Decision Making: Applications in Management and Engineering
Vol. 4, Issue 1, 2021, pp. 104-126.
ISSN: 2560-6018
eISSN: 2620-0104
DOI: https://doi.org/10.31181/dmame2104104s

CAREER SELECTION OF STUDENTS USING HYBRIDIZED DISTANCE MEASURE BASED ON PICTURE FUZZY SET AND ROUGH SET THEORY

Rekha Sahu [1], Satya R. Dash [2] and Sujit Das [3*]

[1] School of Computer Engineering, KIIT University, Bhubaneswar, India
[2] School of Computer Applications, KIIT University, Bhubaneswar, India
[3] Department of Computer Science and Engineering, National Institute of Technology Warangal, Warangal, India

Received: 20 October 2020;
Accepted: 2 February 2021;
Available online: 13 February 2021.

Original scientific paper

Abstract: *Since the future of the society depends upon the role of students, so suitable career selection methods for the students are considered to be an important problem to explore. It is assumed that if a student has the required capability and positive attitudes towards a subject, then the student will achieve more in that subject. To consider the uncertain issues involved with students' career selection, picture fuzzy set (PFS) and rough set based approaches are proposed in this study as they are found to be appropriate due to their inherent characteristics to deal with incomplete and imprecise information. For the purpose of selecting a suitable career, the article analyzes student's features in terms of career, memory, interest, knowledge, environment and attitude. We propose two hybridized distance measures using Hausdorff, Hamming and Euclidian distances under picture fuzzy environment where the evaluating information regarding students, subjects and student's features are given in picture fuzzy numbers. Then we present an algorithmic approach using the proposed distance measures and rough set theory. We apply rough set theory to determine whether a particular subject is suitable for a student even if there is controversy to select a stream. Lower and higher approximation with boundary region of rough set theory is used to manage the inconsistent situations. Finally, two case studies are demonstrated to validate the applicability of the proposed idea.*

***Key words**: Picture fuzzy set, distance measure, rough set, students' career*.

* Corresponding author.
E-mail addresses: sahu_r@rediffmail.com; (R. Sahu), sdashfca@kiit.ac.in; (SR. Dash), sujit.das@nitw.ac.in (S. Das).

1. Introduction

Selection of the subject for a better career of a student is a vital task since it is concerned with future employment which influences the whole life of the student and ultimately leads to social development. It is a delegate decision for the students as they are ambitious about it (Batool et al., 2020, Van Dinh et al., 2019, Pratiwi et al., 2020, McKenzie et al., 2020, Wang et al., 2020, Babajide et al., 2020, Orewere et al., 2020). Some students are in confusion or careless to choose a stream. The major confusion is that whether the stream is suitable for their future establishment and their eligibilities are sufficient or not for analyzing and studying efficiently and sincerely. Sometimes many students have no faith in them as they do not know clearly what they will study or what is the content of a particular subjects or stream. As found in various literature related to different colleges and streams (Wen et al., 2018, Nehmeh et al., 2018), we have observed that students often face difficulty in choosing proper stream and make it as a career. To solve the students' career selection problem, many authors have contributed in the last few decades. We have summarized some of the significant contributions which are narrated below. In (Wen et al., 2018), authors have discussed the career choice issue related to choosing accountant as the career, which influenced the researchers to search the various methods for choosing a career for the student by selecting an appropriate subject. The survey of 216 students has shown that both of the internal and external factors influence the career selection process (Babajide et al., 2020). A survey of three hundred students has shown that parents' economic status or social class, financial support, decision making and learning abilities have a huge impact on career decision of students (Batool et al., 2020). Research has also revealed that in order to fulfil parents' expectation and for family or cultural values, few students choose their career in medical school in spite of having lower academic performance (Griffin et al., 2019) which in future hampers their career. Students' engagement, family encouragement, family capital and various scientific matters also encourage the students to study science (Silseth et al., 2018) and select their career accordingly. Although the scientific beliefs, teachers' and parents' expectations, sense of encouragements, and academic prediction, motivate the students to choose science as their career but there is an explicit gap between the motivation factors and the career selection which has been illustrated in (Ramentol et al., 2019). In the rural areas, the environment such as family poverty and rurality often influence the choices of students' career (Carrico et al., 2019). In (Holloway-Friesen et al., 2018), the study found that academic persistence, pursuit of career goals, and high career expectations are significantly influenced by the college environment. Again, it is observed that the environmental impact in terms of more guidance and counselling centers influence the student to choose the right career (Orewere et al., 2020). Hence it can be concluded that the impact of environmental factors like the parent, teacher, family, place, and the institute is a major issue to choose a career. Along with the environmental factors, few other factors are also there. The study of 502 students showed that the different factors like self-efficacy, outcome expectation and career intention have more impact on career choice (Pratiwi et al., 2020, McKenzie et al., 2020). A survey on student mentioned in (Madden et al., 2018) specified that hardness and softness of a subject do not matter for a student if he/she has an interest in that subject. Interest in choosing profession influence the performance in the service and activities (Alkaya et al., 2018). Goel et al. (2018) investigated that the decisions to join the medical profession are mainly dependent on the factors like scientific (interest in medicine), socials (respect/prestige) and humanitarian (desire to help others). Hannula et al. (2002) observed four ways to evaluate the attitude of a student which are emotions aroused

in the situation, emotions associated with the stimuli, expected consequences and related situation to personal values. Positive attitude towards a subject plays an important role, where the positive attitude influences the expected achievement (Guido et al., 2018, Burns et al., 2018). Based on the above discussion, it can be concluded that extrinsic motivation, intrinsic interest and perceived support and encouragement to a particular subject strongly aspire the students to study that subject. Hence by analyzing the related concerns, we find that the specific characteristics of a student may be the cause for the success in a stream. Various studies mentioned above unpack that attitude, knowledge, interest, career, memory and environment are might be considered to be the key factors for students' success.

Two computational intelligence techniques, i.e., picture fuzzy set (PFS) and rough set (RS) theory have a big role to predict the career for students and other decision-making process (Kumbhar et al., 2020, Si et al., 2020, Das et al., 2018, De et al., 2019, Si et al., 2019). In (Dutta, 2018), PFS is proposed for medical investigation and diagnosis. PFS is an extension of intuitionist fuzzy set (IFS) to deal with uncertainty in the situations involving more answers of the types: yes, abstain, no (Cuong et al., 2013), whereas a rough set is a pair of crisp sets, i.e., the lower approximation of set and upper approximation of the original set and these sets may contain fuzzy values (Pawlak, 1995). A detailed study on picture fuzzy set is found in (Cuong, 2014). Picture fuzzy clustering algorithm can be developed for exploiting and investigating hidden knowledge from data. Hierarchical picture clustering is proposed in (Son, 2016) which is an integration of generalized picture distance measure and hierarchical picture fuzzy clustering. The PFS are also useful for computational intelligence problem (Cuong et al., 2013). In the other side, a probability-based rough set theory is proposed in (Ramentol et al., 2019) to predict how likely a student is to succeed in the academic year. An algorithm for decision making based on rough fuzzy set with $\boldsymbol{\delta}$-clustering and upper-lower ($\boldsymbol{\gamma}$, $\boldsymbol{\delta}$) – approximation is modelled in (Ramentol et al. 2019). The intuitionistic fuzzy rough set (Tan et al., 2018) is a combination is of intuitionistic fuzzy (IF) and RS theory, where IF relations are defined and characterized by the lower and upper approximations of the rough set. Then the measures are developed to evaluate approximation quality and ability of classification. The notion of picture fuzzy rough soft set is introduced in (Cuong et al., 2018) which the combination of PFS and RS, and this formulation is used for classification, decision making, and knowledge discovery. In (Cuong et al., 2018), picture fuzzy rough soft sets and picture fuzzy dynamic systems are introduced and these are the extensions of PFS with its applications. Van et al. (2019) defined the distance measure between PFS with similarity and dissimilarity which are useful for image segmentation, decision making and pattern recognition.

As found in literature, none of the researchers has contributed to the student career consideration based on the hybridization of distance measure using PFS and RS. Choosing a stream as the career is a research area of applied fuzzy set theory. After having the basic knowledge of different subjects in school, the student is in a situation to make decisions on the career for whole life by choosing a stream. But due to the imprecise and incomplete nature of information regarding different streams, the concept of a fuzzy set is inevitable for decision-making purpose. In (Dutta, 2018), authors discussed the application of PFS in medical diagnosis, but as per our knowledge, a very few researchers have applied fuzzy set theory in career selection although the deciding on choosing stream is a vital and critical task (Wen et al., 2018). Our proposed model, PFS with hybrid distance measures, has the intension to find the subject having the minimum distance from the student and help to decide to choose the right stream. Some students have positive potential towards a particular subject like a student has high caliber in computer science and so the degree of positive

potential is to be taken into consideration when choosing a stream as a career. Again, some students have less or negative potential towards a particular subject like having high caliber with literature or art may have less or negative knowledge in mathematics. Hence, the degree of negative potential towards the stream is to be considered. It is observed in a career choice that some calibers may have a neutral effect on acquiring subject, i.e., the degree of neutrality can be considered for the subjects like physics, mathematics, or computer science, say neither negative nor positive. Therefore, it is convenient to consider the degree of neutrality in career considering.

