Incorporating Game Theory with Soft Sets

ChenQi Li, JingTao Yao
Department of Computer Science, University of Regina, Canada
{clg185, jingtao.yao}@uregina.ca

***Abstract*—Solving uncertainty issues is a challenge for decision-making. Soft set theory aims to aid complex decision-making when multiple uncertainty variables are involved. We introduce a game-theoretic soft set model to handle the fusion of uncertain data from different information resources and resolve conflicts among parameters in a soft set. The model is utilized to solve three-way classification problems by establishing measurement thresholds for parameters. The experiment shows that the model can strike a balance among different parameters, resulting in a decrease in misclassification error in classification. Furthermore, the extent of the decrease can be fine-tuned by adjusting the ratio between the cost for misclassification error and the cost for undecided error.**

***Index Terms*—soft sets, decision-making, game theory, uncertainty, three-way classification**

I. Introduction

It is challenging to make decisions in an environment involving uncertainty. To handle this challenge, several theories have been proposed and widely used in the fields of economics, engineering, social science, social contexts and others [10]. These include the theory of probability, the theory of fuzzy sets [20], the theory of rough sets [19] and others. However, they all face the difficulties due to the inadequacy of parametrization tools [10]. In other words, the subsets of the universe cannot be conveniently specified by parameters. Molodtsov initiated the theory of soft sets to solve these difficulties [10]. The theory has been widely employed to solve decision-making problems in various real-life domains, including forecasting [14], uncertainty measure [3], classification [9], and so on. The optimal object was recognized based on a set of inputs from multi-observers [13]. A significant portion of soft set research, as we have mentioned so far, has been dedicated to calculating score values or choice values for objects by aggregating evaluations of parameters. We address conflicts among different parameters by games without forming score values or choice values. Our aggregation operates on sets instead of real numbers.

We aim to solve three-way classification problems by reconciling conflicting opinions from multiple experts. Game theory has demonstrated its ability to address such complex situations where multiple entities engage in cooperative or competitive interactions [18]. It has been utilized to analyze the trade-off between node lifetime and communication reliability, as well as the impact of jamming on these factors [11]. Additionally, game theory is employed to analyze the behaviors of mobile users and service providers in a network constrained by limited resources [8]. Moreover, when it comes to determining region inclusion thresholds, game has been proven to be effective. Game-theoretic rough sets (GTRS) study the trade-offs between classification approximation measures as well as between region inclusion parameters to determine optimal thresholds [2] [7]. GTRS find applications in diverse domains such as image classification for medical diagnosis [15], spam email detection [21], classification of news articles into satirical, legitimate, or questionable content [23], and sentiment analysis [4]. The model has also proven effective in handling missing values, particularly in clustering corticosteroid responsiveness [6]. Another game-theoretic model, game-theoretic shadowed sets, focuses on adjusting initial thresholds by striking a balance between elevation error and reduction error [22].

Existing research on combining soft sets and game theory mainly uses soft sets as a representation of games by parametrizing strategy sets. For instance, N-soft set theory is combined with game theory to address missing ratings in classification, achieving a trade-off between classification effectiveness and generality [1]. Intuitionistic neutrosophic soft sets are employed to aid decision-making in a game [5]. However, the resolution of conflicts among parameters of a soft set by forming a game has not been thoroughly addressed. We propose game-theoretic soft sets (GTSoft) to fill the gap.

In data mining, information is often collected from various sources, exhibiting inconsistencies and conflicting opinions. Consequently, the process of decision-making encounters significant challenges. Our model addresses this issue by effectively consolidating these different data sources and facilitating decision-making in a classification context.

The rest of the paper is organized as follows: the background knowledge of soft sets and game theory is discussed in Section II. Section III provides details about GTSoft. Section IV elaborates on the game formulation and an iterative learning process. Section V shows an experiment and its results. Section VI summarizes this study.

II. Preliminaries

A. Soft Sets

The effective parametrization tools of soft sets empowers us to represent data from various perspectives, with each perspective defined by a parameter.

Definition 1. *Let U be the universe and E a set of parameters. Let $P(U)$ denote the power set of U and $A \subset E$. A pair (S, A) is called a soft set over U, where S is a mapping given by [10]:*

$$S\colon A \to P(U).$$

In other words, a soft set over U is a parametrized family of subsets of U.

B. Three-way Classification

When addressing classification problems involving uncertainty, the information accessible to a single expert is often incomplete, and objects with identical information available to the expert may belong to opposing classes. Due to the incompleteness and non-deterministic nature, it is difficult for experts to determine whether certain objects are positive or negative. A boundary class is introduced for objects that cannot be assigned to either category due to a high degree of uncertainty [19]. Objects in the universe are classified into three distinct regions: positive, negative, and boundary regions. These regions are disjoint, and their union encompasses the universe. We refer to this type of classification as three-way partition or three-way classification [19].

