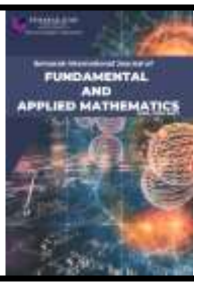
Semarak International Journal of Fundamental and Applied Mathematics 6, Issue 1 (2025) 67-86

Semarak International Journal of Fundamental and Applied Mathematics

Journal homepage:
https://semarakilmu.my/index.php/sijfam/index
ISSN: 3030-5527



Exploring Airline Passenger Satisfaction Behaviours Using a Supervised Reverse Analysis Logic Mining Model with Weighted Random 2-Satisfiability

Nur 'Afifah Rusdi[1], Nor Syafiqah Warianto[2], Mohd Shareduwan Mohd Kasihmuddin[2,*], Yueling Guo[3], Nur Ezlin Zamri [4]

[1] Institute of Engineering Mathematics, Universiti Malaysia Perlis, 02600 Arau, Perlis, Malaysia
[2] School of Mathematical Sciences, Universiti Sains Malaysia, 11800 USM, Penang, Malaysia
[3] School of Science, Hunan Institute of Technology, 421002 Hengyang, China
[4] Department of Mathematics and Statistics, Faculty of Science, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

ARTICLE INFO

Article history:
Received 18 March 2025
Received in revised form 19 April 2025
Accepted 21 May 2025
Available online 30 June 2025

Keywords:
Airline satisfaction; weighted random 2 satisfiability; discrete hopfield neural network; supervised logic mining; knowledge extraction

ABSTRACT

Passenger satisfaction plays a crucial role in the profitability and service quality of the airline and travel industries as it directly influences customer loyalty and overall business success. To maintain a competitive edge, airlines collect extensive amounts of customer feedback. However, identifying the most impactful factors remains a challenge, leading to inefficiencies in improving service quality. Traditional approaches struggle to extract meaningful patterns, making it essential to develop more effective analytical methods. To address this challenge, this study proposes artificial neural network through logic mining approach to provide knowledge extraction to analyze factors contribute to airline passenger satisfaction during in-flight services, focusing on enhancing both accuracy and interpretability. In this study, Weighted Random 2-satisfiability logic was used to represent the attributes and Discrete Hopfield Neural Network is employed as a computational model. The proposed model also implemented a statistical correlation test during the pre-processing phase to select optimal attributes. To improve the efficiency of the proposed model, the metaheuristic algorithm is introduced at two critical stages. In the logic phase, the Artificial Bee Colony algorithm is employed to ensures the optimal distribution of negative literals within the logical clauses while in the learning phase, Election algorithm is implemented to enhance the process of obtaining optimal synaptic weights. Experimental results show that the model performs strongly, achieving an Accuracy of 84.85% and a Specificity of 95.39%. In conclusion, the findings validate the effectiveness of the proposed logic mining model in extracting knowledge from airline passenger satisfaction data. By enhancing accuracy and interpretability, this approach serves as a valuable tool for airlines to refine service quality, make informed decisions, and enhance passenger satisfaction over time.

* *Corresponding author.*
E-mail address: shareduwan@usm.my

https://doi.org/10.37934/sijfam.6.1.6786

1. Introduction

Recent advancements in data mining have introduced logic mining as an alternative method for uncovering patterns and knowledge from datasets using logical rules. Existing research on logic mining has shown promising results in improving model transparency and interpretability for classification tasks. According to Zamri *et al*., [1], the logic mining model consists of three main components including Satisfiability (SAT), Discrete Hopfield Neural Network (DHNN) and Reverse Analysis method (RA). SAT was first introduced by Abdullah [2] as a symbolic rule to govern the neurons in DHNN and addressed the issue of black-box model without compromising network behavior. In this work, Horn Satisfiability (HornSAT) was proposed to represent the neuron in DHNN. However, the structure of HornSAT which consist of disjunction of literals with at most one positive literal in each clause lead to potential redundancy of literals in HornSAT. When this SAT is implemented in DHNN, this results in ineffective neuron connections which reduces the functionality of the neurons.