In this article, we present a PFS-based approach which uses the hybridized distance measure for choosing a suitable and appropriate career for a student, where the rough set is used to avoid any kinds of confusion in choosing a stream. When students complete school education, they need to choose a suitable and particular stream as their career or to fulfil their ambitions. To find out the student's ability and then assign a stream to him/her according to suitability and appropriateness is a major challenge. To resolve the issue, we have quantified two qualitative concepts i.e., the requirement of the student to understand the subjects and the student's abilities towards the subject. Hence, we have considered the subjects with its features concerning the student's interest, knowledge, memory, career, attitude and environment impact as the general requirement. We have also considered a student's features as interest, knowledge, memory, career, attitude and environment impact towards the specific stream. With these two quantities, we investigate the distance between the student and the subjects using the proposed hybridized distance measures and then select the subject as the career which has a minimum distance from the student. In the process, we may face some situations as defined below.

a) Eligible for more than one stream
b) Neutral for more than one stream.
c) Perfect for one subject.
d) Not eligible for the subjects.

There will be no problem for case (c), but for cases (a), (b) and (d), the decision is critical. These cases can be concluded when choosing a stream using Rough Set (RS) theory. For Example, if computer science, mathematics, physics have the same distance and also minimum distance, there will be fuzziness or inconsistency. Again, for choosing computer science with a specialty in data science as a career, the student must have caliber on computer science, mathematics, statistics and in this case fuzziness or inconsistency may arise, which is resolved using rough set theory.

We have proposed two hybridized distance measures namely hybridization of Hausdorff and Hamming distance measures, and hybridization of Hausdorff and Euclidean distance measures based on Hausdorff, Hamming and Euclidian distance for measuring the distance between student's features and stream's features. The student's features are summarized based on the degree of positive potential, the degree of neutral potential and the degree of negative potential towards the subject whereas subject's features are identified with interest, subject knowledge, memory, career, attitude and environmental impact as the potential's requirement. Also, the refusal degree in sense of neither positive, negative and neutral is taken into consideration. Again, for interest, subject knowledge, memory, career, attitude and environment impact, the degree of positive potential, the degree of neutral potential, the degree of negative potential and the refusal degree are considered. Thus, to choose a career we have taken the degree of positive potential, the negative potential, the degree of neutrality and refusal degree of student, interest, subject knowledge,

memory, career, attitude and environment impact. For each student, we have a four-dimensional vector of information towards a particular subject and for each subject, we have a set of four-dimensional vectors where the set contents six elements. We have illustrated the proposed methods using two case studies. The workflow diagram of our proposed model is depicted in figure 1.

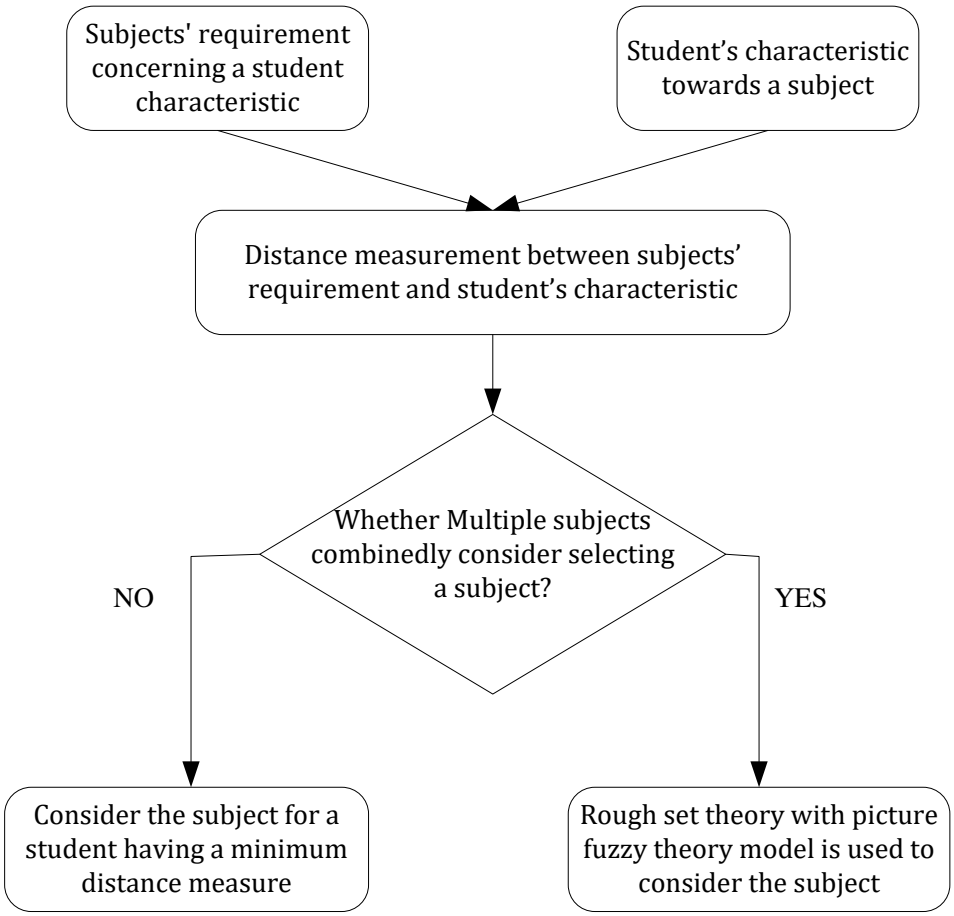

Figure 1. *Workflow diagram of the proposed model*.

Rest of this paper is structured as follows. We have noted the basic concepts used in the proposed models in Section 2. Section 3 presents the proposed model followed by its application in deciding a stream as a career of a student in Section 4, where two case studies are illustrated for choosing the stream. The first case study is implemented using PFS and the second case study is implemented using both PFS and RS. Comparative study is given in Section 5. Finally, in Section 6, we have given our conclusion with possible future works.

2. Preliminaries

In this section, we have discussed, some basic concepts related to this paper.

2.1. Picture Fuzzy Set (PFS) (Cuong, 2014)

PFS A on X is an object of the form

$$A = \{(x,\ \mu_A(x), \tau_A(x),\ \gamma_A(x)) : x \epsilon X\},$$

where $\mu_A(x), \tau_A(x)$ and $\gamma_A(x)$ belongs to the close interval $[0, 1]$and $\mu_A(x)$, $\tau_A(x)\ and\ \gamma_A(x)$ are satisfied with the conditions

$0 \le \mu_A(x) + \tau_A(x) + \gamma_A(x) \le 1, \rho_A(x) = 1 - \left(\mu_A(x) + \tau_A(x) + \gamma_A(x)\right).$

Here $\mu_A(x)$, be the degree of positive membership of x in A, $\tau_A(x)$, be the degree of neutral membership of x in A, and $\gamma_A(x)$, be the degree of negative membership of x in A, and $\rho_A(x)$ be the refusal degree of x in A.

2.2. Picture Distance Measure (Dutta, 2018)

For P, Q ∈ PFS(X), d (P, Q) is called the picture fuzzy distance measure if it satisfies the following criteria:

(i) $d(P, Q) \ge 0$ & $d(P, Q) \le 1$,

(ii) $d(P, Q) = d(Q, P)$,

(iii) $d(P, Q) = 0 \ <=> P = Q$,

(iv) $\mu_{PQ} \times d(P, Q) + \mu_{PR} \times d(P, R) \ge \mu_{QR} \times d(Q, R),\ for\ R \in PFS(X)$.

The symbol "X" indicates the arithmetic product, $\mu_{PQ}, \mu_{PR} and\ \mu_{QR}$ are composition operations of P, Q, and R, and the min-max composition formulae to calculate $\mu_{PQ}, \mu_{PR} and\ \mu_{QR}$ are as follows:

$$\mu_{PQ} = \min_i \left\{\max\{\mu_P(x_i), \mu_Q(x_i)\}\right\}$$

$$\mu_{QR} = \min_i \left\{\max\{\mu_Q(x_i), \mu_R(x_i)\}\right\}$$

$$\mu_{PR} = \min_i \{\max\{\mu_P(x_i), \mu_R(x_i)\}\}$$

2.3. Hamming Distance (Tugrul et al., 2017)

For P ∈ PFS(X), Q ∈ PFS(Y), d (P, Q) is called the picture fuzzy Hamming distance measure defined as follows.

$$P = \{(x,\ \mu_P(x), \tau_P(x),\ \gamma_P(x)) : x \epsilon X\} \quad \& \quad Q = \{(y,\ \mu_Q(y), \tau_Q(y),\ \gamma_Q(y)) : y \epsilon Y\}$$

$$d(P, Q) = |x - y| + |\mu_P(x) - \mu_Q(y)| + |\tau_P(x) - \tau_Q(y)| + |\gamma_P(x) - \gamma_Q(y)|$$

2.4. Euclidean Distance (Tugrul et al., 2017)

For P ∈ PFS(X), Q ∈ PFS(Y), d (P, Q) is called the picture fuzzy Euclidian distance measure defined as follows.

$$P = \{(x,\ \mu_P(x), \tau_P(x),\ \gamma_P(x)) : x \epsilon X\} \ \&\ Q = \{(y,\ \mu_Q(y), \tau_Q(y),\ \gamma_Q(y)) : y \epsilon Y\}$$

$$d(P, Q) = \sqrt{(x - y)^2 + (\mu_P(x) - \mu_Q(y))^2 + (\tau_P(x) - \tau_Q(y))^2 + (\gamma_P(x) - \gamma_Q(y))^2}$$

2.5. Hausdorff Distance (Aspert et al., 2002)

For P ∈ PFS(X), Q ∈ PFS(Y), Hausdorff distance from set P to set Q is a maximum function, defined as

$$h(P,Q) = \max_{a\in P}\{\min_{b\in Q}\{ d(a,b)\}\}$$

where a and b are points of sets P and Q respectively and d (a, b) is any distance metric between the points of P and Q.

