

MDPI

Article

# A Fuzzy Knowledge Graph Pairs-Based Application for Classification in Decision Making: Case Study of Preeclampsia Signs

Hai Van Pham 1,\* D, Cu Kim Long 1,2,\*, Phan Hung Khanh 1 and Ha Quoc Trung 2

- School of Information Communication Technology, Hanoi University of Science and Technology (HUST), Hanoi 100000, Vietnam
- <sup>2</sup> Information Technology Center, Ministry of Science and Technology (MOST), Hanoi 100000, Vietnam
- \* Correspondence: haipv@soict.hust.edu.vn (H.V.P.); longck.2006@gmail.com (C.K.L.)

Abstract: Problems of preeclampsia sign diagnosis are mostly based on symptom data with the characteristics of data collected periodically in uncertain, ambiguous, and obstetrician opinions. To reduce the effects of preeclampsia, many studies have investigated the disease, prevention, and complication. Conventional fuzzy inference techniques can solve several diagnosis problems in health such as fuzzy inference systems (FIS), and Mamdani complex fuzzy inference systems with rule reduction (M-CFIS-R), however, the computation time is quite high. Recently, the research direction of approximate inference based on fuzzy knowledge graph (FKG) has been proposed in the M-CFIS-FKG model with the combination of regimens in traditional medicine and subclinical data gathered from medical records. The paper has presented a proposed model of FKG-Pairs3 to support patients' disease diagnosis, together with doctors' preferences in decision-making. The proposed model has been implemented in real-world applications for disease diagnosis in traditional medicine based on input data sets with vague information, quantified by doctor's preferences. To validate the proposed model, it has been tested in a real-world case study of preeclampsia signs in a hospital for disease diagnosis with the traditional medicine approach. Experimental results show that the proposed model has demonstrated the model's effectiveness in the decision-making of preeclampsia signs.

Keywords: fuzzy knowledge graph; FKG-Pairs; disease diagnosis; preeclampsia; decision making



Citation: Pham, H.V.; Long, C.K.; Khanh, P.H.; Trung, H.Q. A Fuzzy Knowledge Graph Pairs-Based Application for Classification in Decision Making: Case Study of Preeclampsia Signs. *Information* **2023**, *14*, 104. https://doi.org/10.3390/ info14020104

Academic Editors: Shlomo Berkovsky, Sidong Liu and Cristián Castillo Olea

Received: 16 December 2022 Revised: 1 February 2023 Accepted: 2 February 2023 Published: 7 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

# 1. Introduction

Recently, preeclampsia signs have several typical clinical symptoms such as high blood pressure, proteinuria, and edema. Severe cases may be accompanied by convulsions and narcotisms [1–3]. Preeclampsia is a sort of multi-organ dysfunction related to pregnancy, accounting for about 2–10% during the entire pregnancy [1,3,4]. Over the past two decades, the preeclampsia rate has increased by about 25%, especially in the early preeclampsia group. In Asia, a statistic from 2001 to 2014 shows that preeclampsia increased quickly, from 0.5% to 0.8% during the entire pregnancy [1]. In Viet Nam, the preeclampsia rate before 34 weeks is 0.43%, the preeclampsia rate from 34 to 37 weeks is 0.7%, and the preeclampsia rate after 37 weeks is 1.68% compared to the entire pregnancy. A series of studies from 2012 to 2016 in Hue showed the preeclampsia rate is about 2.8–5.5% [2]. In applied AI healthcare applications, many studies of preeclampsia pathology in recent years have focused on the field of disease occurrence prediction, disease progression prediction, and pregnancy outcomes diagnosis as well as preeclampsia prophylaxis.

Preeclampsia is a pathology with multi-variable evidence for both mother and fetus. This causes of death maternal and perinatal mortality worldwide. Maternal mortality associated with hypertension in general pregnancy and preeclampsia accounts for about 14%. Death rate maternal mortality is associated with an increase in hypertension during

Information 2023, 14, 104 2 of 20

pregnancy in the range of 12.9–16.1% [1,4]. Besides, the effects of preeclampsia are given after birth, in relation to subsequent births, and are a risk factor for later cardiovascular diseases [5]. Despite efforts in the management of the prenatal phase, preeclampsia is still one of the disease burdens in maternal and child healthcare. To reduce the effects of preeclampsia, many investigations have been given to aim disease forecasting and prophylaxis, optimally export prophylaxis presents the disease, prevents severe progression, and prevents complications. Since 2011, WHO has issued recommendations for forecasting and prophylaxis of preeclampsia pathology [1]. Many organizations (Federation Obstetrics and Gynecology International (FIGO), American Society of Obstetricians and Gynecologists (ACOG), Healthcare Institute British National and Clinical Health (NICE), Canadian Society of Obstetricians and Gynecologists (SOGC)) and other specialized associations have also provided guidance on forecasting and redundancy preeclampsia [6–9]. With the motivation to apply information technology in early disease diagnosis, the authors have endeavored to deploy an application based on regimes in traditional medicine to solve the problem of preeclampsia sign diagnosis in pregnant women.

To handle problems in disease diagnosis, a knowledge graph-based approach is considered to support doctors in disease diagnosis [10–12]. KG is considered a powerful technique to support decision support systems, predictive analysis systems, and recommendation systems. It can be combined with other techniques to find the output labels of new samples. However, knowledge graphs face difficulties in representing knowledge and making approximate inferences based on input data sets with unclear information (such as subclinical symptom data with amplitude factor in the medical sector).

To solve the limitations of the knowledge graph, several techniques based on fuzzy inference systems are applied to build real-world applications. It has received much attention from many researchers all over the world, such as fuzzy inference systems (FIS) [13–18], complex fuzzy inference systems (CFIS) [19–21], and Mamdani-type complex fuzzy inference systems (M-CFIS) [22], M-CFIS-R [23]. These techniques cannot generate the output labels when new samples are not in the fuzzy rule base. Furthermore, these techniques can also show that the computation time is still quite high. Nevertheless, the M-CFIS-FKG still has low accuracy with incomplete information input data sets. Recently, the research direction of approximate inference based on fuzzy knowledge graph (FKG) has been proposed in the M-CFIS-FKG model [24] with the combination of regimens in traditional medicine and subclinical data gathered from electric medical records. This helps FKG to overcome the previous works' drawbacks.

From the above limitations of the M-CFIS-FKG, a new model (so-called FKG-Pairs) was proposed in [25]. It is considered an extension of FKG in the M-CFIS-FKG model [24]. It improved the single-pairs FKG (FKG-Pairs1) by using combinations of attribute pairs to compute the weights and inference of the output label (e.g., double-pairs FKG (FKG-Pairs2), triple-pairs FKG (FKG-Pairs3), quadruple-pairs FKG (FKG-Pairs4), and quintuplepairs FKG (FKG-Pairs5)). These methods have been applied to improve the inference performance of decision-making systems in terms of accuracy. In related works [26–28], the investigations have proposed Decision Support System to apply ontology-based for diabetic patients [26], Fuzzy Knowledge applied to give decision-making in diagnosis Decision Support System [27]. Further investigations have focused on human hearing abilities classification [28], applied U-Net with Deep Learning of glaucoma [29], and multiple Machine Learning techniques for decision making of chronic kidney disease [30], Prediction and analysis using dynamic neural network, genetic algorithm [31], and medical system has been applied to use deep learning techniques for clinic diagnosis [32]. In clinical decision support systems, knowledge reasoning and linguistics can be used in decision-making [33], multi-attribute decision-making [34,35], and group decision-making [36] for significant contributions to medical diagnosis. As mentioned in the related works, with incomplete information input data sets, conventional methods or fuzzy inference techniques have not been considered fully effective solutions. Studying to deploy easy and convenient FKG-Pairs-based applications is to meet the requirements for real-world problems.

Information 2023, 14, 104 3 of 20

The paper has presented a proposed model of FKG-Pairs3 to support patients' disease diagnosis, together with doctors' preferences in decision-making. The proposed model has been tested in a real-world case study of preeclampsia signs in a hospital for disease diagnosis with the traditional medicine approach. The **main contributions** of this study are as follows:

- Proposing the FKG-Pairs3-based preeclampsia sign diagnosis model. This proposed
  model is used to quantify incomplete and vague information input data sets including
  qualitative and quantitative factors, quantified from doctor's preferences.
- Applying FKG-Pairs3 for consideration of patient's symptom pairs to approximate reasoning to find the output labels of new samples, matched with the class of classification in decision-making problems with incomplete information input data sets.
- Implementing the proposed model for the development of real-world applications
  using datasets collected from medical records in accordance with the preferences of
  doctors through a case study of preeclampsia sign diagnosis for support of traditional
  medicine inpatient medical records.

The rest of this paper is organized as follows. Part 2 presents the background knowledge of the research process. Part 3 describes the detailed proposed model in medical diagnosis including the problem statement and the process of building and applying the FKG-Pairs to solve the problem. Experimental results and performance evaluation of the application are shown in Part 4. The final part is to give conclusions and future investigations.

## 2. Research Background

# 2.1. Fuzzy Sets

Fuzzy sets were first introduced by Zadeh in 1965 [37], introduced as a new mathematical tool for solving problems with ambiguous, uncertain information. Unlike normal sets, which evaluate the membership of the set according to the binary logic "an element belongs or not to the set", fuzzy logic [38,39] evaluates the membership of a part element through a membership function  $\mu \to [0,1]$ , which represents the membership of an element to a set.

