Operational Research in Engineering Sciences: Theory and Applications
Vol. 5, Issue 1, 2022, pp. 41-55
ISSN: 2620-1607
eISSN: 2620-1747
DOI: https://doi.org/10.31181/oresta190222061s

CRITERIA SELECTION AND DECISION MAKING OF HOTELS USING DOMINANCE BASED ROUGH SET THEORY

Haresh Kumar Sharma [1*], Aarti Singh [2], Dixy Yadav[1], Samarjit Kar[3]

[1] Department of Mathematics, Shree Guru Gobind Singh Tricentenary University, Gurugram, India
[2] Department of Management, Fore school of management, New Delhi, India
[3] Department of Mathematics, National Institute of Technology Durgapur, West Bengal, India

Received: 16 June 2021
Accepted: 24 December 2021
First online: 19 February 2022

Research article

Abstract: *Accommodation is one of the necessities of tourists and travel agencies' significant responsibilities. With the growing competition and profit-making various tour organising companies have started providing attractive accommodation options to the travellers to win their choices. Present research performs a case study on accommodation providing hotels through designing a strategy to enhance their profit earrings by welcoming more and more tourists. The methodology comprises rough set theory (RST) using the Dominance Based rough set theory (DRST) on the collected data of selected variables such as location, facility, value for money, etc., of hotels. Correspondingly, if and then decision rule has been used to classify these variables. The statistical methods regression analysis has also been used to define each variable's relationship and influence on concerned authorities' decision-making. The results show that hotels and tourists can benefit from the proposed strategy and help in decision making by understanding tourist behaviour, increasing profit, improving services, and quality of hotels.*

Keywords: *Hotel criteria, Dominance-based rough set theory, regression analysis, decision making.*

1. Introduction

The Indian tourism industry has been growing rapidly in the past decades. The tourism places attract tourists from all over the world, which makes Indian Tourism a direct contributor to the Economy. According to the Indian Ministry of Tourism Annual report, the tourism industry contributed 6.23% to National Gross Domestic Product (GDP) in the year 2018-2019, where the tourism industry growth rate is

* Corresponding author.
Hareshshrm@gmail.com (H.K. Sharma), singh.aartij@gmail.com (A. Singh), y.dixy12@gmail.com (D. Yadav), kar_s_k@gmail.com (S. Kar)

increased from 3.0 to 14.12% percent from the year 2014 to 2019 respectively. This growth rate supports the rise in competition in hotels which are one of the main contributors to the tourism industry. The tourism industry became another crucial source of foreign exchange and new job creation by providing 8.78% of jobs in India. The global trends show that the Indian tourism sector is one of the fast-growing industries which will proliferate in future (Sharma and Kalotra, 2016). In 2019, Travel & Tourism Competitiveness Index (2019) had confirmed that India secures at 34th place in the travel and tourism business. In contrast, in terms of cost-effectiveness and business environment, it lies at 13 and 39 positions in worldwide competitiveness.

There has been a progressive growth of tourism and hospitality management worldwide in the past decade (Mohajerani and Miremadi, 2012). As well as growing competition in the tourism business, management systems are trying to create equilibrium between the ethics of the business world and customer accommodation without compromising the quality of services to the customers in the hotels business (Sohrabi et al., 2012). In other words, priority must be given to customer satisfaction. With increased competition in the hotel and tourism industry, the hotel management system must find the opportunity and threats of the quality of service they provided to their customers (Chu and Choi, 2000). The hotel business can proliferate only when the hotel offers high-quality services to their customers, which promotes long-term relationships among customers and the hotel management system (Martin, 1986; Croby et al., 1990).

Further, consider the creating steadiness of actual customer state of mind with customer ratings, i.e., establishing the linkage among actual customer ratings given by the customers to hotels management system with genuine customer sentiments (Geetha et al., 2017). As tourism is considered an essential business activity for the hotel and tourism industry thus hotel management and tour agencies should introduce new advancements initiatives like adequate and flexible customer services for promoting business and attracting more customers (Hsieh and Lin, 2010). It shows that customer satisfaction is a vital measurement and essential to hotels. Thus to maintain customer services and to satisfy customers, hotel management and tourism agencies have to keep their adequate flexibility in their services and also introduce promotional activities which can attract maximum customers (Sohrabi et al., 2012).

In literature, many studies have been conducted to analyse to explore the quality of hotels by using various research methodologies like factor analysis, descriptive statistics, and regression techniques (Ren et al., 2016; Xu and Li 2016; Lahap et al., 2016; Li et al.,2017; Lai and Hitchcock, 2017; Patiar et al., 2017). Hua and Yang (2017) applied econometric models to identify factors of crime on the overall hotel performance of Houston hotels. Alptekin and Büyüközkan (2011) identify influencing factors for the hotel industry by using exploratory factor analysis mixed with fuzzy logic. The regression model has been developed to analyse the effect of localised competition on the hotel industry by considering demographic variables, prize and population density as independent variables (Joel and Mezias, 1992).

