Wireless Personal Communications (2022) 125:3443–3460
https://doi.org/10.1007/s11277-022-09719-6

Predicting Consumer Recommendation Decisions from Online Reviews: A Rough Set Approach

Praphula Kumar Jain[1,2] · **Anand Prakash**[3]

Accepted: 14 April 2022 / Published online: 13 May 2022

Abstract
Predictive recommendations based on online reviews are considered one of the recent sentiment analysis tasks that classify the emotion expressed in online reviews as recommendations and not recommendations. The consumer recommendations help the upcoming consumer and service provider both; a consumer can refer to it before making purchase decisions. The service provider may utilize it in quality improvements of service or products. This study implemented the Rough Set Theory (RST) for drawing consumer recommendations using online reviews. An RST is a mathematical tool based on knowledge & information to predict/approximate decisions. We have classified the various service aspects as attributes and used RST to predict the reduct and core characteristics to classify flight recommendations. We have also investigated the effect of each service provider's service and redundant service(s). Experimental results showed that implementing RST for predictive recommendations is the best option because of its improved accuracy (96.70%) and faster classification process, incredibly, with a vast volume of data.

Keywords Recommendations · Rough set · Approximation · Prediction · Sentiment analysis · Airline

1 Introduction

With the growth of the social media applications such as Twitter, Facebook, Instagram and Telegram has emerged into a vast volume of user-generated content information in the form of online reviews and opinions related to various products or services [1]. Peoples like to post their experiences, sentiments, views, opinion, and choices to their sensing and

✉ Praphula Kumar Jain
praphulajn1@gmail.com

Anand Prakash
anand.p.pal@gmail.com

1 Department of Computer Science & Engineering, Indian Institute of Technology (Indian School of Mines), Dhanbad, JH 826004, India

2 Department of Computer Engineering & Applications, GLA University, Mathura, UP, India

3 Department of Mathematics, Shri Vaishnav Vidyapeeth Vishwavidyalaya, Indore, MP, India

findings related to the services, products, or events [2, 3]. Sometimes, they recommend services or products for upcoming consumers based on their experience [4]. Their observations can be positive, negative, or neutral. In practice, this sentiment may be utilized for finding user interest and recent trends [5]. For the same product, one user may have positive, while others may have negative emotions; this is a challenging task in sentiment analysis [6]. Based on the consumer's experience, they are explicitly advising whether they will recommend a particular service or not.

This vast volume of online reviews and opinions required an intelligent model to evaluate and classify it positively, negatively, or neutral. The thoughtful approach for assessing these opinions is called sentiment analysis (SA) [7]. The application of SA in a predictive recommendation decision is a recent trend in the research from the last few years. In the predictive recommendation decisions task, we are classifying the consumer sentiment as a recommendation and not recommendations decisions [8, 9]. For instance, we may draw a reviewer's attention using online reviews, whether they recommend particular products or not.

Consumer sentiment or emotion classification is challenging because shared online reviews and experiences are not proper. Hence, the analysis of unstructured data and its evaluation to determine the consumer's sentiment needs exceptional machine learning (ML) or deep learning (DL) approaches for their classifications. Consumer sentiment analysis is one of the recent research areas that gained attention in current years. Many industries utilize sentiment analysis and predictive recommendations to gain significant benefits, including better consumer relations, new consumer attraction, and business growth. Sentiment analysis can be done based on machine learning or a lexicon-based approach. Nowadays, sentiment analysis tasks are considered as text summarizations, document classification, and information retrieval.

In this study, we implemented the concept of RST for predictive consumer recommendations based on online reviews. We collected Skytrax data from http://www.airlinequality.com, reviewed by many airline travelers from 2015 to 2019. For the predictive recommendations, we generated rules based on rough set theory. Our findings show an essential contribution to the existing literature in the field of recommendation prediction. Further, this research proves significant implications for the service providers to improve their products or services. Additionally, this work is also providing valuable guidelines for forthcoming consumers before making their purchase decisions.

RST for predictive consumer recommendations is the first attempt presented in this paper to the best of our knowledge. Previous work only classifies the sentiment into positive, negative, and neutral classes, but in this research, we have identified the predicted consumer recommendations based on the previous consumer online reviews. Our experiments analysis presents excellent observations in the predictive recommendations.

The remaining article is structured as follows: Sect. 2 discusses the literature review related to this study. The methodology is illustrated in the Sect. 3. The experimental results are discussed in the Sect. 6. The conclusions of this study are mentioned in the Sect. 7.

2 Related Work

Recommendations based on online reviews are essential aspects of consumers' decision-making, as previous consumer decisions are more reliable for forthcoming consumers than those provided by service provider organizations [10]. An online review sometimes helps consumers avoid selecting a particular service that may not suit their expectations [11]. Reviewers who had a high fan following are mostly those who are aware of their fans either to items or services that must be avoided or to the high-quality service or items to purchase [12].

Online consumer reviews are mostly free from any rules to follow while writing their experience, and reviewers can write in a way that better presents their feelings [13]. Reviewers are not bound with any specific tone; they may select a positive or negative manner while mentioning their experience regarding various service aspects, like their experience with the staff, service waiting time, and offered facilities. However, some reviewers externally said their recommendation decisions about service to their readers by using sentences like"I recommend this" or"I will prefer this again," which are often considered in reviews to express to peers that the particular offerings are a good experience [14]. Such endorsements are currently found in related studies as drawings of self-identified scores [15], and recommendation decisions [16]. Although there is no rule as to whether recommendations decisions should be mentioned explicitly or implicitly inside the reviews, it must help draw consumers' recommendation decisions [16]. The recent research article grounds the identification of explicit recommendation decisions on the reviewers posting that the goal is clear to upcoming consumers. However, consumers' understandings and intentions are well known to influence clear ideas so that such a view reflects how consumers rate various service aspects based on their experiences. Consumer sentiment is an essential indicator of consumer satisfaction, precisely the consumers' intention towards the staff's attitude.

