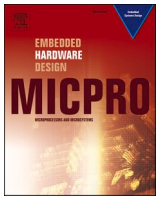
Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Microprocessors and Microsystems 97 (2023) 104758

Contents lists available at ScienceDirect

Microprocessors and Microsystems

journal homepage: www.elsevier.com/locate/micpro



A novel technique for the detection of Covid-19 patients with the applications of three-way decisions using variance-based criterion

K M Remesh, Latha R Nair *

Division of Computer Science and Engineering, School of Engineering, CUSAT, Kochi, India

ARTICLE INFO

Keywords:
Three-way decisions
Conditional probability
Optimal pair of thresholds
Probabilistic rough sets
Within region variance
Between region variance

ABSTRACT

Everyone is making constant efforts to establish an effective diagnostic approach, therapy and control of the spread of the pandemic. Due to a flexible formulation, the parameters prior to the normal distributions and explicitly formulate assumptions on the transition probabilities between these categories over time. The spread of the COVID-19 pandemic represents a serious threat for scientists and academics, health professionals and even governments today. The Hospital wards are classified into Intensive Care Unit (ICU), Regular Wards (RW) with Recovered (R) and Deceased (D).. The formulation may be truncated to include particular hypotheses with an epidemiological interpretation. The principles of Three-Way Decision Theory could be used to anticipate and diagnose COVID-19 patients were classified into one of three zones based on their symptoms: Positive, Negative, or Boundary, and treatment are recommended if necessary. The thresholds that distinguish the three zones are determined using a variance-based criterion. Examine the impact of nonpharmaceutical interventions and the findings from data gathered during the second wave of the pandemic in Trivandrum, India.The Three-Way Decision Theory model has a good fit and gives good predictive performance, especially for RW and ICU patients, according to suitable discrepancy metrics that were created to assess and compare models. 95 percent accuracy increased and calculated values for 10 days to demonstrate the temporal aspects of the expected daily reproduction number R

1. Introduction

Researchers and scientists were attempting to expand an effective diagnosis mechanism and develop a vaccine to combat the COVID-19 pandemic's rapid and unpredictable expansion. By broadening their technological interpretations and realistic solutions to Machine Learning, Data Science, and Artificial Intelligence (AI) to control and prevent the COVID-19 outbreak. Deep learning has made significant improvements in the medical industry in recent years, with promising results and improved accuracy [1].

The information in real-world data is scant and ambiguous. Insufficient and uncertain information can be effectively dealt with using Z Pawlak's rough set theory, which he presented in the early 1990s. It can be considered an in-depth integrated management technique in a large variety of AI systems [2]. When there is adequate proof and knowledge, definitive judgments are taken. In the absence of it, cautious decisions are made by avoiding errors caused by wrong acceptance and rejection [3]. When deciding on an uncertain condition, it's common to choose a criterion as a balance between errors and judgments. Uncertainty is resolved by applying a probability function to a decision, such as agree or disagree. A threshold is used in a bidirectional probability evaluation method to make one of the two decisions, accept or reject.

There is no tolerance for inaccuracy with Pawlak rough set approximations. In other words, an object in the complement of the upper approximation is certain not to belong to the set while an object in the lower approximation is assured to belong to the set. Rough set generalisations have been introduced in a variety of ways as a result of this. Probabilistic rough sets are quantitative expansions of Pawlak rough sets, which are qualitative rough sets with specified levels of error introduced [4,5,6]. Inaccuracy in the lower and upper approximations, or alternatively in the probabilistically positive, boundary, and negative regions, is tolerable in probabilistic rough set models. Based on the ideas of rough membership functions and rough inclusion, probabilistic rough set approximations can be developed . Both ideas can be explained in terms of a posteriori or conditional probability. To create probabilistic or parameterized approximations, threshold values—also known as parameters—are applied to a rough membership function or a rough inclusion.

* Corresponding author.
E-mail address: latharnair482@gmail.com (L.R. Nair).

https://doi.org/10.1016/j.micpro.2023.104758
Received 7 May 2022; Received in revised form 30 October 2022; Accepted 2 January 2023
Available online 4 January 2023

Although some authors had studied a number of different specific models of probabilistic rough sets, Yao et al. were the first to put forth a more general model based on Bayesian decision theory known as the decision-theoretic rough set (DTRS) model. Variable precision rough sets, Bayesian rough sets, parameterized rough sets, game-theoretic rough sets, variable-precision rough sets, stochastic dominance-based rough sets, Naive Bayesian rough sets, information-theoretic rough sets, confirmation-theoretic rough sets, and numerous other varieties of probabilistic rough set approximations are among the other probabilistic models [4,5].

Within a three-way probabilistic decision model, there are three decision areas. The three-way decision making method, an additional zone of choice, the region without commitment, is added to the two decision-making zones [7] . In a bidirectional decision-making model, a unique threshold crosses the universe. Considering that in a three-way decision process, a group of criteria (α,β) such that $0 \leq \beta < \alpha \leq 1$ is expended for sorting the world. If the probability of the hypothesis was less than the threshold then the negative region is chosen. If the probability is between the two thresholds then the boundary region or no-commitment area is chosen 8]. If the likelihood is above the criteria, the positive area was selected. Selecting the best pair of thresholds minimizes the number of errors due to inappropriate acceptance and rejection. An evaluation function that reduces or maximizes finding the best groups of criteria [9]. These assessment functions may be used to quantify the cost, quality, and other factors [7]. In this study, variance-based criteria are utilized to determine thresholds and trisect the decision space for detection and remedial activities using a data set of COVID-19 symptoms.