In the literature, there are several studies of rough set theory and its application in diverse domains. Stević et al. (2017) formulated a multicriteria decision model with eight criteria and eight alternatives for an internal transport logistics of a paper manufacturing company. They used the simple additive weighting (SAW) method and rough numbers, which is used for ranking the potential solutions and selecting the most suitable one. Roy et al. (2019) has proposed an integrated uneven number based COPRAS Model to evaluate the ranking of Delhi hotels. Sharma et al. (2019) has offered a rough set based double exponential smoothing model for forecasting air passengers data. Žižović & Pamucar (2019) has suggested level based weight assessment (LBWA) based multicriteria decision-making model for the investigation of criteria weightage. Popov (2020) applied Johnson–Kendall–Roberts (JKR) theory to find the relation between smooth and rough elastic bodies. Božanić et al. (2020) used a rough interval-based Level Based Weight Assessment and Multi Attributive Ideal-Real Comparative Analysis method (LBEA-IR-MAIRCA) model to determine constructive elements of new weapons. Pamucar et al. (2022) has utilised FUll COnsistency Method (FUCOM) and Multi-Attributive Ideal-Real Comparative Analysis (MAIRCA) methods as integrated rough group analysis for and prioritisation of railway infrastructure project evaluation. Sharma et al. (2021) hybrid rough set model-based analysis has been performed to forecast the sugarcane yield of India. Kazemitash et al. (2021) has used the data of Biofuel Company's supplier selection for the information system performance calculation by the integration of rough set theory through the Best-Worst method (BWM). The authors have also employed the rough BWM to determine the weight values of the criteria. Hu et al. (2021) proposed the weighted neighbourhood rough set (WNRS) and accordingly introduced a unique attribute reduction technique. Subsequently, Yu et al. (2021) demonstrated that the concept refinement in topology is too abstract to elucidate the variability of the rough set model along with the variation in granules. Here, the authors proposed two novel granule cover refinements, including point-set topology and rough set theory. Ye et al. (2021) also introduced a novel decision-making method based on a fuzzy rough set. They applied the technique in a real-world scenario to illustrate the feasibility of the proposed method. After that, Kusunoki et al. (2021) considered two parametric dominance-based rough set approaches (DRSA) and offered variable precision DRSA (VP-DRSA) and variable consistency DRSA (VC-DRSA). Following this, Błaszczyński et al. (2021) examined a new data set for auto loan applications using a technique not yet explored for financial fraud prediction, namely the Dominance-based Rough Set Balanced Rule Ensemble (DRSA-BRE).

Pawlak (1982) established an effective method known as Rough Set theory for extracting the facts from the information system. However, the traditional rough set methodology is not adequate to study the relationship among preference order arising from attributes like debt ratio (Blaszczynski et al. 2007), service strategies, product quality, and business indicators (Couto and Gaiado, 2015). Therefore, this study proposes applying the Dominance-based Rough Set theory (DRST) to solve preference-ordered situations. According to Greco et al. (2000), DRST approach has been anticipated to solve the preference-ordered situations in data mining. It is a powerful tool for attribute reduction in the qualitative-based data set. The dominance based rough set theory has been successfully employed in a variety of areas. Chakhar and Saad (2012) proposed a DRST approach to study groups in the multicriteria class study. The dominance-based rough set methodology has been used to develop the model for limiting the speed of vehicles in speed-controlling zones (Augeri et al., 2015). Chakhar et al. (2016) suggested that DRST has been used to derive rules in multicriteria group decision-making based on several case studies. Sawicki and Zak (2014) have reported that DRST based analysis is done on transportation problems by producing decision rules depending on customer view and expectations. Moreover, it has also been used in different uncertain multicriteria decision-making applications (Kazemitash et al., 2021; Pamučar and Janković, 2020; Pamučar et al., 2018; Đalić et al. 2020).

The study has been organised as follows. The basic concepts and some related properties of DRST are discussed in section' Dominance based RST'. A case study of hotel data and analysis of hotel data using DRST for multicriteria decision model is presented in case study section. The comparison purpose statistical analysis of hotel data is discussed in regression analyses section. Finally, the result and discussion, conclusion, and future scope of our study are stated in section result and discussion, and conclusion'.

2. Dominance based rough set theory (DRST)

DRST extends the Classical Rough Set theory (CRST) introduced by Pawlak in 1982. The multicriteria decision representation used in this research applies the concept of DRST. Thus, the RST methodology is an efficient mathematical mechanism to dealing with uncertainty and vagueness. However, Classical rough set theory (CRST) is restricted to sort problems where the preferences-orders in the set of attributes (criteria) are considered. These are the inconsistencies generated due to the violation of the dominance principle that eventually cannot be handled by the model. Hence in case of such inconsistencies, some methodological changes to CRST are required. Greco et al. (2000) have proposed an expansion of the RST depending on the dominance concept that would allow it to handle the inconsistency. This idea relies on replacing the indiscernibility relation for a dominance relation in the rough approximation theory of the decision category.

2.1. Information system

Sample The information concerning the objects is often structured in the form of an information table whose different rows mention distinct actions (objects) and whose columns mention the other criteria or attributes considered.

Formally, an information table is structured in a 4-tuple information system $S = (U, Q, V, F)$, Where U is a non-empty finite set of objects (universe) and $Q = \{q_1, q_2, \dots \dots, q_m\}$ is a non-empty finite set of attributes or criteria such that $q: U \rightarrow V_q$ for every $q \in Q, V_q$ is the domain of the attributes or criteria q. $V = \cup_{q \in Q} V_q$ and $f: U \times Q \rightarrow V$ is the information function determined such that $f(x, q) \in V_q$ for every attributes q $\in Q; x \in U$. The set Q is often separated into a set C ≠ $\emptyset$ of condition attributes, and a set D ≠ $\emptyset$ of decision attributes such that $C \cap D = \emptyset$ and $C \cup D = Q$. In such a situation, S is called an information table.

