

Feature Selection and Ranking Method based on Intuitionistic Fuzzy Matrix and Rough Sets

Bich Khue Vo
University of Finance and Marketing
Ho Chi Minh City, Viet Nam
Email: votbkhue@gmail.com
and vokhue@ufm.edu.vn

Hung Son Nguyen
Institute of Computer Science
University of Warsaw
ul. Banacha 2,02-927 Warszawa, Poland
Email: son@mimuw.edu.pl

Abstract—In this paper we propose a novel rough-fuzzy hybridization technique to feature selection and feature ranking problem. The idea is to model the local preference relation between pair of features by intuitionistic fuzzy values and search for a feature ranking that is consistent with those constraints. We apply the techniques used in group decision making where constraints are presented in form of intuitionistic fuzzy preference relation. The proposed method has been illustrated by some simple examples and verified on a benchmark dataset.

Index Terms—Intuitionistic fuzzy matrix, feature ranking, reducts, rough sets.

I. INTRODUCTION

DECISION making plays an extremely important role in the real life problems. With strong development of datasets having a great deal of features, recognizing preferred features is a challenging task. In recent years, rough-fuzzy hybridization becomes a hot trend with great success in machine learning, data mining, decision making, etc. [1].

Usually, a group decision-making process must collect all decision-makers's opinions, establish a suitable method for measuring them, obtain the final scores of all alternatives, and then rank them. Decision makers have to rank the alternatives, find the most preferred feature in order to make decision. So, preference relations are effective techniques to gather the overviews of a group of decision makers. Recently, many researchers have developed many methods of preference relations [2], [3], [4], [5], [6][7]. To solve decision-making problems with uncertain information or not precise judgments, preference relations with Zadeh's fuzzy sets are proposed. There are two number values to measure the degree of membership and nonmembership in Zadeh's fuzzy theory with sum is one, but ignore the decision maker's hesitation in the decision making process. The Atanassov's intuitionistic fuzzy sets consider fully expressing affirmation, negation and hesitation. Particularly, the usage of intuitionistic fuzzy preference relation to affirmative, negative and hesitant characteristics makes the research problem more and more attractive and competitive.

Real world problems may have numerous irrelevant features, and in such cases feature selection can help decision makers choose important information hidden in the full dataset. Feature selection (FS) method belongs to one of the

three main groups: embedded, filter or wrapper methods and can be defined as selecting a subset of available features in a dataset that is associated with the response variables by excluding irrelevant and unnecessary features. An alternative to FS for dimensionality reduction is feature extraction (FE) in which original features are united and then projected into a new feature space with smaller dimensionality. In this paper, we interpret the feature selection and feature ranking as decision making problems and apply the recent techniques for solving it.

In the process of group making decisions, there are often many different opinions of defining the degree of certainty, uncertainty or hesitation among decision makers (DMs). Therefore, it's necessary to define an intuitionistic fuzzy preference relation which have the capability of representing all selections. The consistency of intuitionistic fuzzy preference relations (IFPRs) and the priority weights of DMs gathered from these preference relations play a vital role in group making decision to lead to the most best result. In some works, the consequences of additive consistent and multiplicative consistent IFPRs on priority weights is examined and considered to calculate the priority weights. Numerical analyses have shown that the ranking of the individual priority weights do not differ seriously despite of the different priority weight vectors of the individual priority weights. The intuitionistic fuzzy preference relation is introduced as an indispensable tool for enabling decision-makers to judge the superiority or inferiority of one object to another, in the presence of fuzziness. Ranking methods for alternatives with intuitionistic fuzzy information are expressed straightforwardly and efficiently to get the solution of group decision problems.

The paper is organized as follows: in Section II we recall some basic notions in intuitionistic fuzzy set theory and rough set theory. Section III describes ranking methods that are consistent or semi-consistent with one or more intuitionistic fuzzy preference relations. In Section IV, we present a rough-fuzzy hybridization method for feature ranking and illustrate the proposed method on the base of some simple examples. The results of experiments on the accuracy of the feature ranking methods in the context of classification task are reported in Section V. The conclusions and plan for future research are presented in Section VI.

II. BASIC NOTIONS

In this section we present some fundamental knowledge about intuitionistic fuzzy set, intuitionistic fuzzy relation and feature reduction in rough set theory.

A. Intuitionistic fuzzy sets and Intuitionistic fuzzy matrices

Fuzzy set (FS) theory which was introduced by Zadeh in 1975 [8] just considers the problems with the degree of membership and non-membership without mentioning the degree of hesitation of no decision-making. The Atanassov's intuitionistic fuzzy set (IFS) theory [9] considers fully expressing affirmation, negation and hesitation of decision-makers. Therefore, with real-life situations, IFS theory solves the problems more successfully than FS theory. In this part, some basic notions related to IFS are recalled.

Definition 1 (IFS [9]). Let $X = \{x_1, x_2, \dots, x_m\}$ be a finite universal set. An intuitionistic fuzzy set in X is a set $A = \{(x_i, \mu_A(x_i), \nu_A(x_i)) : x_i \in X\}$, where $\mu_A(x_i), \nu_A(x_i) \in [0, 1]$ and $0 \leq \mu_A(x_i) + \nu_A(x_i) \leq 1$ for any $x_i \in X$. The functions $\mu_A : X \rightarrow [0, 1]$ and $\nu_A : X \rightarrow [0, 1]$ are called the membership and non-membership functions, respectively.

The value $\pi_A(x_i) = 1 - \mu_A(x_i) - \nu_A(x_i)$ is called the intuitionistic index of the element x_i in the set A . It describes a degree of hesitation (or uncertainty) whether x_i is in A or not. For any $x_i \in X$, we have $0 \leq \pi_A(x_i) \leq 1$.

The class of IFS in a universe X is denoted by $\mathcal{IFS}(X)$.

Let \mathcal{F} be the set of tuples (a_1, a_2) , where $a_1, a_2 \in [0, 1]$ and $a_1 + a_2 \leq 1$, i.e.,

$$\mathcal{F} = \{(a_1, a_2) \in [0, 1]^2 : a_1 + a_2 \leq 1\}.$$

A partial order $\leq_{\mathcal{F}}$ over \mathcal{F} defined by:

$$(a_1, a_2) \leq_{\mathcal{F}} (b_1, b_2) \Leftrightarrow a_1 \leq b_1 \text{ and } a_2 \geq b_2.$$

The elements of \mathcal{F} are called the intuitionistic fuzzy values (IFV), of which $(0, 1)$ is the least element and $(1, 0)$ is the greatest element. The operations in $(\mathcal{F}, \leq_{\mathcal{F}})$ are defined by

$$(a_1, a_2) \vee (b_1, b_2) = (\max\{a_1, b_1\}, \min\{a_2, b_2\}).$$

$$(a_1, a_2) \wedge (b_1, b_2) = (\min\{a_1, b_1\}, \max\{a_2, b_2\})$$

$$(a_1, a_2)^c = (a_2, a_1).$$

Each IFS of a universe X is in fact a map from X to \mathcal{F} . If A and B are two IFSs defined by (μ_A, ν_A) and (μ_B, ν_B) , correspondingly, then the union, intersection and complement are defined as follows:

$$\begin{aligned} (\mu_{A \cup B}, \nu_{A \cup B}) &= (\mu_A, \nu_A) \vee (\mu_B, \nu_B) \\ &= (\max\{\mu_A, \mu_B\}, \min\{\nu_A, \nu_B\}) \end{aligned}$$

$$\begin{aligned} (\mu_{A \cap B}, \nu_{A \cap B}) &= (\mu_A, \nu_A) \wedge (\mu_B, \nu_B) \\ &= (\min\{\mu_A, \mu_B\}, \max\{\nu_A, \nu_B\}) \end{aligned}$$

$$(\mu_{A^c}, \nu_{A^c}) = (\mu_A, \nu_A)^c = (\nu_A, \mu_A)$$

$$A \subseteq B \Leftrightarrow (\mu_A, \nu_A) \leq_{\mathcal{F}} (\mu_B, \nu_B)$$

Relations (between two sets X and Y) in traditional set theory are defined as subsets of the Cartesian product $X \times Y$.