Recently, the theory of three-way decision based on probabilistic rough set over two universes has been employed for conflict analysis because probabilistic rough sets can be used to incorporate the structure and characteristics of a conflict situation. Three distinct attitudes—favorable, opposed, and neutral—represent the objects (agents), disputes (problems), and the interaction between the objects and disputes, respectively [8]. For any business in the financial industry, risk management includes the critical responsibility of credit scoring. The application of three-way decisions using probabilistic rough sets to credit scoring, which reduces decision-relevant costs, is also found. One of the more recent sentiment analysis jobs is predictive recommendations based on internet reviews, which categorises the emotion represented in online reviews as recommendations, not recommendations, or cannot be predicted. The theory of three-way decisions with probabilistic rough sets are used in predictive analysis.

Algorithm 1

Table 1: Optimal_Thresholds_by_Variance_Based_Criteria.

Input: A set of training examples U and the target concept C ⊆ UStep_size by which α and β are to be modifiedOutput: A pair of optimal thresholds (α, β)Begin :

1 Identify equivalent classes of U

2 For every equivalent class i, repeat steps 3 to 4

3 Determine the probability of the equivalent classes :

$$\Pr([x]_i) = \frac{\textit{Number of elements in the class}}{\textit{Total number of elements in the data set}}$$

1 Determine the conditional probabilities of the class :

$$\Pr(C/[x]_i) = \frac{\textit{Number of elements in the class that belongs to the concept}}{\textit{Total number of elements in the class}}$$

1 Sort $\Pr([x]_i)$, $\Pr(C|[x]_i)$ *and* $\Pr(C^c|[x]_i)$ in descending order of$\Pr(C|[x]_i)$

2 Initialize $\alpha = 1, \beta = 0, \max = 0$;

3 Do

4 Compute the probabilities of the three regions; $POS_{(\alpha,\beta)}(C)$, $NEG_{(\alpha,\beta)}(C)$, $BND_{(\alpha,\beta)}(C)$

5 Compute the mean conditional probability of the three regions

6 Compute the within region variance of the three regions

7 Compute the overall within region variance of the three regions

8 Compute the mean conditional probability of all equivalence classes

9 Compute the between region variance of the three regions

10 Compute the overall between region variance of the three regions

11 Compute the discriminant criterion:

12 If (Discrimenent criterion > max)

max = Discrimenent Criterion$\alpha' = \alpha\beta' = \beta$ end

1 Modify $\alpha = \alpha - step\_size, \beta = \beta + step\_size$

2 **While ((β< 0.5) and (α ≥ 0.5))**

3 Return (α' , β')

End.

2. Related study

Deep Neural Networks (DNN) were used to create a way for identifying COVID-19 patients as a substitute to the good screening method to analyze their X-rays [10]. In their study [11], X-ray images to construct a deep characteristic by Support Vector Machine (SVM) to detect COVID-19 infection in patients. Deep features were extracted from 13 pre-trained CNN models and sent to an SVM classifier to classify. In the image classification method, the SVM-based methodology achieved 93.4 percent accuracy. Rough Set Theory [12] is concerned with knowledge that is imperfect, unclear, imprecise, inconsistent, and uncertain. The indiscernibility relation is a fundamental idea in Rough Set Theory. When the knowledge described by two items is the same, they are said to be indiscernible [13].

Game-Theoretic Rough-Sets to analyze the connection across modification in probability criteria and the impact on the Gini coefficient of rough group areas in his work [14]. A competition game across areas was designed, and an interactive educational process was used to reach effective criteria. In their study [15] presented a broad framework for calculating the suitable group of criteria as an optimization issue. A decision-theoretic rough sets method was established, and a data-theoretic method was proposed that employs Shannon Entropy as a result of variation and determines Shannon entropy minimum, and the pair of thresholds obtained having considered the optimal pair of thresholds. A heuristic technique for detecting the optimal pair of thresholds is also devised using the gradient-descent method. In their paper [6] provided an information-theoretic approach for determining a pair of thresholds utilizing Shannon entropy as a result of the variation. A task to finding the best pair of criteria was defined as minimizing the Shannon entropy of the three areas. In their study [16], use games between multiple criteria of equivalence classes to decide whether or not to include the equivalence class in a region. Game-theoretic roughest models are used to tackle challenges that arise when rough-group-based three-way planning was generalized to many parameter planning. An objective function is visualized with a Chi-square statistic that assesses the goodness of a trisection [17]. The trisection and classification have a strong association when the Chi-square statistic is at its maximum value. In their paper [18], proposed variance-based criteria as an assessment function for identifying the best group of criteria in three-way options. For several pairs of thresholds, different-based criteria - within areas different, between areas deviation - are established, as well as the pair with the largest discriminant criterion is the best group of criteria.

In the early 1980s developed a rough group theory to examine data stored in an information table [19]. It's a mathematical method for working with hazy, inaccurate, inconsistent, and ambiguous data. Extended likelihood notions to the lowest and highest estimate of rough group theory in 1987 to overcome the error tolerance of traditional rough set theory [20]. A concept to the idea of a three-way option arose from the representation of a set with three areas in a rough set assumption. The main premise of this theory is that the cosmos is separated into three groups-wise disjoint areas: negative, positive, and border areas [21]. Each object in the universe is evaluated by an objective function, which compares its value to the thresholds and assigns it to the proper zone. The object is in the positive region the value was higher than or comparable to the upper criteria; the object was in the negative region the score was smaller than or comparable to the lower criteria; otherwise, it is in the boundary region. The most difficult aspect of three-way decisions is in interpreting and identifying the objective functions and threshold values [22–29].