2.2. RST with dominance relation

If the scale of the condition attribute is arranged in increasing or decreasing preference, then it is called criterion. Alternatively, it is known as regular condition attributes. DRST exponents suppose that the preference increases with the value of $f(\cdot, q)$ for every criterion $q \in C$. We also suppose that the set of decision attribute

D (perhaps a singleton {d}) create a parting of universe U into a set of decision classes, let $Cl = \{Cl_t, t \in T\}$, $T =$ {1,....,n} be a finite set of classes of universe U such that every $x \in U$ belongs to one and only one class $Cl_t \in Cl$. We assume that classes are preference-ordered, i.e., for all $r, s \in T$, such that $r > s$, the objects from Cl_r are more preferred to the objects from Cl_s. Suppose P ⊆ C is a subset of condition attributes. The dominance relation ∇_P allied with P is described for every pair of Objects x and y so; $x \nabla_P y \Leftrightarrow f(x,q) \succcurlyeq f(y,q), \forall q \in P$.

The letter "$\succcurlyeq$" should be changed with "$\preccurlyeq$" for criteria according to the decreasing preference. We associate pair of a sets with every object $x \in U$: (i) P-dominating set $\nabla_P^+(x) = \{y \in U: y \nabla_P x\}$ having objects that dominate x and (ii) P-dominated set $\nabla_P^-(x) = \{y \in U: x \nabla_P y\}$ having objects dominated by x. These pair of sets are familiar with approximate decision classes.

The P-lower approximation of $Cl_t^{\geq}$ (upward union), ($\underline{P}(Cl_t^{\geq})$), is constituted of total objects x from U such that all members y, contain at least the similar assessment on all of the examined criteria from P, also member of a class Cl_t or better. In another way, if any object y has at least as good an analysis based on the criteria from P as object x member of $\underline{P}(Cl_t^{\geq})$, then indeed, y is a member of a class Cl_t or preferable class. The P-upper approximation of $Cl_t^{\geq}$ (upward union), which involves all objects with a P-dominating set, is allocated to a class at least as good as Cl_t.

Similarly, the P-lower and P-upper $Cl_t^{\leq}$ approximation with respect to $P \subseteq C$respectively represented as $\underline{P}(Cl_t^{\leq})$ and $\overline{P}(Cl_t^{\leq})$, are defined as:

$$\underline{P}(Cl_t^{\leq}) = \{x \in U: \nabla_P^-(x) \subseteq Cl_t^{\leq}\}, \quad (1)$$

$$\overline{P}(Cl_t^{\leq}) = \bigcup_{x \in Cl_t^{\leq}} \nabla_P^-(x) = \{x \in U: \nabla_P^+(x) \cap Cl_t^{\leq} \neq \emptyset\}. \quad (2)$$

2.3. Accuracy of approximation and quality of classification

For all $t \in \{1, \dots, n\}$ and each, $P \subseteq C$ we described the accuracy of the approximation of $Cl_t^{\geq}$ and $Cl_t^{\leq}$, respectively, as follow:

$$\alpha_P(Cl_t^{\geq}) = \frac{card(\underline{P}(Cl_t^{\geq}))}{card(\overline{P}(Cl_t^{\geq}))}, \alpha_P(Cl_t^{\leq}) = \frac{card(\underline{P}(Cl_t^{\leq}))}{card(\overline{P}(Cl_t^{\leq}))} \quad (3)$$

The coefficient $\gamma_P(Cl) = \frac{card(U - ((\cup_{t \in T} Bn_P(Cl_t^{\leq})) \cup (\cup_{t \in T} Bn_P(Cl_t^{\geq}))))}{card(U)}$ (4)

is known as the quality of approximation of partition Cl using attribute set $P \subseteq C$.

2.4. Decision rules

On the foundation of the approximations found by the use of the dominance relation, it is viable to set off a generalised explanation of the preferential knowledge contained in the information table, such description of the preferential knowledge we can write in the form of "if ..., then..." decision rules. The algorithms for induction related to regulations are acquired by using 4eMka2 software (Poznan University of Technology, Poland, Laboratory of Intelligent Decision Support System 2006).

All the decision rules can be considered in the following three ways:

1. $D_{\geq}$ -decision rules which are having the following form:

If $f(x,q_1) \geq r_{q1}$ and $f(x,q_2) \geq r_{q2}$ and .... $f(x,q_p) \geq r_{qp}$, then $x \in Cl_t^{\geq}$,

These decision rules are assisted by the member of the universe that belongs to the P-lower approximation $Cl_t^{\geq}$.

2. $D_{\leq}$ -decision rules which have the following form:

If $f(x,q_1) \leq r_{q1}$ and $f(x,q_2) \leq r_{q2}$ and .... $f(x,q_p) \leq r_{qp}$, then $\in Cl_t^{\leq}$,

These decision rules are assisted by a member of the universe that belongs to the P-lower approximation of $Cl_q^{\leq}$.