Over the years, researchers have proposed different variants of SAT in DHNN which can be categorized as systematic SAT and non-systematic SAT. The systematic SAT is commonly known as kSAT which consist of strictly k literals in each clause. For example, Kasihmuddin *et al.,* [3] and Mansor *et al.*, [4] proposed 2 Satisfiability (2SAT) and 3 Satisfiability (3SAT) in DHNN. When considering different logical structure, 3SAT outperforms 2SAT in obtaining higher global minima ratio. This observation indicates that having a higher number of literals in each clause increases the possibility for a clause to be satisfied. Despite demonstrating good compatibility with DHNN and producing higher global minima solutions, systematic SAT lacks flexibility in generating diverse final neuron states [5]. The restricted literals within the clause make it easier to approximate the synaptic weights, resulting in similar patterns of final neuron states. This, in turn, leads to the overfitting issue.

To address this issue, Sathasivam *et al.,* [5] proposed non-systematic SAT in DHNN known as Random k Satisfiability (RANkSAT) which consists of first order and second order clauses. Initially, RAN2SAT showed promising potential by offering SAT with a non-restrictive k, promoting more variations of the synaptic weights in terms of magnitude. However, the positive and negative literals are generated randomly, similar to previous SATs. As a result, RAN2SAT overlooks the impact of different distributions of positive and negative literals in the logical rule on neuron connections in DHNN. Additionally, researchers often neglect negative literals because they are typically associated with faults or errors in the SAT formula [6]. This reduces the ability to fully understand neuron behaviours in DHNN and results in poorly governed neuron connections due to the omission of negative literals within the logical rule. Consequently, this leads to suboptimal synaptic weights, particularly in terms of negativity. Recognizing this gap, the role of negative literals are crucial to improve the expressivity and interpretability of SAT in DHNN. Therefore, a key research gap in the literature is the lack of an approach that dynamically distributes negative literals within SAT clauses to better control neuron behaviour and enhance interpretability.

With the development of new formulation of SAT in DHNN, various logic mining models based on Reverse Analysis method has been proposed in literature. The implementation of logic mining models has been shown to improve interpretability and accuracy in addressing real-life classification tasks. Earlier in 2020, Alway *et al.,* [7] utilized 2 Satisfiability Reverse Analysis method (2SATRA) to analyze the price trends of palm oil in relation to other commodities. In this study, a real-life dataset of Malaysia's commodity prices, consisting of seven types of commodities, will be utilized, with each commodity represented as a neuron in DHNN. Although 2SATRA employed only Exhaustive Search (ES) as the learning algorithm, the proposed model successfully generated multiple induced logic rules and identified key components influencing palm oil price trends. Besides that, Jamaludin *et al.,*

[8] introduced the Energy-Based *k*-Satisfiability Reverse Analysis (E*k*SATRA) method to identify relationships between factors influencing systematic recruitment in the electronic (E) recruitment dataset. In this study, E*k*SATRA only focused on induced logic that achieve global minimum energy using different $kSAT$ logical rules. Experimental simulations with varying numbers of neurons demonstrated the effectiveness and reliability of E*k*SATRA in identifying the dominant attributes influencing positive recruitment in Malaysia's insurance industry. Given the strong potential of logic mining models in classification and knowledge extraction, further improvements and advancements are needed to enhance their ability to handle classification tasks across diverse real-world datasets.

For example, a new logic mining model based on 2SATRA with Genetic Algorithm (GA) was proposed in handling online shoppers' purchasing intentions. In this study, 2SAT logic was used to represent the shopping attribute and GA was implemented to optimize the logical rule throughout the learning phase to reduce the learning complexity of the proposed model. The induced logic obtained from this approach serves as a valuable tool for knowledge extraction in the online shopping industry, enabling businesses to enhance sales strategies and improve customer service. Notably, all the previously mentioned logic mining models only considered systematic SAT and random attribute selection method to select attributes to represent the dataset. Although the induced logic generated by systematic SAT is considered ideal, the restricted literals and randomize positive and negative literals within the clauses may reduce the interpretability of the output. Additionally, the use of randomized attribute selection increases the chances of selecting insignificant attributes. As a result, the quality of the induced logic may be compromised, limiting its effectiveness in generalizing the dataset.