3. Proposed method

3.1. Probabilistic rough sets

To some extent, probabilistic rough sets allow overlap across equivalence classes of Pawlek classical rough set. Given that the object x is in [x] and [x] was the correlation category, the transition probability P (C/[x]) quantifies an evaluation of items to a concept C. The items were in the negative zone if the transition probability of the notion was less than or comparable to a certain threshold. The object is in the positive zone if it was highest than or comparable to another threshold; otherwise, the object is in border areas if it is between the two thresholds. The following are the three rough set regions expressed in Eqs. 1-3:

$$\mathrm{POS}_{(\alpha,\beta)}(\mathrm{C}) = \{\mathrm{x} \in \mathrm{U} | \mathrm{P}(\mathrm{C}/[\mathrm{x}]) \geq \alpha\} \tag{1}$$

$$\mathrm{NEG}_{(\alpha,\beta)}(\mathrm{C}) = \{\mathrm{x} \in \mathrm{U} | \mathrm{P}(\mathrm{C}/[\mathrm{x}]) \leq \beta\} \tag{2}$$

$$\mathrm{BND}_{(\alpha,\beta)}(\mathrm{C}) = \{\mathrm{x} \in \mathrm{U} | \beta < \mathrm{P}(\mathrm{C}/[\mathrm{x}]) < \alpha\} \tag{3}$$

Where α, β are the lower and upper thresholds, and POS, BND, and NEG represent the negative, positive, and border areas correspondingly. U is the universal set, which is non-empty and limited.

3.2. Quality of thresholds

A formulation of an objective function is used to determine thresholds. To discover the best threshold values, the optimization algorithm is maximized or minimized explained in Algorithm 1. The quality of the trisection is assessed using objective functions. The properties of objective functions that can be utilized to determine optimal thresholds include cost, Gini index, information entropy, accuracy, precision, and so on shown in Table 1.

Let Q(α, β) denote the overall performance of the three areas' thresholds, and let Q_{POS} (α, β,), Q_{NEG} (α, β), and Q_{BND} (α, β) denote the positive, negative, and border areas' quality, correspondingly. The following formula could be used to calculate the overall grade of the three areas in Eq. (4):

$$\mathrm{Q}(\alpha,\beta) = \mathrm{w1}\ \mathrm{Q}_{\mathrm{POS}}(\alpha,\beta) + \mathrm{w2}\ \mathrm{Q}_{\mathrm{NEG}}(\alpha,\beta) + \mathrm{w3}\ \mathrm{Q}_{\mathrm{BND}}(\alpha,\beta) \tag{4}$$

where w1, w2, and w3 were weights assigned to positive, negative, and border areas, respectively, to signify their relative importance. As a result, the problem of identifying the optimal combination of criteria was framed as an optimization process.

$$\underset{(\alpha,\beta)}{argmin} Q(\alpha,\beta) \quad or \quad \underset{(\alpha,\beta)}{argmax} Q(\alpha,\beta) \tag{5}$$

3.3. The variance of conditional probabilities as a measure of uncertainty

Identify the best criteria, entropy, Gini index, and Chi-square statistics were utilized as measures of uncertainty. The spread in conditional probabilities among data, known as variance, can be used as a criterion for finding the best thresholds. To find the best pair of thresholds, two kinds of differences are examined: within-region different and between-region different.

3.3.1. Within region variance

In a probabilistic framework, let P(C|[x]) be the likelihood probability of a region separated by the thresholds (α, β). The probability of the positive areas was calculated as Eq. (6):

$$P\left(POS_{(\alpha,\beta)}(C)\right) = \frac{|\ POS_{(\alpha,\beta)}(C)\ |}{|\ U\ |} \tag{6}$$

Negative and boundary region probability is computed in the same way.

The average Likelihood probability of the positive areas was the average transition probability score of the equivalent categories to the positive areas and was calculated as Eq. (7):

$$\mu_{POS_{(\alpha,\beta)}(c)} = \sum_{\forall [x] \in POS_{(\alpha,\beta)}(C)} \frac{P(C\,|[x])\ \times P([x])}{P\left(POS_{(\alpha,\beta)}(C)\right)} \tag{7}$$

where the probability of the equivalent class P([x]) is computed as Eq. (8):

$$P([x]) = \frac{|\ [x]|}{|\ U\ |} \tag{8}$$

The mean conditional probabilities of negative and boundary regions are calculated in this way. There could be several such equivalent classes, and the average likelihood probability of all the equivalent categories was calculated as follows in Eq. (9):

$$\mu = \sum_{\forall [x] \in U} P(C|[x])\ \times P([x]) \tag{9}$$

The difference in an area is the distribution of its conditional probability. The variance to a region, say the positive region, can now be calculated as follows in Eq. (10):

$$\sigma^2_{POS_{(\alpha,\beta)}(C)} = \sum_{\forall [x] \in POS_{(\alpha,\beta)}(C)} \frac{\left(P\left(C\middle|[x] - \mu_{POS_{(\alpha,\beta)}(c)}\right)^2 \times P([x])\right)}{P\left(POS_{(\alpha,\beta)}(C)\right)} \tag{10}$$

Negative and boundary region variances are computed in the same way.

