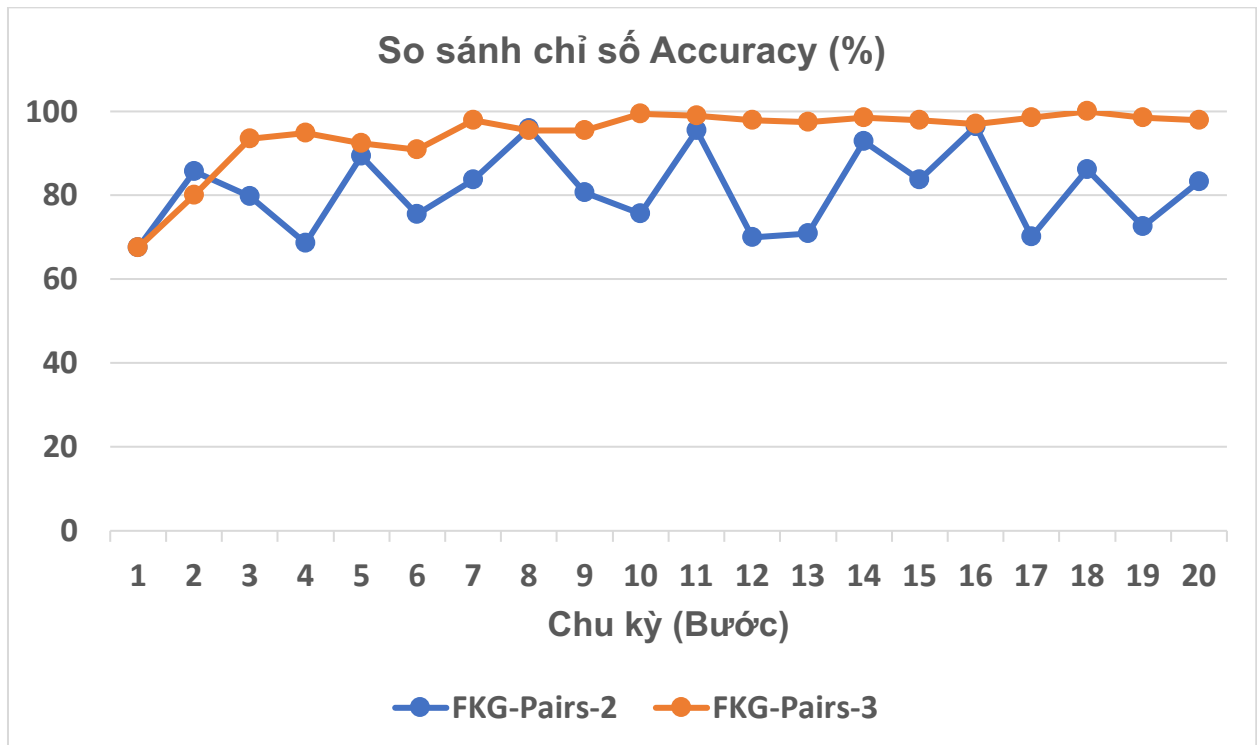
### 8.2.1. Accuracy và Caculating time

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
ACCURACY	FKG-Pairs-2	67.51	85.71	79.7	68.53	89.34	75.51	83.76	95.92	80.71	75.63	95.43	69.9	70.92	92.89	83.76	96.45	70.05	86.22	72.59	83.25	81,189
	FKG-Pairs-3	67.51	80.1	93.4	94.92	92.39	90.82	97.97	95.41	95.43	99.49	98.98	97.96	97.45	98.48	97.97	96.95	98.48	100	98.48	97.91	94,505
CACULATING TIME	FKG-Pairs-2	0.57	0.57	0.56	0.59	0.56	0.57	0.74	0.8	0.84	0.65	0.76	0.68	0.59	0.61	0.57	0.61	0.58	0.6	0.58	0.6	0,6315
	FKG-Pairs-3	0.74	0.92	1.04	1.04	1.34	1.65	1.39	1.48	1.59	1.8	1.97	2.6	2.43	2.82	3.02	3.58	3.56	4.04	4.2	4.75	2,298