The current study on sentiment analysis is based on various machine learning techniques. Pang et al. [17], used a Naíve base (NB), Supprot Vector Machine (SVM), and maximum entropy to evaluate the emotions of film reviews. Their findings show that the SVM algorithm outperforms in dealing with emotions classification for film reviews data. Goldberg and Zhu [18], implemented a graph-based approach to classify sentiment as positive and negative based on ratings. Wang et al. [19] presented the sentiment analysis for the textual data. They executed their study using various features such as emoji, positive features, and negative features and implemented an SVM-based model for mixed feature-based sentiment analysis. The existing literature found that machine learning techniques outperform in sentiment analysis for labeled corpus [20].

Presently many researchers introduced various deep learning techniques for sentiment analysis and obtained better results. The Recursive neural tensor network technique implemented by Socher et al. [21] presented an emotion tree library, which creates a binary tree based on positive and negative sentiment and results in existing models for movie reviews data. The CharSCNN [22] techniques utilized two CNN layers to scrap the various word features and evaluated semantic information to enhance sentiment analysis results for short text data like Twitter. Irsoy and Cardie [23] utilized RNN based on time series, which improved the accuracy of sentiment analysis tasks. Ta et al. [24] implemented a Tree-based LSTM, which had achieved better results for the sentiment analysis task. Baziotis et al. [25] implemented an attention mechanism to LSTM, which results in better accuracy in the study of sentiment classification of SemEval-2017 for Twitter.

Considering various service aspect ratings of online reviews, to utilize sentiment extracted from it, predicting consumer recommendation decisions using rough set theory is presented in this research. This study is the first attempt to implement a rough set theory for predicting recommendation decisions to the best of our knowledge. The findings of this study are helpful for forthcoming consumers and service providers in making their decisions.

3 Research Methodology

Customer recommendation is an essential process for the growth of any business. It is a straightforward approach to know the level of customer satisfaction and improve the quality of individual facilities necessary for business growth. Nowadays a vast data is collected from customer reviews. To know the exact meaning of such collected data, we need to extract the rules of the data. Thus, we use Rough Set Method (RSM) to extract the flight recommendation data set rules. The proposed rough set rule classification framework is given in Fig. 1.

There are two stages: (i) training stage (ii) testing stage. In the training stage, 196 rules were extracted from the training data set. Such rules were used to classify the testing stage. In the testing stage, we fed these 196 rules to predict the flight recommendation rules. We have compared the four different classifiers adopting the same framework.

3.1 Data Collection

Flight recommendation data set is collected from the source http://www.airlinequality.com. This web page contains Airline Reviews and Ratings from all over the world. We have taken a data set of flight recommendations of 2000 customers, having a variety of services, origins, and destinations. Customers' reviews were collected randomly from 2010 to 2020 over seven services with a decision class; we denote each service as an attribute in the data set. The attributes (Cabin Flown, Seat Comfort, Cabin Staff Rating, Food Beverage Ratings, In-flight Entertainment Ratings, Value Money Ratings) have attribute values ranging from 1 to 5 point rating scale, where one means the highly dissatisfied and five means highly satisfied by the service. The attribute (Overall Rating) has attribute values ranging from 1 to 10 point scale, where 1 means the poor quality/

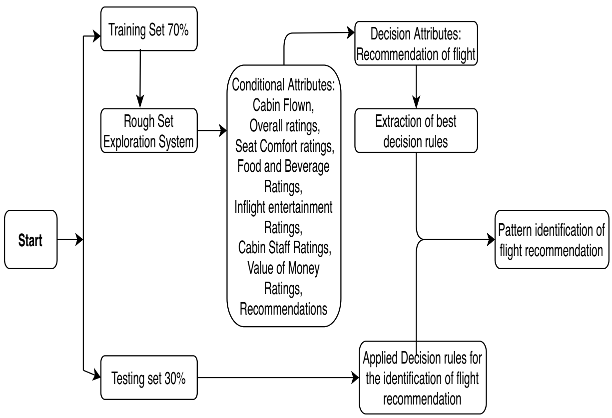


Fig. 1 Flow diagram of rough set model for flight recommendation data set classification

service and 10 means the excellent quality/service. The decision attribute has 0 and 1; 0 means flight is not recommended, and 1 means the flight is recommended.

3.2 Training for Flight Recommendation Data Set

A flight recommendation data set of 2000 customers has been taken with 8 conditional attributes (having alpha-numeric values) and one decision attribute (having numeric values) to identify the hidden pattern of flight recommendation in the data set. We divide the data set into two parts, one for training purposes and the other for testing purposes. In the flight recommendation training data set, the RST is applied to extract the rules. We performed this experiment in Rough Set Exploration System 2.2.2 (RSES 2.2, Warsaw University 2005) software [26], for this experiment, we required a PC with at least 128 MB Ram, 3 MB free Disc space, Windows XP, or any newer version.

3.3 Testing for the Flight Recommendation Data Set

The flight recommendation data set is trained by the RST; for training purposes, we use 70% of data with the same attributes and decision values. Then the rule generated by the training data set (see Table 1) is feed to the test data set (30% of data set) for the recommendation of flight, results are listed in Table 8.