2.6. Rough Set (RS)Theory (Pawlak, 1995)

Let (U, A) be the Information System (IS), where U be a set of objects and A be a finite set of attributes such that, ∀ α ϵ A, α: U -> V_α, where V_α is the value set of α and U ≠ Ø, A ≠ Ø. T = (U, AU{γ}) is the decision system, where the attributes contained in A are condition attributes and γ is the decision attribute. The RS theory deals with imperfect knowledge which is expressed by boundary region of a set, and defined with topological operations, interior and closure approximation.

2.7. The indiscernible relation

R⊆ X × X is a binary relation satisfying reflexive, symmetric and transitive property. For x ∈X, the equivalence class is $[x]_R$ = {y| x R y for y ∈ X}.
IS = (U, A) is an information system and for B ⊆A, the associated equivalence relation is defined as:

$IND_{IS}(B) = \{(x, x') \in U^2 \mid \forall\, a \in B, a(x) = a(x')\}$,

where IND_{IS} (B) is called the B- indiscernible relation. The equivalence classes of the B- indiscernible relation are denoted by $[x]_B$.

2.8. Set Approximation

T= (U, A), B ⊆ A & X ⊆ U is the decision system (DS) and we have.

a. X is approximated by constructing the B-lower approximation and B-upper approximations of X using the information of B as

$\underline{B}X = \{x|[x]_B \subseteq X\}$, and $\bar{B}X = \{x|[x]_B \cap X \neq \emptyset\}$ respectively.

b. B-boundary region of X, $BN_B(X) = \bar{B}X - \underline{B}X$ consists of the objects which are not classified into X in B.
c. B-outside region, U-$\bar{B}$X consists of the objects which are not belonging to X.
d. If boundary reason is non-empty, then we have the rough set.

2.9. Rough membership function

The rough membership function quantifies the degree of relative overlap between the set X and the equivalence class R(x) to which x belongs and is defined as follows.

$\mu_X^R: U \to [0,1]$, where $\mu_X^R(x) = \frac{|X\cap R(x)|}{|R(x)|}$, and |X| denotes the cardinality of X.

The membership function of the rough set is expressed as

(i) The conditional probability that x belongs to X given R.
(ii) A degree that x belongs to X given information about x expressed by R.

Using rough membership function, approximations and boundary region of a set are as follows.

$\underline{R}(X) = \{x \in U | \mu_X^R(x) = 1\}$.

$\bar{R}(X) = \{x \in U | \mu_X^R(x) = 0\}$.

$RN_R(X) = \{x \in U | 0 < \mu_X^R(x) < 1\}$.

2.10. Dealing with inconsistency situations

If the inconsistency situations are shown then it may be solved using one of the following actions.

a) Consult the expert for taking actions.

b) Make different tables for the conflicting situation.

c) The examples having with less support should be removed.

d) Basing on upper approximation set and lower approximation set, the quality method can be used to solve inconsistency.

e) The method of generating new decision attributes.

3. Proposed method

This section presents the proposed method. Initially, we present two hybridized distance measure namely hybridization of Hausdorff and Hamming distance and hybridization of Hausdorff and Euclidean distance measures. Then we present an algorithmic approach using these two distance measures. The hybridization of Hausdorff and Hamming distance measure $d_1(A, B)$, and hybridization of Hausdorff and Euclidean distance measure $d_2(A, B)$ are defined as below.
For A, B ϵ PFS(X),

$$d_1(A,B) = \frac{\left(\frac{1}{N}\sum_{i=1}^{N}\frac{\Delta\mu_i+\Delta\tau_i+\Delta\gamma_i+\Delta\rho_i}{4}+\max(\Delta\mu_i,\Delta\tau_i,\Delta\gamma_i,\Delta\rho_i)\right)}{\left(\frac{1}{N}\sum_{i=1}^{N}\frac{\Delta\mu_i+\Delta\tau_i+\Delta\gamma_i+\Delta\rho_i}{4}+\max(\Delta\mu_i,\Delta\tau_i,\Delta\gamma_i,\Delta\rho_i)\right)+\frac{\sum_{i=1}^{N}(\max\{\varphi_i^A,\varphi_i^B\}+|\varphi_i^A-\varphi_i^B|}{N}+1} \quad (1)$$

For A, B ϵ PFS(X),

$$d_2(A,B) = \frac{\left(\frac{1}{N}\sum_{i=1}^{N}\frac{\Delta\mu_i^2+\Delta\tau_i^2+\Delta\gamma_i^2+\Delta\rho_i^2}{4}+\max(\Delta\mu_i^2,\Delta\tau_i^2,\Delta\gamma_i^2,\Delta\rho_i^2)\right)}{\left(\frac{1}{N}\sum_{i=1}^{N}\frac{\Delta\mu_i^2+\Delta\tau_i^2+\Delta\gamma_i^2+\Delta\rho_i^2}{4}+\max(\Delta\mu_i^2,\Delta\tau_i^2,\Delta\gamma_i^2,\Delta\rho_i^2)\right)+\frac{1}{N}\sum_{i=1}^{N}[\max\{\varphi_i^A,\varphi_i^B\}+|\varphi_i^A-\varphi_i^B|^2]^{\frac{1}{2}}+1} \quad (2)$$

$$\Delta\mu_i = |\mu_A(x_i) - \mu_B(x_i)|, \Delta\tau_i = |\tau_A(x_i) - \tau_B(x_i)|, \Delta\gamma_i = |\gamma_A(x_i) - \gamma_B(x_i)|,$$
$$\Delta\rho_i = |\rho_A(x_i) - \rho_B(x_i)|, \emptyset_i^A = |\mu_A(x_i) + \tau_A(x_i) + \gamma_A(x_i)|,$$
$$i = 1,2,\dots,n$$

A distance measure is an objective score that summarizes the relative difference between two objects in a problem domain. Hausdorff distance is the greatest of all distances from a point in one set to the closest point in the other set. Hamming distance calculates the sum or the average differences between the two values. To calculate distance from a vector data to set of vectors data, we have used a set of distance values derived using Hamming distance and a final distance is concluded using Hausdorff distance. Thus, we have the hybridization for x ∈ X,

$$d(x,Y) = f\,(|x-y|, max|x-y|), \qquad y \in Y.$$

Euclidean distance calculates the distance between two real-valued vectors and it is the square root of the sum of the squared differences between the two vectors. Using Hamming distance, from a vector data to set of vectors data, we have a set of distance values and the final distance is concluded using Hausdorff distance. Thus, we have the hybridization for x ∈ X as

$$d(x,Y) = f\,((x-y)^2, \max((x-y)^2)), \qquad y \in Y.$$

where X and Y are a set of vectors.

Algorithmic approach

Step 1. Degree of measurements with PFS is noted for each subject. For each subject, interest, subject knowledge, memory, career, attitude, environment impact requirement are the characteristics. Each characteristic value is quantified with four grounds i.e., the degree of positive potential, the degree of neutral potential, the degree of negative potential and the refusal degree. Thus, we have

$$A = \{(x, \mu_A(x), \tau_A(x), \gamma_A(x)) : x \epsilon\, X\},$$

where X be the set of subjects, $\mu_A(x), \tau_A(x), \gamma_A(x)$ represent the degree of positive potential, the degree of neutral potential, and the degree of negative potential towards the stream respectively. The values of $\mu_A(x), \tau_A(x), \gamma_A(x)$ lie in the close interval [0, 1] and satisfy the following condition.

$$0 \le \mu_A(x) + \tau_A(x) + \gamma_A(x) \le 1$$

The refusal degree for the stream x denoting $\rho_A(x)$ as follows.

$$\rho_A(x) = 1 - (\mu_A(x) + \tau_A(x) + \gamma_A(x))$$

Step 2. Degree of measurements with PFS is noted for the students. Thus, we have

$$B = \{(s, \mu_B(s), \tau_B(s), \gamma_B(s)) : s \in S\}$$

where, S is the set of students, $\mu_B(s), \tau_B(s), \gamma_B(s)$ represent the degree of positive potential, the degree of neutral potential, and the degree of negative potential towards the subject respectively.