It is quite natural to define intuitionistic fuzzy relations as IFSs in $X \times Y$. If $X = \{x_1, \dots, x_m\}$ and $Y = \{y_1, \dots, y_n\}$, then any intuitionistic fuzzy relation in $X \times Y$ can be represented by an $m \times n$ matrix $R = (\rho_{ij})_{m \times n}$, where $\rho_{ij} = (\mu_{ij}, \nu_{ij}) \in \mathcal{F}$ is the IFV describing the membership and non-membership of (x_i, y_j) to this relation.

Definition 2 (Intuitionistic fuzzy matrices - IFM). Any matrix P of order $n \times m$ with values from $\mathcal{F} = \{(a_1, a_2) \in [0, 1]^2 : a_1 + a_2 \leq 1\}$ is called Intuitionistic Fuzzy Matrices. An IFM is said to be square intuitionistic fuzzy matrix (SIFM) if the number of rows is equal to the number of columns. Moreover:

- 1) An identity IFM \mathbb{I} of order n is the square intuitionistic fuzzy matrix (SIFM) of order n with all diagonal entries $(1, 0)$ and non-diagonal entries $(0, 1)$.
- 2) A null intuitionistic fuzzy matrix (IFM) \mathbb{O} of order n is the square intuitionistic fuzzy matrix (SIFM) of order n with all entries $(0, 1)$.

The concepts of intuitionistic fuzzy relation and intuitionistic fuzzy matrix (IFM) have been studied by many authors [10], [11], [12]. IFM is a generalization of Fuzzy Matrix and has been useful in dealing with decision-making, clustering analysis, relational equations, etc.

Since IFM is an extension of FM by replacing the values from $[0, 1]$ by IFV, i.e. elements of $\mathcal{F} = \{(a_1, a_2) \in [0, 1]^2 : a_1 + a_2 \leq 1\}$, and the fuzzy operations \vee and \wedge were extended for the elements of \mathcal{F} . Most of operations on fuzzy matrices can be also extended for IFMs. In particular, if $A = (a_{ij}), B = (b_{ij}) \in \mathcal{F}_{m \times n}$, where $a_{ij} = (\mu_{ij}^a, \nu_{ij}^a), b_{ij} = (\mu_{ij}^b, \nu_{ij}^b)$, for $i = 1, \dots, m; j = 1, \dots, n$ then

- disjunction and conjunction are defined by:

$$A \vee B = (a_{ij} \vee b_{ij})_{m \times n} = (\delta_{ij})_{m \times n}$$

$$\text{where } \delta_{ij} = (\min\{\mu_{ij}^a, \mu_{ij}^b\}, \max\{\nu_{ij}^a, \nu_{ij}^b\}).$$

$$A \wedge B = (a_{ij} \wedge b_{ij})_{m \times n} = (\gamma_{ij})_{m \times n}$$

$$\text{where } \gamma_{ij} = (\max\{\mu_{ij}^a, \mu_{ij}^b\}, \min\{\nu_{ij}^a, \nu_{ij}^b\}).$$

- Comparison: $A \leq B \Leftrightarrow a_{ij} \leq b_{ij}$ for all $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n$.
- The tranpose of $A = (a_{ij})_{m \times n}$ is $A^T = (x_{ij})_{n \times m}$, where $x_{ij} = a_{ji}$.
- The composition of two relations or the product of two matrices $A \in \mathcal{F}_{m \times n}$ and $C \in \mathcal{F}_{n \times l}$ is the matrix $D = (d_{ij}) \in \mathcal{F}_{m \times l}$, where

$$\begin{aligned} d_{ij} &= \bigvee_{k=1}^n (a_{ik} \wedge c_{kj}) = \left(\bigvee_{k=1}^n (\mu_{ik}^a \wedge \mu_{kj}^c), \bigwedge_{k=1}^n (\nu_{ik}^a \vee \nu_{kj}^c) \right) \\ &= \left(\max_{k=1, \dots, n} \{\min\{\mu_{ik}^a, \mu_{kj}^c\}\}, \min_{k=1, \dots, n} \{\max\{\nu_{ik}^a, \nu_{kj}^c\}\} \right). \end{aligned}$$

This operation is denoted by $D = A \circ C$.

B. Feature selection and feature ranking problem

In machine learning, a classification task is defined as the problem of learning the partition of a set of objects into subsets called decision classes (or briefly classes). The partition should

be expressed in terms of object features. Let U be a set of objects. Any function $a : U \rightarrow V_a$, where V_a is the domain (the set of possible values) of a , is called a feature or attribute for U . If a is a measurement such as a person's weight, height, blood pressure or the weather temperature, i.e. V_a is a real interval, then a is called the numeric or quantitative feature. Otherwise, if the values in V_a are not comparable, or if they can not be ordered in a linear order, then a is called categorical, symbolic or qualitative feature.

Decision table is a tuple $T = (U, A \cup \{d\})$, where $U = \{u_1, u_2, \dots, u_n\}$ is a finite set of objects, $A = \{a_1, a_2, \dots, a_m\}$ is a finite set of features called conditional attributes and d is the decision attribute, i.e. the attribute defining the partition of objects into decision classes. Any decision table $T = (U, A \cup \{d\})$ stores a training data set $D = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_n, y_n)\}$ for machine learning algorithms, where $\mathbf{x}_i = (a_1(u_i), \dots, a_m(u_i))$ and $y_i = d(u_i)$.

The goal of the feature selection (FS) problem for decision table $T = (U, A \cup \{d\})$ is to determine the minimal subsets of A satisfying a particular classification performance requirement. Feature ranking problem is to order the feature in a ranking list so that the more important features are at the beginning of the ranking list while the less important features are at the end of the ranking list.

C. Rough Sets and feature selection problem

In rough set theory, the FS problem is formulated as a problem of searching for reducts. Intuitively, reducts of a decision table $T = (U, A \cup \{d\})$ are the minimal subsets (with respect to inclusion) of the set of all attributes A that guarantee the classification performance of the reduced decision table. In the pioneering paper in rough set theory [13], [14], [15], the classification performance was defined in terms of the discernibility between the objects. Formally, for any subset $B \subset A$ and two objects $u, v \in U$, we say that

$$\begin{aligned} B \text{ discerns } u \text{ and } v \text{ (or } u, v \text{ are discernible by } B) \\ \Leftrightarrow \text{there exists } b \in B \text{ such that } b(u) \neq b(v) \end{aligned}$$

The set $B \subset A$ is called the reduct of decision table T if

- For any $u, v \in U$, if u, v are discernible by A and discernible by $\{d\}$ then u, v are discernible by B ;
- No proper subset of B satisfies the previous condition.

This original concept of reducts has been generated by many researchers [16], [17], [18]. However, the solution space for the problem of searching for the reduct with minimal cardinality is $2^m - 1$ where $m = |A|$.

The first approach to minimal reduct problem has been proposed in [14]. For the given decision table $T = (U, A \cup \{d\})$ with $U = \{u_1, u_2, \dots, u_n\}$ and $A = \{a_1, a_2, \dots, a_m\}$, the authors constructed a discernibility matrix $\mathbb{M}(T)$, which is in fact a function $\mathbb{M}(T) : U \times U \rightarrow \mathcal{P}(A)$, where $\mathcal{P}(A)$ denotes the power set of A . For each two objects $u_i, u_j \in U$, we denote $\mathbb{M}(T)(u_i, u_j) = S_{ij}$, where

$$S_{ij} = \begin{cases} \emptyset & \text{if } d(x_i) = d(x_j) \\ \{a \in A : a(x_i) \neq a(x_j)\} & \text{otherwise} \end{cases}$$

The example of discernibility matrix is presented in Table II.