3. $D_{\geq\leq}$ -decision rules which have the following form:

If $f(x,q_1) \geq r_{q1}$ and $f(x,q_2) \geq r_{q2}$ and .... $f(x,q_k) \geq r_{qk}$, and $f(x,q_{k+1}) \leq r_{qk+1}$ and ..... $f(x,q_p) \leq r_{qp}$, then $x \in Cl_t \cup Cl_{t+1} \cup \ldots \cup Cl_s$,

These decision rules assisted by a member of the universe that belongs to the boundary region of the union of classes $Cl_s^{\leq}$ and $Cl_t^{\geq}$, where $P = \{q_1, q_2, \ldots \ldots, q_p\} \subseteq C$, $(r_{q1}, r_{q2}, \ldots \ldots, r_{qp}) \in V_{q1} \times V_{q2} \times \ldots \ldots \times V_{qp}$, and $t \in \{1, \ldots, n\}$.

3. Case Study

The hospitality industry is one of the major contributors of growth among the all-service sector industries in India. Since, India is a country of diversity with its rich culture and heritage, hence the tourism contributes a significant source of foreign exchange. As, tourism is the integral part which has a considerable effect on the hotel industry. This indicates that the digital advancement in tourism sector also affect hospitality industry. The digital enhancement in tourism of India through digital tools used for planning, booking, and experiencing a journey have significant effect over hospitality industry. The empirical study focuses on the Indian hospitality industry includes data collected from various online platforms in the hotels. Since customer satisfaction harms the hotel industry, the possibility of getting a hotel that satisfies customers' needs is maximised by selecting specific attributes which are related to the hotel industry. The following study scrutinises the influence of overall rating (O) on location (LO), hospitality (HT), facilities (FT), sanitation and Cleanliness (SC), the value of money (V), food quality (FD), and price (PR) using both Indian and international tourists' hotel data. Criteria descriptions are listed in table 1. The study's objectivity has been kept in mind, and all variables are used according to data availability. Online reviews play an essential role in the hotel selection process as websites provide customer reviews based on their personal experiences with provided hotel services. These websites give the travelers an overall idea to select the best hotel which satisfies their needs based on others' experiences. Sometimes decision-making becomes difficult as there are different reviews based on one's perspective. The data related to the hospitality industry are extracted from tourism websites. The presented approach assists the hotel selection process based on the influence of overall rating on location, hospitality, facilities provided, sanitation and cleanliness, the value of money, food, and price. The proposed study is used to select the best hotel based on existing data.

Table 1. Criteria Description

Criteria	Description
Location (LO)	The geographical location of the hotel has been considered according to the convenience of tourists.
Hospitality (HT)	It includes a friendly and generous welcome and entertainment for tourists.
Facilities (FT)	It includes a Travel desk, eating place, parking, pieces of equipment, or services provided to tourists for their stay.
Sanitation and Cleanliness (SC)	Sanitation and Cleanliness include the sanitary condition of a hotel.
Value for money (V)	A beneficial combination of sustainability, cost, and quality to meet tourist requirements.
Food Quality (FD)	The acceptable standard quality of food served.
Price (PR)	Convenient fare according to traveler and hotel management
Overall Rating (O)	It includes the net classification of hotels based on the different quality scale.

The objective of this case study is to extract the decision rules to show the hotel features and classify the different characteristics of the tourist industry. It has been found that the Rough set theory is the most suitable approach for criteria selection in decision-making problems. For this study, data has been collected from the best tourism website (https://www.makemytrip.com), and it will help the tourism management for analysis of significant criteria of the hotel industry. The model must provide relevant information to hotel management for improvement of their service quality.

4. DRST analysis

Based on several studies such as (Geetha et al., 2017; Li et al. (2017) of hotel tourism, and expert interviews of hotel managers and their management teams, tourism and travelling management of India has conclusively given higher priority to the eight essential criteria/attributes given in section 3 of 609 best Indian hotels. Because according to experts, these selected eights criteria are preferred mainly by the maximum tourists while making their hotel selection decision. In eight attributes, seven attributes are called condition criteria, and another one is decision criteria were investigated for analysis. In this study, we have applied the DRST technique for rule generation. DRST toolkit 4eMka2 software from Poland, Laboratory of Intelligent Decision Support System 2006, is used for constructing the decision rules.

4.1. Accuracy approximation and quality of classification

Table 2, provides approximation accuracy for all decision classes, as approximation sets (specifically lower and upper approximation) and accuracy of approximation has been already explained in section 2.2 and 2.3. The selected criteria can be adequate to approximate the classification if the classification quality and accuracy of the approximation. The class "at most medium" means class related to "overall hotel rating will be medium and lower values". The decision class "at most good" contains the two classes, which are "good" and "medium". Further, the decision class" at least good" represents the class "overall hotel rating will be good or excellent". Finally, the decision class "at least excellent" consists of only one class, i.e. overall rating of the hotel be will be excellent.