The values of $\mu_B(s), \tau_B(s), \gamma_B(s)$ lie in the close interval [0, 1] and satisfy the following condition.

$$0 \le \mu_B(s) + \tau_B(s) + \gamma_B(s) \le 1$$

The refusal degree for the subject s denoting $\rho_B(s)$ is evaluated as follows.

$$\rho_B(s) = 1 - (\mu_B(s) + \tau_B(s) + \gamma_B(s))$$

Step 3. Hybridized distance measures i.e., Hybridization of Hausdorff and Hamming distance measures and Hybridization of Hausdorff and Euclidean distance measures are calculated between student and subject. The subject features in terms of PFS are defined in step 1 and student caliber towards the subject in terms of PFS is defined in step 2.

Step 4. Step 3 is repeated for all the subjects. Then the subject having minimum distance is noted.

Step 5. If the subject distance measure from the student is sufficient to decide the career, then the subject having minimum distance is chosen for the career. Otherwise, we have to follow the next step.

Step 6. If multiple subjects have the minimum distance or more than one subject is required to consider for choosing a career, we use RS theory for solving the fuzziness or inconsistency.

Step 6.1. All subjects (say set S) are approximated by constructing the B-lower approximation and B-upper approximations of S according to distance measures set B (the information as the distance measures from student to subjects) which are respectively stated as $\underline{B}S = \{x|[x]_B \subseteq S\}$ and $\bar{B}S = \{x|[x]_B \cap S \neq \emptyset\}$.

Step 6.2. B-boundary region of S, $BN_B(S) = \bar{B}S - \underline{B}S$ consists of the subjects which are not classified into S in B.

Step 6.3. B-outside region, U-$\bar{B}$S consists of the subjects which are not suitable for a student.

Step 6.4. RS theory is used when the boundary region is non-empty, where we the membership function values of the subjects are considered using probability and subjects with high membership function values are selected.

4. Case study

We have explained two case studies. In case study 1, we have illustrated the distance between students and subjects, then find out the minimum distance for choosing a stream. In case study 2, we have illustrated the inconsistency situation for choosing and to solve its rough set theory is implemented.

Case study 1:

This case study is based on students' distance measures from a stream for selecting the career using two different approaches i.e., Hybridization of Hausdorff and Hamming distance measures and Hybridization of Hausdorff and Euclidean distance measures. Steps to be followed for the purpose is as follows.

Step 1. Find the required calibers for the subjects.

Step 2. Find the calibers of the student towards different subjects.

Step 3. Measure the distance between different subjects from the student using the approaches Hybridization of Hausdorff & Hamming distance measures and Hybridization of Hausdorff & Euclidean distance measures.

Step 4. A suitable stream is selected for the student having minimum distance measure.

In this case study, we have taken five subjects such as computer science(x_1), physics (x_2), chemistry (x_3), mathematics (x_4) and statistic (x_5) with their features as interest (y_1), subject knowledge (y_2), memory (y_3), career (y_4), attitude (y_5) and environment impact (y_6) as a potential requirement towards respective subjects as summarized in table-1. The evaluating values of the features concerning the subjects are represented in terms of PFS as $A = \{(x, \mu_A(x), \tau_A(x), \gamma_A(x)) : x \epsilon X\}$,
where X be the set of streams, $\mu_A(x), \tau_A(x), \gamma_A(x)$ represent the degree of positive potential, the degree of neutral potential, and the degree of negative potential towards the stream respectively. The values of $\mu_A(x), \tau_A(x), \gamma_A(x)$ lie in the close interval [0, 1] and satisfy the following condition.

$$0 \leq \mu_A(x) + \tau_A(x) + \gamma_A(x) \leq 1$$

Again, we have evaluated the refusal degree for the stream x denoting $\rho_A(x)$ as follows.

$$\rho_A(x) = 1 - (\mu_A(x) + \tau_A(x) + \gamma_A(x))$$

Table 1 presents the potentials of the stream towards the students on basis of features $y_1, y_2, y_3, y_4, y_5 and\ y_6$ where each cell value of the table is the element of picture fuzzy set A and table 1 is called as feature-subject picture fuzzy relation R: S→ X.

Then, we have considered four students $s_1, s_2, s_3 and\ s_4$ and their potentials towards the subjects $x_1, x_2, x_3, x_4 and\ x_5$ in terms of PFS B as follows.

$$B = \{(s, \mu_B(s), \tau_B(s), \gamma_B(s)) : s \in S\}$$

Where, S is the set of students, $\mu_B(s), \tau_B(s), \gamma_B(s)$ represent the degree of positive potential, the degree of neutral potential, and the degree of negative potential towards the stream respectively.

Table 1. Feature-subject picture fuzzy relation

Relation R: (X Y) → [0, 1]	y_1 (Interest)	y_2 (Subject knowledge)	y_3 (Memory)	y_4 (Career)	y_5 (Attitude)	y_6 (Environment Impact)
x_1(Computer Science)	(0.4,0,0)	(0.3,0.2,0.4)	(0.1,0.35,0.5)	(0.4,0.3,0.2)	(0.1,0.25,0.5)	(0.7,0.1,0)
x_2 (Physics)	(0.7,0,0)	(0.2,0.4,0.35)	(0,0.4,0.5)	(0.7,0.1,0)	(0.1,0.3,0.5)	(0.4,0.2,0.3)
x_3 (Chemistry)	(0.3,0.4,0.3)	(0.6,0.2,0.1)	(0.2,0.3,0.4)	(0.2,0.35,0.3)	(0.1,0.2,0.6)	(0.2,0.1,0.3)
x_4 (Mathematics)	(0.1,0.3,0.5)	(0.2,0.4,0.3)	(0.8,0,0)	(0.2,0.4,0.3)	(0.2,0.35,0.3)	(0.6,0,0.2)
x_5 (Statistic)	(0.1,0.3,0.5)	(0,0.5,0.35)	(0.2,0.3,0.5)	(0.2,0.35,0.4)	(0.8,0,0.1)	(0,0.5,0.35)

The values of $\mu_B(s), \tau_B(s), \gamma_B(s)$ lie in the close interval [0, 1] and satisfy the following condition.

$$0 \leq \mu_B(s) + \tau_B(s) + \gamma_B(s) \leq 1$$

The refusal degree for the stream x denoting $\rho_B(s)$ is evaluated as follows.

$$\rho_B(s) = 1 - (\mu_B(s) + \tau_B(s) + \gamma_B(s))$$

$\Delta\gamma_1 = |\gamma_A(x_1) - \gamma_B(x_1)| = |0.2 - 0.1| = 0.1$

$\Delta\rho_1 = |\rho_A(x_1) - \rho_B(x_1)| = |0.1 - 0.5| = 0.4$

For y_5,

$\Delta\mu_1 = |\mu_A(x_1) - \mu_B(x_1)| = |0.1 - 0.4| = 3,$

$\Delta\tau_1 = |\tau_A(x_1) - \tau_B(x_1)| = |0.25 - 0.4| = 0.15$

$\Delta\gamma_1 = |\gamma_A(x_1) - \gamma_B(x_1)| = |0.5 - 0.1| = 0.4$

$\Delta\rho_1 = |\rho_A(x_1) - \rho_B(x_1)| = |0.15 - 0.5| = 0.35$

For y_6,

$\Delta\mu_1 = |\mu_A(x_1) - \mu_B(x_1)| = |0.4 - 0.4| = 0,$

$\Delta\tau_1 = |\tau_A(x_1) - \tau_B(x_1)| = |0.3 - 0.4| = 0.1$

$\Delta\gamma_1 = |\gamma_A(x_1) - \gamma_B(x_1)| = |0.2 - 0.1| = 0.1$

$\Delta\rho_1 = |\rho_A(x_1) - \rho_B(x_1)| = |0.1 - 0.5| = 0.4$

$$\frac{1}{6}\left(\sum_{i=1}^{6} \frac{\Delta\mu_i + \Delta\tau_i + \Delta\gamma_i + \Delta\rho_i}{4} + \max(\Delta\mu_i, \Delta\tau_i, \Delta\gamma_i, \Delta\rho_i)\right)$$

$$= \frac{1}{6}\left\{\frac{0 + .4 + .1 + .5}{4} + \max(0, .4, .1, .5) + \frac{.1 + .2 + .3 + 0}{4}\right.$$
$$+ max(.1, .2, .3, 0) + \frac{.3 + .05 + .4 + 0}{4} + \max(0.3, .05, 0.4, 0)$$
$$+ \frac{0 + .1 + .1 + .4}{4} + \max(0, .1, .1, .4) + \frac{.3 + .15 + .4 + .35}{4}$$
$$\left. + \max(.3, .15, .4, .35) + \frac{0 + .1 + .1 + .4}{4} + \max(0, .1, .1, .4)\right.$$
$$= \frac{1}{6}\{.25 + .5 + .15 + .3 + .1875 + .4 + .15 + .4 + .3 + .4 + .15 + .4\}$$
$$= 0.5979166667$$

$$\frac{\sum_{i=1}^{6}(\max\{\varphi_i^A, \varphi_i^B\} + |\varphi_i^A - \varphi_i^B|)}{6}$$