One can notice that a subset of attributes $B \subset A$ discerns a pair of objects $u_i, u_j \in U$ if and only if $B \cap S_{ij} \neq \emptyset$.

Let us notice that for any $u_i, u_j \in U$ we have $S_{ij} = S_{ji}$. That's why, in case of the decision table with two decision classes, the discernibility matrix can be simplified into $p \times q$ matrix where p and q are the cardinalities of the two decision classes.

Since minimal reduct calculation problem is NP-hard, many heuristics algorithms are using random permutations of features [16], [19] as a nondeterministic policy and the algorithm is searching for the reduct according to the attribute order defined by the given permutations.

III. FROM INTUITIONISTIC FUZZY PREFERENCE RELATION TO RANKING.

In this Section we present a method of using Intuitionistic Fuzzy Sets to approximate the concept of fuzzy preference [20] [21], [22].

Definition 3 (intuitionistic fuzzy preference relation). *An intuitionistic fuzzy preference relation B on $X = \{x_1, \dots, x_n\}$ is defined as a matrix $B = (b_{ij})_{n \times n}$, where $b_{ij} = (\mu_{ij}, \nu_{ij})$ for all $i, j = 1, 2, \dots, n$ is an intuitionistic fuzzy value, composed by the certainty degree μ_{ij} to which x_i is preferred to x_j and the certainty degree ν_{ij} to which x_i is non-preferred to x_j , and $\pi_{ij} = 1 - \mu_{ij} - \nu_{ij}$ is interpreted as the hesitation degree to which x_i is preferred to x_j . Moreover, μ_{ij} and ν_{ij} satisfy the following conditions:*

$$\mu_{ij} + \nu_{ij} \leq 1, \quad \mu_{ij} = \nu_{ji}, \quad \mu_{ii} = \nu_{ii} = 0.5$$

for all $i, j = 1; 2; \dots, n$.

Usually, the intuitionistic fuzzy preference relation expresses the opinions of the decision makers about each pair of choices (alternatives), but we would like to convert this relation into a linear order (a ranking list). We can do it by assigning a weight w_i to the i -th choice so that the higher weight means the more preferred choice. Without loss of generality, we can assume that the weight vector can be determined in form of a probability vector, i.e. a vector $\mathbf{w} = (w_1, w_2, \dots, w_n)^T$ such that $w_i \in [0, 1]$ for $i = 1, \dots, n$ and $\sum_{i=1}^n w_i = 1$. We can define the concept of consistent preference relation as follows:

Definition 4 (Additive consistent preference relation). *An intuitionistic fuzzy preference relation $B = ((\mu_{ij}, \nu_{ij}))_{n \times n}$ on $X = \{x_1, \dots, x_n\}$ is an additive consistent preference relation if there exists a probability vector $\mathbf{w} = (w_1, w_2, \dots, w_n)^T$ satisfying the condition:*

$$\mu_{ij} \leq 0.5(w_i - w_j + 1) \leq 1 - \nu_{ij} \quad (1)$$

for all $1 \leq i < j \leq n$.

It is obvious that not every intuitionistic fuzzy preference relation is also an additive consistent relation. In this case the

condition in Eq. (1) can be relaxed by introducing the non-negative deviation variables l_{ij} and r_{ij} for $1 \leq i < j \leq n$ such that

$$\mu_{ij} - l_{ij} \leq 0.5(w_i - w_j + 1) \leq 1 - \nu_{ij} + r_{ij} \quad (2)$$

for all $1 \leq i < j \leq n$. As the deviation variables l_{ij} and r_{ij} become smaller, B becomes closer to an additive consistent intuitionistic fuzzy preference relation. Therefore, in order to find the smallest deviation variables one can developed the following linear optimization model [23], [21]:

Model (A1):

$$\delta = \min \sum_{i=1}^{n-1} \sum_{j=i+1}^n (l_{ij} + r_{ij})$$

$$s.t. \left\{ \begin{array}{l} 0.5(w_i - w_j + 1) + l_{ij} \geq \mu_{ij} \\ 0.5(w_i - w_j + 1) - r_{ij} \leq 1 - \nu_{ij} \\ l_{ij}, r_{ij} \geq 0 \\ w_i \geq 0 \quad \text{for } i = 1, \dots, n, \\ \sum_{i=1}^n w_i = 1 \end{array} \right\} \quad (*)$$

where (*) must be true for all $1 \leq i < j \leq n$.

Let δ^o be the optimal value and let l_{ij}^o and r_{ij}^o for $1 \leq i < j \leq n$ be optimal deviation values of the optimization model (A1). One can see that if $\delta^o = 0$ then B is an additive consistent intuitionistic fuzzy preference relation. Otherwise, we can improve the additive consistency of B by defining the new intuitionistic fuzzy preference relation $\hat{B} = ((\hat{\mu}_{ij}, \hat{\nu}_{ij}))_{n \times n}$, where

$$\hat{\mu}_{ij} = \begin{cases} \mu_{ij} - l_{ij}^o & \text{if } i < j \\ 0.5 & \text{if } i = j \\ \hat{\nu}_{ij} & \text{if } i > j \end{cases} \quad \hat{\nu}_{ij} = \begin{cases} \nu_{ij} - r_{ij}^o & \text{if } i < j \\ \nu_{ij} = 0.5 & \text{if } i = j \\ \hat{\mu}_{ij} & \text{if } i > j \end{cases}$$

Based on matrix \hat{B} we can calculate the priority weight vector $\mathbf{w} = (w_1, \dots, w_n)^T$ by establishing the weight intervals $[w_k^-, w_k^+]$ for each $k = 1, \dots, n$. In order to do that, we solve the following optimization models

Model (A2): for each $k = 1, 2, \dots, n$:

$$(w_k^-, w_k^+) = (\min w_k, \max w_k)$$

$$s.t. \left\{ \begin{array}{l} 0.5(w_i - w_j + 1) \geq \hat{\mu}_{ij} \\ 0.5(w_i - w_j + 1) \leq 1 - \hat{\nu}_{ij} \\ w_i \geq 0 \quad \text{for } i = 1, \dots, n, \\ \sum_{j=1}^n w_j = 1. \end{array} \right\} \quad (*)$$

where (*) must be true for all $1 \leq i < j \leq n$.

It has been shown [23] that if \hat{B} is additive consistent then Model (A2) will return an unique solution for the considered optimization problem.

Definition 5 (Multiplicative consistent preference relation:). An intuitionistic fuzzy preference relation $B = ((\mu_{ij}, \nu_{ij}))_{n \times n}$

on $X = \{x_1, \dots, x_n\}$ is an multiplicative consistent preference relation if there exists a probability vector $\mathbf{w} = (w_1, w_2, \dots, w_n)^T$ satisfying the condition:

$$\mu_{ij} \leq \frac{w_i}{w_i + w_j} \leq 1 - \nu_{ij} \text{ for all } 1 \leq i < j \leq n. \quad (3)$$

Checking for the multiplicative consistency is quite similar to the additive consistency. In this case, we can establish the optimization model (M1). In contrast to model (A1), this model is nonlinear.

Model (M1):

$$\delta = \min \sum_{i=1}^{n-1} \sum_{j=i+1}^n (l_{ij} + r_{ij})$$

$$s.t. \left\{ \begin{array}{l} \frac{w_i}{w_i + w_j} + l_{ij} \geq \mu_{ij} \\ \frac{w_i}{w_i + w_j} - r_{ij} \leq 1 - \nu_{ij} \\ l_{ij}, r_{ij} \geq 0 \\ w_i \geq 0 \quad \text{for } i = 1, \dots, n, \\ \sum_{i=1}^n w_i = 1 \end{array} \right\} \quad (*)$$

where (*) must be true for all $1 \leq i < j \leq n$.