As a result, the three areas' within-region variance is as follows in Eqs. (11)-13:

$$\sigma^2_{W_{POS}}(\alpha,\beta) = \sigma^2_{POS_{(\alpha,\beta)}(C)} \tag{11}$$

$$\sigma^2_{W_{NEG}}(\alpha,\beta) = \sigma^2_{NEG_{(\alpha,\beta)}(C)} \tag{12}$$

$$\sigma^2_{W_{BND}}(\alpha,\beta) = \sigma^2_{BND_{(\alpha,\beta)}(C)} \tag{13}$$

These probabilities can be used as the weights for the determination of the overall quality

As a result, the three areas' within-region variance is as follows in Eq. (14):

$$\begin{aligned}\sigma^2_w(\alpha,\beta) = {} & P\left(POS_{(\alpha,\beta)}(C)\right) \times \sigma^2_{W_{POS}}(\alpha,\beta) + P\left(NEG_{(\alpha,\beta)}(C)\right) \\ & \times \sigma^2_{W_{NEG}}(\alpha,\beta) + P\left(BND_{(\alpha,\beta)}(C)\right) \times \sigma^2_{W_{BND}}(\alpha,\beta)\end{aligned} \tag{14}$$

The total within-region entropy is a measure of the contractual probability's total distribution throughout each location. In each region, the conditional probabilities' overall spread is measured by the equation above. Reduced fluctuation in conditional probability values will result in compact regions containing equivalence classes [7]. Within-region variance with a minimum value indicates that equivalence classes within each region have less fluctuation in conditional probability values. As a result, in this example, the optimization problem is in Eq. (15):

$$\underset{(\alpha,\beta)}{arg\ min}\ \sigma^2_w(\alpha,\beta) \tag{15}$$

3.3.2. Between region variance

The difference between the region-mean of that region and the overall mean helps determine the between-region variance of that region. The three regions' between-region variances are calculated as follows in Eqs. 16-18:

$$\sigma^2_{B_{POS}}(\alpha,\beta) = \left(\mu_{POS_{(\alpha,\beta)}(c)} - \mu\right)^2 \tag{16}$$

$$\sigma^2_{B_{NEG}}(\alpha,\beta) = \left(\mu_{NEG_{(\alpha,\beta)}(c)} - \mu\right)^2 \mathrm{v} \tag{17}$$

$$\sigma^2_{B_{BND}}(\alpha,\beta) = \left(\mu_{BND_{(\alpha,\beta)}(c)} - \mu\right)^2 \quad (18)$$

As a result, the overall between-region variance is calculated as follows in Eq. (19):

$$\sigma^2_B(\alpha,\beta) = P\left(POS_{(\alpha,\beta)}(C)\right) \times \sigma^2_{B_{POS}}(\alpha,\beta) + P\left(NEG_{(\alpha,\beta)}(C)\right) \times \sigma^2_{B_{NEG}}(\alpha,\beta) + P\left(BND_{(\alpha,\beta)}(C)\right) \times \sigma^2_{B_{BND}}(\alpha,\beta) \quad (19)$$

As with the computation of within-region variance, the probabilities of the three areas are formed as weights. The stretch of the areas means the overall average or the means of the total population is referred to as between-region variance. The aforementioned equation depicts how the region's means are distributed overall in relation to the population's overall mean (or global mean). The regions in conditional probability will be clearly demarcated and discernible as this is maximized [11]. A maximum value of the between-region variance means the regions are well separated or distinguishable and thus, the optimization criterion is in Eq. (20):

$$\underset{(\alpha,\beta)}{argmax}\ \sigma^2_B(\alpha,\beta) \quad (20)$$

As shown below, the rate of between-region variance to within-areas differences produces a discriminant criterion that can be used to determine optimal thresholds in Eq. (21).

$$\underset{(\alpha,\beta)}{arg\ max}\frac{\sigma^2_B(\alpha,\beta)}{\sigma^2_w(\alpha,\beta)} \quad (21)$$

Since the thresholds (α, β) determine the regions, they also affect the regions' mean and variance. The division between the positive-boundary is controlled by threshold α, and the division between the negative-boundary is controlled by β, taking the normal condition of $0 \le \beta < 0.5 \le \alpha \le 1$ into consideration. As a result, the threshold affects both the mean and variation of the positive zone as well as the mean and variance of the negative region[7].

This work proposes an algorithm for determining the optimal groups of criteria based on variance; between-region variance, and within-region variance. The proposed algorithm is given with a set of training examples and the target concept and the output of the algorithm is a pair of optimal thresholds. Initially the values of α and β are 1 and 0. By adjusting α and β by a step size, this approach establishes the discriminant criterion to different groups of criteria as well as identifies the couple to the highest ratio as the best group of criteria such that $0 \le \beta < \alpha \le 1$. The algorithm calculates mean conditional probability, within region variance, overall within region variance, between region variance, and overall between region variance of all three regions from the probabilities of the three regions. It also computes discremenent criterion and updates to the maximum value through iterations. The values of α and β are updated until the values reach 0.5.