## 2.2. Fuzzy Inference System

A fuzzy inference system (FIS) is a popular computational framework based on the concept of fuzzy set theory, often applied when building decision support systems in case the input information is not clear [16]. The general framework of the FIS as shown in Figure 1 can be summarized with three main parts: fuzzification, knowledge base, and defuzzification. The FIS has the basic structure as follows:

- Fuzzification: It is responsible for converting input values into language values.
- The knowledge base consists of two parts: The database (definition of fuzzy set membership functions used in fuzzy rules) and the set of rules (including IF-THEN structural fuzzy rules).
- Engine: Perform inference operations in the fuzzy rule base.

Defuzzification: It is responsible for converting the fuzzy result values of the fuzzy inference system into clear values.

Fuzzy inference is the process of finding a conclusion for a set of input values, based on the synthesized fuzzy rule system. Fuzzy inference methods are regularly referred to as FIS, CFIS, and M-CFIS. These inference systems, also known as classical inference methods, have been widely used in automatic control systems and decision support systems. The FKG [24] is known as a new, efficient, and more accurate inference method than previous fuzzy inference methods. Fuzzy inference systems are classified into three main methods: Mamdani fuzzy inference systems, Takagi-Sugeno fuzzy inference systems, and Tsukamoto fuzzy inference systems.

Information 2023, 14, 104 4 of 20

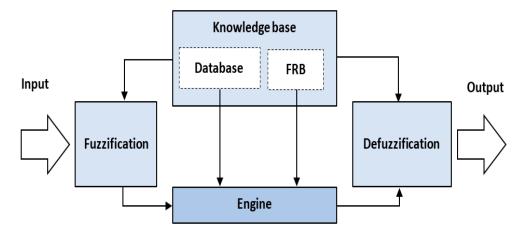


Figure 1. The general framework of the FIS.

#### 2.3. Knowledge Graph

Knowledge graph (KG) was first introduced by Google in 2012 [40]. The main purpose of KG is to analyze to maximize the value of knowledge, detect and avoid errors, and to be able to infer new conclusions from existing data. The selection of new entity representations and their relationships through the KG model can gain a lot of useful information and can be more supportive for practical applications. It is for this reason that KG is researched, proposed, and applied by the community of researchers in many practical problems, especially in models with approximate reasoning. It is considered a powerful technique to support decision support systems, predictive analysis systems, and recommendation systems [41–43].

#### 2.4. Fuzzy Knowledge Graph

The FKG was first proposed in 2020 [24] to solve the limitations of KG in representing knowledge and making approximate inferences based on input data sets with unclear or incomplete information by using linguistic labels for the attributes in the training set connected to the output labels. It represents inference through natural law where the impact of language labels is capable of generating corresponding output labels. By accumulating single events (or single pairs in the FKG), it can determine the final output of a new sample. The FKG has two main phases including representation and approximate reasoning.

In the representation phase, the weights of edges are calculated and are briefly summarized as follows. Firstly, for edges connecting among vertices or input attributes' labels on FKG, the weight  $A_{ii}^t$  of these edges is calculated by using Equation (1):

$$A_{ij}^{t} = \frac{\left| X_{i} \text{ is related to } X_{j} \text{ in rule } t \right|}{|R|} \tag{1}$$

where:  $X_i, X_j$  are input attribute vertices,  $1 \le i \le j \le m$ ,  $R_t$  is the  $t^{th}$  rule,  $t = \overline{1, k}$ .

Secondly, for edges connecting the input attribute label and the output label on FKG, the weight  $B_{ij}^t$  of these edges is calculated by using Equation (2):

$$B_{il}^{t} = \left(\sum A_{ij}^{t}\right) \times \frac{|X_{i} \text{ is related to the label l in rule t}|}{|R|}$$
 (2)

where:  $X_i$  is input attribute vertex,  $1 \le i < j < m$ ,  $R_t$  is the  $t^{th}$  rule,  $t = \overline{1, k}$ , l is output label vertex and  $l = \overline{1, C}$ .

The results of the two sets of weights are stored in an adjacency matrix, which can be used in the FISA algorithm in the next phase.

In the approximate reasoning phase, the FISA algorithm [24] is applied to approximate reasoning and find the output labels of the new records.

Information 2023, 14, 104 5 of 20

# 2.5. Approximate Reasoning and Decision Making

Approximate reasoning is defined as a tool for inferences from propositions of unspecified meaning through fuzzy logic [43]. Usually, the approximate inference method has a lower accuracy than the conventional inference techniques based on clear data, but the advantage of approximation reasoning unclear data with language variables. In [24], it is applied in the FISA algorithm to find the output labels of the new records.

Decision-making is considered the core of decision-support systems. It supports leaders and managers make accurate, timely, and effective decisions. It is applied in many sectors in the real world, such as healthcare, finance, stock, transportation, environment, agriculture, business, and other studies [31,42–46].

# 3. The Proposed Model for Preeclampsia Sign Diagnosis in Decision Making

This section describes the preeclampsia sign diagnosis proposed model in detail. The main contents are presented in this section including the problem statement, the proposed model, and a numerical example to elucidate the solving problem.

#### 3.1. Problem Statement

**Input:** Suppose that we have an original database after extracting from pregnancy patients' medical records. By supporting of doctors and obstetricians, the subclinical features are selected to construct the FKG-Pairs (*such as Blood Pressure*, *HGB*, *PLT*, *Urea*, *Creatinine*, *Acid Uric*, *ALT*, *AST*, *Total Protein*, *Albumin*, *LDH*, *Proteinuria and so on*). The training and testing data sets, with splitting 70% and 30% respectively, are obtained after applying the data preprocessing method and the rule-generated mechanism.

A fuzzy rule base of the training data set is described in detail in Table 1. It includes n rules  $(P_1, P_2, P_3, \ldots, P_{n-1}, P_n)$  representing patients' medical records, m input features  $(S_1, S_2, S_3, \ldots, S_{m-1}, S_m)$  representing the symptoms of the preeclampsia, and C output labels  $(1, 2, 3, \ldots, C)$  representing the doctor's diagnosis results.

In addition, the testing data set includes samples after applying the rule-generated mechanism with the IF-THEN structure similar to the rules in Table 1. For instance:

IF  $S_1$  is " $L_1$ " and  $S_2$  is " $L_2$ " and  $S_3$  is " $H_2$ " and  $S_4$  is " $VH_4$ " and . . . and  $S_{m-1}$  is " $H_{m-1}$ " and  $S_m$  is " $L_m$ " THEN Output label = 3.

$P_i$ $S_j$	$S_1$	$S_2$	•••	$S_{m-1}$	$S_m$	Output Labels
$P_1$	$H_1$	$H_2$		$VH_{m-1}$	$H_m$	2
$P_2$	$M_1$	$M_2$	• • •	$M_{m-1}$	$L_m$	1
				•••		
$P_{n-1}$	$M_1$	$M_2$		$M_{m-1}$	$M_m$	1
$P_n$	$L_1$	$M_2$		$L_{m-1}$	$L_m$	3
Patients' symptoms	$\{H_1, M_1, L_1\}$	$\{H_2, M_2\}$		$\{VH_{m-1}, H_{m-1}, M_{m-1}, L_{m-1}\}$	$\{H_m, M_m, L_m\}$	{1, 2, 3,, C}

**Table 1.** The fuzzy rule bases.

**Output:** Find the output label of the new records with subclinical data inputted by doctors or patients based on the preeclampsia sign diagnosis module.

# 3.2. Proposed Model

In this subsection, we have presented the main contents related to the construction of triple-pairs FKG (FKG-Pairs3) for the problem of disease diagnosis based on symptom data, namely: Giving a model for the problem of preeclampsia sign diagnosis in gestational women; Describing steps to follow the proposed model; and give numerical examples to illustrate the proposed model.

Information 2023, 14, 104 6 of 20

# 3.2.1. The Preeclampsia Sign Diagnosis Proposed Model

The preeclampsia sign diagnosis proposed model consists of two phases (preparation phase; diagnosis phase) as shown in Figure 2. In the problem statement, assuming we have a fuzzy rule base of the preeclampsia sign diagnosis problem after applying several steps (including the pre-processing, the rule-generated mechanism, and the training). Then, the preeclampsia sign diagnosis module (based on FKG-Pairs3) is constructed by using the training data set and is validated by using the testing data set. Finally, obstetricians and gestational women can use this module to check the preeclampsia signs by inputting the subclinical data.

# 3.2.2. The Steps to Implement the Application

To implement the preeclampsia sign diagnosis application based on FKG-Pairs3, several steps are obligated strictly as follows:

**Step 1.** Gather data sets to establish the original database.

From the medical records of the patients, the features related to the preeclampsia signs are extracted. Then, the clinical and subclinical signs data are gathered with the doctor's support. The data are stored in a database (considered the original database).

**Step 2.** Prepare the fuzzy rule base.

In this step, we conduct some tasks before constructing the FKG-Pairs3 as follows:

- Conducting the data preprocessing.
- Applying the rule-generated mechanism (herein FIS or M-CFIS).
- Applying the cluster sampling method and splitting the dataset into two parts including the training set and testing set with rates of 70% and 30% respectively.

Step 3. Construct the FKG-Pairs3 based on the training data set.