$$= \frac{1}{6}\{\max(.4, .9) + |.4 - .9| + \max(.9, .9) + |.9 - .9| + \max(.95, .9)$$
$$+ |.95 - .9| + \max(.9, .9) + |.9 - .9| + \max(.85, .9) + |.85 - .9|$$
$$+ \max(.8, .9) + |.8 - .9|\}$$
$$= \frac{1}{6}\{.9 + .5 + .9 + .0 + .95 + .05 + .9 + 0 + .9 + .05 + .9 + .1\}$$
$$= 1.025$$

$$d_1(A,B) =$$

$$\frac{\frac{1}{N}\left(\sum_{i=1}^{N}\frac{\Delta\mu_i+\Delta\tau_i+\Delta\gamma_i+\Delta\rho_i}{4}+\max(\Delta\mu_i,\Delta\tau_i,\Delta\gamma_i,\Delta\rho_i)\right)}{\frac{1}{N}\left(\sum_{i=1}^{N}\frac{\Delta\mu_i+\Delta\tau_i+\Delta\gamma_i+\Delta\rho_i}{4}+\max(\Delta\mu_i,\Delta\tau_i,\Delta\gamma_i,\Delta\rho_i)\right)+\frac{\sum_{i=1}^{N}(\max\{\varphi_i^A,\varphi_i^B\}+|\varphi_i^A-\varphi_i^B|)}{N}+1}$$

$$= \frac{0.5979166667}{0.5979166667+1.025+1} = 0.2278703839$$

Thus, for student s_1 and stream x_1, the Hausdorff and Hamming distance is 0.2278703839.

Similarly, for other students and subjects, we have summarized in table 3.

Table 2. Student-Subject Hybridization of Hausdorff and Hamming distance measures

d1(A, B)	x_1	x_2	x_3	x_4	x_5
s_1	0.22787	0.218241	0.142857	0.193277	0.148241
s_2	0.224556	0.186441	0.172414	0.168111	0.183673
s_3	0.187817	0.201331	0.142857	0.145907	0.165217
s_4	0.170984	0.170887	0.211823	0.210526	0.215686

Now we measure the distance of students' potentials from streams' required potentials using the metric of Hausdorff and Euclidean distance measure.

For A, B belongs to PFS X,

$$d_2(A,B) =$$

$$\frac{\left(\frac{1}{N}\sum_{i=1}^{N}\frac{\Delta\mu_i^2+\Delta\tau_i^2+\Delta\gamma_i^2+\Delta\rho_i^2}{4}+\max(\Delta\mu_i^2,\Delta\tau_i^2,\Delta\gamma_i^2,\Delta\rho_i^2)\right)}{\left(\frac{1}{N}\sum_{i=1}^{N}\frac{\Delta\mu_i^2+\Delta\tau_i^2+\Delta\gamma_i^2+\Delta\rho_i^2}{4}+\max(\Delta\mu_i^2,\Delta\tau_i^2,\Delta\gamma_i^2,\Delta\rho_i^2)\right)+\frac{1}{N}\sum_{i=1}^{N}[\max\{\varphi_i^A,\varphi_i^B\}+|\varphi_i^A-\varphi_i^B|^2]^{\frac{1}{2}}+1}$$

Where, for i = 1 to 4

$$\Delta\mu_i = |\mu_A(x_i)-\mu_B(x_i)|, \Delta\tau_i = |\tau_A(x_i)-\tau_B(x_i)|, \Delta\gamma_i = |\gamma_A(x_i)-\gamma_B(x_i)|, \Delta\rho_i = |\rho_A(x_i)-\rho_B(x_i)|,$$

$$\emptyset_i^A = |\mu_A(x_i)+\tau_A(x_i)+\gamma_A(x_i)|, \emptyset_i^B = |\mu_B(x_i)+\tau_B(x_i)+\gamma_B(x_i)|.$$

For student $s_1 \in S$ let us consider for stream $x_1 \in X$ where x_1 stands for computer science.

$$Q(s_1,x_1) = (0.4,0.4,0.1) \quad where,$$

$$\mu_B(x_1) = 0.4,\ \tau_B(x_1) = 0.4, and\ \gamma_B(x_1) = 0.1, \rho_B(x_1) = 1-(0.4+0.4+0.1) = 0.1$$

$$\emptyset_1^B = |\mu_B(x_1)+\tau_B(x_1)+\gamma_B(x_1)| = 0.9.$$

$$R(x_1,y_1) = (0.4,0,0) \quad where,$$

$$\mu_A(x_1) = 0.4,\ \tau_A(x_1) = 0, and\ \gamma_A(x_1) = 0, \rho_A(x_1) = 1-(0.4+0+0) = 0.6$$

$$\emptyset_{11}^A = |\mu_A(x_1)+\tau_A(x_1)+\gamma_A(x_1)| = 0.4.$$

$$R(x_1,y_2) = (0.3,0.2,0.4) \quad where,$$

$\mu_A(x_1) = 0.3,\ \tau_A(x_1) = 0.2, and\ \gamma_A(x_1) = 0.4, \rho_A(x_1) = 1 - (0.3 + 0.2\ + 0.4) = 0.1$

$\emptyset_{12}^A = |\mu_A(x_1) + \tau_A(x_1) + \gamma_A(x_1)| = 0.9.$

$$R(x_1, y_3) = (0.1, 0.34, 0.5)\quad where,$$

$\mu_A(x_1) = 0.1,\ \tau_A(x_1) = 0.35, and\ \gamma_A(x_1) = 0.5, \rho_A(x_1) = 1 - (0.1 + 0.35\ + 0.5) = 0.05$

$\emptyset_{13}^A = |\mu_A(x_1) + \tau_A(x_1) + \gamma_A(x_1)| = 0.95.$

$$R(x_1, y_4) = (0.4, 0.3, 0.2)\quad where,$$

$\mu_A(x_1) = 0.4,\ \tau_A(x_1) = 0.3, and\ \gamma_A(x_1) = 0.2, \rho_A(x_1) = 1 - (0.4 + 0.3\ + 0.2) = 0.1$

$\emptyset_{14}^A = |\mu_A(x_1) + \tau_A(x_1) + \gamma_A(x_1)| = 0.9.$

$$R(x_1, y_5) = (0.1, 0.25, 0.5)\quad where,$$

$\mu_A(x_1) = 0.1,\ \tau_A(x_1) = 0.25, and\ \gamma_A(x_1) = 0.5, \rho_A(x_1) = 1 - (0.1 + 0.25 + 0.5) = 0.15$

$\emptyset_{15}^A = |\mu_A(x_1) + \tau_A(x_1) + \gamma_A(x_1)| = 0.85.$

$$R(x_1, y_6) = (0.7, 0.1, 0)\quad where,$$

$\mu_A(x_1) = 0.7,\ \tau_A(x_1) = 0.1, and\ \gamma_A(x_1) = 0, \rho_A(x_1) = 1 - (0.7 + 0.1 + 0) = 0.2$