COVID-19 patients are increasing at an exponential rate over the world, putting a strain on healthcare systems. It's impractical to undertake a traditional RT-PCR test on every patient with modest symptoms like fever, sore throat, runny nose, and so on unless the symptoms indicate a higher risk of COVID-19 infection. There are only a few diagnostic kits available, as well as a limited number of hospital beds. COVID-19 infection can be detected based on symptoms, which can aid in the quarantining of suspected patients until test results are available [13]. Patients with COVID-19 symptoms will be partitioned into three zones based on their severity using an appropriate pair of criteria, allowing the government and health care systems to more precisely identify patients who require immediate attention, patients who are not severe, and so on. As a result, effective use of available resources such as RT-PCR and PPE kits [13].

Experiments were carried out using the algorithm described above with 5 different but identical Covid-19 datasets retrieved from Github. All of the datasets' experimental outcomes were outstanding and informative. The results obtained were comparable, and a sample of them is included in this report for analysis. The dataset used to examine the outcomes contains symptoms from around 2500 COVID-19 patients from throughout the world. Fever, Body pain, Sore Throat, Nasal congestion, Severity, Runny Nose, Contact with Covid Patient, Difficulty Breathing, as well as the decision attribute Infected are among the 9 attributes in the dataset. The discriminant criterion for various (α, β) was determined and presented in Table 1 as follows:

Table 1
Discriminant Criteria ratio.

β → α ↓	0.00	0.10	0.20	0.30	0.40
0.60	31.29	33.30	39.80	36.26	32.26
0.70	28.94	32.58	37.81	39.41	33.40
0.80	18.22	21.30	31.00	40.76	38.01
0.90	12.35	14.44	21.91	30.20	29.40
1.00	7.38	9.09	16.12	23.32	24.82

Table 2
The upper and lower boundaries of the calculated retroactive 95 percent forecast.

	S	R	Q	H	ICU	D
S	–	(0,0)	(1218, 3189)	(0, 721)	(0, 3)	(0, 0)
R	–	(62 471, 62 501)	(0, 27)	(0, 0)	(0, 0)	(0, 0)
Q	–	(1271, 4359)	(78, 183, 84, 680)	(33, 1578)	(0, 0)	(0, 0)
EW	–	(0, 509)	(464, 1131)	(22 439, 22 323)	(27, 139)	(122, 284)
ICU	–	(0, 0)	(0, 0)	(0, 0)	(1964, 2076)	(99, 212)
D	–	–	–	–	–	–

Table 1 shows that the discriminant criterion has a maximum value of 40.76 for $\alpha = 0.80$ and $\beta = 0.30$. As a result, the couple ($\alpha = 0.80$, $\beta = 0.30$) is the efficient couple of criteria. According to table 2, if the discriminant criterion value is less than 0.3, the patient is in the negative region and COVID-19 is not infected, if it is between 0.30 and 0.80, the patient is in the COVID-19 suspected region or boundary region, and if it is greater than 0.80, the patient is COVID-19 infected or in the positive region. Out of 2499 samples, there are only 237 cases in the boundary region, with 1167 negative cases and 1095 positive cases. In the positive and negative zones, there are only 30 false-positive cases and 42 false-negative cases, respectively, and the accuracy is 98 percent for both positive and negative cases combined. The specificity is 0.964 and the recall is 0.972.

The results obtained were compared with the results of other two algorithms that uses Shannon Entropy and Chi-square statistic as the objective functions. The values of thresholds obtained in both the cases were alike with the result obtained using this algorithm.

4. Results and discussion

The calculated 95 percent forecast ranges are shown in gray Fig. 1 along with the daily reported and forecasted numbers for each classification with a time frame of 10 days. The expected transformation probabilities for the 25th and 26th of April, as contained in the transformation matrix, are given using their posterior averages. Table 2 lists the appropriate 95 percent upper and lower projected boundaries. Some people were anticipated to go from Q & H (516) and from H to ICU (73). With a confidence range ranging from 123 to 282, it should be predicted that 149 deaths among ICU patients were projected on April 26 opposed to April 25. These projections suggest a median ICU stay of 10 to 22 days. Another interesting observation is that on the same day of 26 April, 73 hospitalized patients (0.33%) are predicted to require intensive care with a credibility interval of 25 to 137 patients (0.11%, 0.62%). On the other hand, 197 hospitalized patients are predicted to die on the same

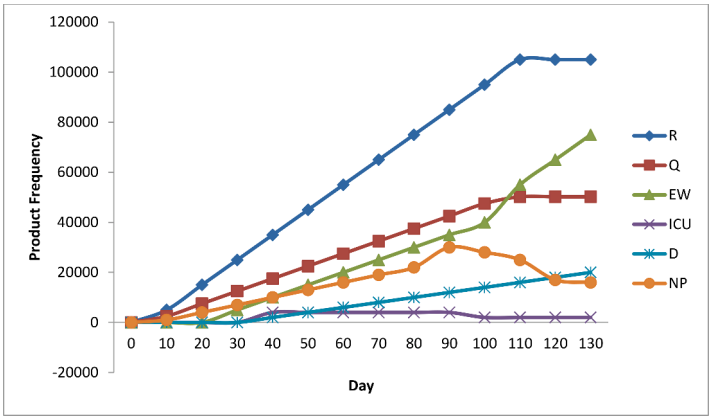


Fig. 1. Forecasts for the next 10 days and verified occurrences (before the vertical dotted lines corresponds to the 25th of April).

Table 3
95 percent forecast ranges and calculated posterior averages.