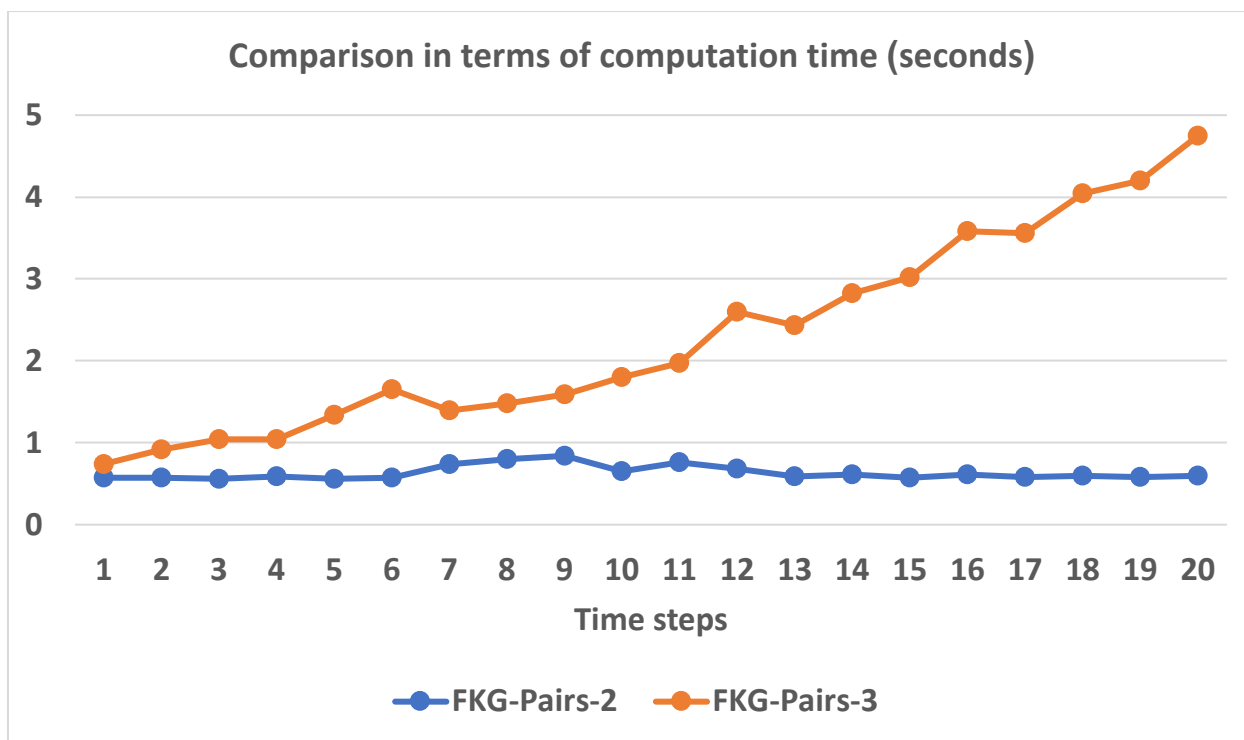
Hình 8.1. Kết quả huấn luyện FKG-Pairs 2 và FKG-Pairs 3

	FKG-Pairs 2	FKG-Pairs 3
Accuracy (%)	81,189	94,505
Time Caculating (s)	0,6315	2,298