Despite the promising results of existing logic mining models, a significant gap exists in addressing the dynamic distribution of negative literals within SAT clauses, which directly impacts the interpretability and flexibility of DHNN models. Additionally, the existing literature often overlooks the selection and proper weighting of attributes, which limits the models' ability to generalize across diverse real-world datasets. This gap is crucial because suboptimal attribute selection can cause models to underperform, while improper weighting may lead to an overemphasis or underemphasis of certain attributes, ultimately compromising the accuracy and interpretability of the induced logic. Considering these challenges, this paper proposes a new logic mining model that incorporates non-systematic SAT with a dynamic distribution of negative literals within clauses to analyse airline passenger satisfaction behaviours. Additionally, supervised learning approach using correlation test will be implemented in screening and selecting significant attributes to represent the dataset and Election Algorithm will be implemented in the learning phase to ensure optimal learning phase of the network. By considering all the elements in the proposed model, the proposed model aims to generate high-quality induced logic capable of extracting valuable insights from the airline passenger satisfaction dataset. Therefore, the main contributions of this paper are outlined as follows:

i) To propose Weighted Random 2 Satisfiability in Discrete Hopfield Neural Network as a neuron representation. The proposed logical rule incorporates a weighted feature and employs the Artificial Bee Colony Algorithm to generate a dynamic distribution of negative literals based on the desired ratio. By applying the proposed logical rule, the neurons in the Discrete Hopfield Neural Network are structured effectively, which leads to improved performance in non-systematic logic.
ii) To propose a logic mining model namely Supervised Weighted Random 2 Satisfiability Reverse Analysis in doing classification task for airline passenger satisfaction dataset. Election Algorithm is implemented in the learning phase to ensure optimal synaptic weights are obtained. Thus, the proposed algorithm enhances the quality of the retrieved induced

logic, with the extracted best-induced logic identifying the key attributes that influence the outcome of the dataset.

iii) To assess the performance of Supervised Weighted Random 2 Satisfiability Reverse Analysis by considering the evaluation of confusion matrix classifications based on the retrieved final induced logic to the actual outcome of airline passenger satisfaction dataset. The optimal induced logic is based on the highest retrieval frequency from all *k*-folds cross validation.

The organization of the paper is as follows: Section 2 presents the methodology, covering the formulation of the proposed logic, its implementation in DHNN, and an explanation of the proposed logic mining model. Section 3 describes the experimental setup while section 4 provides the results and discussion. Finally, Section 5 concludes the paper and suggests directions for future research.

2. Methodology

This section presents the proposed methodology for extracting knowledge related to airline passenger satisfaction dataset. It covers the formulation of Weighted Random 2-Satisfiability (*r*2SAT), the general concepts of *r*2SAT in DHNN (DHNN-*r*2SAT), and the Supervised Weighted Random 2-Satisfiability Reverse Analysis.

2.1 Weighted Random 2 Satisfiability

According to Zamri *et al.,* [9], Weighted Random k Satisfiability (*r*2SAT) is a non-systematic SAT consist of first order and second order clauses. Additionally, the weighted feature in *r*2SAT determine the number of negative literals within the clauses. The general formulation for *r*2SAT is presented as follows:

$$\Gamma_{r2SAT} = \wedge_{i=1}^{u} J_i^{(k=2)} \wedge_{i=1}^{v} J_i^{(k=1)}. \tag{1}$$

Referring to Eq. (1), $J_i^{(k=2)}$ and $J_i^{(k=1)}$ represents the second order and first order clauses respectively with u and v denoting the maximum number of literals within the clauses. The total clauses of Γ_{r2SAT} is denoted as d such that $d = u + v$. In general, the clauses of Γ_{r2SAT} with k order logics can be rewritten as shown in Eq. (2).

$$\Gamma_{r2SAT} = J_1^{(k=2)} \vee J_2^{(k=2)} \vee J_u^{(k=2)} \vee J_1^{(k=1)} \vee J_2^{(k=1)} \vee J_v^{(k=1)}. \tag{2}$$

The literals within the clauses of $J_i^{(k=2)}$ and $J_i^{(k=1)}$ are presented as A_i, B_i and C_i as follows:

$$J_i^{(k=1,2)} = \begin{cases} \left(A_i \vee B_i\right), & k = 2 \\ C_i, & k = 1 \end{cases} \tag{3}$$