Day	EW	PI	ICU	PI
26th June	-475	(-1052, 444)	-77	(-142, -3)
27th June	-469	(-1035, 462)	-74	(-135, -3)
28th June	-462	(-998, 453)	-70	(-129, -2)
29th June	-453	(-974, 452)	-64	(-119, 0)
30th June	-444	(-954, 466)	-59	(-109, 4)
1st July	-436	(-949, 452)	-56	(-106, 6)
2nd July	-398	(-944, 486)	-50	(-101, 7)
3rd July	-371	(-927, 455)	-48	(-97, 5)

day. The estimated posterior means and the 95% predicted interval for the increase in totals for RW and ICU from 26th to 29th April are reported in Table 3. These are of particular interest, since during the first period of the pandemic, there was a significant daily increase in the demand for hospital beds and ICU, especially from the general population at risk of being affected by the virus. Fig. 2 illustrates the calculated values, 95 percent acceptable ranges, and calculated values for 10 days to demonstrate the temporal aspects of the expected daily reproduction number Rt.

4.1. Validation

Designers present findings from the estimation of the proposed models using Trivandrum, India-related information. The findings from Framework are shown in the sections that follow. Model is appropriate for use with exploration information since no restrictions were placed on the probabilities of category changes. Table 4 provides the posterior averages of the projected change probabilities pertaining to the 25th and 26th of April contained in the transition matrix. This demonstrates that the recommended limitations, as previously indicated, are appropriate to meet the epidemiological characteristics of the pandemic. Table 5 reports the calculated retrospective average and the 95% expected CI for the rise in H and ICU aggregates from the 26th to the 29th of April. The largest change between this figure and Table 4 may be seen in the

Table 4
Predicted transfers between groups for the 25th and 26th of April, calculated posterior averages.

	S	R	Q	EW	ICU	D
S	62 122 108	1247	1642	24	4	0
R	0	58 108	2279	0	44	72
Q	0	4 482	78 780	1677	22	333
EW	0	1157	1233	19 619	3	68
ICU	0	2	141	3	2032	1
D	0	0	0	0	0	26 967

Table 5
Calculated subsequent averages and 95% probability ranges .

Day	EW	PI	ICU	PI
26th June	-755	(-1498, -66)	-77	(-135, 12)
27th June	-900	(-1572, -83)	-72	(-129, 24)
28th June	-948	(-1628, 118)	-64	(-129, 39)
29th June	-989	(-1825, -166)	-56	(-127, 58)
30th June	-953	(-1968, -215)	-48	(-133, 105)
1st July	-1152	(-2025, 230)	-35	(-144, 200)
2nd July	-1179	(-2044, -245)	-18	(-156, 260)
3rd July	-1198	(-2044, 251)	-4	(-160, 353)

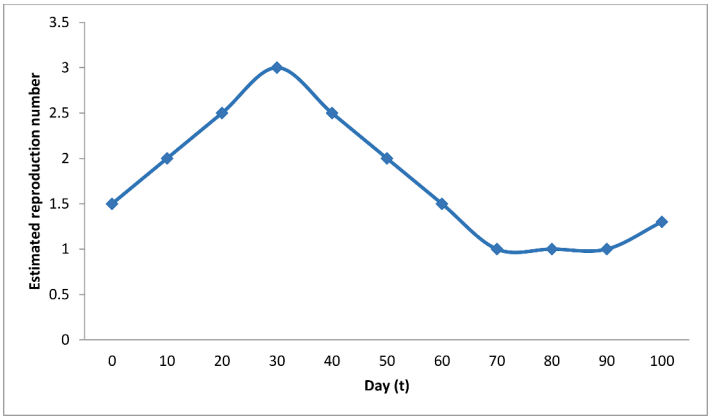


Fig. 2. Calculated or anticipated replication number Rt.

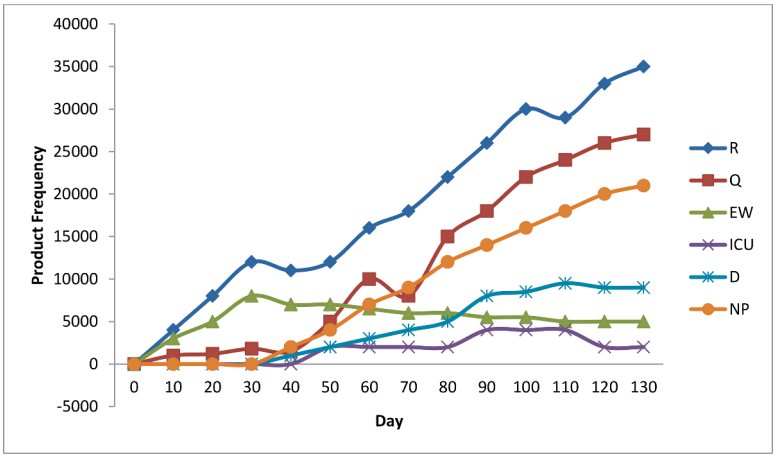


Fig. 3. Estimates for the next five days and reported occurrences (before the vertical dotted lines matching to April 25).