Bảng 8.1. Accuracy và Time caculating lấy ra từ kết quả huấn luyện



Hình 8.2. So sánh chỉ số Accuracy của FKG-Pairs 2 và FKG-Pairs 3



**Hình 8.3: So sánh thời gian tính toán của FKG-Pairs 2 và FKG-Pairs 3**

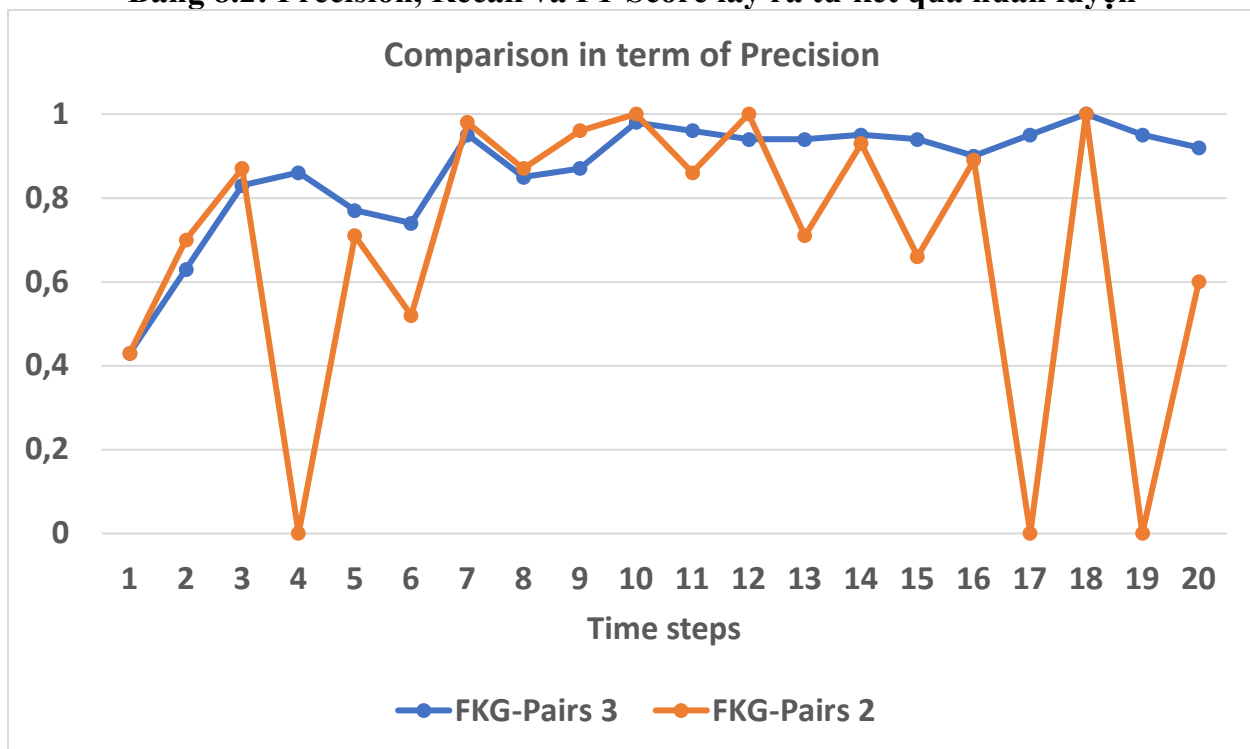
### 8.2.2. Precision, Recall và F1-Score

Steps	Precision		Recall		F1-Score	
	FKG-Pairs 3	FKG-Pairs 2	FKG-Pairs 3	FKG-Pairs 2	FKG-Pairs 3	FKG-Pairs 2
1	42,86	42,86	100	100	60	60
2	62,86	70,21	100	100	77,2	82,5
3	82,67	86,67	100	41,94	90,51	56,53
4	86,11	0	100	0	92,54	0
5	77,27	70,83	100	100	87,18	82,92
6	74,29	52	100	100	85,25	68,42
7	94,67	97,56	100	56,34	97,26	71,43
8	85,25	86,67	100	100	92,04	92,86
9	87,32	96,15	100	40,32	93,23	56,81
10	98	100	100	2,04	98,99	4
11	96,49	85,94	100	100	98,21	92,44
12	93,75	100	100	1,67	96,77	3,29
13	93,65	71,43	98,33	8,33	95,93	14,92
14	94,64	93,33	100	79,25	97,25	85,72
15	93,94	65,96	100	100	96,88	79,49
16	90,48	89,06	100	100	95	94,21
17	95,16	0	100	0	97,52	0
18	100	100	100	52,63	100	68,96
19	94,64	0	100	0	97,25	0
20	92,31	60	100	100	96	75
Avengers	86,818	68,4335	99,9165	59,126	92,2505	54,475