Let δ^* be the optimal value and let l_{ij}^* and r_{ij}^* for $1 \leq i < j \leq n$ be optimal deviation values of the optimization model (M1). One can see that if $\delta^* = 0$ then B is an multiplicative consistent intuitionistic fuzzy preference relation. Otherwise, we can improve the multiplicative consistency of B by defining the new intuitionistic fuzzy preference relation $B^* = ((\mu_{ij}^*, \nu_{ij}^*))_{n \times n}$, where

$$\mu_{ij}^* = \begin{cases} \mu_{ij} - l_{ij}^* & \text{if } i < j \\ 0.5 & \text{if } i = j \\ \nu_{ij}^* & \text{if } i > j \end{cases} \quad \nu_{ij}^* = \begin{cases} \nu_{ij} - r_{ij}^* & \text{if } i < j \\ 0.5 & \text{if } i = j \\ \mu_{ij}^* & \text{if } i > j \end{cases}$$

Based on matrix B^* we can calculate the priority weight vector $\mathbf{w} = (w_1, \dots, w_n)^T$ by establishing the weight intervals $[w_k^-, w_k^+]$ for each $k = 1, \dots, n$. In order to do that, we solve the following optimization models

Model (M2): for each $k = 1, 2, \dots, n$:

$$(w_k^-, w_k^+) = (\min w_k, \max w_k)$$

$$s.t. \left\{ \begin{array}{l} \frac{w_i}{w_i + w_j} \geq \mu_{ij}^* \\ \frac{w_i}{w_i + w_j} \leq 1 - \nu_{ij}^* \\ w_i \geq 0 \quad \text{for } i = 1, \dots, n, \\ \sum_{j=1}^n w_j = 1. \end{array} \right\} \quad (*)$$

where (*) must be true for all $1 \leq i < j \leq n$.

IV. HYBRID METHOD FOR FEATURE RANKING PROBLEM

In this section we present a rough-fuzzy hybridization technique for searching for the optimal ranking list of features, called RAFAR (Rough-fuzzy Algorithm For Attribute Ranking). We introduce the concept of fuzzy discernibility matrix,

which is a generalization of discernibility matrix in rough set theory, and combine it with the ranking calculation methods from intuitionistic fuzzy preference relations.

A. Construction of IFPR from decision table

The general framework of our proposition is presented in the Fig. 1:

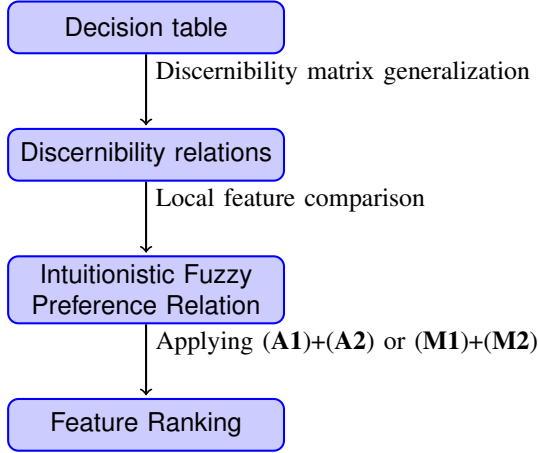


Fig. 1: The general framework of our proposition.

For a given decision table $T = (U, A \cup \{d\})$, where $U = \{u_1, u_2, \dots, u_n\}$ and $A = \{a_1, a_2, \dots, a_m\}$. To simplify the description, let us assume that d is a binary decision attribute, e.g., $V_d = \{-1, 1\}$. The proposed methods also work for the multi-class case.

Let $\mathcal{RED}(T) = \{R \in A : R \text{ is a reduct in } T\}$ denotes the set of all reducts of the decision table T . In [24], the authors classified the attributes into 3 categories:

- Core attributes: the attributes that occur in all reducts:

$$CORE(T) = \bigcap_{R \in \mathcal{RED}} R$$

- Reductive attributes: the attributes that present in at least one reduct:

$$REAT(T) = A - \bigcup_{R \in \mathcal{RED}} R$$

- The attribute is called redundant if it is not a reductive attribute.

Our aim is to generate a feature ranking that at least follows this classification.

Let consider the case of decision table with symbolic values. We define the two decision classes by

$$U_- = \{u \in U : d(u) = -1\}; \quad U_+ = \{u \in U : d(u) = 1\}$$

For each feature $a_k \in A$, we define a function $P_{a_k} : U_+ \times U_- \rightarrow \{0, 1\}$ by

$$P_{a_k}(u_i, u_j) = \begin{cases} 1 & \text{if } a(u_i) \neq a(u_j), \\ 0 & \text{otherwise.} \end{cases}$$

We can see that if a_k is a symbolic feature then P_{a_k} is a relation between U_- and U_+ in the traditional meaning. This relation is called the discernibility relation [14]. Moreover, if $\mathbb{M}(T) = (S_{ij})_{n \times n}$ is the discernibility matrix of T , then

$$P_{a_k}(u_i, u_j) = 1 \text{ if and only if } a_k \in S_{ij}.$$

In case of numeric features, instead of the discernibility relation, we will define a fuzzy discernibility relation. If $a_k \in A$ be a real value feature, we define a fuzzy membership function in $U_+ \times U_-$ for the relation $P_{a_k}^{\alpha, \beta}(u_i, u_j)$ as follows:

$$P_{a_k}^{\alpha, \beta}(u_i, u_j) = \begin{cases} 0 & \text{if } d_k(u_i, u_j) \leq \alpha; \\ 1 & \text{if } d_k(u_i, u_j) \geq \beta; \\ \frac{d_k(u_i, u_j) - \alpha}{\beta - \alpha} & \text{otherwise.} \end{cases}$$

where $0 < \alpha < \beta$ are real parameters and $d_k(u_i, u_j) = |a_k(u_i) - a_k(u_j)|$.

Now, we propose the following method for construction of IFPR over the set of features A . For each $a_k \in A$, we define a function $Score_{a_k} : U_- \times U_+ \rightarrow [0, 1]$ such that

$$Score_{a_k}(u_i, u_j) = \frac{1}{\sum_{a_k \in A} P_{a_k}(u_i, u_j)}$$

Intuitively, the value $Score_{a_k}(u_i, u_j)$ determines the probability that a_k should be selected in order to discern u_i from u_j . For any pair of features $a_k, a_l \in A$, we define the following sets:

$$\begin{aligned} X_{kl} &= \{(u_i, u_j) \in U_+ \times U_- : \\ &\quad Score_{a_k}(u_i, u_j) > Score_{a_l}(u_i, u_j)\} \\ Y_{kl} &= \{(u_i, u_j) \in U_+ \times U_- : \\ &\quad Score_{a_k}(u_i, u_j) = Score_{a_l}(u_i, u_j)\} \\ Z_{kl} &= \{(u_i, u_j) \in U_+ \times U_- : \\ &\quad Score_{a_k}(u_i, u_j) < Score_{a_l}(u_i, u_j)\} \end{aligned}$$

Using those sets, we can calculate the following values

$$\begin{aligned} x_{kl} &= \sum_{(u_i, u_j) \in X_{kl}} Score_{a_k}(u_i, u_j) - Score_{a_l}(u_i, u_j) \\ y_{kl} &= \sum_{(u_i, u_j) \in Y_{kl}} Score_{a_k}(u_i, u_j) \\ z_{kl} &= \sum_{(u_i, u_j) \in Z_{kl}} Score_{a_l}(u_i, u_j) - Score_{a_k}(u_i, u_j) \end{aligned}$$

The discernibility IFPR: $P_{dis} = ((\mu_{kl}, \nu_{kl}))_{m \times m}$ as follows

$$\begin{aligned} \mu_{kl} &= \frac{x_{kl}}{x_{kl} + y_{kl} + z_{kl}} \\ \nu_{kl} &= \frac{z_{kl}}{x_{kl} + y_{kl} + z_{kl}} \end{aligned}$$

B. The illustrated examples

Consider an exemplary decision table shown in Table I. This table was created by taking first 10 objects from the famous “weather data set” and adding one more feature (smog) as the fifth feature.