This step is considered the most important step to implement the preeclampsia sign diagnosis application in the proposed model as shown in Figure 2. There are three main tasks in this step as follows:

Firstly, for the edges connecting among vertices (input features' labels) on FKG, the weight  $\widetilde{A}_{ijhk}^t$  of these edges is calculated by using Equation (3):

$$\widetilde{A}_{ijhk}^{t} = \frac{\left| S_i \to S_j \to S_h \to S_k \text{ in rule t } \right|}{|R|} \tag{3}$$

in which  $t = \overline{1, n}$ ,  $1 \le i \le j < h < k < m - 1$ , |R| is the number of rules in the training data set.

Secondly, for edges connecting the input feature label and the output label on FKG, the weight  $B_{iihl}^t$  of these edges is calculated by using Equation (4):

$$\widetilde{B}_{ijhl}^{t} = \left(\sum \widetilde{A}_{ijhk}^{t}\right) \times \text{MIN}\left(\frac{|S_i \to l \text{ in rule } t|}{|R|}, \frac{|S_j \to l \text{ in rule } t|}{|R|}, \frac{|S_h \to l \text{ in rule } t|}{|R|}\right) \tag{4}$$

in which  $t = \overline{1,n}$ ;  $1 \le i < j < h \le m$ ;  $l = \overline{1,C}$ , |R| is the number of rules in the training data set.

Finally, the FKG-Pairs algorithm in [37] (with k=3 or FKG-Pairs3) is applied to approximate reasoning and find the output labels of the new records. The algorithm is described briefly below.

Calculating the sum of the weights of the edges  $(\widetilde{C})$  connecting from the super-nodes (i.e., each node is the combination among three features' labels) is given to the output label by using Equation (5).

$$\widetilde{C}_{ijhl} = \sum_{t} \widetilde{B}_{ijhl}^{t} \tag{5}$$

in which  $t = \overline{1, n}$ ;  $1 \le i < j < h \le m$ ;  $l = \overline{1, C}$ .

Information 2023, 14, 104 7 of 20

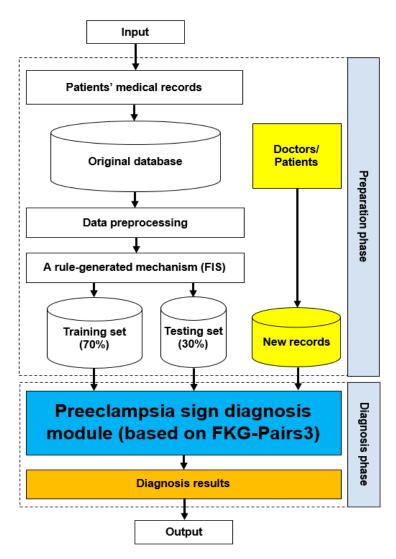


Figure 2. The proposed model of preeclampsia sign diagnosis.

Calculating the membership value  $(\widetilde{D}_l)$  by using the Max-Min operator, given by Equation (6).

$$\widetilde{D}_{l} = Max_{1 \leq i < j < h \leq m} \left( \widetilde{C}_{ijhl} \right) + Min_{1 \leq i < j < h \leq m} \left( \widetilde{C}_{ijhl} \right)$$
(6)

in which  $t = \overline{1, n}$ ;  $1 \le i < j < h \le m$ ;  $l = \overline{1, C}$ .

Finding the labels of the new records by using the *Max* operator, given by Equation (7).

$$Label = p \ if \ \widetilde{D}_p = Max_{l=\overline{1,C}} \Big( \widetilde{D}_l \Big)$$
 (7)

# **Step 4.** Validate the testing data set.

Before applying the preeclampsia sign diagnosis module for new records validated on the testing data set, which supports the application becoming more reliable. In case, the confidence of the diagnosis module cannot meet the doctors' requirements, we have to return to step 1 to get more data from new medical records to enrich the fuzzy rule base.

**Step 5.** Input the subclinical data of a gestational woman.

This step permits the patients (or users) to input the subclinical data to check the preeclampsia signs. Note that the data entered into the program has to meet the system's requirements (i.e., in the valued range).

**Step 6.** Get the diagnosis results.

Information 2023, 14, 104 8 of 20

After processing and approximate reasoning based on the subclinical data entered by patients or users, the application calculates and finds the output label corresponding to the diagnosis result (users can recheck it from obstetricians if possible).

#### 3.2.3. A numerical Example to Illustrate the Proposed Model

To deeply understand the proposed model, a numerical example is given in this subsection.

**Input:** Suppose that we have six rules  $\{R_1, R_2, R_3, R_4, R_5, R_6\}$  representing six gestational women. Each gestational woman has five features  $\{S_1, S_2, S_3, S_4, S_5\}$  representing the preeclampsia signs test results. The above cases of gestational women were examined and diagnosed based on the test results by the doctor. The output labels are 0, 1, and 2 corresponding to the doctor's diagnosis conclusions "Normal", "Preeclampsia" and "Severe Preeclampsia" respectively. After going through the data preprocessing step as well as applying the fuzzification of the input value by using the linguistic variables, a fuzzy rule base system is obtained as shown in Table 2.

Table 2. The fuzzy rule base assumes that the medical examination results of six gestational women
have been concluded to be diagnosed by doctors.

	$S_1$	$S_2$	$S_3$	$S_4$	$S_5$	$S_6$	Output Label
$R_1$	$H_1$	$M_2$	$M_3$	$M_4$	$H_5$	$H_6$	2
$R_2$	$H_1$	$M_2$	$M_3$	$M_4$	$M_5$	$M_6$	1
$R_3$	$VH_1$	$M_2$	$M_3$	$M_4$	$H_5$	$H_6$	2
$R_4$	$M_1$	$M_2$	$M_3$	$M_4$	$M_5$	$M_6$	0
$R_5$	$H_1$	$M_2$	$M_3$	$M_4$	$M_5$	$M_6$	1
$R_6$	$VH_1$	$L_2$	L <sub>3</sub>	$M_4$	$M_5$	$M_6$	2

In addition, we have also a new gestational woman after applying the rule-generated mechanism with the IF-THEN structure similar to the rules in Table 2. For instance: **IF**  $S_1$  is " $H_1$ " and  $S_2$  is " $H_2$ " and  $S_3$  is " $M_3$ " and  $S_4$  is " $M_4$ " and  $S_5$  is " $H_5$ " and  $S_6$  is " $H_6$ " **THEN** the output label = ?

**Output:** Find the output label of several new gestational women based on the fuzzy rule base in Table 2.

To solve the requirements given in the output, two steps are performed as follows:

**Step 1:** Constructing the FKG-Pairs3 based on the fuzzy rule base as shown in Table 2. *Firstly*, we calculate the set of weights  $\widetilde{A}_{ijhk}^t$  by applying Equation (3). With the rule  $R_1$ , we have:

$$\begin{split} \widetilde{A}_{1234}^1 &= \frac{|H_1 \to M_2 \to M_3 \to M_4|}{|R|} = \frac{1}{6} = 0.17, \ \widetilde{A}_{1235}^1 = \frac{|H_1 \to M_2 \to M_3 \to H_5|}{|R|} = \frac{1}{6} = 0.17, \\ \widetilde{A}_{1236}^1 &= \frac{|H_1 \to M_2 \to M_3 \to H_6|}{|R|} = \frac{1}{6} = 0.17, \ \widetilde{A}_{1245}^1 = \frac{|H_1 \to M_2 \to M_4 \to H_5|}{|R|} = \frac{1}{6} = 0.17, \\ \widetilde{A}_{1246}^1 &= \frac{|H_1 \to M_2 \to M_4 \to H_6|}{|R|} = \frac{1}{6} = 0.17, \ \widetilde{A}_{1256}^1 = \frac{|H_1 \to M_2 \to H_5 \to H_6|}{|R|} = \frac{1}{6} = 0.17, \\ \widetilde{A}_{1345}^1 &= \frac{|H_1 \to M_3 \to M_4 \to H_5|}{|R|} = \frac{1}{6} = 0.17, \ \widetilde{A}_{1346}^1 = \frac{|H_1 \to M_3 \to M_4 \to H_6|}{|R|} = \frac{1}{6} = 0.17, \\ \widetilde{A}_{1356}^1 &= \frac{|H_1 \to M_3 \to H_5 \to H_6|}{|R|} = \frac{1}{6} = 0.17, \ \widetilde{A}_{1456}^1 = \frac{|H_1 \to M_4 \to H_5 \to H_6|}{|R|} = \frac{1}{6} = 0.17, \\ \widetilde{A}_{2345}^1 &= \frac{|M_2 \to M_3 \to M_4 \to H_5|}{|R|} = \frac{1}{3} = 0.33, \ \widetilde{A}_{2346}^1 = \frac{|M_2 \to M_3 \to M_4 \to H_6|}{|R|} = \frac{1}{3} = 0.33, \\ \widetilde{A}_{2356}^1 &= \frac{|M_2 \to M_3 \to H_5 \to H_6|}{|R|} = \frac{1}{3} = 0.33, \ \widetilde{A}_{2456}^1 = \frac{|M_2 \to M_4 \to H_5 \to H_6|}{|R|} = \frac{1}{3} = 0.33, \\ \widetilde{A}_{3456}^1 &= \frac{|M_3 \to M_4 \to H_5 \to H_6|}{|R|} = \frac{1}{3} = 0.33 \end{split}$$

By the same calculation, we obtain the weights  $\widetilde{A}_{ijhk}^t$  for six rules in Table 3 as follows.