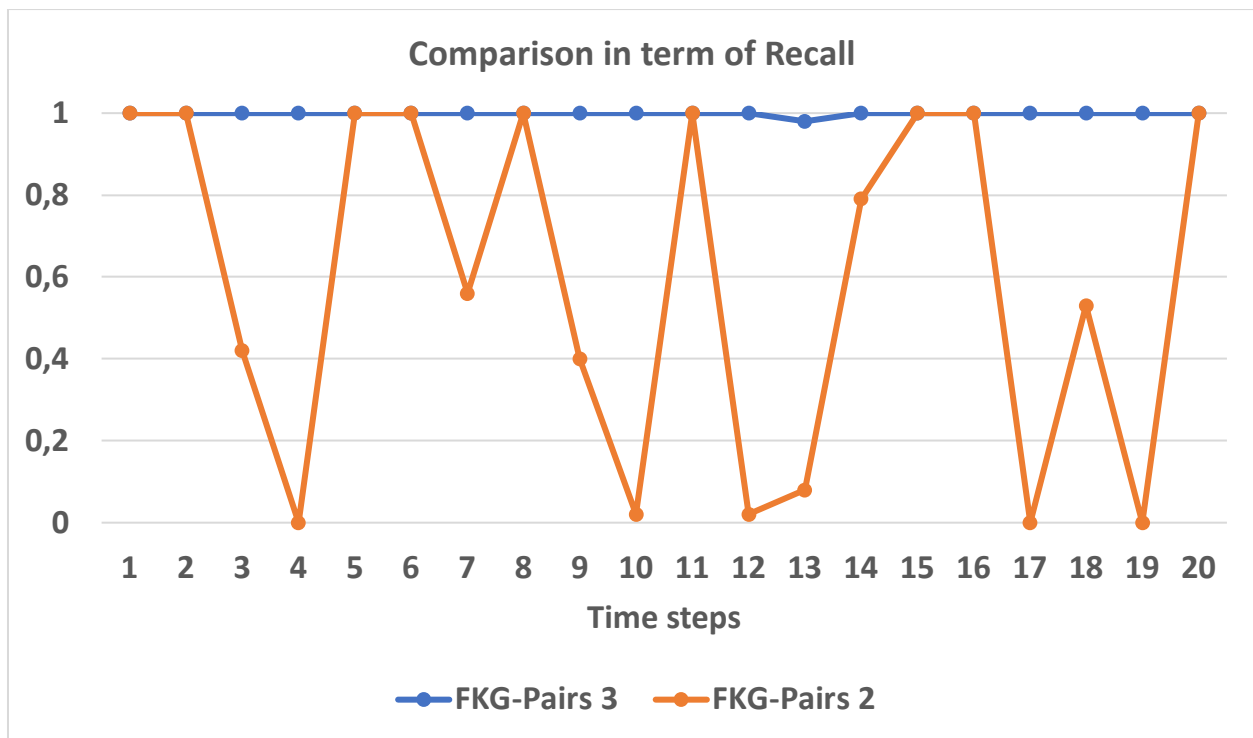
**Hình 8.4: Kết quả huấn luyện FKG-Pairs 2 và FKG-Pairs 3**

	FKG-Pairs 2	FKG-Pairs 3
Precision (%)	68,4335	86,818
Recall (%)	59,126	99,9165
F1-Score (%)	54,475	92,2505

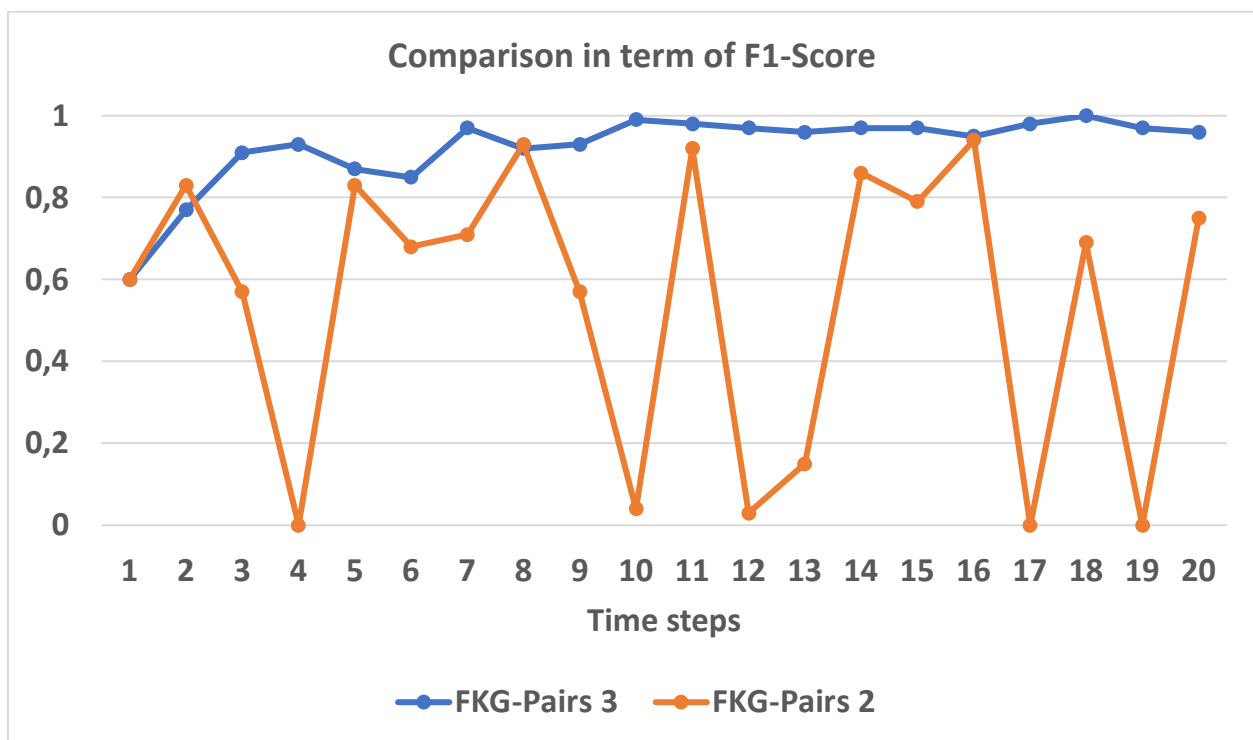
**Bảng 8.2: Precision, Recall và F1-Score lấy ra từ kết quả huấn luyện**