Experts translate their information and knowledge into a measurement function and a pair of thresholds, based on which they can derive three-way classification according a set of rules.

Definition 2. *Let $F(x)$ be the measurement function $F\colon O \to [0,1]$ where O is the object set. A three-way classification/partition over O induced by a pair of thresholds (α, β) with $0 \le \beta < \alpha \le 1$ can be expressed by:*

$$POS_{(\alpha,\beta)}(F) = \{x \in O | F(x) \ge \alpha\}, \tag{1a}$$
$$BND_{(\alpha,\beta)}(F) = \{x \in O | \beta < F(x) < \alpha\}, \tag{1b}$$
$$NEG_{(\alpha,\beta)}(F) = \{x \in O | F(x) \le \beta\}. \tag{1c}$$

C. Game Theory

Game theory is a mathematical tool to structure and analyze complicated decision-making problems involving multiple entities competing against or cooperating with each other in an interactive environment [18]. We confine this study to non-zero-sum competitive games. One entity's payoff not only depends on its own choice but also depends on other entities' choices. Moreover, all entities are considered as rational players [16]. They compete to reach their maximum possible benefits. A game provides an analytical tool to reach a trade-off among different entities. The formal definition of a game gives a clear picture of the players, the choices that are available to the players and how the associated payoffs are calculated [16].

We can use G=(P, S, U) to denote a game,

- P is a set of players and $P = \{p_1, p_2, \ldots, p_n\}$.
- S_i is a set of strategies that are available to player p_i and $i = 1, 2, \ldots, n$.
- Each player p_i chooses a strategy $s_i \in S_i$ where $i = 1, 2, \ldots, n$. The combination of strategies chosen by all players, denoted as $(s_1, s_2, \ldots, s_n)$, is referred to as a strategy profile.
- S is a strategy profile set which is the Cartesian product of all strategy sets, and $S = \{S_1 \times S_2 \times \ldots \times S_n\}$.
- ψ is a set of payoff functions and $\psi = \{u_1, u_2, \ldots, u_n\}$, where $u_i : S \to \mathbb{R}$ and it specifies the numerical payoff for player p_i with respect to a strategy profile.

In a game, player p_i chooses a strategy $s_i \in S_i$ and his payoff u_i is not only depending on s_i but also depending on all the choices s_j's made by other players where $j \ne i, j = 1, 2, \ldots, n$. Moreover, each player aims to achieve the largest possible payoff by choosing a strategy. To settle the conflicts and competition among players, we analyze the payoffs of all possible strategy profiles and find the best one so that all the players can benefit the most or reach a balanced trade-off.

In non-cooperative games, the goal is to find the best fit for all players—a Nash equilibrium solution. A Nash equilibrium is defined as follows [12]:

$$u_i(s_1^*, s_2^*, \ldots, s_i, \ldots, s_n^*) \le u_i(s_1^*, s_2^*, \ldots, s_i^*, \ldots, s_n^*) \tag{2}$$
$$\text{for } i = 1, 2, \ldots, n \text{ and } s_i \ne s_i^*.$$

The strategy profile $(s_1^*, s_2^*, \ldots, s_i^*, \ldots, s_n^*)$ is the most balanced trade-off among all players, and no player can achieve a better payoff by unilaterally deviating to another strategy. There can be more than one Nash equilibrium strategy profile in a game.

III. Game-Theoretic Soft Sets

We aim to solve three-way classification problems involving multiple experts, each offering distinct three-way classification outcomes. Relying solely on one expert inevitably results in high classification error or cost. However, by consolidating all opinions, we aim to reduce error or cost. Our approach involves representing classifications as a soft set. We then introduce a set of rules and a pair of thresholds (α, β) for three-way classification using soft sets. Notice that, determining values for (α, β) introduces conflicts among parameters. To resolve these conflicts, we employ a game-theoretic framework and an iterative learning process to find the optimal values.

A. Three-way Classification Using Soft Sets

Let O be the set of objects that need to be classified and the universe U be all possible three-way classifications over O. The parameter set is denoted as A, where each parameter $\epsilon_i \in A$ represents an expert. Parameter ϵ_i is associated with a measurement function $F_{\epsilon_i}(x)$ and a pair of thresholds $(\alpha_{\epsilon_i}, \beta_{\epsilon_i})$. Therefore, for each parameter ϵ_i, there exists a corresponding three-way partition induced by $F_{\epsilon_i}(x)$ and $(\alpha_{\epsilon_i}, \beta_{\epsilon_i})$. We define a soft set ($\phi$, A) over U, where $\phi(\epsilon_i)$ corresponds to a subset of U that contains a three-way partition.