4 Classifier Algorithms

4.1 Local Transfer Function Classifier

Local Transfer Function Classifier (LTF-C) has a three layered network structure with four main parts: (i) changing position of reception field, (ii) changing their sizes, (iii) insertion of new hidden neurons and (iv) removal of unnecessary ones during the training [27, 28]. LTF-C mapping of vector $(x_1, x_2, \ldots, x_n)$ of a pattern X (where X is a vector space) is given as:

$$y'_l = \sum_{i=1}^{m} w'_{il}\ exp\left(-\sum_{j=1}^{n}\left(\frac{w_{ij}-x_j}{r_{ij}}\right)^2\right),$$

where y'_l is the response of the l-th neuron in the output layer, w'_{il} is the weight of the i-th neuron, r_{ij} is the radii of the i-th neuron. It is a simple and fast algorithm [27].

4.2 Decomposition Tree

A tree decomposition of a graph $G(V, E)$ is a tree T, defined as follows:

1. Each vertex i of T is labeled by a subset $Bi \subset V$ of vertices of G, referred to as a “bag”.

Table 1 Decision Rules

S. No.	Matches	Decision Rules
1	156	(cabin flown = Economy) & (overall rating = 1) => (recommended = 0 [156])
2	122	(cabin staff rating = 5) & (food beverages rating = 5) & (overall rating = 10) =>(recommended = 1 [122])
3	88	(cabin flown = Economy) & (seat comfort rating = 1) & (overall rating = 1) =>(recommended = 0[88])
4	85	(cabin flown = Economy) & (cabin staff rating = 4) & (value money rating = 5) =>(recommended = 1[85])
5	81	(cabin staff rating = 5) & (cabin flown = Business Class) & (value money rating = 5) =>(recommended = 1[81])
6	78	(cabin flown = Economy) & (seat comfort rating = 4) & (cabin staff rating = 5) & (value money rating=5) = > (recommended = 1[78])
7	75	(cabin flown = Economy) & (cabin staff rating = 5) & (seat comfort rating = 5) & (value money rating = 5) = > (recommended = 1[75])
8	70	(cabin flown = Economy) & (in-flight entertainment rating = 1) & (food beverages rating = 1) & (seat comfort rating = 1) = > (recommended = 0[70])
9	59	(value money rating = 4) & (food beverages rating = 4) & (cabin staff rating = 4) = >(recommended = 1[59])
10	55	(cabin flown = Economy) & (inflight entertainment rating = 1) & (food beverages rating = 1) & (cabin staff rating = 1) & (overall rating = 1) =>(recommended = 0[55])
11	52	(cabin flown = Economy) & (cabin staff rating = 4) & (seat comfort rating = 4) & (value money rating = 4) => (recommended = 1[52])
12	52	(cabin staff rating = 5) & (food beverages rating = 4) & (seat comfort rating = 5)=>(recommended = 1[52])
13	47	(cabin flown = Economy) & (value money rating = 2) & (seat comfort rating = 1)=>(recommended = 0[47])
14	42	(value money rating = 4) & (seat comfort rating = 4) & (overall rating = 7)=>(recommended = 1[41])
15	39	(cabin flown = Business Class) & (value money rating = 1)=>(recommended = 0[39])
16	24	(cabin flown = Economy) & (value money rating = 1) & (overall rating = 2)=>(recommended = 0[24])
17	23	(cabin flown = Economy) & (food beverages rating = 3) & (value money rating = 4) & (overall rating = 7)=> (recommended = 1[23])
18	20	(cabin flown = Economy) & (value money rating = 4) & (cabin staff rating = 5) & (overall rating = 8) = >(recommended = 1[20])

Table 1 (continued)

S. No.	Matches	Decision Rules
19	20	(cabin flown = Economy) & (value money rating = 4) & (cabin staff rating = 5) & (overall rating = 9) = >(recommended = 1[20])
20	19	(cabin flown = Economy) & (value money rating = 2) & (seat comfort rating = 2) & (cabin staff rating=1) = > (recommended = 0[19])

2. Each edge of G is in a subgraph induced by at least one of the Bi (i.e. is in at least one of the “bags”of T).
3. The subtree of T consisting of all “bags”containing u is connected, for all vertices u in G.

It is important to note that a graph may have several different tree decompositions (tree decompositions are generally not unique). Similarly, the same tree decomposition can be valid for several different graphs.

4.3 k-NN

k-NN is a non parametric, instance based classifier algorithm, introduced by Fix and Hedges in 1951 [29, 30]. k-NN learn by a group of k-samples that are nearest neighbour's to the unknowns [31]. k-NN applies the Euclidean distance or Cosine similarity to find the k- nearest neighbours for the classification of test tuples [32]. Euclidean distance of two tuples $X_1 = \{x_{11}, x_{12}, x_{13}, \dots, x_{1n}\}$ and $X_2 = \{x_{21}, x_{22}, x_{23}, \dots, x_{2n}\}$ is given as:

$$d(X_1, X_2)^2 = \sum_{i=1}^{n} (x_{1i} - x_{2i})^2.$$

The cosine similarity of two vectors, say x and y is given as $cos(x, y) = \frac{(x,y)}{\|x\|.\|y\|}$; where $\|x\|$ is the Euclidean norm of vector $x = \{x_1, x_2, x_3, \dots, x_n\}$ and calculates as $\sqrt{x_1^1 + x_2^2 + x_3^2 + \cdots + x_n^2}$.