**Hình 8.5: So sánh Precision của FKG-Pairs 2 và FKG-Pairs 3**



**Hình 8.6: So sánh Recall của FKG-Pairs 2 và FKG-Pairs 3**



**Hình 8.7: So sánh F1-Score của FKG-Pairs 2 và FKG-Pairs 3**



### 8.3. Đánh giá kết quả thực nghiệm

#### 8.3.1. Mô hình FKG-Pairs 2

**Ưu điểm:** Thời gian tính toán rất nhanh.

**Nhược điểm:** Độ chính xác, Precision, Recall và F1-Score đều ở mức trung bình hoặc thấp.

**Phù hợp cho:** Các ứng dụng yêu cầu tốc độ tính toán cao và chấp nhận hiệu suất dự đoán ở mức trung bình.

#### 8.3.2. Mô hình FKG-Pairs 3

**Ưu điểm:** Độ chính xác, Precision, Recall và F1-Score đều rất cao, cho thấy hiệu suất tổng thể tốt.

**Nhược điểm:** Thời gian tính toán dài hơn so với FKG-Pairs 2.

**Phù hợp cho:** Các ứng dụng yêu cầu độ chính xác và hiệu suất cao, có thể chấp nhận thời gian tính toán dài hơn.

### 8.4. Giao diện Ứng dụng chẩn đoán bệnh tim bằng FKG-Pairs

#### 8.4.1. Màn hình chính

Màn hình chính của ứng dụng khá đơn giản, chỉ bao gồm chức năng chính của ứng dụng là chẩn đoán bệnh, và hướng dẫn sử dụng chi tiết ứng dụng được triển khai dưới dạng nút bấm chuyển hướng như Hình 1.

#### 8.4.2. Chức năng chẩn đoán bệnh

Sau khi bấm vào nút chẩn đoán bệnh, ứng dụng sẽ chuyển hướng đến màn hình nhập thông tin sơ bộ của bệnh nhân (Hình 3), sau khi nhập các thông tin vào ô tương ứng, bấm để chuyển hướng đến màn hình nhập các chỉ số xét nghiệm đã thu được của bệnh nhân (Hình 4), tại đây, tiếp tục nhập các chỉ số xét nghiệm để tiến hành chẩn đoán ( lưu ý các chỉ số đánh dấu (\*) là chỉ số không được để trống ), kết quả chẩn đoán sẽ được hiển thị như trên Hình 5, bao gồm kết quả chẩn đoán, kèm theo khuyến nghị của bác sĩ đối với bệnh nhân, tại đây, người dùng ứng dụng có thể có lựa chọn lưu trường hợp bệnh nhân mới vào cơ sở dữ liệu của ứng dụng (Hình 6), đồ thị tri thức mờ sẽ được cập nhật định kỳ để tăng độ chính xác của mô hình chẩn đoán.