Information 2023, 14, 104 9 of 20

$ ilde{A}$	$R_1$	R <sub>2</sub>	$R_3$	$R_4$	$R_5$	$R_6$
$\widetilde{A}_{1234}^t$	0.17	0.33	0.17	0.17	0.33	0.17
$\widetilde{A}_{1235}^t$	0.17	0.33	0.17	0.17	0.33	0.17
$\widetilde{A}_{1236}^t$	0.17	0.33	0.17	0.17	0.33	0.17
$\widetilde{A}_{1245}^t$	0.17	0.33	0.17	0.17	0.33	0.17
$\widetilde{A}_{1246}^t$	0.17	0.33	0.17	0.17	0.33	0.17
$\widetilde{A}_{1256}^t$	0.17	0.33	0.17	0.17	0.33	0.17
$\widetilde{A}_{1345}^t$	0.17	0.33	0.17	0.17	0.33	0.17
$\widetilde{A}_{1346}^t$	0.17	0.33	0.17	0.17	0.33	0.17
$\widetilde{A}_{1356}^t$	0.17	0.33	0.17	0.17	0.33	0.17
$\widetilde{A}_{1456}^t$	0.17	0.33	0.17	0.17	0.33	0.17
$\widetilde{A}_{2345}^t$	0.33	0.33	0.33	0.17	0.33	0.17
$\widetilde{A}_{2346}^t$	0.33	0.33	0.33	0.17	0.33	0.17
$A_{2356}^{t}$	0.33	0.33	0.33	0.17	0.33	0.17
$\widetilde{A}_{2456}^t$	0.33	0.33	0.33	0.17	0.33	0.17
$\widetilde{A}_{3456}^t$	0.33	0.33	0.33	0.17	0.33	0.17

**Table 3.** Results of weight matrix calculation  $\widetilde{A}$ .

*Secondly*, after calculating the set of weights  $\widetilde{A}_{ijhk}^t$ , we calculate the set of weights  $\widetilde{B}_{ijhl}^t$  by applying the Equation (4). With the rule  $R_1$ , it is expressed by:

$$\begin{split} \widetilde{B}_{1231}^1 &= \left(\widetilde{A}_{1234}^1 + \widetilde{A}_{1235}^1 + \widetilde{A}_{1236}^1 + \widetilde{A}_{1245}^1 + \widetilde{A}_{1246}^1 + \widetilde{A}_{1256}^1 + \widetilde{A}_{1345}^1 + \widetilde{A}_{1346}^1 + \widetilde{A}_{1356}^1 \right. \\ &\quad + \widetilde{A}_{1456}^1 + \widetilde{A}_{2345}^1 + \widetilde{A}_{2346}^1 + \widetilde{A}_{2356}^1 + \widetilde{A}_{2456}^1 + \widetilde{A}_{3456}^1 \right) \\ &\quad \times MIN \left( \frac{|H_1 \to 1|}{|R|}, \frac{|M_2 \to 1|}{|R|}, \frac{|M_3 \to 1|}{|R|} \right) = 3.33 \times MIN \left( \frac{1}{3}, \frac{1}{3}, \frac{1}{3} \right) \\ &= 3.33 \times \frac{1}{3} = 1.11 \end{split}$$

By similar calculation, we find  $\widetilde{B}_{1241}^1 = \widetilde{B}_{1251}^1 = \widetilde{B}_{1261}^1 = \widetilde{B}_{1341}^1 = \widetilde{B}_{1351}^1 = \widetilde{B}_{1361}^1 = \widetilde{B}_{1361}^1 = \widetilde{B}_{1451}^1 = \widetilde{B}_{1461}^1 = \widetilde{B}_{1561}^1 = \widetilde{B}_{2341}^1 = \widetilde{B}_{2351}^1 = \widetilde{B}_{2361}^1 = \widetilde{B}_{2451}^1 = \widetilde{B}_{2461}^1 = \widetilde{B}_{3451}^1 = \widetilde{B}_{3461}^1 = \widetilde{B}_{3561}^1 = \widetilde{B}_{4561}^1 = 1.11.$ 

 $\widetilde{B}_{4561}^{1}=1.11.$ After applying Equation (4) for six rules, we obtain results of the entire weighted matrix  $\widetilde{B}_{iihl}^{t}$ , given in Table 4.

**Table 4.** Results of weight matrix calculation  $\widetilde{B}$ .

$\tilde{B}$	$R_1$	$R_2$	$R_3$	$R_4$	$R_5$	$R_6$
$\widetilde{B}_{1231}$	1.11	1.67	0.56	0.42	1.67	0.42
$\widetilde{B}_{124l}$	1.11	1.67	0.56	0.42	1.67	0.42
$\widetilde{B}_{1251}$	1.11	1.67	0.56	0.42	1.67	0.42
$\widetilde{B}_{126l}$	1.11	1.67	0.56	0.42	1.67	0.42
$\widetilde{B}_{1341}$	1.11	1.67	0.56	0.42	1.67	0.42
$\widetilde{B}_{1351}$	1.11	1.67	0.56	0.42	1.67	0.42
$\widetilde{B}_{1361}$	1.11	1.67	0.56	0.42	1.67	0.42

Information 2023, 14, 104 10 of 20

$\tilde{B}$	$R_1$	$R_2$	$R_3$	$R_4$	$R_5$	$R_6$
$\widetilde{B}_{1451}$	1.11	1.67	0.56	0.42	1.67	0.42
$\widetilde{B}_{1461}$	1.11	1.67	0.56	0.42	1.67	0.42
$\widetilde{B}_{1561}$	1.11	1.67	0.56	0.42	1.67	0.42
$\widetilde{B}_{2341}$	1.11	1.67	1.11	0.42	1.67	0.42
$\widetilde{B}_{2351}$	1.11	1.67	1.11	0.42	1.67	0.42
$\widetilde{B}_{2361}$	1.11	1.67	1.11	0.42	1.67	0.42
$\widetilde{B}_{2451}$	1.11	1.67	1.11	0.42	1.67	0.42
$\widetilde{B}_{2461}$	1.11	1.67	1.11	0.42	1.67	0.42
$\widetilde{B}_{2561}$	1.11	1.67	1.11	0.42	1.67	0.42
$\widetilde{B}_{3451}$	1.11	1.67	1.11	0.42	1.67	0.42
$\widetilde{B}_{3461}$	1.11	1.67	1.11	0.42	1.67	0.42
$\widetilde{B}_{3561}$	1.11	1.67	1.11	0.42	1.67	0.42
$\widetilde{B}_{4561}$	1.11	1.67	1.11	0.42	1.67	0.42

Table 4. Cont.

*Finally*, the FKG-Pairs3 module is built by applying the FKG-Pairs algorithm in [37] (with k = 3 or FKG-Pairs3) to approximate reasoning and find the output labels.

After calculating the weights  $A_{ijhk}^t$  and  $\widetilde{B}_{ijhl}^t$  for six rules in the fuzzy rule base, we calculate the sum of the weights of the edges  $(\widetilde{C})$  connecting from the super-nodes (i.e., each node is the combination among three feature labels) to the output label by using Equation (5). Let's consider the rule  $R_1$ :

• With label l = 0, we have:

$$\begin{split} \widetilde{C}_{1230} = \widetilde{C}_{1240} = \quad \widetilde{C}_{1250} = \widetilde{C}_{1260} = \widetilde{C}_{1340} = \widetilde{C}_{1350} = \widetilde{C}_{1360} = \widetilde{C}_{1450} = \widetilde{C}_{1460} = \widetilde{C}_{1560} = \widetilde{C}_{2340} \\ = \widetilde{C}_{2350} = \widetilde{C}_{2360} = \widetilde{C}_{2450} = \widetilde{C}_{2460} = \widetilde{C}_{2560} = \widetilde{C}_{3450} = \widetilde{C}_{3460} = \widetilde{C}_{3560} \\ = \widetilde{C}_{4560} = 0 \end{split}$$

• With label l = 1, we have:

$$\begin{split} \widetilde{C}_{1231} = \widetilde{C}_{1241} = & \quad \widetilde{C}_{1251} = \widetilde{C}_{1261} = \widetilde{C}_{1341} = \widetilde{C}_{1351} = \widetilde{C}_{1361} = \widetilde{C}_{1451} = \widetilde{C}_{1461} = \widetilde{C}_{1561} = \widetilde{C}_{2341} \\ = & \quad \widetilde{C}_{2351} = \widetilde{C}_{2361} = \widetilde{C}_{2450} = \widetilde{C}_{2461} = \widetilde{C}_{2561} = \widetilde{C}_{3451} = \widetilde{C}_{3461} = \widetilde{C}_{3561} \\ = & \quad \widetilde{C}_{4561} = 0 \end{split}$$

• With label l = 2, we have:

$$\begin{array}{ll} \widetilde{C}_{1232} = \widetilde{C}_{1252} &= \widetilde{C}_{1262} = \widetilde{C}_{1342} = \widetilde{C}_{1352} = \widetilde{C}_{1362} = \widetilde{C}_{1452} = \widetilde{C}_{1462} = \widetilde{C}_{1562} = \widetilde{C}_{2352} = \widetilde{C}_{2362} \\ &= \widetilde{C}_{2452} = \widetilde{C}_{2462} = \widetilde{C}_{2562} = \widetilde{C}_{3450} = \widetilde{C}_{3452} = \widetilde{C}_{3462} = \widetilde{C}_{3562} = \widetilde{C}_{4562} \\ &= \widetilde{B}_{1232}(Rule\ R_1) = 1.11 \end{array}$$

$$\widetilde{C}_{2342} = \widetilde{B}_{2342}(\textit{Rule } \textit{R}_1) + \widetilde{B}_{2342}(\textit{Rule } \textit{R}_3) = 2.22$$

By the same calculation with the rules  $R_2$ ,  $R_3$ ,  $R_4$ ,  $R_5$ ,  $R_6$ , we obtain results of the set of weights  $(\widetilde{C})$ , given in Table 5.