4.4 Cross-validation Method

k-fold cross-validation method [33] splits data in k random subsets of approximately equal size, say $D_1, D_2, \dots, D_k$. A subset D_k is tested k-times randomly, each times (say, $t \in 1, 2, \dots, k$) trained by $k-1$ subsets. The k-fold cross-validation method accuracy is the overall number of True classification, divided by the number of instances in the data set. Suppose $D_{(i)}$, is test set with $x_i = \{v_i, y_i\}$ instance, then accuracy can be calculates as

$$Accuracy_{(cv)} = \frac{1}{n} \sum_{(v_i, y_i) \in D} \delta\big(I(D/D_{(i)}, v_i), y_i\big)$$

5 Rough Set Method

Rough set theory (RST) is a branch of mathematics, developed by Pawlak in 1982 [34]. RST is a perfect tool to handle vague information, imperfect knowledge, rule extraction and data mining etc. in the data set [35, 36]. An information system $IS = (U, T)$, where U and T are non-empty finite sets. U is a set of objects, we call this as universe and T is a set of attributes. Every attributes $t \in T$, having values is called domain of attributes. Further, attributes are divided into conditional and decision attributes. Every subset S of T has a binary relation $R(S)$ on U, is known as indiscernibility relation. This relation is defined by $(x, y) \in R(S)$ iff $t(x) = t(y)$ for every $t \in T$, where $t(x)$ defines values of attribute t for element x. Here a partition of U determined by S is indicated by $U/R(S)$, where $R(S)$ is an equivalence relation. If $(x, y) \in R(S)$ this implies that x and y are indiscernible with respect to S. We call "all the equivalence classes of $R(S)$" as granules of S. An information system of data set denoted as $IS = (U, C, D)$, C a condition and D a decision attributes respectively, is known as a decision table. Here, C and D are disjoint sets. Suppose our data set or information system is $IS = (U, T), U \,\&\, S \subseteq T$. Here, we need to describe the set V in terms of attribute values of S, for this we define two sets S_X i.e. lower approximation and S^X i.e. upper approximation.

$$S_X = \cup_{x \in U}\{S(x) : S(x) \subseteq X\}$$

$$S^X = \cup_{x \in U}\{S(x) : S(x) \cap X \neq \emptyset\}$$

And the boundary of the set X, under S-granules $= S^X - S_X$. If the boundary region of X is non-empty then X is roughly defined with respect to S i.e. X is rough otherwise X is crisp.

5.1 Rule Generation for Decision or Calculation of Decision Rule

Let $IS = (U, T)$ be an information system with every $S \subseteq T$, we associate a set of rules $R(S)$. These rules are developed from attribute value pair (t, s), where $t \in T$ and $s \in S_t$. In an information system (IS), a rule is a statement $\chi \rightarrow \psi$ means if χ then ψ , where $\chi \in E(C)$ & $\psi \in R(D)$ i.e. where χ belongs to condition attribute and ψ to the decision attribute. Here we say that χ and ψ are condition and decision of the rule, respectively. Our rule $\chi \rightarrow \psi$ is true in an information system if $\|\chi\|_{IS} \subseteq \|\psi\|_{IS}$ where $\|.\|$ denotes the cardinality. Number of such rules i.e. $\chi \in \psi$ in IS, will be calculated as $card(\|\chi \wedge \psi\|_{IS})$ and known as support of the rule. Where $(\|\chi \wedge \psi\|_{IS}) = \|\chi\|_{IS} \cap \|\psi\|_{IS}$.

For distribution of probability of an element or object x of $U \neq \emptyset$ we define $prob(x) = \dfrac{1}{card(U)}$, where U is universe of our information system and probability distribution of $X \subseteq U$ is $prob = \dfrac{card(X)}{card(U)}$. Now we define probability of any rule associated with information system by $\delta_{IS}(\chi) = prob(\|\psi\|_{IS})$ and cardinal probability of any rule set associated with information system (IS) $\delta_{IS}(\chi) = prob(\|\psi\|_{IS}/\|\psi\|_{IS})$ we call this as certainty factor [37]. The confidence coefficients of a rule $\chi \rightarrow \psi$ is calculated as

$$\delta_{IS}(\psi/\chi) = \frac{card(\|\chi \wedge \psi\|_{IS})}{card(\|\chi\|_{IS})}$$

where $\|\chi\|_{IS} \neq \emptyset$. The confidence will be 100% if $\chi \rightarrow \psi$ is true, i.e. $\delta_{IS}(\psi/\chi) = 1$. If $\delta_{IS}(\psi/\chi) = 1$, then $\chi \rightarrow \psi$ is precise rule, if $0 < \delta_{IS}(\psi/\chi) < 1$ then decision is imprecise. To evaluate the quality of decision rule, we calculate the conditional probability that χ is true in IS, given that ψ is true in IS, as $\delta_{IS}(\chi/\psi) = Prob_U(\|\chi\|_{IS}/\|\psi\|_{IS})$ will be known as coverage factor [reference Tsumoto] [38]. This implies

$$\delta_{IS}(\chi/\psi) = \frac{card(\|\chi \wedge \psi\|_{IS})}{card(\|\psi\|_{IS})}$$

To know the robustness of a rule $\chi \rightarrow \psi$ in IS, we will calculate the number η of IS as:

$$\eta_{IS}(\chi, \psi) = \frac{supp_{IS}(\chi, \psi)}{card(U)} = \delta_{IS}(\psi/\chi) \cdot \delta_{IS}(X).$$

5.2 Decision Algorithm

Let $Dec(IS) = \{\chi \to \psi\}_{i=1}^{n}, n \geq 2$ is a collection of rules in our decision table of $IS = (U, T, S)$