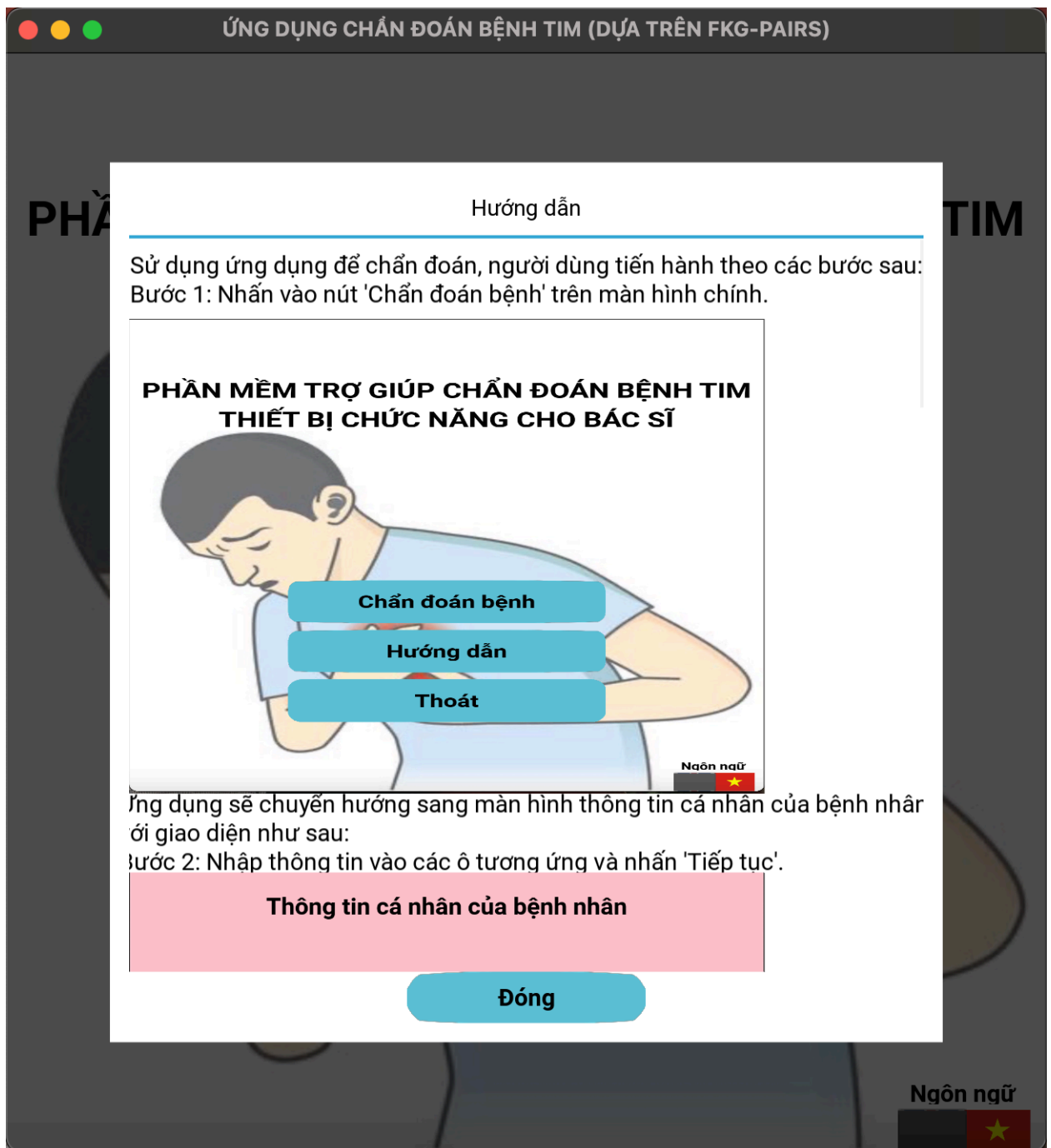
#### 8.4.3. Màn hình hướng dẫn

Trong màn hình này (Hình 2), sẽ là hướng dẫn sử dụng ứng dụng tương tự như đã trình bày ở phần trên.