Based on a pair of thresholds (α, β) with $0 \le \beta < \alpha \le 1$, three-way classification using soft set (ϕ, A) is defined as follows:

$$(P)\ POS = \cap\{x | F_{\epsilon_i}(x) \ge \alpha, \epsilon_i \in A\}, \tag{3a}$$
$$(N)\ NEG = \cap\{x | F_{\epsilon_i}(x) \le \beta, \epsilon_i \in A\}, \tag{3b}$$
$$(B)\ BND = \cup\{x | \beta < F_{\epsilon_i}(x) < \alpha, \epsilon_i \in A\}. \tag{3c}$$

TABLE I
THREE-WAY CLASSIFICATIONS FROM DIFFERENT EXPERTS

	α_{e_i}	β_{e_i}	POS	NEG	BND
e_1	0.6	0.2	$\{h_2, h_6\}$	$\{h_1, h_3, h_5\}$	$\{h_4\}$
e_2	0.65	0.3	$\{h_6\}$	$\{h_1, h_3, h_5\}$	$\{h_2, h_4\}$
e_3	0.8	0.1	$\{h_2, h_3, h_4, h_6\}$	$\{h_1\}$	$\{h_5\}$
e_4	0.7	0.2	$\{h_2, h_6\}$	$\{h_1, h_3, h_4, h_5\}$	$\emptyset$

The objects with measurements of all parameters, as given by measurement functions, greater than or equal to α are assigned to the positive region, the objects with measurements less than or equal to β are put into the negative region, and the remaining objects are assigned to the boundary region. Consequently, a final three-way partition {POS, NEG, BND} is obtained.

B. An Example of Three-way Classification Using Soft Sets

An example is given for illustration. Let O be a set of six candidates under consideration, and A be a set of parameters that characterize the evaluation of candidates from a group of experts, represented as e_i for i = 1 to 4. We have:

$$O = \{h_1, h_2, h_3, h_4, h_5, h_6\},$$
$$A = \{e_1, e_2, e_3, e_4\}.$$

Expert e_i uses a measurement function $F_i(x)$ and a pair of thresholds (α_i, β_i) for three-way partition. Function $F_i(x)$ assigns evaluation scores to the candidates and is defined as follows:

$$F_{e_1} = \{h_1 : 0.2, h_2 : 0.7, h_3 : 0.2, h_4 : 0.3, h_5 : 0.2, h_6 : 0.9\},$$
$$F_{e_2} = \{h_1 : 0.2, h_2 : 0.6, h_3 : 0.1, h_4 : 0.6, h_5 : 0.2, h_6 : 0.7\},$$
$$F_{e_3} = \{h_1 : 0.1, h_2 : 0.8, h_3 : 0.8, h_4 : 0.8, h_5 : 0.7, h_6 : 0.9\},$$
$$F_{e_4} = \{h_1 : 0.2, h_2 : 0.8, h_3 : 0.1, h_4 : 0.2, h_5 : 0.0, h_6 : 0.8\}.$$

For each parameter, the three-way classification derived based on rules in Definition 2 are shown in TABLE I. Specifically, for expert e_1, with $\alpha_{e_1} = 0.6$ and $\beta_{e_1} = 0.2$, candidates h_2 and h_6 are assigned to the positive region as their scores provided by e_1 are greater than α_{e_1}, h_1, h_3 and h_5 to the negative region, h_4 to the boundary region.

A soft set (ϕ, A) is defined over all possible three-way classifications {POS, NEG, BND}, where A is the parameter set and ϕ is the mapping function. For each $e_i \in A$, $\phi(e_i)$ specifies a single-element set with a particular three-way classification. For instance, $\phi(e_1) = \{a\}$ where $a = \{\{h_2, h_6\}, \{h_1, h_3, h_5\}, \{h_4\}\}$.

To consolidate all three-way partitions, we set the thresholds $\alpha = 0.6$ and $\beta = 0.2$. According to partition rules in Formula (3), the final partition contains the following three sets:

$$POS = \{h_2, h_6\},\ NEG = \{h_1\},\ BND = \{h_4, h_3, h_5\}.$$

The challenge lies in determining the values for (α, β). Since parameters of a soft set compete with each other to influence the final three-way classification decision. Game theory is employed to find the optimal values for (α, β), by achieving a balance among parameters through data exploration. We will discuss it in the next subsection.