From the values  $(\widetilde{C}_{ijhl})$  in Table 5, we continue to compute the membership value  $(\widetilde{D}_l)$  for each label (l = 0, 1, 2) by using the Max - Min operator, given by Equation (6). For instance, in rule  $R_1$  we have:

Information 2023, 14, 104 11 of 20

$$\begin{split} \widetilde{D}_0 &= Max(\widetilde{C}_{1230},\widetilde{C}_{1240},\widetilde{C}_{1250},\,\widetilde{C}_{1260},\widetilde{C}_{1340},\widetilde{C}_{1350},\widetilde{C}_{1360},\widetilde{C}_{1450},\widetilde{C}_{1460},\widetilde{C}_{1560},\widetilde{C}_{2340},\\ \widetilde{C}_{2350},\widetilde{C}_{2360},\widetilde{C}_{2450},\widetilde{C}_{2460},\widetilde{C}_{2560},\widetilde{C}_{3450},\widetilde{C}_{3460},\widetilde{C}_{3560},\widetilde{C}_{4560}) + Min\,(\widetilde{C}_{1230},\widetilde{C}_{1240},\widetilde{C}_{1250},\,\widetilde{C}_{1260},\widetilde{C}_{1340},\widetilde{C}_{1350},\widetilde{C}_{1360},\widetilde{C}_{1450},\widetilde{C}_{1460},\widetilde{C}_{1560},\widetilde{C}_{2340},\widetilde{C}_{2350},\widetilde{C}_{2360},\widetilde{C}_{2450},\widetilde{C}_{2460},\widetilde{C}_{2560},\widetilde{C}_{3450},\widetilde{C}_{3450},\widetilde{C}_{3560},\widetilde{C}_{4560}) = 0 \end{split}$$

$$\begin{split} \widetilde{D}_{1} &= \textit{Max}(\widetilde{C}_{1231}, \widetilde{C}_{1241}, \widetilde{C}_{1251}, \ \widetilde{C}_{1261}, \widetilde{C}_{1341}, \widetilde{C}_{1351}, \widetilde{C}_{1361}, \widetilde{C}_{1451}, \widetilde{C}_{1461}, \widetilde{C}_{1561}, \widetilde{C}_{2341}, \\ \widetilde{C}_{2351}, \widetilde{C}_{2361}, \widetilde{C}_{2451}, \widetilde{C}_{2461}, \widetilde{C}_{2561}, \widetilde{C}_{3451}, \widetilde{C}_{3461}, \widetilde{C}_{3561}, \widetilde{C}_{4561}) + \textit{Min} \ (\widetilde{C}_{1231}, \widetilde{C}_{1241}, \widetilde{C}_{1251}, \ \widetilde{C}_{1261}, \widetilde{C}_{1341}, \widetilde{C}_{1351}, \widetilde{C}_{1361}, \widetilde{C}_{1451}, \widetilde{C}_{1461}, \widetilde{C}_{1561}, \widetilde{C}_{2341}, \widetilde{C}_{2351}, \widetilde{C}_{2361}, \widetilde{C}_{2451}, \widetilde{C}_{2461}, \\ \widetilde{C}_{2561}, \widetilde{C}_{3451}, \widetilde{C}_{3461}, \widetilde{C}_{3561}, \widetilde{C}_{4561}) &= 0 \end{split}$$

$$\begin{split} \widetilde{D}_2 &= Max(\widetilde{C}_{1232},\widetilde{C}_{1242},\widetilde{C}_{1252},\ \widetilde{C}_{1262},\widetilde{C}_{1342},\widetilde{C}_{1352},\widetilde{C}_{1362},\widetilde{C}_{1452},\widetilde{C}_{1462},\widetilde{C}_{1562},\widetilde{C}_{2342},\\ \widetilde{C}_{2352},\widetilde{C}_{2362},\widetilde{C}_{2452},\widetilde{C}_{2462},\widetilde{C}_{2562},\widetilde{C}_{3452},\widetilde{C}_{3462},\widetilde{C}_{3562},\widetilde{C}_{4562}) + Min\ (\widetilde{C}_{1232},\widetilde{C}_{1242},\widetilde{C}_{1252},\ \widetilde{C}_{1262},\widetilde{C}_{1342},\widetilde{C}_{1352},\widetilde{C}_{1362},\widetilde{C}_{1452},\widetilde{C}_{1462},\widetilde{C}_{1562},\widetilde{C}_{2342},\widetilde{C}_{2352},\widetilde{C}_{2362},\widetilde{C}_{2452},\widetilde{C}_{2462},\widetilde{C}_{2562},\widetilde{C}_{2362},\widetilde{C}_{3452},\widetilde{C}_{2462},\widetilde{C}_{2562},\widetilde{C}_{3452},\widetilde{C}_{3462},\widetilde{C}_{3562},\widetilde{C}_{4562}) = 2.22 + 1.11 = 3.33 \end{split}$$

**Table 5.** Results of calculating the set of weights  $\widetilde{C}$ .

	$\tilde{c}$	$R_1$	R <sub>2</sub>	R <sub>3</sub>	$R_4$	$R_5$	$R_6$
	$\widetilde{C}_{1230}$	0.00	0.00	0.00	0.42	0.00	0.00
	$\widetilde{C}_{1240}$	0.00	0.00	0.00	0.42	0.00	0.00
	$\widetilde{C}_{1250}$	0.00	0.00	0.00	0.42	0.00	0.00
	$\widetilde{C}_{1260}$	0.00	0.00	0.00	0.42	0.00	0.00
	$\widetilde{C}_{1340}$	0.00	0.00	0.00	0.42	0.00	0.00
	$\widetilde{C}_{1350}$	0.00	0.00	0.00	0.42	0.00	0.00
	$\widetilde{C}_{1360}$	0.00	0.00	0.00	0.42	0.00	0.00
	$\widetilde{C}_{1450}$	0.00	0.00	0.00	0.42	0.00	0.00
el 0	$\widetilde{C}_{1460}$	0.00	0.00	0.00	0.42	0.00	0.00
Label 0	$\widetilde{C}_{1560}$	0.00	0.00	0.00	0.42	0.00	0.00
	$\widetilde{C}_{2340}$	0.00	0.00	0.00	0.42	0.00	0.00
	$\widetilde{C}_{2350}$	0.00	0.00	0.00	0.42	0.00	0.00
	$\widetilde{C}_{2360}$	0.00	0.00	0.00	0.42	0.00	0.00
	$\widetilde{C}_{2450}$	0.00	0.00	0.00	0.42	0.00	0.00
	$\widetilde{C}_{2460}$	0.00	0.00	0.00	0.42	0.00	0.00
	$\widetilde{C}_{2560}$	0.00	0.00	0.00	0.42	0.00	0.00
	$\widetilde{C}_{3450}$	0.00	0.00	0.00	0.42	0.00	0.00
	$\widetilde{C}_{3460}$	0.00	0.00	0.00	0.42	0.00	0.00
	$\widetilde{C}_{3560}$	0.00	0.00	0.00	0.42	0.00	0.00
	$\widetilde{C}_{4560}$	0.00	0.00	0.00	0.42	0.00	0.00
	$\widetilde{C}_{1231}$	0.00	3.33	0.00	0.00	3.33	0.00
	$\widetilde{C}_{1241}$	0.00	3.33	0.00	0.00	3.33	0.00
	$\widetilde{C}_{1251}$	0.00	3.33	0.00	0.00	3.33	0.00

*Information* **2023**, 14, 104

 Table 5. Cont.