## PHẦN MỀM TRỢ GIÚP CHẨN ĐOÁN BỆNH TIM THIẾT BỊ CHỨC NĂNG CHO BÁC SĨ



Hình 1: Màn hình chính



Hình 2: Màn hình hướng dẫn sử dụng

ỨNG DỤNG CHẨN ĐOÁN BỆNH TIM (DỰA TRÊN FKG-PAIRS)

### Thông tin cá nhân của bệnh nhân

1. Tuổi

20

(Tuổi)

2. Nghề nghiệp

Sinh viên

3. Chiều cao

173

(Cm)

4. Cân nặng

72

(Kg)

ĐÓNG

TIẾP TỤC

Hình 3: Màn hình nhập thông tin cá nhân

ỨNG DỤNG CHẨN ĐOÁN BỆNH TIM (DỰA TRÊN FKG-PAIRS)

Kết quả xét nghiệm của bệnh nhân

1. Tuổi

20

(Năm)

2. Giới tính

1

3. Loại đau thắt ngực

1

4. Huyết áp khi nghỉ ngơi

110

(mmHg)

5. Mức cholesterol

105

(mg/dl)

6. Mức đường huyết lúc đói

115

(mg/dl)

7. Điện tâm đồ

1

8. Nhịp tim tối đa

120

(Nhịp)

9. Đau thắt ngực

1

10. Suy giảm ST

2

11. Độ dốc ST

3

(Độ)

12. Lượng mạch máu chính

150

(Số lượng)