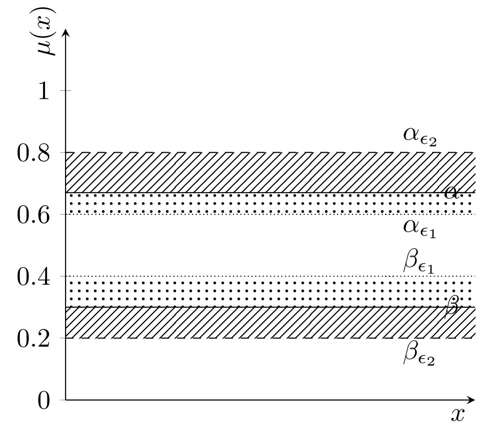

Fig. 1. A two-player game

C. Problem Decomposition

Each parameter $\epsilon_i \in A$ is represented as a player. For simplicity, we use a two-player game to illustrate, where $A = \{\epsilon_1, \epsilon_2\}$. The thresholds $(\alpha_{\epsilon_1}, \beta_{\epsilon_1})$ are associated with ϵ_1 and $(\alpha_{\epsilon_2}, \beta_{\epsilon_2})$ care associated with ϵ_2. In the example, we assume that $\alpha_{\epsilon_1} < \alpha_{\epsilon_2}$ and $\beta_{\epsilon_1} > \beta_{\epsilon_2}$. This assumption can be changed as long as the $\min\{\alpha_{\epsilon_1}, \alpha_{\epsilon_2}\}$ is greater than $\max\{\beta_{\epsilon_1}, \beta_{\epsilon_2}\}$.

Referring to Fig. 1, the dotted lines correspond to α_{ϵ_1} and β_{ϵ_1}, while the dashed lines correspond to α_{ϵ_2} and β_{ϵ_2}. Additionally, the bold lines represent α and β. The consolidation problem is decomposed into two parts—agreement and disagreement.

We confine ourselves to the scenarios where both players can only reach a consensus about region assignment under the following conditions:

- if $F_{\epsilon_1}(x) \geq \max\{\alpha_{\epsilon_1}, \alpha_{\epsilon_2}\}$ and $F_{\epsilon_2}(x) \geq \max\{\alpha_{\epsilon_1}, \alpha_{\epsilon_2}\}$, object x is assigned to the positive region,
- if $F_{\epsilon_1}(x) \leq \min\{\beta_{\epsilon_1}, \beta_{\epsilon_2}\}$ and $F_{\epsilon_2}(x) \leq \min\{\beta_{\epsilon_1}, \beta_{\epsilon_2}\}$, object x is assigned to the negative region,
- if $\max\{\beta_{\epsilon_1}, \beta_{\epsilon_2}\} < F_{\epsilon_1}(x) < \min\{\alpha_{\epsilon_1}, \alpha_{\epsilon_2}\}$ and $\max\{\beta_{\epsilon_1}, \beta_{\epsilon_2}\} < F_{\epsilon_2}(x) < \min\{\alpha_{\epsilon_1}, \alpha_{\epsilon_2}\}$, object x is assigned to the boundary region.

We also assume that the conflicting zones consist of only two parts:

- The upper shadowed area in Fig. 1, where $\{x | \beta_{\epsilon_2} < F_{\epsilon_2}(x) < \alpha_{\epsilon_2}, F_{\epsilon_1}(x) \geq \alpha_{\epsilon_1}, \beta_{\epsilon_2} < \alpha_{\epsilon_1}\}$
- The lower shadowed area in Fig. 1, where $\{x | \beta_{\epsilon_2} < F_{\epsilon_2}(x) < \alpha_{\epsilon_2}, F_{\epsilon_1}(x) \leq \beta_{\epsilon_1}, \beta_{\epsilon_1} < \alpha_{\epsilon_2}\}$

Player ϵ_2 prefers to assign the objects in the upper shadowed area to the boundary region whereas ϵ_1 advocates for these objects to be placed in the positive region. Additionally, ϵ_2 prefers to put the objects in the lower shadowed area into the

TABLE II
THE PAYOFF TABLE FOR A GAME

		ϵ_1		
		no change	$\uparrow\alpha$	$\downarrow\beta$
	no change	$\langle u_1, u_2\rangle$	$\langle u_1, u_2\rangle$	$\langle u_1, u_2\rangle$
	$\downarrow\alpha$	$\langle u_1, u_2\rangle$	$\langle u_1, u_2\rangle$	$\langle u_1, u_2\rangle$
ϵ_2	$\uparrow\beta$	$\langle u_1, u_2\rangle$	$\langle u_1, u_2\rangle$	$\langle u_1, u_2\rangle$

boundary region whereas ϵ_1 advocates for these objects to be placed in the negative region.