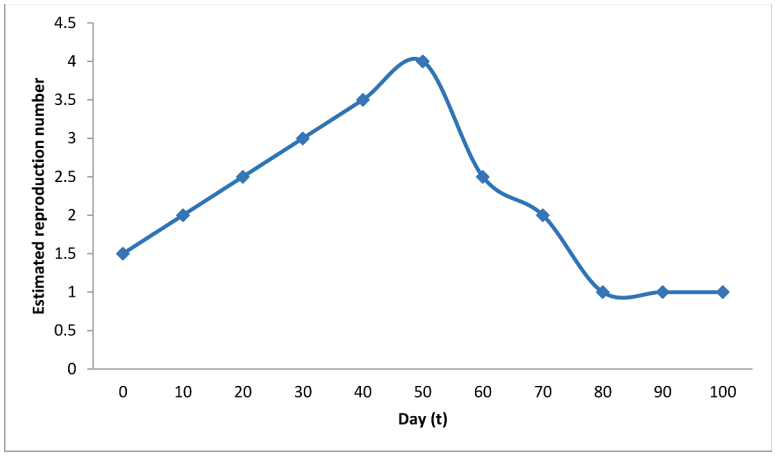


Fig. 4. Calculated and verified replication number Rt for Trivandrum (from the vertical line).

estimated rates for classification H, which are wider than anticipated. The calculated 95 percent forecast ranges are shown in gray in Fig. 3 along with the daily reported and forecasted numbers with a time horizon of 10 days.

A dynamics of Rt as determined by the calculated framework are shown in Fig. 4. During the initial phases of the pandemic, researchers discovered that, on aggregate, this value rises over duration. A few days after the NPIs were issued, on the 11th day, or the 5th of March, it starts to decline.

It can be employed in a variety of other circumstances where the presumption regarding the order of contingency tables of the "transformation rates" between two successive time events was acceptable, COVID-19 context. For instance, when performed on customer data, they could be utilised in the study of the transformations across classifications of malignant tumors, such as in the tumor, node, metastasis categorization, or for the assessment of the transformations between levels of intensity of different diseases. This approach could be helpful in understanding the dynamics of the deceased, which vary significantly among Italian areas. Researchers are optimistic that our idea would improve the platform's aggressive public health campaign planning and help people avoid the onset of serious sickness.

5. Conclusion

Due to the rapid spread of COVID-19, health workers and scientists are concerned about how to rapidly and accurately diagnose infected people based on their symptoms and recommend isolation or therapy. COVID-19 positive samples, negative samples, and abstain or doubtful cases were categorized as COVID-19 positive samples, negative cases, abstain or questionable cases using optimal pair thresholds. In three-way decision theory, the variance-based discriminant criterion can utilize an evaluation process to select optimal couples of criteria. COVID-19 patients are efficiently classified into three classes using the algorithm provided in this research. Early need detection is important for policy makers and doctors. Our proposed model used to estimate the daily reproduction number after epidemiological hypotheses are incorporated and it evaluated using more conventional SEIR models. It is more exploratory when compared to SEIR models because it needs only fewer presumptions and hypotheses. It set minimal constraints for knowledge-driven that estimates more consistent. The proposed model estimated using data at the regional level, allowing for a comparison of transition rates between categories between areas and a the time spent.

Declaration of Competing Interest

The authors declare no conflict of interest.

Data availability

No data was used for the research described in the article.

References

[1] Harsh Panwar, P.K. Gupta, Mohammad Khubeb Siddiqui, Ruben Morales-Menendez, Vaishnavi Singh, Application of Deep Learning For Fast Detection of COVID-19 in X-Rays Using Ncovnet, Elsevier Ltd., 2020, https://doi.org/10.1016/j.chaos.2020.109944, 0960-0779/©.

[2] Yuhua Qian, Jiye Liang, Yiyu Yao, Chuangyin Dang, MGRS: a multi-granulation rough set, Inf. Sci. (Ny) 180 (2010) 949–970. ELSEVIER.

[3] Xiaofei Deng, Yiyu Yao, A multifaceted analysis of probabilistic three-way decisions, Fundam. Inform. 132 (2014) 291–313, https://doi.org/10.3233/FI-2014-1045, 291IOS press.

[4] Xiaofei Deng, Three-way Classification Models, University of Regina, 2015. PhD Thesis.

[5] K.M. Remesh, Latha R Nair, Determination of optimal pair of thresholds in three-way decisions using objective functions, Int. J. Adv. Trends in Comput. Sci. Eng. (2020), https://doi.org/10.30534/ijatcse/2020/205942020.

[6] Yiyu Yao, An outline of a theory of three-way decisions, in: RSCTC 2012 LNCS (LNAI), 7413, Springer, Heidelberg, 2012, pp. 1–17.

[7] Cong Gao, Yiyu Yao, Determining thresholds in three-waydecisions with chi-square statistic. Springer international publishing AG 2016V. Flores et al. (Eds.), IJCRS (2016) 272–281, https://doi.org/10.1007/978-3-319-47160-0 25. LNAI 99202016.

[8] Bingzhen Sun, Xiangstang Chen, Liye Zhang, Weimin Ma, Three-way decision making approach to conflict analysis and resolution using probabilistic rough set over two universes, Inf. Sci. (Ny) 506 (2020) 809–822. Elsevier.

[9] Yu Fang, Cong Gao, Yiyu Yao, Granularity-driven Sequential three-way decisions: a cost-sensitive approach to classification, Inf. Sci. (Ny) 507 (2020) 644–664. Elsevier.

[10] Bing Zhou, Yiyu Yao, Feature selection based on confirmation-theoretic rough sets, RSCTC (2014) 181–188. LNAI 8536Springer Switzerland 2014.