TABLE I: The decision table extended by one new feature.

T_1	a_1	a_2	a_3	a_4	a_5	dec
ID	outlook	temp.	hum.	windy	smog	play
1	sunny	hot	high	FALSE	no	no
2	sunny	hot	high	TRUE	no	no
3	overcast	hot	high	FALSE	yes	yes
4	rainy	mild	high	FALSE	yes	yes
5	rainy	cool	normal	FALSE	no	yes
6	rainy	cool	normal	TRUE	no	no
7	overcast	cool	normal	TRUE	no	yes
8	sunny	mild	high	FALSE	no	no
9	sunny	cool	normal	FALSE	no	yes
10	rainy	mild	normal	FALSE	no	yes

The simplified version \mathbb{M}'_1 of the discernibility matrix for T_1 is presented in Table II. One can see that this new decision table has exactly 2 reducts: $R_1 = \{a_1, a_2, a_4\}$ and $R_2 = \{a_1, a_3, a_4\}$. According to [24], the features a_1 and a_4 are called the *core attributes* and a_5 is the *redundant attribute*.

TABLE II: The simplified discernibility matrix for decision table from Table I.

$\mathbb{M}'(T_1)$	1	2	6	8
3	a_1, a_5	a_1, a_4, a_5	a_1, a_2, a_3, a_4, a_5	a_1, a_2, a_5
4	a_1, a_2, a_5	a_1, a_2, a_4, a_5	a_2, a_3, a_4, a_5	a_1, a_5
5	a_1, a_2, a_3	a_1, a_2, a_3, a_4	a_4	a_1, a_2, a_3
7	a_1, a_2, a_3, a_4	a_1, a_2, a_3	a_1	a_1, a_2, a_3, a_4
9	a_2, a_3	a_2, a_3, a_4	a_1, a_4	a_2, a_3
10	a_1, a_2, a_3	a_1, a_2, a_3, a_4	a_2, a_4	a_1, a_3

Since all features of T_2 are symbolic, the discernibility relations are presented in Table III. The corresponding *Score* functions for features are showing in Table IV.

Now we can calculate the IFPR P_{dis} . Let consider the two features a_1 and a_2 :

$$X_{12} = \{(3, 1), (3, 2), (4, 8), (7, 6), (9, 6), (10, 8)\}$$

$$Z_{12} = \{(4, 6), (9, 1), (9, 2), (9, 8), (10, 6)\}$$

$$Y_{12} = \text{the rest of features.}$$

The sums of scores in previous sets are:

$$x_{12} = 0.5 + 0.33 + 0.5 + 1 + 0.5 + 0.5 = 3.33$$

$$z_{12} = 0.25 + 0.5 + 0.33 + 0.5 + 0.5 = 2.083$$

$$y_{12} = 3.54$$

Therefore:

$$\mu_{12} = \nu_{21} \approx \frac{3.33}{3.33 + 2.083 + 3.54} \approx 0.3759$$

$$\nu_{12} = \mu_{21} \approx \frac{2.083}{3.33 + 2.083 + 3.54} \approx 0.2349$$

The IFPR P_{dis} for decision table T_1 is presented as follows:

$$\begin{pmatrix} (0.50, 0.50) & (0.38, 0.23) & (0.45, 0.19) & (0.51, 0.23) & (0.62, 0.04) \\ (0.23, 0.38) & (0.50, 0.50) & (0.23, 0.08) & (0.41, 0.25) & (0.61, 0.19) \\ (0.19, 0.45) & (0.08, 0.23) & (0.50, 0.50) & (0.39, 0.36) & (0.61, 0.33) \\ (0.23, 0.51) & (0.25, 0.41) & (0.36, 0.39) & (0.50, 0.50) & (0.55, 0.28) \\ (0.04, 0.62) & (0.19, 0.61) & (0.33, 0.61) & (0.28, 0.55) & (0.50, 0.50) \end{pmatrix}$$

TABLE III: The discernibility relation for symbolic features from decision table T_1 .

(u_i, u_j)	P_{a_1}	P_{a_2}	P_{a_3}	P_{a_4}	P_{a_5}
(3,1)	1	0	0	0	1
(3,2)	1	0	0	1	1
(3,6)	1	1	1	1	1
(3,8)	1	1	0	0	1
(4,1)	1	1	0	0	1
(4,2)	1	1	0	1	1
(4,6)	0	1	1	1	1
(4,8)	1	0	0	0	1
(5,1)	1	1	1	0	0
(5,2)	1	1	1	1	0
(5,6)	0	0	0	1	0
(5,8)	1	1	1	0	0
(7,1)	1	1	1	1	0
(7,2)	1	1	1	0	0
(7,6)	1	0	0	0	0
(7,8)	1	1	1	1	0
(9,1)	0	1	1	0	0
(9,2)	0	1	1	1	0
(9,6)	1	0	0	1	0
(9,8)	0	1	1	0	0
(10,1)	1	1	1	0	0
(10,2)	1	1	1	1	0
(10,6)	0	1	0	1	0
(10,8)	1	0	1	0	0

TABLE IV: The *Score* functions for features from T_1 .

(u_i, u_j)	$Score_{a_1}$	$Score_{a_2}$	$Score_{a_3}$	$Score_{a_4}$	$Score_{a_5}$
(3,1)	0.5	0	0	0	0.5
(3,2)	0.33	0	0	0.33	0.33
(3,6)	0.2	0.2	0.2	0.2	0.2
(3,8)	0.33	0.33	0	0	0.33
(4,1)	0.33	0.33	0	0	0.33
(4,2)	0.25	0.25	0	0.25	0.25
(4,6)	0	0.25	0.25	0.25	0.25
(4,8)	0.5	0	0	0	0.5
(5,1)	0.33	0.33	0.33	0	0
(5,2)	0.25	0.25	0.25	0.25	0
(5,6)	0	0	0	1	0
(5,8)	0.33	0.33	0.33	0	0
(7,1)	0.25	0.25	0.25	0.25	0
(7,2)	0.33	0.33	0.33	0	0
(7,6)	1	0	0	0	0
(7,8)	0.25	0.25	0.25	0.25	0
(9,1)	0	0.5	0.5	0	0
(9,2)	0	0.33	0.33	0.33	0
(9,6)	0.5	0	0	0.5	0
(9,8)	0	0.5	0.5	0	0
(10,1)	0.33	0.33	0.33	0	0
(10,2)	0.25	0.25	0.25	0.25	0
(10,6)	0	0.5	0	0.5	0
(10,8)	0.5	0	0.5	0	0

Now we can use the models (A1) and (A2) to find the feature ranking that is additively consistent with P_{dis} . As a result we receive:

$$(w_1, w_2, w_3, w_4, w_5) = (0.468, 0.214, 0.214, 0.104, 0)$$

This means a_1 is the best and a_5 is the worst feature.

Let us consider the decision table T_2 , which is almost the same as T_1 . The only difference is that, a_2 (temperature) and a_3 (humidity) are numeric features.