IV. Game Formulation and Equilibrium Learning

To settle the conflict in the upper shadowed area, a new upper threshold, α is introduced, which both players agree to use for three-way classification. Consequently, objects in the upper sliced area are assigned to the positive region and objects in the upper dotted area are assigned to the boundary region. In addition, each player desires the consolidated three-way classification to closely align with their individually derived three-way classification, indicating a preference for α to be as close as possible to their original corresponding upper threshold. This same rationale extends to the lower threshold, β. Consequently, we can formulate a game in which players compete to adjust the values for α and β. Thus, the strategy set for player ϵ_1 could be $S_1 = \{no\ change, \uparrow\alpha, \downarrow\beta\}$ and the strategy set for player ϵ_2 could be $S_2 = \{no\ change, \downarrow\alpha, \uparrow\beta\}$. We use $\uparrow$ to indicate an increase in the following threshold and $\downarrow$ to denote a decrease.

A game $G = (P, S, U)$ is formulated between the players as follows:

- The player set $P = \{\epsilon_1, \epsilon_2\}$.
- The strategy profile set $S = \{(s_1, s_2)|\ s_1 \in S_1, s_2 \in S_2\}$.
- Payoff functions $\psi = \{u_1, u_2\}$ where $u_1 : S \rightarrow \mathbb{R}$ and $u_2 : S \rightarrow \mathbb{R}$.

A. Analyzing Payoffs

A cost function is introduced to quantify the cost associated with a three-way classification. The cost is associated with classification error, specifically two types. Misclassification error refers to the number of objects incorrectly classified in the positive or negative region, while undecided error refers to the number of objects classified into the boundary region. Let λ_m be the cost of misclassification error for an object x and λ_u the cost of undecided error for x. Let y denote the true label for x and h(x) denote the assigned label derived from $(\alpha_{\epsilon_i}, \beta_{\epsilon_i})$ and measurement function F_{ϵ_i} according Definition 2. The total cost of the three-way classification are as follows:

$$C(\alpha_{\epsilon_i}, \beta_{\epsilon_i}, F_{\epsilon_i}) = \sum_{x\in O} \mathbf{1}\{h(x) \neq y \wedge h(x) \neq 2\} \cdot \lambda_m + \sum_{x\in O} \mathbf{1}\{h(x) \neq y \wedge h(x) = 2\} \cdot \lambda_u. \quad (4)$$

Where the first term sums the cost of misclassification error for all objects while the second term sums the cost of undecided error. The indicator function $\mathbf{1}\{condition\}$ counts the number of objects satisfying the given condition. Moreover, we label the positive, negative, and boundary regions as 1, 0, and 2, respectively. The players' goal is to reduce their own cost by adjusting the values of the thresholds. Therefore, we can define the payoff functions as the reduced cost after seeking the compromise, expressed as follows:

$$u_1(\alpha, \beta) = C(\alpha_{\epsilon_1}, \beta_{\epsilon_1}, F_{\epsilon_1}) - C(\alpha, \beta, F_{\epsilon_1}), \quad (5)$$

$$u_2(\alpha, \beta) = C(\alpha_{\epsilon_2}, \beta_{\epsilon_2}, F_{\epsilon_2}) - C(\alpha, \beta, F_{\epsilon_2}). \quad (6)$$

Where F_{ϵ_1} and F_{ϵ_2} are the measurement functions associated with $\epsilon 1$ and $\epsilon 2$ respectively. The definition of payoff functions manifests that each player aims for the consolidated three-way classification to closely align with their individually derived three-way classification.

B. Payoff Tables and Nash Equilibria

A payoff table is used to analyze the two-player game as shown in TABLE II. Each column in the table corresponds to a potential strategy for player ϵ_1, drawn from his strategy set S_1. Likewise, each row represents a potential strategy for player ϵ_2 from S_2. We use u_1 and u_2 to denote payoffs for the strategy profile corresponding to each row and column. Subsequently, we input the payoff tuple $\langle u_1, u_2\rangle$ into the respective cell.

The equilibrium of the game can be obtained by analyzing the payoff table. The strategy profile (s_1^*, s_2^*) satisfying the condition in Formula (2) is the equilibrium. Both players cannot find a better strategy within their strategy sets given the other player's choice. Notice that, there could be more than one Nash equilibrium solution.

C. Iterative Learning

A game equilibrium can only identify the optimal strategy profile within the currently defined strategy sets. The existence of a global equilibrium may extend beyond the current strategy sets. To address this, we can iterate the game to gradually approach the most balanced solution. The iterative learning process is as follows:

1) Set the initial values.
   $\alpha = \frac{1}{2}(\alpha_1 + \alpha_2)$
   $\beta = \frac{1}{2}(\beta_1 + \beta_2)$
2) Calculate the payoffs for all the strategy profiles and plug the results into the payoff table.
3) Obtain the equilibrium by analyzing the payoff table.
4) Reset (α, β) as the resulting thresholds from the previous equilibrium and update the strategy sets for both players.
5) Iterate step 2 to 4 until the game equilibria meet the stopping criteria.