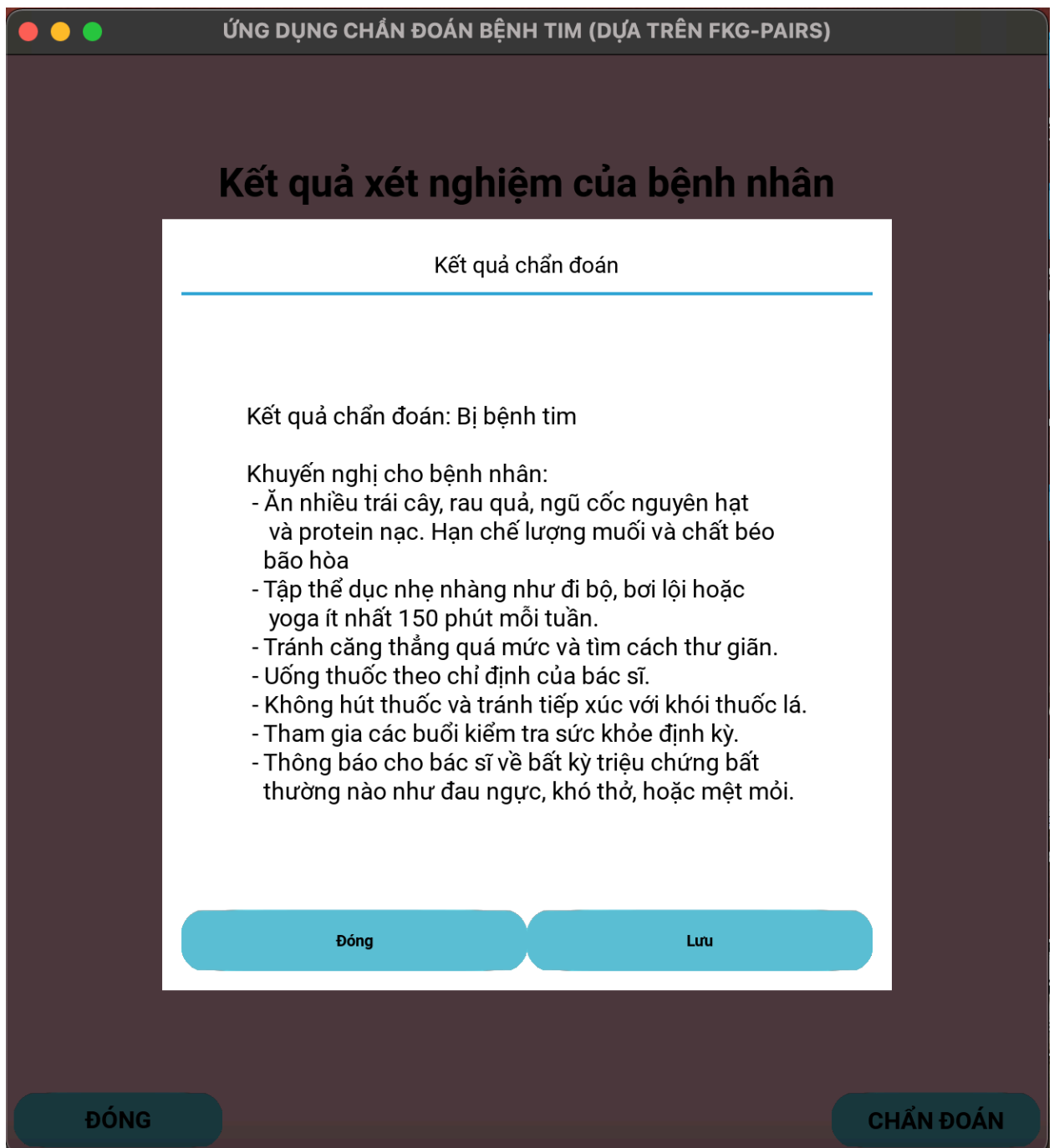
13. Chỉ số Thalassemia

1

ĐÓNG

CHẨN ĐOÁN

Hình 4: Màn hình nhập kết quả xét nghiệm của bệnh nhân



Hình 5: Màn hình in ra kết quả chẩn đoán và khuyến nghị

ỨNG DỤNG CHẨN ĐOÁN BỆNH TIM (DỰA TRÊN FKG-PAIRS)

### Kết quả xét nghiệm của bệnh nhân

1. Tuổi	20	(Năm)
2. Giới tính	1	
3. Loại đau thắt ngực	1	
4. Huyết áp khi nghỉ ngơi	110	(mmHg)
5. Mức cholesterol	105	(mg/dl)
6. Mức đường huyết		(mg/dl)
7. Điện tâm đồ		
8. Nhịp tim		(Nhịp)
9. Đa lượng		
10. Suy tim		
11. Độ dốc ST	3	(Độ)
12. Lượng mạch máu chính	150	(Số lượng)
13. Chỉ số Thalassemia	1	

Thông báo

Xác nhận lưu thông tin bệnh nhân ?  
Nhấn 'Có' để tiếp tục.

CóKhông

ĐÓNG

CHẨN ĐOÁN

Hình 6: Chức năng lưu thông tin

ỨNG DỤNG CHẨN ĐOÁN BỆNH TIM (DỰA TRÊN FKG-PAIRS)

### Kết quả xét nghiệm của bệnh nhân

1. Tuổi	20	(Năm)
2. Giới tính	1	
3. Loại đau thắt ngực	1	
4. Huyết áp khi nghỉ ngơi	110	(mmHg)
5. Mức cholesterol	105	(mg/dl)
6. M		(ng/dl)
7. Điện		
8. Nh		(Nhịp)
9. Đa		
10. Su		
11. Độ dốc ST	3	(Độ)
12. Lượng mạch máu chính	150	(Số lượng)
13. Chỉ số Thalassemia	1	

Thông báo

Lưu thông tin bệnh nhân thành công!

Đóng

ĐÓNG

CHẨN ĐOÁN

Hình 7: Màn hình hiển thị lưu thông tin thành công

## 9. Kết luận

Trong bản báo cáo này, em đã trình bày về quá trình nghiên cứu, tìm hiểu, và thiết kế ứng dụng hỗ trợ chẩn đoán bệnh tim dựa trên đồ thị tri thức mờ dạng cặp, thông qua quá trình thực hiện, em đã dùng phương pháp toán học và phương pháp thực nghiệm để đánh giá 3 mô hình tri thức mờ dạng cặp, đồng thời em cũng đã hiệu chỉnh được về các kiến



thức mới liên quan đến đồ thị tri thức, hệ suy diễn mờ, cũng như trau dồi các kỹ năng lập trình và thiết kế ứng dụng.

## 10. Tài liệu tham khảo

- [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.Phạm, 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.combiomed.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] Phạm Minh Chuẩn, Trần Mạnh Tuấn, Cù Kim Long, Nguyễn Hồng Tân (2023) ỨNG DỤNG ĐỒ THỊ TRI THỨC MỜ TRONG HỖ TRỢ CHẨN ĐOÁN CHO BỆNH NHÂN BỊ ĐÁI THÁO ĐƯỜNG. <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](https://link.springer.com/chapter/10.1007/978-981-15-8411-4_239)
- [10] Boya Cheng, Yuan Zhang, Cai Dejun, 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-cr2w3ri1f>