The discernibility relations for a_1, a_4, a_5 remain unchanged. As an example, for a_2 , we use the fuzzy discernibility relation $P_{a_2}^{\alpha, \beta}$ with $\alpha = 2$ and $\beta = 12$ and for a_3 , we use the

TABLE V: The decision table with numeric features

T_2	a_1	a_2	a_3	a_4	a_5	dec
ID	outlook	temp.(F)	hum.(%)	windy	smog	play
1	sunny	85.0	85.0	FALSE	no	no
2	sunny	80.0	90.0	TRUE	no	no
3	overcast	83.0	86.0	FALSE	yes	yes
4	rainy	70.0	96.0	FALSE	yes	yes
5	rainy	68.0	80.0	FALSE	no	yes
6	rainy	65.0	70.0	TRUE	no	no
7	overcast	64.0	65.0	TRUE	no	yes
8	sunny	72.0	95.0	FALSE	no	no
9	sunny	69.0	70.0	FALSE	no	yes
10	rainy	75.0	80.0	FALSE	no	yes

fuzzy discernibility relation $P_{a_3}^{\alpha, \beta}$ with $\alpha = 5$ and $\beta = 15$. The fuzzy discernibility relations as well as the corresponding *Score* functions for features are presented in Table VI and Table VII.

TABLE VI: The fuzzy discernibility relations for features from decision table T_2 .

(u_i, u_j)	P_{a_1}	P_{a_2}	P_{a_3}	P_{a_4}	P_{a_5}
(3,1)	1	0	0	0	1
(3,2)	1	0.1	0.5	1	1
(3,6)	1	1	1	1	1
(3,8)	1	0.9	0	0	1
(4,1)	1	1	1	0	1
(4,2)	1	0.8	1	1	1
(4,6)	0	0.3	0	1	1
(4,8)	1	0	1	0	1
(5,1)	1	1	1	0	0
(5,2)	1	1	1	1	0
(5,6)	0	0.1	0	1	0
(5,8)	1	0.2	1	0	0
(7,1)	1	1	1	1	0
(7,2)	1	1	1	0	0
(7,6)	1	0	0.4	0	0
(7,8)	1	0.6	1	1	0
(9,1)	0	1	1	0	0
(9,2)	0	0.9	1	1	0
(9,6)	1	0.2	0	1	0
(9,8)	0	0.1	1	0	0
(10,1)	1	0.8	0.8	0	0
(10,2)	1	0.3	1	1	0
(10,6)	0	0.8	0.3	1	0
(10,8)	1	0.1	0	0	0

The *Score* functions can be used to construct the IFPR in the same way as previously. As the result we receive the following matrix:

$$\begin{pmatrix} (0.50, 0.50) & (0.55, 0.22) & (0.36, 0.24) & (0.51, 0.25) & (0.65, 0.06) \\ (0.22, 0.55) & (0.50, 0.50) & (0.16, 0.41) & (0.44, 0.43) & (0.64, 0.28) \\ (0.24, 0.36) & (0.16, 0.41) & (0.50, 0.50) & (0.47, 0.29) & (0.66, 0.20) \\ (0.25, 0.51) & (0.43, 0.44) & (0.29, 0.47) & (0.50, 0.50) & (0.56, 0.25) \\ (0.06, 0.65) & (0.28, 0.64) & (0.20, 0.66) & (0.25, 0.56) & (0.50, 0.50) \end{pmatrix}$$

And the optimal coefficients for this matrix are:

$$(w_1, w_2, w_3, w_4, w_5) = (0.327, 0.227, 0.322, 0.124, 0)$$

We can see that in this case the feature a_3 become almost important as the feature a_1 , and the redundant feature a_5 is still located at the end of the ranking.

TABLE VII: The *Score* functions for features from T_2 .

(u_i, u_j)	$Score_{a_1}$	$Score_{a_2}$	$Score_{a_3}$	$Score_{a_4}$	$Score_{a_5}$
(3,1)	0.5	0	0	0	0.5
(3,2)	0.278	0.028	0.139	0.278	0.278
(3,6)	0.2	0.2	0.2	0.2	0.2
(3,8)	0.345	0.310	0	0	0.345
(4,1)	0.25	0.25	0.25	0	0.25
(4,2)	0.208	0.167	0.208	0.208	0.208
(4,6)	0	0.130	0	0.435	0.435
(4,8)	0.33	0	0.33	0	0.33
(5,1)	0.33	0.33	0.33	0	0
(5,2)	0.25	0.25	0.25	0.25	0
(5,6)	0	0.091	0	0.901	0
(5,8)	0.455	0.091	0.455	0	0
(7,1)	0.25	0.25	0.25	0.25	0
(7,2)	0.33	0.33	0.33	0	0
(7,6)	0.714	0	0.286	0	0
(7,8)	0.278	0.167	0.278	0.278	0
(9,1)	0	0.5	0.5	0	0
(9,2)	0	0.310	0.345	0.345	0
(9,6)	0.455	0.091	0	0.455	0
(9,8)	0	0.091	0.901	0	0
(10,1)	0.385	0.308	0.308	0	0
(10,2)	0.303	0.091	0.303	0.303	0
(10,6)	0	0.381	0.143	0.476	0
(10,8)	0.909	0.091	0	0	0

C. Simplified ranking method.

The presented above method for feature ranking has quite high computational complexity. The time complexity of this proposition is $O((n \cdot m)^2)$, where n is the number of objects and m is the number of attributes. In this Section we propose a heuristic solution called sRAFAR (simplified Rough-fuzzy Algorithm For Attribute Ranking), which is applicable for the data sets with larger number of objects. The idea is to generate a simplified IFPR instead of the full method presented in Section IV.A. In particular, for any continuous feature $a_k \in A$, we discretize its domain into k equal length intervals and use the binary discernibility relation for the discretized feature.

$$A_k = \{(u_i, u_j) \in U_+ \times U_- : u_i, u_j \text{ are discerned by } a_k\}$$

$$= \{(u_i, u_j) \in U_+ \times U_- : P_{a_k}(u_i, u_j) = 1\}$$

Then the simplified IFPR: $P_s = ((\mu'_{kl}, \nu'_{kl}))_{m \times m}$ can be defined by

$$\mu'_{kl} = P(A_k - A_l | A_k \cup A_l) = 1 - \frac{|A_l|}{|A_k \cup A_l|}$$

$$\nu'_{kl} = P(A_l - A_k | A_k \cup A_l) = 1 - \frac{|A_k|}{|A_k \cup A_l|}$$

Thus, μ'_{kl} is the probability that a pair of objects is discernible a_k but not discernible by a_l , and ν'_{kl} is the probability that a pair of objects is discernible a_l but not discernible by a_k . We have the following theorem:

Theorem 1. For any pair of features $a_k, a_l \in A$, the values $|A_k \cup A_l|, |A_k|, |A_l|$ can be calculated in time $O(n)$, where n is the number of objects. Therefore, the heuristic IFPR can be calculated in $O(n \cdot m^2)$, where m is the number of features.

The proof of this fact is similar to the properties of the MD-heuristic in [24].

V. EXPERIMENT RESULTS

In this section, we present the application of our feature ranking methods for the WDBC data set [25]. The WDBC dataset contains features extracted from digitized image of a fine needle aspirate of a breast mass which describes the characteristics of the cell nuclei in the image. This dataset consists of 569 instances with 30 attributes and two decision classes. The features are encoded by V_1, V_2, \dots, V_{30} . We will compare the quality of feature ranking lists generated by:

- RAFAR: Rough-fuzzy Algorithm For Attribute Ranking;
- sRAFAR: simplified version of RAFAR;
- Random Forest Feature Importance; ¹
- No ranking, i.e. using the original feature list: V_1, V_2, \dots, V_{30} .

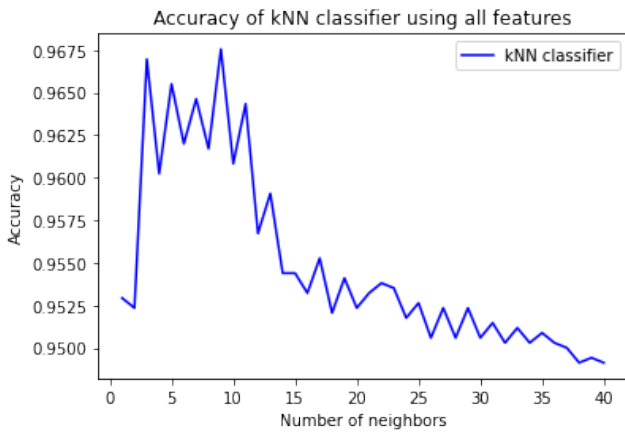


Fig. 2: The accuracy of kNN classifier for different values of k . The highest accuracy is achieved for $k = 9$.