The stopping criteria for a game vary with the application context. The possible stopping criteria can be as follows:

- The new iteration of the game can not improve payoffs further.
- The resulting values for α and β violate the constraint $0 \leq \beta < \alpha \leq 1$ or the original assumption $\alpha_{\epsilon_1} \leq \alpha \leq \alpha_{\epsilon_2}$ and $\beta_{\epsilon_2} \leq \beta \leq \beta_{\epsilon_1}$ (in our example).
- No equilibrium exists in the game.

- The times of iteration exceed the maximum count.

We compare the payoffs of the previous iteration with the payoffs at the current equilibrium. We stop the game if both players lose. Alternatively, if one player gains and the other player loses, we check if the gain of one player cannot exceed the loss of the other.

V. Experiments

We use a generated dataset to demonstrate the GTSoft model. Subsequently, we evaluate and analyze the experimental results.

A. Data Generation

We generated 600 objects with two parameters ϵ_1 and ϵ_2. Assume $(\alpha_{\epsilon_1}, \beta_{\epsilon_1})$ for ϵ_1 is (0.75, 0.4) and $(\alpha_{\epsilon_2}, \beta_{\epsilon_2})$ for ϵ_2 is (0.9, 0.25). For parameter ϵ_1, we employ a normal distribution to generate 300 measurement values centered near α_{ϵ_1}, and use another normal distribution to generate 300 measurement values centered near β_{ϵ_1}, the indices of values are used to indicate the corresponding object. In a similar manner, we generate 600 measurement values for parameter ϵ_2. The data is labeled based on certain probabilities. We set λ_u as 1 and λ_m as 2. We split the entire dataset into a training set and a test set with a ratio of 75% to 25%.

B. Learning α and β to Find Equilibria

After preparing the data and setting the cost parameters, we initiate the game on the training set by setting the initial α to 0.825 and β to 0.325. We define a strategy set for ϵ_1 as {no change, decrease α by $1/2(\alpha - \alpha_{\epsilon_1})$, decrease β by $1/3(\beta - \beta_{\epsilon_1})$ }. Likewise, the strategy set for ϵ_2 is defined as {no change, decrease α by $1/3(\alpha - \alpha_{\epsilon_2})$, decrease β by $1/2(\beta - \beta_{\epsilon_2})$}. Next, we use the payoff functions to calculate payoffs for the entire strategy profile set and identify the equilibrium. Afterward, we reset the α and β to the results obtained from the previous iteration. We then update the strategy sets and iterate the game. After each iteration, the equilibrium payoffs, along with the resulting α and β values, and the corresponding two types of error, are all recorded in TABLE III.

TABLE III
ITERATIVE LEARNING WITH THE INITIAL $\alpha = 0.825$ AND $\beta = 0.325$.

Iteration	u_1	u_2	α	β	Misclassification error	Undecided error
1	-4	-2	0.8250	0.2875	38	153
2	-1	-3	0.8250	0.2687	29	169
3	0	-2	0.8250	0.2593	26	174
4	0	-1	0.8250	0.2546	25	176
5	0	0	0.8125	0.2546	28	169
6	1	0	0.8104	0.2546	28	168
7	1	0	0.8104	0.2546	28	168

In the first 4 iterations, α remains unchanged, while β decreases after each iteration. During this phase, the misclassification error decreases with each iteration, but the undecided error increases. From the 5th to the 6th iteration, β begins to decrease while α remains constant. During this period, the misclassification error starts to increase, but the undecided error decreases. After the 6th iteration, the game begins to converge, and the payoffs cannot be further improved. This marks the end of the game. The final values of α and β learned from the training set are 0.8104 and 0.2546, respectively.

C. Results and Analysis

Then we use the learned α and β to classify the test set based on rules in Formula (3). The misclassification error and undecided error of GTSoft are 8 and 50 respectively. In contrast, if the three-way classification decision is made by a single expert based on rules in Definition 2, for ϵ_1, misclassification error and undecided error are 25, 21 respectively. For ϵ_2, misclassification error and undecided error are 28, 14 respectively. GTSoft reduces the misclassification error by 69.8% while increasing the undecided error by 185.7% compared to the corresponding average error of both experts. We summarize these results in TABLE IV.