	$\tilde{c}$	$R_1$	$R_2$	$R_3$	$R_4$	$R_5$	$R_6$
	$\widetilde{C}_{1261}$	0.00	3.33	0.00	0.00	3.33	0.00
	$\widetilde{C}_{1341}$	0.00	3.33	0.00	0.00	3.33	0.00
	$\widetilde{C}_{1351}$	0.00	3.33	0.00	0.00	3.33	0.00
	$\widetilde{C}_{1361}$	0.00	3.33	0.00	0.00	3.33	0.00
-	$\widetilde{C}_{1451}$	0.00	3.33	0.00	0.00	3.33	0.00
	$\widetilde{C}_{1461}$	0.00	3.33	0.00	0.00	3.33	0.00
	$\widetilde{C}_{1561}$	0.00	3.33	0.00	0.00	3.33	0.00
Label 1	$\widetilde{C}_{2341}$	0.00	3.33	0.00	0.00	3.33	0.00
La	$\widetilde{C}_{2351}$	0.00	3.33	0.00	0.00	3.33	0.00
	$\widetilde{C}_{2361}$	0.00	3.33	0.00	0.00	3.33	0.00
	$\widetilde{C}_{2451}$	0.00	3.33	0.00	0.00	3.33	0.00
	$\widetilde{C}_{2461}$	0.00	3.33	0.00	0.00	3.33	0.00
	$\widetilde{C}_{2561}$	0.00	3.33	0.00	0.00	3.33	0.00
	$\widetilde{C}_{3451}$	0.00	3.33	0.00	0.00	3.33	0.00
	$\widetilde{C}_{3461}$	0.00	3.33	0.00	0.00	3.33	0.00
	$\widetilde{C}_{3561}$	0.00	3.33	0.00	0.00	3.33	0.00
	$\widetilde{C}_{4561}$	0.00	3.33	0.00	0.00	3.33	0.00
	$\widetilde{C}_{1232}$	1.11	0.00	0.56	0.00	0.00	0.42
	$\widetilde{C}_{1242}$	1.11	0.00	0.56	0.00	0.00	0.42
	$\widetilde{C}_{1252}$	1.11	0.00	0.56	0.00	0.00	0.42
	$\widetilde{C}_{1262}$	1.11	0.00	0.56	0.00	0.00	0.42
	$\widetilde{C}_{1342}$	1.11	0.00	0.56	0.00	0.00	0.42
	$\widetilde{C}_{1352}$	1.11	0.00	0.56	0.00	0.00	0.42
	$\widetilde{C}_{1362}$	1.11	0.00	0.56	0.00	0.00	0.42
	$\widetilde{C}_{1452}$	1.11	0.00	0.56	0.00	0.00	0.42
el 2	$\widetilde{C}_{1462}$	1.11	0.00	0.56	0.00	0.00	0.42
Label 2	$\widetilde{C}_{1562}$	1.11	0.00	0.56	0.00	0.00	0.42
	$\widetilde{C}_{2342}$	2.22	0.00	2.22	0.00	0.00	0.42
	$\widetilde{C}_{2352}$	1.11	0.00	1.11	0.00	0.00	0.42
	$\widetilde{C}_{2362}$	1.11	0.00	1.11	0.00	0.00	0.42
	$\widetilde{C}_{2452}$	1.11	0.00	1.11	0.00	0.00	0.42
	$\widetilde{C}_{2462}$	1.11	0.00	1.11	0.00	0.00	0.42
	$\widetilde{C}_{2562}$	1.11	0.00	1.11	0.00	0.00	0.42
	$\widetilde{C}_{3452}$	1.11	0.00	1.11	0.00	0.00	0.42
	$\widetilde{C}_{3462}$	1.11	0.00	1.11	0.00	0.00	0.42
	$\widetilde{C}_{3562}$	1.11	0.00	1.11	0.00	0.00	0.42
	$\widetilde{C}_{4562}$	1.11	0.00	1.11	0.00	0.00	0.42

Information 2023, 14, 104 13 of 20

By using the Max operator, given by Equation (7), the label of rule  $R_1$  is **2** because of  $\widetilde{D}_2 = Max(\widetilde{D}_0, \widetilde{D}_1, \widetilde{D}_2)$ . Similarly, we obtain the labels of rules  $R_2$ ,  $R_3$ ,  $R_4$ ,  $R_5$ ,  $R_6$  being **1**, **2**, **0**, **1**, **2** respectively.

Step 2: Validating with the new records.

*Case 1:* The new record is in the fuzzy rule base system in Table 1, the output label is the same as the result of one of the six rules.

Case 2: The new record is not in the fuzzy rule base system in Table 1. For instance, IF  $S_1$  is " $H_1$ " and  $S_2$  is " $H_2$ " and  $S_3$  is " $M_3$ " and  $S_4$  is " $M_4$ " and  $S_5$  is " $H_5$ " and  $S_6$  is " $H_6$ " THEN the output label = ?. The output label of this new record is found by applying Equations (5)–(7). We have:

• With label l = 0,

$$\widetilde{C}_{1230} = \widetilde{C}_{1240} = \quad \widetilde{C}_{1250} = \widetilde{C}_{1260} = \widetilde{C}_{1340} = \widetilde{C}_{1350} = \widetilde{C}_{1360} = \widetilde{C}_{1450} = \widetilde{C}_{1460} = \widetilde{C}_{1560} = \widetilde{C}_{2340}$$

$$= \widetilde{C}_{2350} = \widetilde{C}_{2360} = \widetilde{C}_{2450} = \widetilde{C}_{2460} = \widetilde{C}_{2560} = \widetilde{C}_{3450} = \widetilde{C}_{3460} = \widetilde{C}_{3560}$$

$$= \widetilde{C}_{4560} = 0$$

• With label l = 1,

$$\widetilde{C}_{1231} = \widetilde{C}_{1241} = \quad \widetilde{C}_{1251} = \widetilde{C}_{1261} = \widetilde{C}_{1341} = \widetilde{C}_{1351} = \widetilde{C}_{1361} = \widetilde{C}_{1451} = \widetilde{C}_{1461} = \widetilde{C}_{1561} = \widetilde{C}_{2341}$$

$$= \widetilde{C}_{2351} = \widetilde{C}_{2361} = \widetilde{C}_{2450} = \widetilde{C}_{2461} = \widetilde{C}_{2561} = \widetilde{C}_{3451} = \widetilde{C}_{3461} = \widetilde{C}_{3561}$$

$$= \widetilde{C}_{4561} = 0$$

• With label l = 2,

$$\begin{array}{l} \widetilde{C}_{1232} = \widetilde{C}_{1252} = \widetilde{C}_{1262} = \widetilde{C}_{2342} = \widetilde{C}_{2352} = \widetilde{C}_{2362} = \widetilde{C}_{2452} = \widetilde{C}_{2462} = \widetilde{C}_{2562} = \widetilde{C}_{3450} = 0 \\ \widetilde{C}_{1342} = \widetilde{C}_{1352} = \widetilde{C}_{1362} = \widetilde{C}_{1452} = \widetilde{C}_{1462} = \widetilde{C}_{1562} = 0.56 \\ \widetilde{C}_{3452} = \widetilde{C}_{3462} = \widetilde{C}_{3562} = \widetilde{C}_{4562} = 1.11 \end{array}$$

Therefore,  $\widetilde{D}_0 = \widetilde{D}_1 = 0$ ,  $\widetilde{D}_2 = 1.11$ . The result of the new record: **IF**  $S_1$  is " $H_1$ " and  $S_2$  is " $H_2$ " and  $S_3$  is " $H_3$ " and  $S_4$  is " $H_4$ " and  $H_4$ " and  $H_5$ " and  $H_5$ " and  $H_6$ " **THEN** the output label = **2**. The output label is 2 corresponding to the doctor's diagnosis conclusion "Severe Preeclampsia".

# 4. Experimental Results

# 4.1. Experiments

To simulate the proposed model in real-world applications, the data set was gathered from pregnant women who came for regular check-ups to monitor the fetus at the National Hospital of Obstetrics and Gynecology in Vietnam. After receiving expert comments from the obstetrician, the preeclampsia data set includes 210 samples with 19 test parameters, such as Blood Pressure, Hemoglobin (HGB), Platelet Count (PLT), Urea, Creatinine, Acid Uric, Alanine Aminotransferase (ALT), Aspartate Aminotransferase (AST), Total Protein, Albumin, Lactate Dehydrogenase (LDH), Proteinuria and so on (in detail given in Table 6). In this data set, according to the diagnostic conclusion (in which 118 women with signs of normal pregnancy, 60 preeclampsia women, and 32 severe preeclampsia women).

**Table 6.** List of features in the preeclampsia dataset.

No.	Feature's Name	Domain
1	Pregnant Woman's Age	18–66 years old
2	Fetus's Age	15–40 weeks
3	Occupation	Officer, Teacher, Doctor, Worker, Farmer, Freelancer, and so on.
4	Number of Pregnancies	0–9 times
5	Pregnant Woman's Height	1.40–1.90 m

Information 2023, 14, 104 14 of 20

Table 6. Cont.

No.	Feature's Name	Domain
6	Pregnant Woman's Weight	45–95 kg
7	Upper Blood Pressure	90–129 mmHg
8	Lower Blood Pressure	60–84 mmHg
9	Hemoglobin (HGB)	120–160 g/L
10	Platelet Count (PLT)	150–450 g/L
11	Urea	2.5–6.7 mmol/L
12	Creatinine	50.4–98.1 μmol/L
13	Acid Uric	150–350 μmol/L
14	Alanine Aminotransferase (ALT)	<31/37 Ul/L
15	Aspartate Aminotransferase (AST)	<31/37 Ul/L
16	Total Protein	64–83 g/L
17	Albumin	35–52 g/L
18	Lactate Dehydrogenase (LDH)	<247 U/L
19	Proteinuria	0.1–0.25 g/L
Output	Output labels (Diagnostic results)	0: Normal 1: Preeclampsia 2: Severe preeclampsia

The preeclampsia application is built using Python 3.10 language installed on a laptop (ASUS Intel(R) Core (TM) i5-8300H CPU @ 2.30 GHz).

# 4.2. Evaluation Method

To evaluate the proposed model-based system's performance, the parameters are used including the accuracy and calculation time, specifically as follows:

*Accuracy* is evaluated by the ratio of the number of correctly classified samples over the total number of performed samples, estimated by Formula (8).

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{8}$$

#### where:

• *TP*: True Positive

• *TN*: True Negative

• *FP*: False Positive

• *FN*: False Negative

Time is estimated by the total execution time of the classification system (unit: seconds).

#### 4.3. Test Results in Simulations

After implementing the application based on the above data set, the authors established three different scenarios, namely:

- Scenario 1: the systematic random sampling method and the splitting method with training set (70%) and testing set (30%).
- Scenario 2: the systematic random sampling method and the splitting method with training set (10%) and testing set (90%).
- Scenario 3: the systematic random sampling method and the splitting method with training set (5%) and testing set (95%).

Information 2023, 14, 104 15 of 20

The comparison criteria include accuracy and computation time as given in Section 4.2. The experimental results are shown in Figures 3 and 4.