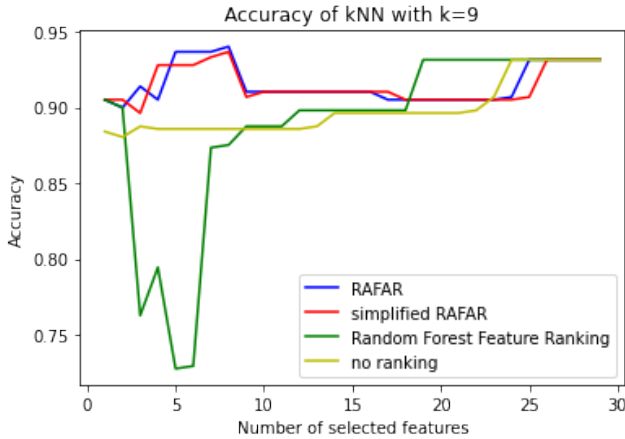


Fig. 3: Comparison of ranking lists with respect to kNN classifier with $k = 9$.

In order to analyze the quality of a ranking list of features (attributes), we select a classifier (classification algorithm) and apply it to the sub-dataset restricted to the first m features

¹https://scikit-learn.org/stable/auto_examples/inspection/plot_permutation_importance_multicollinear.html

for $m = 1, 2, \dots, 30$. For a fixed value of m we evaluate the accuracy of the classifier using 5-fold-cross-validation technique.

The first classifier in our experiment is kNN. Figure 2 presents the accuracy of kNN on the whole data set for different values of the parameter k . We can see that the optimal value of k for kNN classifier equals 9. Therefore we will select kNN with $k = 9$ in the first experiment.

In Fig. 3 the accuracy of kNN with $k = 9$ using first m features of the ranking list for $m = 1, 2, \dots, 30$ is presented. One can see that the accuracy of ranking lists generated by both of our algorithms outperform the other ranking lists.

In Fig. 4 the accuracy comparison of decision tree classifier using first features in the ranking lists is presented. We can see that in this case, the ranking list generated by the RAFAR algorithm seems to be best, especially when we want to use up to 17 features.

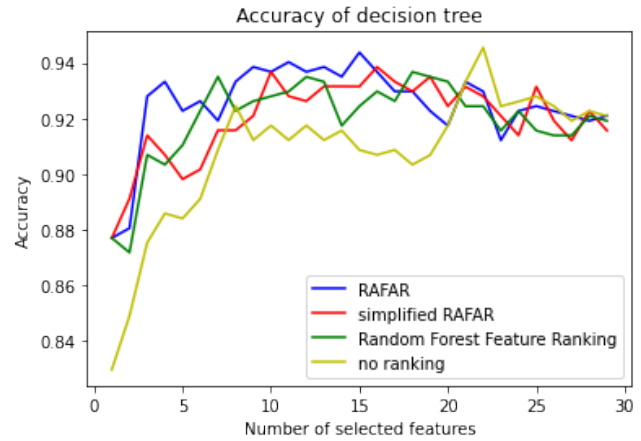


Fig. 4: Comparison of ranking lists with respect to decision tree.

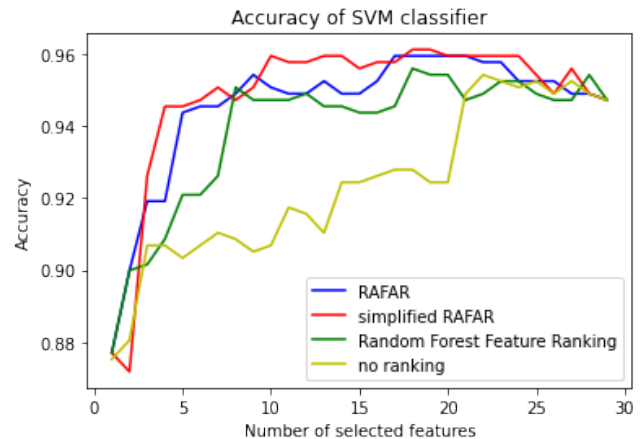


Fig. 5: Comparison of ranking lists with respect to SVM classifier.

Figures 5 present the accuracy comparison for SVM classifier. In this case we notice the fact that the ranking list generated by the sRAFAR algorithm is the best one.

VI. CONCLUSIONS AND FUTURE RESEARCH

In this paper we proposed a new method for feature ranking. We constructed the Intuitionistic Fuzzy Preference Relation (IFPR) for the set of features and searched for the optimal feature ranking that is consistent with the IFPR. All experiments are showing that the proposed rough-fuzzy algorithms for attribute ranking outperform the state of the art method (Random Forest Feature Ranking is the main feature ranking method in the Python scikit-learn library <https://scikit-learn.org/stable/>). We can conclude that the proposed methods are promising and should be thoroughly investigated.

One of the future research direction is multi-criteria feature ranking instead of a single preference relation defined on the based of discernibility power of the features as t has been proposed in RAFAR. The general framework is shown in Fig 1. This idea is motivated by the real life decision making process, where a decision is usually made by a group of experts, $E_k (k = 1, 2, \dots, m)$ with different weights $\lambda = (\lambda_1, \dots, \lambda_m)$, where $\sum_{k=1}^m \lambda_k = 1$ and $\lambda_k \geq 0$ for $k = 1, \dots, k$. In such cases, the individual preference relations of the experts are aggregated to derive a collective preference relation. Let $B^{(k)} = ((\mu_{ij}^{(k)}, \nu_{ij}^{(k)}))_{n \times n}$ be the intuitionistic fuzzy preference relation of the expert E_k , the aggregated preference relation \bar{B} is defined by $\bar{B} = \sum_{k=1}^m \lambda_k \cdot B^{(k)}$. In other words $\bar{B} = ((\bar{\mu}_{ij}, \bar{\nu}_{ij}))_{n \times n}$, where

$$\bar{\mu}_{ij} = \sum_{k=1}^m \lambda_k \mu_{ij}^{(k)}; \quad \bar{\nu}_{ij} = \sum_{k=1}^m \lambda_k \nu_{ij}^{(k)};$$

Theorem 2. *If $B^{(k)}$ are intuitionistic fuzzy preference relation of the expert E_k for $k = 1, \dots, m$ and the weight vector $\lambda = (\lambda_1, \dots, \lambda_m)$ is a probabilistics vector, i.e. $\sum_{k=1}^m \lambda_k = 1$ and $\lambda_k \geq 0$ for $k = 1, \dots, k$, then \bar{B} is also an intuitionistic fuzzy preference relation.*

In such situations, we can apply both ranking methods (i.e. either the models (A1), and (A2) for the additive consistency requirement or the models (M1) and (M2) for the multiplicative consistency requirement) for the collective intuitionistic fuzzy preference relation \bar{B} .

Following this idea, in case of feature ranking problem, we can create more IFPR with different aspects and include them into the calculation process. For example, another preference relation could be calculate on the base of the class homogeneity of features.

We also plan to verify the accuracy of RAFAR and sRAFAR for bigger and more challenging data sets.

ACKNOWLEDGMENT

The authors would like to thank the Vietnam Institute for Advanced Study in Mathematics (VIASM) for the support during the time they had been visiting and working at the Institute.