TABLE IV
PERFORMANCE ON TEST DATA WITH $\lambda_m : \lambda_u = 2 : 1$

	ϵ_1	ϵ_2	GTSoft	Change
Misclassification error	25	28	8	↓ 69.8%
Undecided error	21	14	50	↑185.7%
Cost	71	70	66	↓ 6.4%

At the application level, users typically have a clear understanding of the extent to which they wish to reduce misclassification error and the level of tolerance they can afford for undecided error. It is often the case that no single expert can meet the user's specified targets precisely. Consequently, users may opt for the adoption of GTSoft. To effectively utilize this model, users should establish the appropriate ratio between λ_w and λ_u. The average errors made by experts can be readily computed, allowing us to ascertain the reduction or increase in error rates that users aim to achieve. We employ percentage-based measurements to guide users in making adjustments to the $\lambda_w : \lambda_u$ ratio to align with their desired targets.

GTSoft significantly reduces misclassification error while only slightly decreasing the overall cost with a ratio of λ_m to λ_u being 2:1. However, it substantially increases the size of the boundary region, as illustrated in TABLE IV. Furthermore, adjusting the ratio of λ_m to λ_u to 3:1 accentuates this trend, as evident in TABLE V. In this case, the model not only further reduces misclassification error and the overall cost, but also further increases the size of the boundary region.

If we change the ratio to 1:1, the decrease in misclassification error becomes less significant compared to the 3:1 ratio. The increase in the size of the boundary region also becomes more moderate, accompanied by a slight rise in cost, as indicated in TABLE V. Additionally, we observed that setting the ratio to 1:4 further reduces misclassification error compared to the 1:1 ratio. However, It also leads to a further increase in the size of the boundary region and cost.

In summary, GTSoft effectively decreases misclassification error while increasing undecided error. The extent of this reduction or increase can be adjusted by setting different ratios for λ_m to λ_u. It is evident that the reduction in

TABLE V
PERFORMANCE ON TEST DATA WITH DIFFERENT RATIOS

$\lambda_m : \lambda_u$	Misclassification error	Undecided error	Cost
3:1	$\downarrow$ 96.2%	$\uparrow$ 271.4%	$\downarrow$ 29.9%
2:1	$\downarrow$ 69.8%	$\uparrow$ 185.7%	$\downarrow$ 6.4%
1:1	$\downarrow$ 13.2%	$\uparrow$ 31.4%	$\uparrow$ 4.5%
1:4	$\downarrow$ 43.4%	$\uparrow$ 134.3%	$\uparrow$ 85.5%

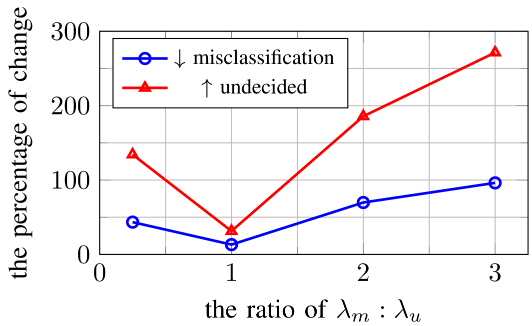

Fig. 2. Two Types of Error with Respect to the Ratio of $\lambda_m : \lambda_u$

misclassification error is positively related to the increase in undecided error, as shown in Fig. 2. Moreover, there is a trade-off between the two types of error, and the model cannot simultaneously decrease both. GTSoft is well-suited for situations where there is a higher tolerance for undecided error compared to misclassification error.

VI. Conclusion

We employed GTSoft to address three-way classification problems characterized by conflicting opinions among multiple experts. By leveraging the robust parametrization capability of soft sets, we represented different classification outcomes. To reach a consensus, we introduced a set of rules and a pair of thresholds (α, β) for three-way classification using soft sets. Determining the optimal values for (α, β) led to conflicts among parameters, and we addressed this using a game-theoretical approach. The iterative learning process converges within a few iterations. GTSoft is capable of reducing misclassification error, although this reduction comes at the expense of increased undecided error. The experiments also showed that by adjusting the ratio of cost parameters associated with the two types of error, we can meet various application specifications regarding these errors. GTSoft is most suitable for situations with a high tolerance for avoiding commitment to any decision but a very low tolerance for making incorrect decisions. In data mining, GTSoft may serve as a valuable tool for integrating data from diverse sources and addressing the challenges inherent in managing inconsistencies and conflicting opinions. Real-life applications need to be explored in future studies. However, the persistent issue of the increasing size of the boundary region remains unresolved. Another concern is the potential occurrence of local minima. Further research could explore more sophisticated models to address these challenges.