Table 2. Accuracy of approximation

	At most medium	At most Good	At least Good	At least Excellent
Lower approximation	10	17	99	20
Upper approximation	510	589	599	592
Boundary	510	572	500	572
Accuracy of approximation	0.636	0.0290	0.17	0.0340
Quality of classification		0.049		

Table 3. Certain decision rules of hotel data set

Decision Rules	Support
If (food ≥ Excellent) & (hospitality ≥ excellent) Then (overall rating≥ good)	61
If (food ≥ Excellent) & (facilities ≥ medium) Then (overall rating ≥ good)	83
If (food ≥ excellent) & (sanitation and Cleanliness≥ excellent) Then (overall rating ≥ good)	89
If (food ≥ Excellent) & (price ≥ medium) & (value for money ≥ excellent) Then (overall rating ≥ excellent)	18
If (food ≥ Excellent) & (price ≥ high) Then (overall rating ≥ excellent)	6
If (food ≤ poor) & (price ≤ low) & (location ≤ bad) Then (overall rating ≤ good)	11
If (hospitality ≤ poor) & (facilities ≤ medium) Then (overall rating ≤ good)	6
If (sanitation and cleanliness ≤ poor) Then (overall rating ≤ good)	5
If (food ≤ poor) & (price ≤ low) & (facilities ≤ good) & location ≤good) Then (overall rating ≤ medium)	5
If (value for money ≤ poor) & (facilities ≤ good) Then (overall rating ≤ medium)	6
If (location ≤ bad) & (sanitation and cleanliness ≤ good) Then (overall rating ≤ medium)	3

As clarified in the section as mentioned above 2.4, the decision rules were formed by analysing the training data of dominance-based rough set theory. These rules were applied to relationships among conditions and decision attributes.

Furthermore, 11 certain decision rules were obtained from the information system. Total 5 decision rules are found to be more accurate since support is greater than 10. Based on these decisions rule, we can analyse which criteria are significant for hotel management. The estimated results of reduced rules are presented in Table 3.

Table 3 shows the 11 minimum cover rules generated from the hotel data set. The minimal cover certain decision rules can be written in the form of IF-THEN statement. Here is some example to illustrate IF-THEN rules:

IF food is excellent AND hospitality is excellent, THEN the decision criteria overall rating will be perfect.

From table 3, it is clear that if the hotel's food is excellent and Cleanliness is excellent, then the overall rating will be excellent with maximum support of 89 (cf. rule 3). It means that food and cleanliness are essential factors for travellers. If the hotel's food is excellent and facilities are medium, and above medium then the overall rating will be excellent with support of 83 (cf. rule 2). If food quality is excellent and hospitality is best then the overall rating will be excellent with support of 61 (cf. rule 1). These decision rules indicate that food, Cleanliness, hospitality, and facilities are essential attributes for travelers. Therefore, it can be suggested that most tourists select their hotel based on food, Cleanliness, hospitality, and facilities. The different stages of analysis are depicted in Figure 2.

5. Regression analysis

By analysing the literature review (Sheather, S., 2009; Ren et al., 2016; Patiar et al., 2017; Hua and Yan., 2017), the regression model is obtained by using the following framework:

$$\text{Overall rating (O)} = \alpha+\alpha 1\ LO+\alpha 2\ HT+\alpha 3\ FT+\alpha 4\ SC-\alpha 5\ V-\ \alpha 6\ FD+\alpha 7\ PR+\ \varepsilon \qquad (5)$$

Where ε is the error, α, α1, …, α7 are the coefficients of considered variables (LO, HT, …, PR), O is the overall rating; LO is the location of the hotel, HT is the hospitality, FT is the facilities provided by the hotel, SC is the sanitation and Cleanliness, V is the value of money, FD is food quality, and PR is for the price of the hotel's room. The estimated regression results are described in Table 3. The acquired result indicates that the hotel's location, hospitality, sanitation and cleanliness, and performance and effectiveness of money charges, i.e. the value of money, has a significant positive effect on overall ratings of the hotel. Whereas, facilities provided, i.e. physical characteristics associated with a hotel-like travel desk, eating place, parking, etc., food quality and hotel price don't seem to have a significant effect on the overall ratings of the hotel. Moreover, the F- statistics results confirmed that the regression model is essential for criterion for hotel selection process since the p-value is 2.2e-16 ≈0.000, which is significant. Estimating sturdiness of the model by using R2, which is 0.8459, i.e. all variables have an approximate 84.59% effect on overall ratings of the hotel for criterion for the hotel selection process, which is considerably good. Therefore, the considered regression model is relevant for the empirical study. Also, from figure1, it is clear that the relationship between overall rating with location, hospitality, facilities, sanitation, and cleanliness, the value of money, food, and the price is linear. The linear line indicates is that the best-fitted model with the curve for the multivariate analysis. Our data are independent and follow Gaussian distribution, then the model is accepted within the robustness test. Figure 1, shows a normal probability plot to decide whether it is reasonable to consider. The accuracy measure derived using regression analysis of hotel data set is sampled from a population, follows a normal distribution.

The regression equation for the variables is:

Overall rating = 0.538+0.1854 Location+0.6706 Hospitality+0.3821 Facalities+0.2193Sanitation and cleanliness-0.01933 Value of Money-0.1549 Food+0.000005004 Price+ ε (6)

Table 4. Regression analysis results

Variables	Coefficients	Standard error	t- value	Pr(>|t|)
Constant	5.388e-01	7.654e-02	7.039	5.29e-12 ***
Location	1.854e-01	1.922e-02	9.646	<2e-12***
Hospitality	6.706e-01	1.995e-02	33.619	<2e-16***
Facilities	3.821e-03	2.380e-03	1.606	0.1088
Sanitation and cleanliness	2.193e-01	2.071e-02	10.589	<2e-16***
Value of money	-1.933e-01	1.954e-02	-9.938	<2e-16***
Food	-1.549e-01	1.439e-02	-1.076	0.2822
Price	5.004e-06	2.784e-06	1.797	0.0728
Adjusted R2	0.8459			
F-statistics	477.8			