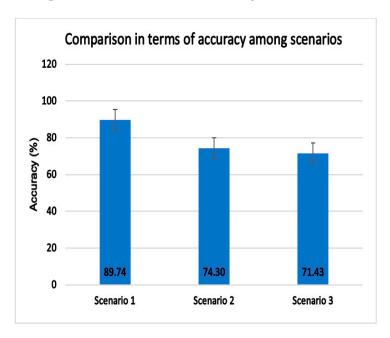


Figure 3. Comparison in terms of accuracy among scenarios.

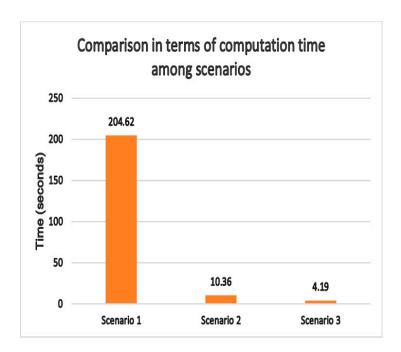


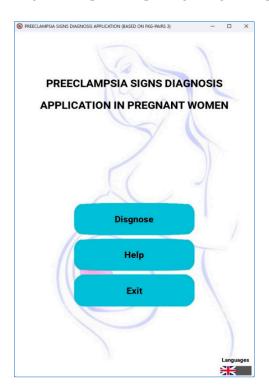
Figure 4. Comparison in terms of computation time among scenarios.

As shown in the results of Figure 3, it is clear to see that the accuracy of scenario 1 is significantly higher than that of the other two scenarios. For example, the accuracy of scenario 1 is **15.44**% higher than scenario 2 and **18.31**% higher than scenario 3. It proves that the accuracy depends on the number of samples in the training data set.

However, the data in Figure 4 shows that the computation time of scenario 1 is also much higher than that of the other scenarios. Specifically, the computation time of scenario 1 is approximately **20 times** higher than that of scenario 2, and nearly **50 times** that of scenario 3. This demonstrated that the data set splitting method contributed significantly to improving the application's performance in terms of time consumption.

Information 2023, 14, 104 16 of 20

After building and evaluating the model's confidence, the authors proceed to build a simple application to diagnose the preeclampsia signs of the built model. Several illustrated images of the preeclampsia sign diagnosis application have been shown in Figures 5–8.



**Figure 5.** The main screen of the application.

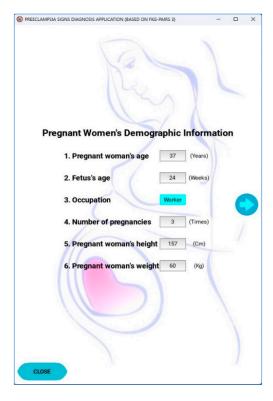
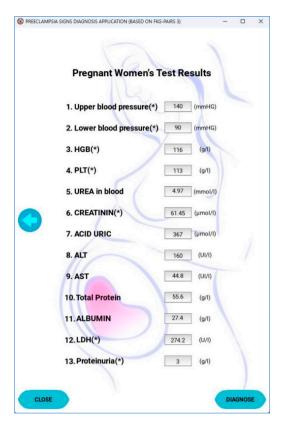


Figure 6. Screen for entering demographic information.

Information 2023, 14, 104 17 of 20



**Figure 7.** Screen for entering subclinical test results.

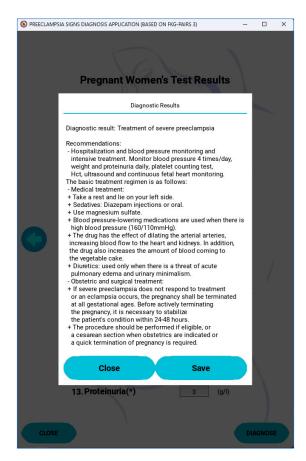


Figure 8. Diagnostic results screen and recommendations.

Information 2023, 14, 104 18 of 20

#### 5. Conclusions

This paper has proposed the FKG-Pairs3-based preeclampsia signs in the diagnosis by combining regimens in traditional medicine and the information in medical records. The proposed model has been performed as a case study on the diagnosis of preeclampsia signs in gestational women based on the clinical and subclinical data collected from pregnant women attending routine antenatal care. The proposed model has been tested for an accuracy performance of **89.74**% with the implementation designed scenario, in which the systematic random sampling method and the splitting method with training set (70%) and testing set (30%).

With complete information input data sets, the accuracy of the preeclampsia sign diagnosis system will continue to improve. However, the limitation of the proposed model is identified with two drawbacks in extreme cases as follows:

- Firstly, with large input data sets, the computation time is high based on the traditional data set splitting method (e.g., in scenario 1).
- Secondly, with too-small training data sets, the accuracy is low (e.g., in scenario 2 and scenario 3).

To expand this research work, the investigation will continue to study the approach of using the FKG-Pairs3 proposed model combined with Q-learning techniques in reinforcement learning to improve the accuracy of a system in extreme cases in which the training data set is much smaller than the testing data set (e.g., in cases scenarios 2 and 3 where the training sets only 10% and 5%, respectively).

#### 6. Patents

This section is not mandatory but may be added if there are patents resulting from the work reported in this manuscript.

**Author Contributions:** Conceptualization, methodology, C.K.L. and H.V.P.; software, P.H.K. and C.K.L.; validation, H.V.P., C.K.L. and H.Q.T.; formal analysis, C.K.L. and H.Q.T.; investigation, C.K.L. and P.H.K.; data curation, P.H.K.; writing—original draft preparation, C.K.L. and P.H.K.; writing—review and editing, H.V.P. and H.Q.T.; project administration, H.V.P.; All authors have read and agreed to the published version of the manuscript.

**Funding:** This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 102.05-2019.316.

**Data Availability Statement:** The dataset and source codes of this paper can be downloaded at: https://github.com/FKGHUST/Preeclampsia.git (accessed on 15 December 2022).

**Conflicts of Interest:** The authors declare no conflict of interest. This research does not involve any human or animal participation. All authors have checked and agreed with the submission.

## References

- 1. World Health Organization. WHO Recommendations for Prevention and Treatment of Pre-Eclampsia and Eclampsia; World Health Organization: Geneva, Switzerland, 2011.
- 2. Nguyen, T.H.; Bui, T.C.; Vo, T.M.; Tran, Q.M.; Luu, L.T.; Nguyen, T.D. Predictive value of the sFlt-1 and PlGF in women at risk for preeclampsia in the south of Vietnam. *Pregnancy Hypertens* **2018**, *14*, 37–42. [CrossRef] [PubMed]
- 3. Masini, G.; Foo, L.F.; Tay, J.; Wilkinson, I.B.; Valensise, H.; Gyselaers, W.; Lees, C.C. Preeclampsia has two phenotypes which require different treatment strategies. *Am. J. Obstet. Gynecol.* **2021**, 226, S1006–S1018. [CrossRef] [PubMed]
- 4. American College of Obstetricians and Gynecologists. Hypertension in Pregnancy: Executive Summary. *Obstet. Gynecol.* **2013**, 122, 1122–1131. [CrossRef] [PubMed]
- 5. Wright, D.; Syngelaki, A.; Akolekar, R.; Poon, L.C.; Nicolaides, K.H. Competing risks model in screening for preeclampsia by maternal characteristics and medical history. *Am. J. Obstet. Gynecol.* **2015**, 213, 62.e1–62.e10. [CrossRef]
- 6. American College of Obstetricians and Gynecologists. ACOG Practice Bulletin No. 202: Gestational Hypertension and Preeclampsia. *Obstet. Gynecol.* **2019**, *133*, e1–e25.
- 7. SOGC guideline, Diagnosis, Evaluation, and Management of the Hypertensive Disoder of Pregnancy: Executive Summary. *J. Obstet. Gynaecol. Can.* **2014**, *36*, 416–438. [CrossRef]

Information 2023, 14, 104 19 of 20

8. NICE. Hypertension in pregnancy: The management of hypertensive disorders during pregnancy. NICE Clin. Guidel 2019. No. 107.