1. if $\chi \to \psi, \chi' \to \psi' \in Dec(IS)$, then either $\chi = \chi'$ or $\|\chi \wedge \chi'\|_{IS} = \emptyset$ and either $\psi = \psi'$ or $\|\psi \wedge \psi'\|_{IS} = \emptyset$, then $Dec(IS)$ is collection of mutually independent decision rules in information system.
2. if $\|\vee_{i=1}^{n} \chi_i\|_{IS} = U$ and $\left\|\vee_{i=1}^{n} \psi_i\right\|_{IS} = U$, then collection of rules i.e. $Dec(IS)$ covers U.
3. if $\chi \to \psi \in Dec(IS)$ and $Supp_{IS}(\chi, \psi) \neq 0$, implies that rule $\chi \to \psi$ is acceptable in the information system(IS).
4. If $\cup_{x \in \frac{U}{S} T_*(X) = } \|\vee_{\chi \to \psi \in Dec^+(IS)} \chi\|_{IS}$, $Dec^+(IS)$ is the set of all certain rules from $Dec(IS)$. Then the collection of such decision rules protect the uniformity of decision table $IS = (U, T, S)$.

The collection of such rules is decision algorithm in IS. We calculate efficiency [39] of decision table by number η as

$$\eta(Dec(IS)) = \sum_{\chi \to \psi \in Dec(IS)} Max\{\delta_{IS}(\chi, \psi)\}_{\psi \in D(\chi)}$$

Where $D(\chi) = \{\psi : \chi \to \psi \in Dec(IS)\}$.

5.3 Extracted Rules and Their Counts

We have taken a data set of 2000 customer reviews of a flight recommendation, with 7 conditional and 1 decision attribute (see Sect. 3.2). On the basis of quality of each service a customer gives reviews and these reviews helps to predict the rule patterns of the data. We have extracted 196 rules from the data set, such rules find the pattern of the flight recommendation class. Here decision attribute has two classes 0 and 1, 1 means flight is recommended and 0 means fight is not recommended. We are exploring rule 6 to know about decision class 1 and rule 10 to know about decision class 0 from the extracted decision rules, see Table 1.

Rule 6 from the extracted rule table is: (cabin flown=Economy) & (seat comfort rating=4) & (cabin staff rating=5) & (value money rating=5)=> (recommended=1[78]). Rule 6 means the customer recommends the flight if cabin flown is economy and seat comfort rating is 4 and cabin staff rating is 5 and value money rating is 5. There is 78 supports are available in the data set for this rule pattern. Rule 10 from the extracted rule table is: (cabin flown=Economy) & (in-flight entertainment rating=1) & (food beverages rating=1) & (cabin staff rating=1) & (overall rating=1)=>(recommended=0[55]). Rule 10 means the customer does not recommends the flight if cabin flown is economy and in-flight entertainment rating is 1 and food beverage rating is 1 and cabin staff rating is 1 and overall rating is 1. 55 supports are available in the data set for this rule pattern.

The extracted rules and statistics of the rough set model are given in Table 2. It shows that 123 rules produce decision class 0 and 74 number of rules produce decision class 1. Total support of decision class 0 is 1056, and support of decision class 1 is 944. It means there are 123 such rules in which the customer does not recommend the flight,

Table 2 Statistics of extracted rules

Number of Rule- 196 Number of attributes = 8		Decision class	Count of decision class	Support of decision rule
Support of rules:	Minimal = 1 Maximal = 156 Mean = 9.8	0	123	1056
Length of rule premises:	Minimal = 2 Maximal = 7 Mean = 6	1	74	944

Fig. 2 Number of rules support decision class

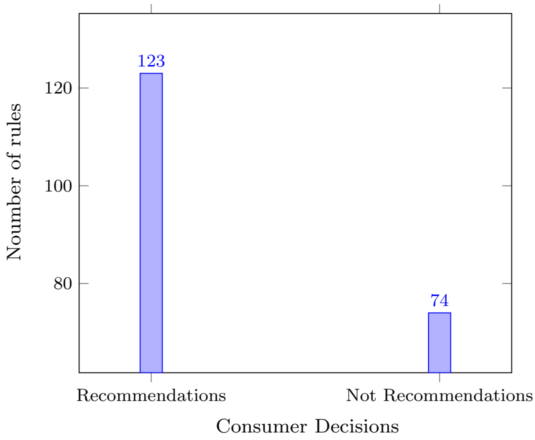
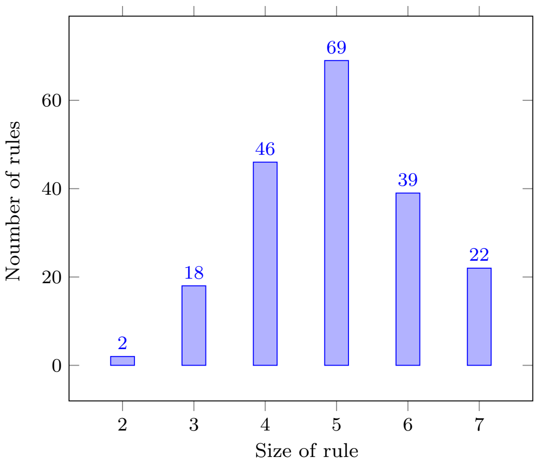


Fig. 3 Rule length in data set



and there are 74 rules in which the customer recommends the flight. Figure 2 shows the number of rules supporting decision classes of flight recommendation data set and length of rules are given in Fig. 3.