References

[1] S. M. Abbas, K. A. Alam, and K.-M. Ko, "A three-way classification with game-theoretic n-soft sets for handling missing ratings in context-aware recommender systems," in *2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. IEEE, 2020, pp. 1–8.

[2] N. Azam and J. T. Yao, "Analyzing uncertainties of probabilistic rough set regions with game-theoretic rough sets," *International journal of approximate reasoning*, vol. 55, no. 1, pp. 142–155, 2014.

[3] N. Bhardwaj and P. Sharma, "An advanced uncertainty measure using fuzzy soft sets: Application to decision-making problems," *Big Data Mining and Analytics*, vol. 4, no. 2, pp. 94–103, 2021.

[4] Y. X. Chen and J. T. Yao, "Sentiment analysis using part-of-speech-based feature extraction and game-theoretic rough sets," in *2021 International Conference on Data Mining Workshops (ICDMW)*. IEEE, 2021, pp. 110–117.

[5] S. Debnath, "Application of intuitionistic neutrosophic soft sets in decision making based on game theory," *International Journal of Neutrosophic Science*, vol. 14, no. 2, pp. 83–97, 2021.

[6] R. Hellali, Z. C. Dagdia, and K. Zeitouni, "Clustering corticosteroids responsiveness in sepsis patients using game-theoretic rough sets," in *17th FedCSIS*, 2023, pp. 539–550.

[7] J. P. Herbert and J. T. Yao, "Game-theoretic rough sets," *Fundamenta Informaticae*, vol. 108, no. 3-4, pp. 267–286, 2011.

[8] C. A. Kamhoua, N. Pissinou, K. Makki, K. Kwiat, and S. S. Iyengar, "Game theoretic analysis of users and providers behavior in network under scarce resources," in *2012 International Conference on Computing, Networking and Communications (ICNC)*. IEEE, 2012, pp. 1149–1155.

[9] S. A. Lashari and R. Ibrahim, "A framework for medical images classification using soft set," *Procedia Technology*, vol. 11, pp. 548–556, 2013.

[10] D. Molodtsov, "Soft set theory—first results," *Computers & mathematics with applications*, vol. 37, no. 4-5, pp. 19–31, 1999.

[11] C. Pielli, F. Chiariotti, N. Laurenti, A. Zanella, and M. Zorzi, "A game-theoretic analysis of energy-depleting jamming attacks," in *2017 International Conference on Computing, Networking and Communications (ICNC)*. IEEE, 2017, pp. 100–104.

[12] S. Rao, "Game theory approach for multiobjective structural optimization," *Computers & Structures*, vol. 25, no. 1, pp. 119–127, 1987.

[13] A. R. Roy and P. Maji, "A fuzzy soft set theoretic approach to decision making problems," *Journal of computational and Applied Mathematics*, vol. 203, no. 2, pp. 412–418, 2007.

[14] S. Sabeena Begam and J. Vimala, "Application of lattice ordered multi-fuzzy soft set in forecasting process," *Journal of Intelligent & Fuzzy Systems*, vol. 36, no. 3, pp. 2323–2331, 2019.

[15] S. Singh and J. T. Yao, "Pneumonia detection with game-theoretic rough sets," in *2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA)*. IEEE, 2021, pp. 1029–1034.

[16] T. Turocy and B. Von Stengel, "Game theory. encyclopedia of information systems, vol. 2," 2002.

[17] T. L. Vincent, "Game theory as a design tool," *J. Mech. Des.*, vol. 105, no. 2, pp. 165–170, 1983.

[18] J. Von Neumann and O. Morgenstern, *Theory of games and economic behavior (60th Anniversary Commemorative Edition)*. Princeton university press, 2007.

[19] Y. Y. Yao, "Three-way decisions with probabilistic rough sets," *Information sciences*, vol. 180, no. 3, pp. 341–353, 2010.

[20] L. A. Zadeh, "Fuzzy sets," *Information and control*, vol. 8, no. 3, pp. 338–353, 1965.

[21] Y. Zhang, P. F. Liu, and J. T. Yao, "Three-way email spam filtering with game-theoretic rough sets," in *2019 International conference on computing, networking and communications (ICNC)*. IEEE, 2019, pp. 552–556.

[22] Y. Zhang and J. T. Yao, "Game theoretic approach to shadowed sets: a three-way tradeoff perspective," *Information Sciences*, vol. 507, pp. 540–552, 2020.

[23] Y. Zhou, Y. Zhang, and J. T. Yao, "Satirical news detection with semantic feature extraction and game-theoretic rough sets," in *Foundations of Intelligent Systems: 25th International Symposium, ISMIS 2020, Graz, Austria, September 23–25, 2020, Proceedings*. Springer, 2020, pp. 123–135.