$\emptyset_{16}^A = |\mu_A(x_1) + \tau_A(x_1) + \gamma_A(x_1)| = 0.8.$

For y_1,

$\Delta\mu_1 = |\mu_A(x_1) - \mu_B(x_1)| = |0.4 - 0.4| = 0,$

$\Delta\tau_1 = |\tau_A(x_1) - \tau_B(x_1)| = |0 - 0.4| = 0.4$

$\Delta\gamma_1 = |\gamma_A(x_1) - \gamma_B(x_1)| = |0 - 0.1| = 0.1$

$\Delta\rho_1 = |\rho_A(x_1) - \rho_B(x_1)| = |0.6 - 0.1| = 0.5$

For y_2,

$\Delta\mu_1 = |\mu_A(x_1) - \mu_B(x_1)| = |0.3 - 0.4| = 0.1,$

$\Delta\tau_1 = |\tau_A(x_1) - \tau_B(x_1)| = |0.2 - 0.4| = 0.2$

$\Delta\gamma_1 = |\gamma_A(x_1) - \gamma_B(x_1)| = |0.4 - 0.1| = 0.3$

$\Delta\rho_1 = |\rho_A(x_1) - \rho_B(x_1)| = |0.1 - 0.1| = 0$

For y_3,

$\Delta\mu_1 = |\mu_A(x_1) - \mu_B(x_1)| = |0.1 - 0.4| = 0.3,$

$\Delta\tau_1 = |\tau_A(x_1) - \tau_B(x_1)| = |0.35 - 0.4| = 0.05$

$\Delta\gamma_1 = |\gamma_A(x_1) - \gamma_B(x_1)| = |0.5 - 0.1| = 0.4$

$\Delta\rho_1 = |\rho_A(x_1) - \rho_B(x_1)| = |0.05 - 0.1| = 0.05$

For y_4,

$\Delta\mu_1 = |\mu_A(x_1) - \mu_B(x_1)| = |0.4 - 0.4| = 0,$

$\Delta\tau_1 = |\tau_A(x_1) - \tau_B(x_1)| = |0.3 - 0.4| = 0.1$

$\Delta\gamma_1 = |\gamma_A(x_1) - \gamma_B(x_1)| = |0.2 - 0.1| = 0.1$

$\Delta\rho_1 = |\rho_A(x_1) - \rho_B(x_1)| = |0.1 - 0.5| = 0.4$

For y_5,

$\Delta\mu_1 = |\mu_A(x_1) - \mu_B(x_1)| = |0.1 - 0.4| = 3,$

$\Delta\tau_1 = |\tau_A(x_1) - \tau_B(x_1)| = |0.25 - 0.4| = 0.15$

$\Delta\gamma_1 = |\gamma_A(x_1) - \gamma_B(x_1)| = |0.5 - 0.1| = 0.4$

$\Delta\rho_1 = |\rho_A(x_1) - \rho_B(x_1)| = |0.15 - 0.5| = 0.35$

For y_6,

$\Delta\mu_1 = |\mu_A(x_1) - \mu_B(x_1)| = |0.4 - 0.4| = 0,$

$\Delta\tau_1 = |\tau_A(x_1) - \tau_B(x_1)| = |0.3 - 0.4| = 0.1$

$\Delta\gamma_1 = |\gamma_A(x_1) - \gamma_B(x_1)| = |0.2 - 0.1| = 0.1$

$\Delta\rho_1 = |\rho_A(x_1) - \rho_B(x_1)| = |0.1 - 0.5| = 0.4$

$$\left(\frac{1}{6}\sum_{i=1}^{N}\frac{\Delta\mu_i^2 + \Delta\tau_i^2 + \Delta\gamma_i^2 + \Delta\rho_i^2}{4} + \max(\Delta\mu_i^2, \Delta\tau_i^2, \Delta\gamma_i^2, \Delta\rho_i^2)\right)$$

$$= \frac{1}{6}\left\{\frac{0 + .16 + .01 + .25}{4} + \max(0, .16, .01, .25)\right.$$
$$+ \frac{.01 + .04 + .09 + 0}{4} + max(.01, .04, .09, 0)$$
$$+ \frac{.09 + .0025 + .16 + 0}{4} + \max(0.09, .0025, 0.16, 0)$$
$$+ \frac{0 + .01 + .01 + .16}{4} + \max(0, .01, .01, .16)$$
$$+ \frac{.09 + .0225 + .16 + .1225}{4} + \max(.09, .0025, .16, .1225)$$
$$+ \frac{0 + .01 + .01 + .16}{4} + \max(0, .01, .01, .16)$$
$$= \frac{1}{6}\{.105 + .25 + .035 + .09 + .063125 + .16 + .045 + .16$$
$$+ .09875 + .16 + .045 + .16\} = .2286458333$$

$$\frac{1}{6}\sum_{i=1}^{N}\left[\max\left\{\varphi_i^A, \varphi_i^B\right\} + \left|\varphi_i^A - \varphi_i^B\right|^2\right]^{\frac{1}{2}}$$

$$= \frac{1}{6}\{\{\max(.4, .9) + |.4 - .9|^2\}^{1/2} + \{\max(.9, .9) + |.9 - .9|^2\}^{1/2}$$
$$+ \{\max(.95, .9) + |.95 - .9|^2\}^{1/2} + \{\max(.9, .9) + |.9 - .9|^2\}^{1/2}$$
$$+ \{\max(.85, .9) + |.85 - .9|^2\}^{1/2} + \{\max(.8, .9) + |.8 - .9|^2\}^{1/2}\}$$
$$= \frac{1}{6}\left\{1.15^{1/2} + .9^{1/2} + .9525^{1/2} + .9^{1/2} + .9025^{1/2} + .91^{1/2}\right\}$$
$$= 0.974941232$$

$$d_2(A, B) =$$

$$\frac{(\frac{1}{N}\sum_{i=1}^{N}\frac{\Delta\mu_i^2+\Delta\tau_i^2+\Delta\gamma_i^2+\Delta\rho_i^2}{4}+\max(\Delta\mu_i^2,\Delta\tau_i^2,\Delta\gamma_i^2,\Delta\rho_i^2))}{(\frac{1}{N}\sum_{i=1}^{N}\frac{\Delta\mu_i^2+\Delta\tau_i^2+\Delta\gamma_i^2+\Delta\rho_i^2}{4}+\max(\Delta\mu_i^2,\Delta\tau_i^2,\Delta\gamma_i^2,\Delta\rho_i^2))+\frac{1}{N}\sum_{i=1}^{N}[\max\{\varphi_i^A,\varphi_i^B\}+|\varphi_i^A-\varphi_i^B|^2]^{\frac{1}{2}}+1}$$

$$=\frac{.2286458333}{.2286458333+0.974941232+1}=0.103760744$$

Thus, for student s_1 and stream x_1, the Hausdorff and Euclidean distance is 0.103760744.

Similarly, for other students and subjects, we have summarized in table 4.

Table 3. Student-Subject Hybridization of Hausdorff and Euclidean distance measures

Distance 's' to x	x_1	x_2	x_3	x_4	x_5
s_1	0. 10376	0.087871	**0.038433**	0.06533	0.051477
s_2	0.114139	0.061914	**0.055452**	0.068413	0.073627
s_3	0.071581	0.075949	**0.041148**	0.047902	0.0591
s_4	0.059973	**0.058362**	0.087786	0.086236	0.090134

Finally, from table 3 and table 4, it is found students s_1, s_2, s_3 have efficiency for chemistry (x_3) where student s_4 have efficiency for physics (x_2) since they have minimum distance in respective subjects as highlighted in boldfaced.

Case study 2:

This case study is the illustration of the extension approach of case study 1. We have used PFS with hybrid distance measure as well as RS theory for choosing a career to consider whether a stream is suitable for a student or not. PFS and hybrid distance measure have used for finding the distance between student's potential from the required potential for a stream. Then the RS theory is implemented to choose the best stream from different options and criteria. First, we have used picture fuzzy set and hybrid distance measure and then RS theory with the distance measures to finalize the suitable stream. The steps have followed is summarized below.

Step 1. Noted the required efficiency for the subjects.

Step 2. Find out the data of students according to their efficiency towards the subjects.

Step 3. Measure the distance between different subjects from the student using the approaches Hybridization of Hausdorff & Hamming distance measures and Hybridization of Hausdorff & Euclidean distance measures.

Step 4. Use RS theory to categorize the students according to their distance measure and also found out the students are in fuzziness.

Steps 5. Solve the fuzziness using the rule (c) of section 2.5.

Step 6. Finalize the stream for the students.

This case study is the study of 10 students 'efficiency towards computer science say x_1. It is found that choosing computer science as stream not only depends upon computer science (x_1), it also depends on efficiency in mathematics say $x_4$and statistic say x_5. It is noted in the previous case study that the requirements of efficiency for $x_1, x_4$and x_5 are defined in table 1. The

efficiency of students towards x_1, x_4 and x_5 of 10 students say $s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8, s_9,$ and s_{10} are summarized in table 5.

Table 4. Students' information towards computer science, mathematics and statistics.

Relation Q (S, X) → [0,1]	x_1 (Computer Science)	x_4(Mathematics)	x_5(Statistic)
s_1	[0.33, 0.2414, 0.32]	[0.2,0.4,0.3]	[0.264,0.386,0.381]
s_2	[0.1,0.1,0.5]	[0.4,0.2,0.2]	[0.4,0.4,0.1]
s_3	[0.23,0.3,0.27]	[0.1,0.4,0.5]	[0.001,0.009,0.001]
s_4	[0.35,0.15,0.2]	[0.4,0.3,0.3]	[0.364,0.356,0.281]
s_5	[0.4,0.4,0.3]	[0.38,0.4,0.2]	[0.001,0.005,0.01]
s_6	[0.2,0.4,0.1]	[0.4,0.4,0.2]	[0.001,0.003,0.0]
s_7	[0.2,0.1,0.2]	[0.4,0.4,0.1]	[0.005,0.004,0.001]
s_8	[0.1,0.5,0.4]	[0.1,0.1,0.4]	[0.001,0.001,0.004]
s_9	[0.8,0.4,0.2]	[0.05,0.1,0.04]	[0.001,0.001,0.0]
s_{10}	[0.9,0.6,0.4]	[0.0,0.95,0.05]	[0.1,0.1,0.1]

Using Hybridization of Hausdorff & Hamming distance measures and Hybridization of Hausdorff & Euclidean distance measures on the tables 1 and 5 (as illustrated in case study 1), the distance of the students towards x_1, x_4 and x_5 are calculated and they are summarized in table 6 and table 7 respectively.