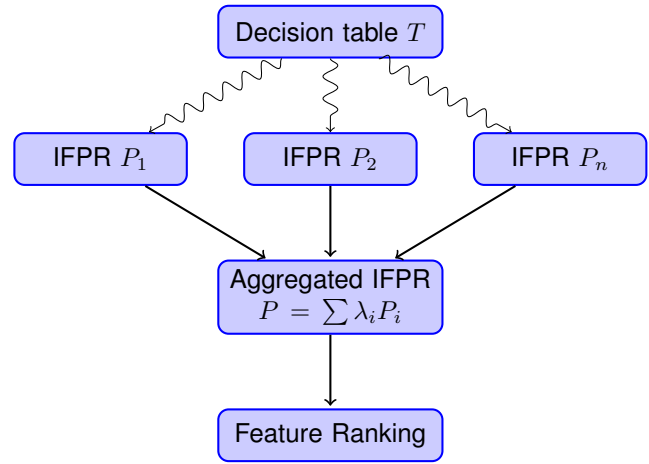


Fig. 6: The general framework for multi-criteria feature ranking.

REFERENCES

- [1] S. K. Pal and A. Skowron, *Rough-Fuzzy Hybridization: A New Trend in Decision Making*, 1st ed. Berlin, Heidelberg: Springer-Verlag, 1999.
- [2] J. Buckley, "Fuzzy hierarchical analysis," *Fuzzy Sets and Systems*, vol. 17, no. 3, pp. 233–247, 1985. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/0165011485900909>
- [3] Y. Dong, Y. Xu, and H. Li, "On consistency measures of linguistic preference relations," *European Journal of Operational Research*, vol. 189, no. 2, pp. 430–444, 2008. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0377221707005619>
- [4] S. Orlovsky, "Decision-making with a fuzzy preference relation," *Fuzzy Sets and Systems*, vol. 1, no. 3, pp. 155–167, 1978. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/0165011478900015>
- [5] J. Tang, F. Meng, and Y. Zhang, "Decision making with interval-valued intuitionistic fuzzy preference relations based on additive consistency analysis," *Information Sciences*, vol. 467, pp. 115–134, 2018. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0020025518305565>
- [6] Z.-J. Wang and X. Tong, "Consistency analysis and group decision making based on triangular fuzzy additive reciprocal preference relations," *Information Sciences*, vol. 361–362, pp. 29–47, 2016. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0020025516303000>
- [7] V. Traneva, S. Tranev, and D. Mavrov, "Interval-valued intuitionistic fuzzy decision-making method using index matrices and application in outsourcing," in *Proceedings of the 16th Conference on Computer Science and Intelligence Systems, Online, September 2-5, 2021*, ser. Annals of Computer Science and Information Systems, M. Ganzha, L. A. Maciaszek, M. Paprzycki, and D. Slezak, Eds., vol. 25, 2021, pp. 251–254. [Online]. Available: <https://doi.org/10.15439/2021F77>
- [8] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, pp. 338–353, 1965. [Online]. Available: <http://www-bisc.cs.berkeley.edu/Zadeh-1965.pdf>
- [9] K. T. Atanassov, "Intuitionistic fuzzy sets," *Fuzzy Sets Syst.*, vol. 20, no. 1, p. 87–96, aug 1986.
- [10] H. Bustince and P. J. Burillo, "Structures on intuitionistic fuzzy relations," *Fuzzy Sets Syst.*, vol. 78, no. 3, pp. 293–303, 1996. [Online]. Available: [https://doi.org/10.1016/0165-0114\(96\)84610-0](https://doi.org/10.1016/0165-0114(96)84610-0)
- [11] M. Pal, S. K. Khan, and A. K. Shyamal, "Intuitionistic fuzzy matrices," *Notes on Intuitionistic fuzzy sets*, vol. 8, no. 2, pp. 51–62, 2002.
- [12] K. T. Atanassov, *Intuitionistic Fuzzy Relations (IFRs)*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 147–193. [Online]. Available: https://doi.org/10.1007/978-3-642-29127-2_8
- [13] Z. Pawlak, "Rough sets," *Int. J. Parallel Program.*, vol. 11, no. 5, pp. 341–356, 1982. [Online]. Available: <https://doi.org/10.1007/BF01001956>

- [14] A. Skowron and C. Rauszer, "The discernibility matrices and functions in information systems," in *Intelligent Decision Support - Handbook of Applications and Advances of the Rough Sets Theory*, ser. Theory and Decision Library, R. Slowinski, Ed. Springer, 1992, vol. 11, pp. 331–362. [Online]. Available: https://doi.org/10.1007/978-94-015-7975-9_21
- [15] Z. Pawlak and A. Skowron, "Rudiments of rough sets," *Information Sciences*, vol. 177, no. 1, pp. 3–27, January 2007.
- [16] D. Slezak, "Rough sets and few-objects-many-attributes problem: The case study of analysis of gene expression data sets," in *Frontiers in the Convergence of Bioscience and Information Technologies 2007, FBIT 2007, Jeju Island, Korea, October 11-13, 2007*, D. Howard and P. Rhee, Eds. IEEE Computer Society, 2007, pp. 437–442. [Online]. Available: <https://doi.org/10.1109/FBIT.2007.160>
- [17] H. S. Nguyen, *Approximate Boolean Reasoning: Foundations and Applications in Data Mining*. Berlin, Heidelberg: Springer-Verlag, 2006, p. 334–506.
- [18] X. Jia, L. Shang, B. Zhou, and Y. Yao, "Generalized attribute reduct in rough set theory," *Knowl. Based Syst.*, vol. 91, pp. 204–218, 2016. [Online]. Available: <https://doi.org/10.1016/j.knosys.2015.05.017>
- [19] D. Slezak and J. Wroblewski, "Order based genetic algorithms for the search of approximate entropy reducts," in *Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing, 9th International Conference, RSFDGrC 2003, Chongqing, China, May 26-29, 2003, Proceedings*, ser. Lecture Notes in Computer Science, G. Wang, Q. Liu, Y. Yao, and A. Skowron, Eds., vol. 2639. Springer, 2003, pp. 308–311. [Online]. Available: https://doi.org/10.1007/3-540-39205-X_45
- [20] Z. Xu and H. Liao, "A survey of approaches to decision making with intuitionistic fuzzy preference relations," *Know.-Based Syst.*, vol. 80, no. C, p. 131–142, may 2015. [Online]. Available: <https://doi.org/10.1016/j.knosys.2014.12.034>
- [21] H. Torun, "Group decision making with intuitionistic fuzzy preference relations," *Knowledge-Based Systems*, vol. 70, 04 2014.
- [22] P. Ren, Z. Xu, and J. Kacprzyk, "Group decisions with intuitionistic fuzzy sets," *Handbook of Group Decision and Negotiation*, pp. 977–995, 2021.
- [23] Z. Xu and R. R. Yager, "Intuitionistic and interval-valued intuitionistic fuzzy preference relations and their measures of similarity for the evaluation of agreement within a group," *Fuzzy Optimization and Decision Making*, vol. 8, pp. 123–139, 2009.
- [24] L. G. Nguyen and H. S. Nguyen, "On elimination of redundant attributes from decision table," in *Federated Conference on Computer Science and Information Systems - FedCSIS 2012, Wroclaw, Poland, 9-12 September 2012, Proceedings*, M. Ganzha, L. A. Maciaszek, and M. Paprzycki, Eds., 2012, pp. 317–322. [Online]. Available: <https://fedcsis.org/proceedings/2012/pliki/324.pdf>
- [25] W. N. Street, W. H. Wolberg, and O. L. Mangasarian, "Nuclear feature extraction for breast tumor diagnosis," in *Biomedical Image Processing and Biomedical Visualization*, R. S. Acharya and D. B. Goldgof, Eds., vol. 1905, International Society for Optics and Photonics. SPIE, 1993, pp. 861 – 870. [Online]. Available: <https://doi.org/10.1117/12.148698>