- 9. Poon, L.C.; Shennan, A.; Hyett, J.A.; Kapur, A.; Hadar, E.; Divakar, H.; McAuliffe, F.; da Silva Costa, F.; von Dadelszen, P.; McIntyre, H.D.; et al. The International Federation of Gynecology and Obstetrics initiative on pre-eclampsia: A pragmatic guide for first-trimester screening and prevention. *Int. J. Gynecol. Obstet.* 2019, 145, 1–33. [CrossRef]
- 10. Yu, T.; Li, J.; Yu, Q.; Tian, Y.; Shun, X.; Xu, L.; Zhu, L.; Gao, H. Knowledge graph for TCM health preservation: Design, construction, and applications. *Artif. Intell. Med.* **2017**, 77, 48–52. [CrossRef]
- 11. Gyrard, A.; Gaur, M.; Shekarpour, S.; Thirunarayan, K.; Sheth, A. Personalized health knowledge graph. In CEUR Workshop Proceedings; NIH Public Access: Bethesda, MD, USA, 2018; Volume 2317.
- 12. Chai, X. Diagnosis Method of thyroid disease combining knowledge graph and deep learning. *IEEE Access* **2020**, *8*, 149787–149795. [CrossRef]
- 13. Troussas, C.; Chrysafiadi, K.; Virvou, M. An intelligent adaptive fuzzy-based inference system for computer-assisted language learning. *Expert Syst. Appl.* **2019**, 127, 85–96. [CrossRef]
- 14. Bakhshipour, A.; Zareiforoush, H.; Bagheri, I. Application of decision trees and fuzzy inference system for quality classification and modeling of black and green tea based on visual features. *J. Food Meas. Charact.* **2020**, *14*, 1402–1416. [CrossRef]
- 15. Tiwari, L.; Raja, R.; Sharma, V.; Miri, R. Fuzzy Inference System for Efficient Lung Cancer Detection. In *Computer Vision and Machine Intelligence in Medical Image Analysis, Advances in Intelligent Systems and Computing*; Springer: Singapore, 2020; Volume 992. [CrossRef]
- 16. Ortega, L.C.; Otero, L.D.; Otero, C. Fuzzy Inference System Framework to Prioritize the Deployment of Resources in Low Visibility Traffic Conditions. *IEEE Access* **2019**, *7*, 174368–174379. [CrossRef]
- 17. Johann, G.; dos Santos, C.S.; Montanher, P.F.; de Oliveira, R.A.P.; Carniel, A.C. Fuzzy inference systems for predicting the mass yield in extractions of chia cake extract. *Softw. Impacts* **2021**, *10*, 100145. [CrossRef]
- 18. Saini, J.; Dutta, M.; Marques, G. Fuzzy Inference System Tree with Particle Swarm Optimization and Genetic Algorithm: A novel approach for PM10 forecasting. *Expert Syst. Appl.* **2021**, *183*, 115376. [CrossRef]
- Man, J.Y.; Chen, Z.; Dick, S. Towards inductive learning of complex fuzzy inference systems. In Proceedings of the NAFIPS 2007: 2007 Annual Meeting of the North American Fuzzy Information Processing Society, San Diego, CA, USA, 24–27 June 2007; pp. 415–420. [CrossRef]
- 20. Tu, C.H.; Li, C. Multiple Function Approximation-A New Approach Using Complex Fuzzy Inference System. In *Intelligent Information and Database Systems*. ACIIDS 2018. Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2018; Volume 10751.
- 21. Ngan, T.T.; Lan, L.T.H.; Tuan, T.M.; Son, L.H.; Tuan, L.M.; Minh, N.H. Colorectal Cancer Diagnosis with Complex Fuzzy Inference System. In *Frontiers in Intelligent Computing: Theory and Applications. Advances in Intelligent Systems and Computing*; Springer: Singapore, 2019; Volume 2013. [CrossRef]
- 22. Selvachandran, G.; Quek, S.G.; Lan, L.T.H.; Son, L.H.; Giang, N.L.; Ding, W.; Abdel-Basset, M.; de Albuquerque, V.H.C. A New Design of Mamdani Complex Fuzzy Inference System for Multiattribute Decision Making Problems. *IEEE Trans. Fuzzy Syst.* **2019**, 29, 716–730. [CrossRef]
- 23. Tuan, T.M.; Lan, L.T.H.; Chou, S.-Y.; Ngan, T.T.; Son, L.H.; Giang, N.L.; Ali, M. M-CFIS-R: Mamdani complex fuzzy inference system with rule reduction using complex fuzzy measures in granular computing. *Mathematics* **2020**, *8*, 707. [CrossRef]
- 24. Lan, L.T.H.; Tuan, T.M.; Ngan, T.T.; Son, L.H.; Giang, N.L.; Ngoc, V.T.N.; Van Hai, P. A New Complex Fuzzy Inference System With Fuzzy Knowledge Graph and Extensions in Decision Making. *IEEE Access* **2020**, *8*, 164899–164921. [CrossRef]
- 25. Long, C.K.; Van Hai, P.; Tuan, T.M.; Lan, L.T.H.; Chuan, P.M.; Son, L.H. A Novel Fuzzy Knowledge Graph Pairs Approach in Decision Making. *Multimedia Tools Appl.* **2022**, *81*, 26505–26534. [CrossRef]
- 26. Sherimon, P.C.; Krishnan, R. OntoDiabetic: An Ontology-Based Clinical Decision Support System for Diabetic Patients. *Arab. J. Sci. Eng.* **2015**, *41*, 1145–1160. [CrossRef]
- 27. Sweidan, S.; El-Sappagh, S.; El-Bakry, H.; Sabbeh, S.; Badria, F.A.; Kwak, K.-S. A Fibrosis Diagnosis Clinical Decision Support System Using Fuzzy Knowledge. *Arab. J. Sci. Eng.* **2018**, *44*, 3781–3800. [CrossRef]
- 28. Ibrahim, I.A.; Ting, H.-N.; Moghavvemi, M. Formulation of a Novel Classification Indices for Classification of Human Hearing Abilities According to Cortical Auditory Event Potential signals. *Arab. J. Sci. Eng.* **2019**, *44*, 7133–7147. [CrossRef]
- 29. Sudhan, M.; Sinthuja, M.; Raja, S.P.; Amutharaj, J.; Latha, G.C.P.; Rachel, S.S.; Anitha, T.; Rajendran, T.; Waji, Y.A. Segmentation and Classification of Glaucoma Using U-Net with Deep Learning Model. *J. Healthc. Eng.* **2022**, 2022, 1601354. [CrossRef]
- 30. Jhou, M.-J.; Chen, M.-S.; Lee, T.-S.; Yang, C.-T.; Chiu, Y.-L.; Lu, C.-J. A Hybrid Risk Factor Evaluation Scheme for Metabolic Syndrome and Stage 3 Chronic Kidney Disease Based on Multiple Machine Learning Techniques. *Healthcare* 2022, 10, 2496. [CrossRef]
- 31. Son, L.H.; Ciaramella, A.; Huyen, D.T.T.; Staiano, A.; Tuan, T.M.; Van Hai, P. Predictive reliability and validity of hospital cost analysis with dynamic neural network and genetic algorithm. *Neural Comput. Appl.* **2020**, *32*, 15237–15248. [CrossRef]
- 32. Ngoc, V.T.N.; Viet, D.H.; Tuan, T.M.; Van Hai, P.; Thang, N.P.; Tuyen, D.N.; Son, L.H. VNU-diagnosis: A novel medical system based on deep learning for diagnosis of periapical inflammation from X-Rays images. *J. Intell. Fuzzy Syst.* **2022**, *43*, 1417–1427. [CrossRef]
- 33. Van Pham, H.; Moore, P.; Cuong, B.C. Applied picture fuzzy sets with knowledge reasoning and linguistics in clinical decision support system. *Neurosci. Inform.* **2022**, 2, 100109. [CrossRef]

Information 2023, 14, 104 20 of 20

34. Garg, H.; Ahmad, A.; Ullah, K.; Mahmood, T.; Ali, Z. Algorithm for multi-attribute decision-making using T-spherical fuzzy Maclaurin symmetric mean operator. *Iran. J. Fuzzy Syst.* **2022**, *19*, 111–124.

- 35. Garg, H.; Deng, Y.; Ali, Z.; Mahmood, T. Decision-making strategy based on Archimedean Bonferroni mean operators under complex Pythagorean fuzzy information. *Comput. Appl. Math.* **2022**, *41*, 152. [CrossRef]
- 36. Van Pham, H.; Khoa, N.D.; Bui, T.T.; Giang, N.T.; Moore, P. Applied Picture Fuzzy Sets for Group Decision-Support in the Evaluation of Pedagogic Systems. *Int. J. Math. Eng. Manag. Sci.* **2022**, *7*, 243–257.
- 37. Zadeh, L. Fuzzy sets. Inf. Control. 1965, 8, 338–353. [CrossRef]
- 38. Chen, X.; Hu, Z.; Sun, Y. Fuzzy Logic Based Logical Query Answering on Knowledge Graphs. *Proc. Conf. AAAI Artif. Intell.* **2022**, 36, 3939–3948. [CrossRef]
- 39. Zadeh, L.A. Approximate reasoning based on fuzzy logic. In Proceedings of the 6th International Joint Conference on Artificial Intelligence, Tokyo, Japan, 20–23 August 1979; Volume 2, pp. 1004–1010.
- 40. Singhal, A. "Introducing the Knowledge Graph: Things, Not Strings", Official Google Blog. Available online: https://blog.google/products/search/introducing-knowledge-graph-things-not/ (accessed on 22 June 2022).
- 41. Paulheim, H. Knowledge graph refinement: A survey of approaches and evaluation methods. *Semantic Web* **2016**, *8*, 489–508. [CrossRef]
- 42. Tian, L.; Zhou, X.; Wu, Y.-P.; Zhou, W.-T.; Zhang, J.-H.; Zhang, T.-S. Knowledge graph and knowledge reasoning: A systematic review. *J. Electron. Sci. Technol.* **2022**, *20*, 100159. [CrossRef]
- 43. Moussa, S.; Kacem, S.B.H. Symbolic approximate reasoning with fuzzy and multi-valued knowledge. *Procedia Comput. Sci.* **2017**, 112, 800–810. [CrossRef]
- 44. Rajabi, E.; Kafaie, S. Knowledge Graphs and Explainable AI in Healthcare. Information 2022, 13, 459. [CrossRef]
- 45. Sachdeva, S.; Bhalla, S. Using Knowledge Graph Structures for Semantic Interoperability in Electronic Health Records Data Exchanges. *Information* **2022**, *13*, 52. [CrossRef]
- 46. Long, C.K.; Van Hai, P.; Tuan, T.M.; Lan, L.T.H.; Ngan, T.T.; Chuan, P.M.; Son, L.H. A novel Q-learning-based FKG-Pairs approach for extreme cases in decision making. *Eng. Appl. Artif. Intell.* **2023**, 120, 105920. [CrossRef]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.