Table 5. Hybridization of Hausdorff & Hamming distance measures

Hamming Distance	x_1 (Computer Science)	x_4 (Mathematics)	x_5 (Statistic)
s_1	0.16637	0.141324	0.16122
s_2	0.210526	0.168111	0.223301
s_3	0.177519	0.20398	0.402435
s_4	0.182283	0.175258	0.189107
s_5	0.215686	0.180608	0.401033
s_6	0.223301	0.185059	0.404215
s_7	0.230769	0.187817	0.402747
s_8	0.230769	0.23445	0.403653
s_9	0.340659	0.343365	0.404718
s_{10}	0.454545	0.347826	0.32299

Table 6. Hybridization of Hausdorff & Euclidean distance measures

Euclidian Distance	x_1(Computer Science)	x_4(Mathematics)	x_5 (Statistic)
s_1	0.05796	0.063871	0.054504
s_2	0.090645	0.04828	0.093741
s_3	0.059574	0.09732	0.356978
s_4	0.054177	0.053721	0.06565
s_5	0.091938	0.060448	0.354451
s_6	0.089632	0.061675	0.360454
s_7	0.110617	0.067327	0.357518
s_8	0.094258	0.111358	0.359446
s_9	0.227657	0.24884	0.361444
s_{10}	0.391903	0.268739	0.216263

Here U = $\{s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8, s_9, s_{10}\}$ be the set of students, A =$\{x_1, x_2, x_3\}$ be the subjects namely computer science, mathematics and statistic respectively, and Y stand for computer science as the decision variable. For $\alpha \in$ A, define α (s_i), i = 1 to 10 as distance measure. Our decision attribute y = computer science has criteria that for α = x_1, $V_\alpha = t_1$, for α = x_4, $V_\alpha = t_2$ and for α = x_5, $V_\alpha = t_3$ where t_1, t_2 and t_3 are the threshold values of corresponding subjects. The relation is defined as R_1: X → Y if $V_y \le t_c$. Vy is the distance measure from X, t_c is threshold value and if V_y is less than equal to t_c then it is interpreted that eligible for computer science.

First considering for Hybridization of Hausdorff & Hamming distance measures, when we have taken as R_1: X → Y with $V_y \le 0.3$, y as computer science, $V_y \le 0.2$, y as mathematics and $V_y \le 0.4$, y as statistics then the conclusion is as follows.

$\underline{B}X = \{s_1, s_2, s_4\}$

$\overline{\boldsymbol{B}}X = \{s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8, s_{10}\}$

$BN_B(X) = \overline{\boldsymbol{B}}X - \underline{B}X = \{s_3, s_5, s_6, s_7, s_8, s_{10}\}$

B-outside region of, U - $\overline{\boldsymbol{B}}X = \{s_9\}$

μ_X^R: $U \to [0,1]$, where $\mu_X^R(x) = \frac{|X \cap R(x)|}{|R(x)|}$, and x ∈ X. $|R(x)| = 3$

$R_*(X) = \{x \in U | \mu_X^R(x) = 1\} = \{s_1, s_2, s_4\}$ $|X \cap R(x)| = 3$

$R^*(X) = \{x \in U | \mu_X^R(x) = 0\} = \{s_1\}$ $|X \cap R(x)| = 0$

$RN_R(X) = \{x \in U | 0 < \mu_X^R(x) < 1\} = \{s_3, s_5, s_6, s_7, s_8, s_{10}\}$ $|X \cap R(x)| = 1 \; or \; 2$

For $s_3, s_8 \; and \; s_{10}$, we have $|X \cap R(x)| = 1$ and for $s_5, s_6 \; and \; s_7$, $|X \cap R(x)| = 2$. Hence the membership function values for $s_3, s_5, s_6, s_7, s_8, and \; s_{10}$ are as follows.

$\mu_X^R(s_3) = \frac{1}{3}, \mu_X^R(s_5) = \frac{2}{3}, \mu_X^R(s_6) = \frac{2}{3}, \mu_X^R(s_7) = \frac{2}{3}, \mu_X^R(s_8) = \frac{1}{3}$ and $\mu_X^R(s_{10}) = \frac{1}{3}$

Again, considering for Hybridization of Hausdorff & Euclidean distance measures, when we have taken as R1: X →Y with $V_y \le 0.2$ y as computer science, $V_y \le 0.09$, y as mathematics and $V_y \le 0.3$, y as statistics then the conclusion is as follows.

$\underline{B}X = \{s_1, s_2, s_4\}$

$\overline{\boldsymbol{B}}X = \{s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8, s_{10}\}$

$BN_B(X) = \overline{\boldsymbol{B}}X - \underline{B}X = \{s_3, s_5, s_6, s_7, s_8, s_{10}\}$

B-outside region of, U-$\overline{\boldsymbol{B}}X = \{s_9\}$

μ_X^R: $U \rightarrow [0,1]$, where $\mu_X^R(x) = \frac{|X \cap R(x)|}{|R(x)|}$, and x ∈ X. $|R(x)| = 3$

$R_*(X) = \{x \epsilon U | \mu_X^R(x) = 1\} = \{s_1, s_2, s_4\}$ $|X \cap R(x)| = 3$

$R^*(X) = \{x \epsilon U | \mu_X^R(x) = 0\} = \{s_9\}$ $|X \cap R(x)| = 0$

$RN_R(X) = \{x \epsilon U | 0 < \mu_X^R(x) < 1\} = \{s_3, s_5, s_6, s_7, s_8, s_{10}\}$ $|X \cap R(x)| = 1 \; or \; 2$

For $s_3, s_8 \; and \; s_{10}$, we have $|X \cap R(x)| = 1$ and for $s_5, s_6 \; and \; s_7$, $|X \cap R(x)| = 2$. Hence the membership function values for $s_3, s_5, s_6, s_7, s_8, and \; s_{10}$ are as follows.

$\mu_X^R(s_3) = \frac{1}{3}, \mu_X^R(s_5) = \frac{2}{3}, \mu_X^R(s_6) = \frac{2}{3}, \mu_X^R(s_7) = \frac{2}{3}, \mu_X^R(s_8) = \frac{1}{3}$ and $\mu_X^R(s_{10}) = \frac{1}{3}$

Thus, in both distance measure illustration, we have found, for students $s_3, s_5, s_6, s_7, s_8, \; and \; s_{10}$ are in inconsistent situations. We may follow the different rules summarized in section 2.4 for the situations. By removing fewer support cases (defined in section 2.2.4 rule c) and following the membership function values for $s_3, s_5, s_6, s_7, s_8, \; and \; s_{10}$, we are in conclusion that $s_5, s_6 \; and \; s_7$ are eligible for computer science whereas $s_3, s_8 \; and \; s_{10}$ are not. Finally, we are in conclusion $s_1, s_2, s_4, s_5, s_6, and \; s_7$ may choose the stream and eligible for computer science whereas $s_3, s_8, s_9, and \; s_{10}$ are not.

5. Comparative study

Fuzzy set-based approaches have also a good contribution on the students' career selection process (Natividad et al., 2019). In this paper, we have attempted to improve the students' career selection process by incorporating more attributes regarding students' career selection which are represented using PFNs. These greater number of attributes are required to predict a more suitable decision in comparison to proposed fuzzy-based approach given in (Natividad et al., 2019). We have interpreted the attributes using four concepts i.e., the degree of positive potential, the degree of neutral potential, the degree of negative potential and the refusal degree for each attribute. In (Nguyen et al., 2018), authors studied fuzzy linguistic approach for multi-criteria decision making by considering the interest of the student but practically along with interest other factors are also there. The proposed approach has considered the other factors also like subject knowledge, memory, career, attitude and environmental impact. In (Peker et al., 2017), the authors have claimed that the students' prior educational successes and teachers' views are combinedly important to identify the students' professional interests and capacities. In the process, the authors proposed a web-based system, namely WEB-CGS, which is modelled using Mamdani fuzzy model where students' interest are interpreted using the traditional methods of question-answering and evaluation by teachers, which may not be always accurate for selecting a career. In (Nie et al., 2018), the authors have worked on the students' information like skill, regularity, economic status and subject interest, and trained that information using machine learning techniques for future forecasting, but the study does not interpret the skill, regularity, economic and interest exclusively which major an issue for accurate prediction. Our proposed method analyses students' information in terms of attitude, knowledge, interest, career, memory and

environment and expressed that information using PFS with the degree of positive potential, the negative potential, the degree of neutrality and refusal degree for both of the students and subjects. Again, RS is implemented in the proposed approach when confusion arises in choosing stream, whereas none of the mentioned methods used it in the same context.

6. Conclusion

Students are in confusion and feel difficulty in choosing the stream as their career after basic schooling. PFS and RS based approaches are useful in selecting the stream that will be appropriate for a student. The hybridization of Hausdorff & Hamming distance measures and hybridization of Hausdorff & Euclidean distance measures are proposed to find the distance between the student and subjects with the attributes interest, subject knowledge, memory, career, attitude and environment impact. The subject having the minimum distance from the student is chosen as the suitable and appropriate for the student. The rough set theory with lower approximation, higher approximation and boundary region is proposed to find out the inconsistency situations when a particular stream is taken into consideration. Thus, we have proposed two models for choosing a career one for selecting a subject and another for selecting a subject in inconsistency situation and for both two case studies are illustrated. Our proposed models have taken the subjects' attributes values with the degree of positive membership, the degree of neutral membership, and the degree of negative membership when considering PFS. Also, the models have considered the students' attributes values with the degree of positive membership, the degree of neutral membership, and the degree of negative membership when considering PFS. Finding the degree of positive membership, the degree of neutral membership and the degree of negative membership values are also challenging jobs. Hence our future work will focus on to make a model to generate the degree of positive, the degree of neutral, and the degree of negative values when considering PFS. Again, it is possible to extend with adding more attributes that influence the students on choosing streams.