5.4 Reduct Set

We get only one reduct set as {*Cabin Flown, Overall Rating, Seat Comfort Rating, Cabin Staff Rating, Food beverage Rating, In-Flight Entertainment Rating, Value Money Rating* } with maximum size 7, containing all attributes and count of each attribute is exactly 1. It shows, there are no redundant attributes in the data set. The mean length and mean of stability coefficients of the reduct are 7 and 0.431, respectively, and the positive region of the reduct is 99.4%.

5.5 Assessment and Accuracy of Extracted Rules

Table 3 shows the assessment of extracted rules. Four possible cases i.e. True positive (TP), False positive (FP), False negative (FN) and True negative (TN) are given in Table 3, with their predicted accuracy. Total accuracy of the extracted rule from the data set = number of correct assessment/number of all assessments.

$$\text{Total accuracy} = \frac{TP + TN}{TP + FP + FN + TN} = \frac{146 + 292}{146 + 3 + 12 + 292} = 96.68\%$$

6 Results

6.1 Experiment Setup

This paper presents knowledge-based learning, a rough set approach for rule extraction of flight recommendation data set. We have extracted 196 rules from the data set of 2000 customers. The performance of ANN, Decomposition Tree, k-NN, *k*-fold cross-validation, and rough set method has been compared based on their accuracy. The experiment has been performed on 2000 customers feedback data sets. Further, the data set were divided into training (70%) and test (30%) data set. Each algorithm was performed using a 10-fold cross-validation protocol; the average of each trial was recorded. Entire experiment has been performed on Intel(R) Core(TM) i3 CPU M380 @ 2.53GHz with 6 GB RAM.

6.2 Discussion

RST finds the pattern of rules in an uncertain data set; hence, we have applied the rough set method to find the ruleset patterns for flight recommendation. Such a rule pattern helps industries to improve service(s) quality. Table 8 shows that RST gives better accuracy than other methods. We have used the Rough Set Exploration System 2.2.2 (RSES 2.2) software to extract the rule pattern of the flight recommendation data set. A view of the extracted rule set is given in Table 1. Total 196 rules are extracted, in which 74 rules recommend the flight, i.e., decision class 1, and 123 rules do not recommend flight, i.e., decision class 0. Decision class 1 is supported by 944 rules, and decision class 0 is supported by 1056 rules, see Table 2. This model predicts consumer behavior accurately up to 96.70%. Here, Rough

Table 3 Assessment of the prediction of different types of decision

Positive outcome	True positive (TP) = 146	False positive (FP) = 3	Accuracy of positive predicted = TP/(TP + FP) = 146/149 = 0.98
Negative outcome	False negative (FN) = 12	True negative (TN) = 292	Accuracy of negative predicted = FN/(FN + TN) = 292/294 = 0.961

Set Exploration System 2.2.2 [26] played a vital role in extracting decision rules from the flight recommendation data set.

We have tested a total of 600 customers feedback, the confusion matrix is given in Table 4 and total sensitivity was found as 96.70%. One hundred forty-six rules are predicted correctly, have been observed as a success (refers in Table 4), in which three rules are observed as a failure. Two hundred ninety-two rules are predicted as a failure and have been observed failure, in which 12 rules are observed as a success. Accuracy of such rules is 96.7% by RSM, while the accuracy of the same data by ANN is 93.70%, Decomposition tree and neural network is 93.6%, see Table 8. The confusion matrix of ANN is given in Table 5, that shows the statistics of True and False predictions.

In decomposition tree method we get total 33 nodes in the tree and Table 6 shows the confusion matrix of decomposition tree with 93.00% accuracy. This matrix shows, 176 rules have been observed as success in which 17 rules observed as a failure. Three hundred thirty-two are predicted as a failure and have been observed as failure in which 17 rules are observed as success. The comparison with the RSMTconfusion matrix (see Table 4) shows the prediction of rules in RSM is more accurate than Decomposition Tree.

We have tested the same data set by k-NN (k-Nearest Neighbour) method. In this method, we split data into ten equal parts, and then each part of the data has been used for

Table 4 Test table using RSM

	Predicted		No. of objects	Accuracy
	0	1		
Actual				
0	146	3	223	0.98
1	12	292	377	0.96

Total number of tested objects: 600

Total accuracy: 0.967

Table 5 Test table using Artificial Neural Network

	Predicted		No. of objects	Accuracy
	0	1		
Actual				
0	211	12	223	0.946
1	26	351	377	0.931

Total number of tested objects: 600

Total accuracy: 0.937

Table 6 Test table using Decomposition Tree

	Predicted		No. of objects	Accuracy
	0	1		
Actual				
0	176	17	223	0.91
1	17	332	377	0.95

Total number of tested objects: 600

Total accuracy: 0.937

Table 7 Test table using Cross validation method

	Predicted		No. of objects	Accuracy
	0	1		
Actual				
0	56	3.1	59.1	0.948
1	4.7	76.2	80.9	0.942

Total number of tested objects: 140

Total accuracy: 0.944

Table 8 The comparative analysis of different methods

S. no.	Methods	Accuracy (%)
1.	**RST**	**96.70**
2.	Cross validation method	94.40
3.	LTF-C Based on ANN	93.70
4.	Decomposition Tree	93.60
5.	k-NN	93.00

Comparative analysis of the accuracy of the flight recommendation by the popular classification algorithms, they are RSM(Rough Set Method), LTF-C: Local Transfer Function Classifier based on artificial neural network architecture, k-NN: k-nearest neighbour, Decomposition Tree and cross validation method

Our main objective is to highlight the accuracy of the proposed method

training and the rest for the test. Distanced-based attribute-weighting has been used with 20 iterations, and we found total accuracy of k-NN is 93.00% in the data set.