Residual standard error: 0.1673 on 601 degree of freedom
Multiple R-squared: 0.8477, p- value < 2.2e-16

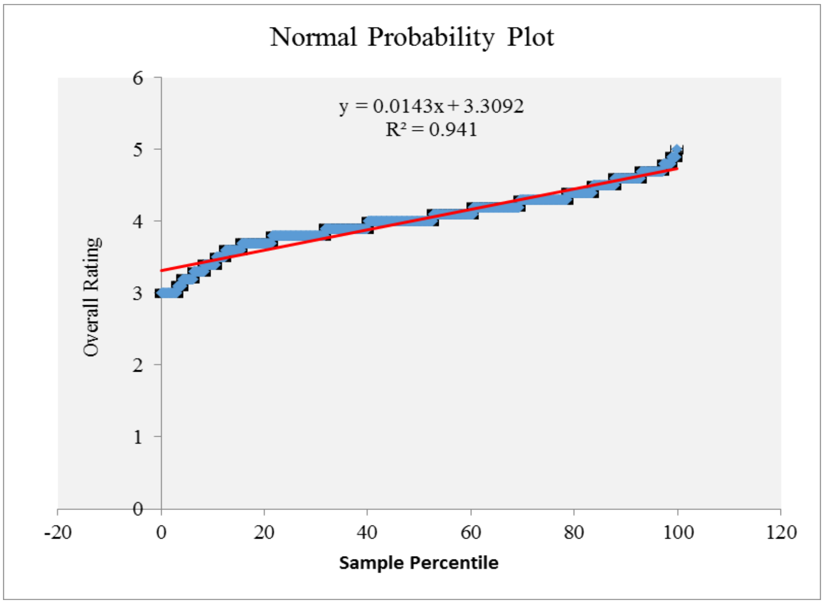

Figure 1. Normal probability plot of statistical analysis

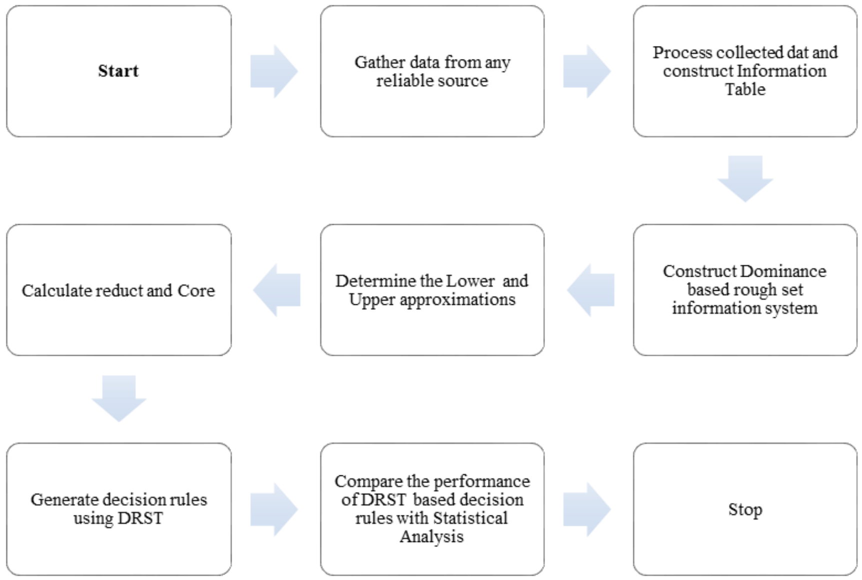

Figure 2. Stages for air transport passengers forecasting

6. Results, discussion and conclusion

This research focuses on hotel selection and estimation through a hybrid method of Dominance rough set theory and regression analysis. This estimated model has been analysed under uncertainty in which DRST is employed in acquiring the information related to significant attributes of the hotel business. Furthermore, a case study on real-life data of Indian hotels has been performed using the DRST approach on the selected attributes.

The foremost suggestion resulting from this study are (i) Food, facilities, Cleanliness, and hospitality are the most significant attributes for any hotel selection as uncovered in decision rule and the expert's opinion based on customer prioritization and feedback. (ii) Hotel management has been turn-up with a clear picture of the hotel's criteria to improve performance according to the current business market. This facilitates the hotel management system to make appropriate decisions regarding the quality and services up-gradation of the hotel. (iii) It can be said that the DRST is a knowledge-based decision-making system that can evaluate the effective and appropriate attributes by the comparison of collected data with secondary data obtained from tourism websites.

Hence, this research leads to a robust hybrid method, 'DRST-Regression', which confirms the accuracy and firmness of the decision making outcomes. It is a unique approach contributed by this study because it gives rise to the most precise and reliable outcomes without any statistical assumptions. Comparatively, DRST-Regression is more preferable to the statistical method due to its dynamic and advanced approach. Therefore, this study resulted in an empirical model that can be preferred over the statistical model because it divulges the consequential decision rules are easy to understand as compared to statistical methods without any distributional assumption.

The main limitations of the DRST are that the approximation sets (upper/lower) depend only on the choice of attributes, which may be regarded as disadvantage, since there may not be enough flexibility for some applications. In future, similar case studies can be considered and analysed using rough sets and different machine learning algorithms, including decision tree, random forest, support vector machine and elastic net.