Author Contributions: Each author has participated and contributed sufficiently to take public responsibility for appropriate portions of the content.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflicts of interest.

References

Alkaya, S. A., Yaman, Ş., & Simones, J. (2018). Professional values and career choice of nursing students. Nursing ethics, 25(2), 243-252.

Aspert, N., Santa-Cruz, D., & Ebrahimi, T. (2002) Mesh: Measuring errors between surfaces using the hausdorff distance. In Proceedings. IEEE international conference on multimedia and expo (ICME), Iausanne, Switzerland, August 26-29, 2002 (Vol. 1, pp. 705-708).

Babajide, O. J., & Esther, O. O. (2020). Factors Influencing Career Choice Among Students Of The Schools Of Nursing And Midwifery In Akure. Lautech Journal Of Nursing (LJN), 7(1), 91-96.

Batool, T., & Raiz, J. (2020). Exploring Parents Involvement in University Students Education. Journal of Business and Social Review in Emerging Economies, 6(1), 187-196.

Burns, M. K., Davidson, K., Zaslofsky, A. F., Parker, D. C., & Maki, K. E. (2018). The relationship between acquisition rate for words and working memory, short-term memory, and reading skills: Aptitude-by-treatment or skill-by-treatment interaction?. Assessment for Effective Intervention, 43(3), 182-192.

Carrico, C., Matusovich, H. M., & Paretti, M. C. (2019). A qualitative analysis of career choice pathways of college-oriented rural central Appalachian high school students. Journal of Career Development, 46(2), 94-111.

Cuong, B. C. (2014). Picture fuzzy sets. Journal of Computer Science and Cybernetics, 30(4), 409-420.

Cuong, B. C., & Kreinovich, V. (2013). Picture Fuzzy Sets-a new concept for computational intelligence problems. Proceedings of the Third World Congress on Information and Communication Technologies (WICT), Hanoi, Vietnam, December 15-18, 2013, pp. 1-6.

Cuong, B. C., & Thong, P. H. (2018). Picture Fuzzy Rough Soft Sets: a new concept for Soft computing problems. Proceedings of the 7th International Conference on High Performance Scientific Computing (HPSC), Hanoi, Vietnam, March 19-23, 2018.

Cuong, B. C., & Thong, P. H. (2018). Two new concepts" Picture Fuzzy Rough Soft Sets" and" Picture Fuzzy Dynamic Systems" in Picture Fuzzy Systems. Proceedings of the 5th NAFOSTED Conference on Information and Computer Science (NICS), Ho Chi Minh City, Vietnam, November 23-24, 2018, pp. 87-92.

Das, S., Malakar, D., Kar, S., Pal, T. (2018). A brief review and future outline on decision making using fuzzy soft set, International Journal of Fuzzy System Applications, 7 (2), 1-43.

De, A., Das, S., Kar, S. (2019). Multiple Attribute Decision Making Based on Probabilistic Interval-valued Intuitionistic Hesitant Fuzzy Set and Extended TOPSIS Method. Journal of Intelligent & Fuzzy Systems, 37 (4), 5229-5248.

Dutta, P. (2018). Medical diagnosis based on distance measures between picture fuzzy sets. International Journal of Fuzzy System Applications (IJFSA), 7(4), 15-36.

Goel, S., Angeli, F., Dhirar, N., Singla, N., & Ruwaard, D. (2018). What motivates medical students to select medical studies: a systematic literature review. BMC Medical Education, 18(1), 1-10.

Griffin, B., & Hu, W. (2019). Parental career expectations: effect on medical students' career attitudes over time. Medical education, 53(6), 584-592.

Guido, R. M. D. (2013). Attitude and motivation towards learning physics. International Journal of Engineering Research & Technology (IJERT), 2(11), arXiv:1805.02293.

Hannula, M. S. (2002). Attitude towards mathematics: Emotions, expectations and values. Educational studies in Mathematics, 49(1), 25-46.

Holloway-Friesen, H. (2018). Acculturation, enculturation, gender, and college environment on perceived career barriers among Latino/a college students. Journal of Career Development, 45(2), 117-131.

Kumbhar K., Das S. (2020). Solving Multi-attribute Decision-Making Problems Using Probabilistic Interval-Valued Intuitionistic Hesitant Fuzzy Set and Particle Swarm Optimization. In: Dutta D., Mahanty B. (eds) Numerical Optimization in Engineering and Sciences. Advances in Intelligent Systems and Computing, 979. Springer, Singapore. https://doi.org/10.1007/978-981-15-3215-3_14

Madden, A. D., Webber, S., Ford, N., & Crowder, M. (2018). The relationship between students' subject preferences and their information behaviour. Journal of Documentation, 74(4), 692-721.

McKenzie, K., Murray, A., Cooper, M., Martin, R., Murray, K., Baguley, C., & Chiscop, A. (2020). An exploration of the factors influencing career choice in mental health. Journal of Clinical Nursing, 29(19-20), 3764-3773.

Natividad, M. C. B., Gerardo, B. D., & Medina, R. P. (2019). A fuzzy-based career recommender system for senior high school students in K to 12 education. MS&E, 482(1), 012025.

Nehmeh, G., & Kelly, A. (2018). Women physicists and sociocognitive considerations in career choice and persistence. Journal of Women and Minorities in Science and Engineering 24(2),95–119.

Nguyen, J., Sánchez-Hernández, G., Armisen, A., Agell, N., Rovira, X., & Angulo, C. (2018). A linguistic multi-criteria decision-aiding system to support university career services. Applied Soft Computing, 67, 933-940.

Nie, M., Yang, L., Sun, J., Su, H., Xia, H., Lian, D., & Yan, K. (2018). Advanced forecasting of career choices for college students based on campus big data. Frontiers of Computer Science, 12(3), 494-503.

Orewere, E. O. R. A., & Ojochogu, D. (2020). The Effect of Guidance and Counselling Services on Students Career Choice in Selected Secondary Schools of Jos Metropolis. International Journal of Education and Evaluation, 6(1), 34-42.

Pawlak, Z. (1995). Vagueness and uncertainty: a rough set perspective. Computational intelligence, 11(2), 227-232.

Peker, M., Gürüler, H., Şen, B., & İstanbullu, A. (2017). A new fuzzy logic-based career guidance system: WEB-CGS. Technical Gazette 24, 6(2017), 1863-1868.

Pratiwi, F., Syakurah, R. A., Yuliana, I., & Siburian, R. (2020, June). Relationships of Self-Efficacy, Outcome Expectation, Career Intention and Career Exploration in Nutrition Science Student's Career Choice. In 2nd Sriwijaya International Conference of Public Health (SICPH 2019) (pp. 302-309). Atlantis Press.

Ramentol, E., Madera, J., & Rodríguez, A. (2019). Early detection of possible undergraduate drop out using a new method based on probabilistic rough set theory. In Uncertainty Management with Fuzzy and Rough Sets (pp. 211-232). Springer, Cham.

Si, A., Das, S., Kar, S. (2019). An approach to rank picture fuzzy numbers for decision making problems, Decision Making: Applications in Management and Engineering, 2(2), 54-64.

Si, A., Das, S., Kar, S. (2020). Extension of TOPSIS and VIKOR Method for Decision-Making Problems with Picture Fuzzy Number. In: Mandal J., Mukhopadhyay S. (eds) Proceedings of the Global AI Congress, Kolkata, September 12-14, 2019, Advances in Intelligent Systems and Computing, vol 1112. pp. 563-577.

Silseth, K. (2018). Students' everyday knowledge and experiences as resources in educational dialogues. Instructional Science, 46(2), 291-313.

Son, L. H. (2016). Generalized picture distance measure and applications to picture fuzzy clustering. Applied Soft Computing, 46(C), 284-295.

Tan, A., Wu, W. Z., Qian, Y., Liang, J., Chen, J., & Li, J. (2018). Intuitionistic fuzzy rough set-based granular structures and attribute subset selection. IEEE Transactions on Fuzzy Systems, 27(3), 527-539.

Tugrul, F., Gezercan, M., & Citil, M. (2017). Application of intuitionistic fuzzy set in high school determination via normalized euclidean distance method. Notes on Intuitionistic Fuzzy Sets, 23(1), 42-47.

Van Dinh, N., Thao, N. X., & Chau, N. M. (2019). Distance and dissimilarity measure of picture fuzzy sets. PROCEEDING of Publishing House for Science and Technology, 1(1), 104-109.

Wang, J., Yang, M., Lv, B., Zhang, F., Zheng, Y., & Sun, Y. (2020). Influencing Factors of 10th Grade Students' Science Career Expectations: A Structural Equation Model. Journal of Baltic Science Education, 19(4), 675-686.

Wen, L., Yang, H., Bu, D., Diers, L., & Wang, H. (2018). Public accounting vs private accounting, career choice of accounting students in China. Journal of Accounting in Emerging Economies, 8(1), 124-140.

Yu, B., Cai, M., Dai, J., & Li, Q. (2020). A novel approach to predictive analysis using attribute-oriented rough fuzzy sets. Expert Systems with Applications, 161, 113644.