In the cross-validation method, we have used a 10-fold cross-validation protocol [40]. Further, we performed 20 trials in the cross-validation method, training and test data set were randomly partitioned at each trial. The obtained accuracy from each trial was averaged, and the averaged accuracy was considered as final accuracy; averaged accuracy is 94.40%. A huge difference can be noticed with RST accuracy. The confusion matrix of cross-validation method is given in Table 7, showing the accuracy and statistics of True & False prediction. When we observe comparative analysis Table 8, RSM accuracy is much higher than the other popular methods.

These results (refer to Table 8) shows the importance of RSM in rule extraction; such study will also be useful in related areas.

7 Conclusions

As per the current research, customer feedback is an essential step for any business/industry growth. Customer feedback helps enterprises improve the quality of a particular service or product. Nowadays, colossal feedback data is generated, so we need to analyze the data. We used flight recommendation data set to know the customer's recommendation from their feedback. This work will help the industry to improve the particular service. We have used the rough set method for the classification of the flight recommendation data set. Rough Set Exploration System 2.2.2 (RSES 2.2) software has been used in this experiment. We have extracted 196 rules from the data set and fed them to train the test data set; the generated decision rules are given in the Table 1.

Further, the performance comparison of five different methods has been investigated. The RST achieved the highest accuracy, 96.70%, on the flight recommendation data set (refer to Table 8). Hence it proves that the RST is a more powerful tool than other approaches to classify patterns in the data set. Such encouraging results show that RST finds better rules to recommend and improve the service(s). We will use more attributes/services and apply the RST to classify the flight recommendation data set in future work.

Funding We did not receive any funding from a private or government body for this research work.

Data Availability http://www.airlinequality.com.

Declarations

Conflict of interest All author's declare that they do not have any conflict of interests.

References

1. Al-Ayyoub, M., Khamaiseh, A. A., Jararweh, Y., & Al-Kabi, M. N. (2019). A comprehensive survey of Arabic sentiment analysis. *Information Processing & Management, 56*(2), 320–342.
2. Song, C., Guo, J., & Zhuang, J. (2020). Analyzing passengers' emotions following flight delays—A 2011–2019 case study on skytrax comments. *Journal of Air Transport Management, 89,* 101903.
3. Korfiatis, N., Stamolampros, P., Kourouthanassis, P., & Sagiadinos, V. (2019). Measuring service quality from unstructured data: A topic modeling application on airline passengers' online reviews. *Expert Systems with Applications, 116,* 472–486.
4. Tansitpong, P. (2020). Determinants of recommendation in the airline industry: An application of online review analysis. In: International conference on decision support system technology, Springer (pp. 125–135).
5. Jain, P. K., Pamula, R., & Srivastava, G. (2021). A systematic literature review on machine learning applications for consumer sentiment analysis using online reviews. *Computer Science Review, 41,* 100413.
6. Huddar, M. G., Sannakki, S. S., & Rajpurohit, V. S. (2019). A survey of computational approaches and challenges in multimodal sentiment analysis. *International Journal of Computational Science and Engineering, 7*(1), 876–883.
7. Liu, B. (2020). *Sentiment analysis: Mining opinions, sentiments, and emotions*. Cambridge University Press.
8. Siering, M., Deokar, A. V., & Janze, C. (2018). Disentangling consumer recommendations: Explaining and predicting airline recommendations based on online reviews. *Decision Support Systems, 107,* 52–63.
9. Kuhn, M., & Johnson, K. (2019). *Feature engineering and selection: A practical approach for predictive models*. CRC Press.
10. Filieri, R., Alguezaui, S., & McLeay, F. (2015). Why do travelers trust tripadvisor? Antecedents of trust towards consumer-generated media and its influence on recommendation adoption and word of mouth. *Tourism Management, 51,* 174–185.
11. Bronner, F., & De Hoog, R. (2011). Vacationers and ewom: Who posts, and why, where, and what? *Journal of Travel Research, 50*(1), 15–26.
12. Chen, H., & Rahman, I. (2018). Cultural tourism: An analysis of engagement, cultural contact, memorable tourism experience and destination loyalty. *Tourism Management Perspectives, 26,* 153–163.