References

Alptekin, G. I., & Büyüközkan, G. (2011). An integrated case-based reasoning and MCDM system for Web based tourism destination planning. Expert Systems with Applications, 38(3), 2125-2132. https://doi.org/10.1016/j.eswa.2010.07.153

Augeri, M. G., Cozzo, P., & Greco, S. (2015). Dominance-based rough set approach: An application case study for setting speed limits for vehicles in speed controlled zones. Knowledge Based Systems, 89, 288-300. https://doi.org/10.1016/j.knosys.2015.07.010

Błaszczyński, J., Adiel T. de Almeida Filho, Matuszyk, A., Szeląg, M., & Słowiński, R. (2021). Auto loan fraud detection using dominance-based rough set approach versus machine learning methods. Expert Systems with Applications, 163, 113740. https://doi.org/10.1016/j.eswa.2020.113740

Božanić, D., Ranđelović, A., Radovanović, M., & Tešić, D. (2020). A hybrid LBWA - IR-MAIRCA multicriteria decision-making model for determination of constructive elements of weapons. facta Universitatis, Series: Mechanical Engineering, 18(3), 399-418. 10.22190/FUME200528033B

Chakhar, S., & Saad, I. (2012). Dominance-based rough set approach for groups in multicriteria classification problems. Decision Support Systems, 54(1), 372-380. https://doi.org/10.1016/j.dss.2012.05.050

Chakhar, S., Ishizaka, A., Labib, A., & Saad, I. (2016). Dominance-based rough set approach for group decisions. European Journal of Operational Research, 251(1), 206-224. https://doi.org/10.1016/j.ejor.2015.10.060

Croby, L. A., Evans, K. R., & Cowels, D. (1990). Relatioship quality in service selling: an interprersonal influence perspective. Journal of Marketing, 54 (2), 68–81.

Couto, & Gaiado, A. B. (2015). Using a dominance based rough set approach for analysing Business indicators, Information Technology and Quantitative Management (ITQM). Procedia Computer Science, 55, 350-359. https://doi.org/10.1016/j.procs.2015.07.062

Geetha, M., Singha, P., & Sinha, S. (2017). Relationship between customer sentiment and online customer ratings for hotels-An empirical analysis. Tourism Management, 61, 43–54. https://doi.org/10.1016/j.tourman.2016.12.022

Greco, S., Matarazzo, B., & Slowinski, R. (2000). Extension of the rough set approach to multicriteria decision support. INFOR: Information Systems and Operational Research, 38(3), 161-195. https://doi.org/10.1080/03155986.2000.11732407

Đalić, I., Stević, Ž., Karamasa, C., & Puška, A. (2020). A novel integrated fuzzy PIPRECIA–interval rough SAW model: Green supplier selection. Decision Making:

Applications in Management and Engineering, 3(1), 126-145. https://doi.org/10.31181/dmame2003114d

Hsieh, L. F., & Lin, L. H. (2010). A performance evaluation model for international tourist hotels in Taiwan—An application of the relational network DEA. International Journal of Hospitality Management, 29(1), 14-24. https://doi.org/10.1016/j.ijhm.2009.04.004

Hua, N., & Yang, Y. (2017). Systematic effects of crime on hotel operating performance. Tourism Management, 60, 257-269. https://doi.org/10.1016/j.tourman.2016.11.022

Hu, M., Tsang, E. C. C., Guo, Y., Chen, D., & Xu, W. (2021). A novel approach to attribute reduction based on weighted neighborhood rough sets. Knowledge-Based Systems, 220, 106908. https://doi.org/10.1016/j.knosys.2021.106908

Joel A. C. Baum, & Mezias, S. (1992). Localised Competition and Organisational Failure in the Manhattan Hotel Industry, 1898-1990. Administrative Science Quarterly, 37(4), 580-604. https://doi.org/10.2307/2393473

Kazemitash, N., Fazlollahtabar, H., & Abbaspour, M. (2021). Rough Best-Worst Method for Supplier Selection in Biofuel Companies based on Green criteria. Operational Research in Engineering Sciences: Theory and Applications, 4(2), 1-12. https://doi.org/10.31181/oresta20402001k

Kusunoki, Y., Błaszczyński, J., Inuiguchi, M., & Słowiński, R. (2021). Empirical risk minimization for dominance-based rough set approaches. Information Sciences, 567, 395-417. https://doi.org/10.1016/j.ins.2021.02.043

Lahap, J., Ramli, N. S., Said, N. M., Radzi, S. M, & Zain R. A. (2016). A Study of Brand Image towards Customer's Satisfaction in the Malaysian Hotel Industry. Procedia - Social and Behavioral Sciences, 224, 149 – 157. https://doi.org/10.1016/j.sbspro.2016.05.430

Lai, I. K. W. & Hitchcock, M. (2017). Sources of satisfaction with luxury hotels for new, repeat, and frequent travelers: A PLS impact-asymmetry analysis. Tourism Management, 60, 107-129. https://doi.org/10.1016/j.tourman.2016.11.011

Li. L., Peng, M., Jiang, N., & Law, R. (2017). An empirical study on the influence of economy hotel website quality on online booking intentions. International Journal of Hospitality Management, 63, 1-10. https://doi.org/10.1016/j.ijhm.2017.01.001

Martin, W.B. (1986). Quality Service: The Restaurant Manager's Bible. School of Hotel Administration, Cornell University, Ithaca, NY.