[11] Nouman Azam, Jing Tao Yao, Variance based determination of three-way decisions using probabilistic rough sets. Springer international publishing AG 2016V. Flores et al. (Eds.), IJCRS (2016) 209–218, https://doi.org/10.1007/978-3-319-47160-0 19. LNAI 99202016.

[12] Nouman Azam, JingTao Yao, Analyzing uncertainties of probabilistic rough set regions with game-theoretic rough sets, Int. J. Approximate Reasoning 55 (2014) 142–155. ELSEVIER.

[13] T.P. Latchoumi, L. Parthiban, Quasi oppositional dragonfly algorithm for load balancing in a cloud computing environment, Wireless Personal Commun. (2021) 1–18.

[14] P.R. Garikapati, K. Balamurugan, T.P. Latchoumi, G. Shankar, A quantitative study of small dataset machining by agglomerative hierarchical cluster and K-medoid. Emergent Converging Technologies and Biomedical Systems, Springer, Singapore, 2022, pp. 717–727.

[15] Z. Pawlak, Rough set theory and its applications, J. Telecommun. Inf. Technol. (2002). March.

[16] Qinghua Zhang, Qin Xie, Guoyin Wang a, a survey on rough set theory and its applications, ScienceDirect, CAAI Trans. Intelligence Technol. 1 (2016), https://doi.org/10.1016/j.trit.2016.11.001, 323e333.

[17] Prabira Kumar Sethy, Santi Kumari Behera, Pradyumna Kumar Ratha, Preesat Biswas, Detection of coronavirus disease (COVID-19) based on deep features and support vector machine, Int. J. Math. Eng. Manag. Sci. Vol. 5 (4) (2020) 643–651, https://doi.org/10.33889/IJMEMS.2020.5.4.052.

[18] T.P. Latchoumi, R. Swathi, P. Vidyasri, K. Balamurugan, Develop new algorithm to improve safety on WMSN in health disease monitoring, in: 2022 International Mobile and Embedded Technology Conference (MECON), IEEE, 2022, pp. 357–362.

[19] B. Karnan, A. Kuppusamy, T.P. Latchoumi, A. Banerjee, A. Sinha, A. Biswas, A. K. Subramanian, Multi-response optimization of turning parameters for cryogenically treated and tempered WC–Co inserts, J. Institution of Eng. (India): Series D (2022) 1–12.

[20] Xiaofei Deng, Yiyu Yao, An information-theoretic interpretation of thresholds in probabilistic rough sets. RSKT 2012, LNAI 7414, © Springer-Verlag, Berlin Heidelberg, 2012, pp. 369–378, 2012.

[21] J.F. Banu, P. Muneeshwari, K. Raja, S. Suresh, T.P. Latchoumi, S. Deepan, Ontology-based image retrieval by utilizing model annotations and content, in: 2022 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence), IEEE, 2022, pp. 300–305.

[22] M.S.R. Naidu, P. Rajesh Kumar, K. Chiranjeevi, Shannon and Fuzzy entropy based evolutionary image thresholding for image segmentation, Alexandria Eng. J. (2017), https://doi.org/10.1016/j.aej.2017.05.024.

[23] K.M. Remesh, Latha R Nair, A novel method to reduce attributes from reducts, Int. J. Res. Appl. Sci. Eng. Technol. (IJRASET) (2019). ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887, Volume 7 Issue II, Feb- Available at, www.ijraset.com.

[24] Yiyu Yao, Decision-theoretic rough set models. rough set and knowledge technology, second international conference, RSKT 2007, LNAI 4481, pp.1–12, 2007.

[25] M.S.R. Naidu, P. Rajeshkumar, K. Chiranjeevi, Shannon and fuzzy entropy-based evolutionary image thresholding for image segmentation, Alexandria Eng. J. (2017) 0168–1110, https://doi.org/10.1016/j.aej.2017.05.024. Faculty of Engineering, Alexandria University. Production and hosting by Elsevier B.V.doi.org/.

[26] Yiyu Yao, Bing Zhou, Naive bayesian rough sets, RSKT (2010) 719–726. LNAI 64012010.

[27] Wentao Li, Witold Pedrycz, Xiaoping Xue, Xiaoyan Zhang, Bingjiao Fan, Bingham Long, Information measure of absolute and relative quantification in the double-quantitative decision-theoretic rough set model, J. Eng. 2018 (16) (2018) 1436–1441, https://doi.org/10.1049/joe.2018.8315, eISSN 2051-3305.

[28] anne-laure boulesteix, Maximally selected chi-square statistics and binary splits of nominal variables, Biometrical J. 48 (2006) 838–848, https://doi.org/10.1002/bimj.200510191, 5.

[29] Yan Zhang, Optimizing gini coefficient of probabilistic rough set regions using game-theoretic rough sets, in: 26th IEEE Canadian Conference of Electrical and Computer Engineering, 2023, 978-1-4799-0033-6/13.

Remesh K M is having 29 years of experience in teaching and presently working as Joint Director with the Department of Technical Education, Kerala State, India. He has published a number of papers in the areas of Data mining and Three-way decisions. His-areas of interest are Machine Learning, Data Ming and Three-way decision theory .

Dr. Latha R Nair has published number of papers in the area of Machine intelligence and Natural Language Processing. She has done extensive research in Malayalam Language Computing. Her areas of interest are Machine Learning, Image Processing and Natural Language processing .