13. Wattanacharoensil, W., Schuckert, M., Graham, A., & Dean, A. (2017). An analysis of the airport experience from an air traveler perspective. *Journal of Hospitality and Tourism Management, 32,* 124–135.
14. Jain, P. K., Yekun, E. A., Pamula, R., & Srivastava, G. (2021). Consumer recommendation prediction in online reviews using cuckoo optimized machine learning models. *Computers & Electrical Engineering, 95,* 107397.
15. Villarroel Ordenes, F., Ludwig, S., De Ruyter, K., Grewal, D., & Wetzels, M. (2017). Unveiling what is written in the stars: Analyzing explicit, implicit, and discourse patterns of sentiment in social media. *Journal of Consumer Research, 43*(6), 875–894.
16. Packard, G., & Berger, J. (2017). How language shapes word of mouth's impact. *Journal of Marketing Research, 54*(4), 572–588.
17. Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment classification using machine learning techniques. arXiv:cs/0205070
18. Goldberg, A. B.,&Zhu, X. (2006). Seeing stars when there aren't many stars: Graph-based semi-supervised learning for sentiment categorization. In: Proceedings of TextGraphs: The first workshop on graph based methods for natural language processing, (pp. 45–52).
19. Wang, Y.-Z., Zheng, X., Hou, D., & Hu, W. (2018). Short text sentiment classification of high dimensional hybrid feature based on svm. *Computer Technology for Development, 28*(2), 88–93.
20. Boiy, E., & Moens, M.-F. (2009). A machine learning approach to sentiment analysis in multilingual web texts. *Information Retrieval, 12*(5), 526–558.
21. Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C.D., Ng, A.Y., & Potts, C. (2013). Recursive deep models for semantic compositionality over a sentiment treebank. In: Proceedings of the 2013 conference on empirical methods in natural language processing, (pp. 1631–1642).
22. Dos Santos, C., & Gatti, M. (2014). Deep convolutional neural networks for sentiment analysis of short texts. In: Proceedings of COLING 2014, the 25th international conference on computational linguistics: Technical papers, (pp. 69–78).
23. Irsoy, O., & Cardie, C. (2014). Opinion mining with deep recurrent neural networks. In: Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), (pp. 720–728).
24. Tai, K.S., Socher, R., & Manning, C.D. Improved semantic representations from tree-structured long short-term memory networks. arXiv:1503.00075.
25. Baziotis, C., Pelekis, N., & Doulkeridis, C. (2017). Datastories at semeval-2017 task 4: Deep lstm with attention for message-level and topic-based sentiment analysis. In: Proceedings of the 11th international workshop on semantic evaluation (SemEval-2017), (pp. 747–754).
26. Rses 2.2 user's guide 2005 warsaw university, http://logic.mimuw.edu.pl/~rses/get.html, Warsaw University (2005)
27. Wojnarski, M. (2003). Ltf-c: Architecture, training algorithm and applications of new neural classifier. *Fundamenta Informaticae, 54,* 89–105.
28. Ouyang, Y., & Ma, J. (2006). Classification of multi-spectral remote sensing data using a local transfer function classifier. *International Journal of Remote Sensing, 27*(24), 5401–5408.
29. Peterson, L. E. (2009). K-nearest neighbor. *Scholarpedia, 4*(2), 1883.
30. Jiawei, H., & Kamber, M. (2001). *Data mining: Concept and techniques* (pp. 223–224). Kaufmann Publishers.
31. Cover, T., & Hart, T. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory, 13*(1), 21–27. https://doi.org/10.1109/TIT.1967.1053964.
32. Adeniyi, D., Wei, Z., & Yongquan, Y. (2016). Automated web usage data mining and recommendation system using k-nearest neighbor (KNN) classification method. *Applied Computing and Informatics, 12*(1), 90–108.
33. Kohavi, R. et al. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Ijcai, Vol. 14, Montreal, Canada (pp. 1137–1145).
34. Pawlak, Z. (1982). Rough sets. *International Journal of Computer & Information Sciences, 11,* 341–356.
35. Pawlak, Z. (1991). *Rough sets-theoretical aspects of reasoning about data, Vol. 9 of Theory and decision library: series D*. Kluwer.
36. Pawlak, Z. (1997). Rough set approach to knowledge-based decision support. *European Journal of Operational Research, 99*(1), 48–57.
37. Łukasiewicz, J. (1970). Die logishen grundlagen der wahrscheinilchkeitsrechnung. In I. Borkowski (Ed.), *Jan łukasiewicz-selected works*. North Holland Publishing Company.

38. Tsumoto, S. (1998). Modelling medical diagnostic rules based on rough sets. In L. Polkowski & A. Skowron (Eds.), *International conference on rough sets and current trends in computing* (pp. 475–482). Berlin: Springer.
39. Pawlak, Z. (2001). Rough sets and decision algorithms. In W. Ziarko & Y. Yao (Eds.), *Rough sets and current trends in computing* (pp. 30–45). Berlin: Springer.
40. Jain, P. K., Quamer, W., Pamula, R., & Saravanan, V. (2021). Spsan: Sparse self-attentive network-based aspect-aware model for sentiment analysis. *Journal of Ambient Intelligence and Humanized Computing*. https://doi.org/10.1007/s12652-021-03436-x.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Praphula Kumar Jain is a Ph.D. researcher at the Department of Computer Science and Engineering, Indian Institute of Technology (ISM), Dhanbad, JH, India. He started working as a Assistant Professor in the Department of Computer Engineering & Applications, GLA University, Mathura, UP, India. He obtained a B.E. degree in Computer Science and Engineering at the Faculty of Computer Science and Engineering, RGPV, Bhopal, MP, India, and an M.Tech degree in Computer Science and Engineering at the Indian Institute of Technology (ISM), Dhanbad, JH, India. His Ph.D. research focuses on Machine Learning methods for various applications. He has 3 years of teaching experience and has published many technical articles in international scholarly journals such as ACM Transactions on Asia and Low Resource Language information processing, Computer Science Review Elsevier, Computers and Electrical Engineering Elsevier, Wireless personal communication springer, International Arab Journal of Information Technology, SN Applied Sciences Springer, and Opsearch Springer. He is also associated with reputed journals like the Journal of supercomputing, AI & Ethics, AI & Society, and conferences like FSDM, GLOBECOM 2020 as a reviewer. He is very much interested in social activity and providing free-of-cost education to underprivileged students.

Anand Prakash is working as an Assistant Professor in the Department of Mathematics at Shri Vaishnav Vidyapeeth Vishwavidyalaya, Indore, M. P. He has done his Ph.D. from the National Institute of Technology, Raipur, CG. His research area includes Rough Set, Fuzzy set, and machine learning.