Mohajerani, P., & Miremadi, A. (2012). Customer Satisfaction Modeling in Hotel Industry: A Case Study of Kish Island in Iran. International Journal of Marketing Studies, 4(3), 134-152. doi:10.5539/ijms.v4n3p134

Pamučar, D., & Janković, A. (2020). The application of the hybrid interval rough weighted Power-Heronian operator in multicriteria decision making. Operational Research in Engineering Sciences: Theory and Applications, 3(2), 54-73. https://doi.org/10.31181/oresta2003049p

Pamucar, D., Macura, D., Tavana, M., Božanić, D., & Knežević, N. (2022). An integrated rough group multicriteria decision-making model for the ex-ante prioritization of infrastructure projects: the Serbian railways case. Socio-Economic Planning Sciences, 79, 101098. https://doi.org/10.1016/j.seps.2021.101098

Pamučar, D., Stević, Ž., & Zavadskas, E. K. (2018). Integration of interval rough AHP and interval rough MABAC methods for evaluating university web pages. Applied soft computing, 67, 141-163. https://doi.org/10.1016/j.asoc.2018.02.057

Patiar, A., Emily, M., Sandra, K., & Cox R. (2017). Students' perceptions of quality and satisfaction with virtual field trips of hotels. Journal of Hospitality and Tourism Management, 31, 134-141. https://doi.org/10.1016/j.jhtm.2016.11.003

Pawlak, Z. (1982). Rough sets. International journal of computer & information sciences, 11(5), 341-356. https://doi.org/10.1007/BF01001956

Popov, V. (2020). Coefficients of restitution in normal adhesive impact between smooth and rough elastic bodies. Reports in Mechanical Engineering, 1(1), 103-109. https://doi.org/10.31181/rme200101103p

Poznan University of Technology, Poland, Laboratory of Intelligent Decision Support System. 4eMka2 software http://idss.cs.put.poznan.pl/site/4emka.html

Ren, L., Qiuc, H., Wangd, P., & Lin, P. M. C. (2016). Exploring customer experience with budget hotels: Dimensionality and satisfaction. International Journal of Hospitality Management, 52, 13–23. https://doi.org/10.1016/j.ijhm.2015.09.009

Roy, J., Sharma, H., Kar, S., Kazimieras, Z. E., & Saparauskas, J. (2019). An extended COPRAS model for multicriteria decision-making problems and its application in web-based hotel evaluation and selection. Economic research – Ekonomska istraživanja, 32 (1), 219-253. https://doi.org/10.1080/1331677X.2018.1543054

Sawicki, P., & Zak, J. (2014). The application of Dominance based rough set theory to evaluation of transportation systems. Procedia -Social and Behavioral Sciences, 111, 1238-1248. https://doi.org/10.1016/j.sbspro.2014.03.727

Sharma, N., & Kalotra, A. (2016). Hospitality Industry in India: A Big Contributor to India's Growth, International Journal of Emerging Research in Management & Technology, 5(6), 202-210.

Sharma, H. K., Kumari, K., & Kar, S. (2019). Short-term Forecasting of Air Passengers based on Hybrid Rough Set and Double Exponential Smoothing Models, Intelligent Automation and Soft Computing, 25(1), 1-13.

Sharma, H. K., Kumari, K., & Kar, S. (2021). Forecasting Sugarcane Yield of India based on rough set combination approach. Decision Making: Applications in Management and Engineering, 4(2), 163-177. https://doi.org/10.31181/dmame210402163s

Sheather, S. (2009). A modern approach to regression with R, Springer-Verlag New York.

Sohrabi, B., Vanani, I. R., Tahmasebipur, K., & Fazli, S. (2012). An exploratory analysis of hotel selection factors: A comprehensive survey of Tehran hotels. International Journal of Hospitality Management, 31(1), 96-106. https://doi.org/10.1016/j.ijhm.2011.06.002

Stević, Ž., Pamučar, D., Kazimieras Zavadskas, E., Ćirović, G., & Prentkovskis, O. (2017). The Selection of Wagons for the Internal Transport of a Logistics Company: A Novel Approach Based on Rough BWM and Rough SAW Methods. Symmetry, 9(11), 264. https://doi.org/10.3390/sym9110264

Travel and tourism website: Retrieved February, 2021. https://www.makemytrip.com.

Travel & Tourism Competitiveness Index (2019). Retrieved at March, 2021. https://reports.weforum.org/travel-and-tourism-competitiveness-report-2019/country-profiles/#economy=IND

Xu Xun., & Li Y. (2016). The antecedents of customer satisfaction and dissatisfaction toward various types of hotels: A text mining approach. International Journal of Hospitality Management, 55, 57–69. https://doi.org/10.1016/j.ijhm.2016.03.003

Ye, J., Zhan, J., & Xu, Z. (2021). A novel multi-attribute decision-making method based on fuzzy rough sets. Computers & Industrial Engineering, 155, 107136. https://doi.org/10.1016/j.cie.2021.107136

Yu, Z., Wang, D., & Wang, P. (2021). A study of interrelationships between rough set model accuracy and granule cover refinement processes. Information Sciences, 578, 116-128. https://doi.org/10.1016/j.ins.2021.07.049

Žižović, M., & Pamucar, D. (2019). New model for determining criteria weights: Level Based Weight Assessment (LBWA) model. Decision Making: Applications in Management and Engineering, 2(2), 126-137. https://doi.org/10.31181/dmame1902102z