Based on Eq. (3), the literals in Γ_{r2SAT} are unique and each literal can be either a positive $\{A_i, B_i, C_i\}$ or a negative literal $\{\neg A_i, \neg B_i, \neg C_i\}$. In this case, each literal can be represented as bipolar values of $\{1, -1\}$. The total number of literals in Γ_{r2SAT} can be denoted as λ_{r2SAT} in Eq. (4), where *u* and *v* represent the total number of second-order and first-order clauses, respectively.

$$\lambda_{r2SAT} = 2u + v \tag{4}$$

Additionally, a key component of *r*2SAT is the controlled distribution of negative literals using a weighted ratio, r. The main idea of introducing r is to promote control and diversification of negative literals in the logic. By defining r, a dynamic number of negative literals can be generated for any given λ_{r2SAT}. The total number of negative literals can be computed based on Eq. (5) as follows:

$$N_v = r\lambda \tag{5}$$

Referring to Eq. (5), λ represents the total number of literals in Γ_{r2SAT} while r is a weighted ratio defined between 0.1 and 0.9 with a step size, $\Delta r = 0.1$. In this context, Γ_{r2SAT} allows randomized allocation of negative literals with respect to λ and r. For example, if $r = 0.5$ and $\lambda = 10$, Γ_{r2SAT} can be generalized as follows:

$$\Gamma_{r2SAT} = (A_1 \vee B_1) \wedge (A_2 \vee \neg B_2) \wedge (A_3 \vee B_3) \wedge (A_4 \vee B_4) \wedge \neg C_1 \wedge C_2 \tag{6}$$

In this case, assigning negative literals effectively before entering DHNN can be challenging, especially when the value of λ is large. To address this, the proposed logic phase provides a systematic approach for generating the appropriate Γ_{r2SAT} structure for any given λ and r. Due to the complexity introduced by high-dimensional data, metaheuristic methods are necessary to support the logic phase. In this study, Artificial Bee Colony (ABC) algorithm is implemented in the logic phase of r2SAT to ensure correct generation of Γ_{r2SAT}. Hence, the generated Γ_{r2SAT} will act as a symbolic rule to represent the neurons in DHNN. The next section explained the process involved in DHNN.

2.2 Weighted Random 2 Satisfiability in Discrete Hopfield Neural Network

The implementation of *r*2SAT as neuron in DHNN can be denoted as DHNN-*r*2SAT model. Generally, DHNN consists of two main phases, the learning phase and testing phase. In the learning phase, the associated logical rule will be learned by the network via Wan Abdullah (WA) method to ensure optimal synaptic weight [2]. In consideration to the WA method, the cost function of DHNN-*r*2SAT can be formulated as shown in Eqs. (7)-(8) as follows:

$$\Phi_{\Gamma_{r2SAT}} = \frac{1}{4}\sum_{i=1}^{u}\left(\prod_{j=1}^{2} Z_{ij}\right) + \frac{1}{2}\sum_{i=1}^{v}\left(\prod_{j=1}^{1} Z_{ij}\right) \tag{7}$$

such that *u* and *v* represent the total number of second-order and first-order clauses, respectively. The inconsistency of Γ_{r2SAT} can be defined as follows:

$$Z_{ij} = \begin{cases} (1 - S_X) & \text{if} \neg X \\ (1 + S_X) & \text{if } X \end{cases} \tag{8}$$

Referring to Eq.(8), *X* consists of arbitrary literals of $\{A_i, B_i, C_i\}$. In this context, the value obtained from Eq. (7) correspond to the number of unsatisfied clauses such that $\Phi_{\Gamma_{r2SAT}} = 0$ indicates that every clause in the Γ_{r2SAT} is satisfied while $\Phi_{\Gamma_{r2SAT}} \neq 0$ depicts the number of unsatisfied clauses. Thus, $\min\left(\Phi_{\Gamma_{r2SAT}}\right)$ will guarantee optimal synaptic weights in DHNN. By considering WA method, the synaptic weight values can be obtained by directly comparing the cost function in Eq. (7) with the Lyapunov energy function of DHNN in Eq. (9). Then, an associative memory feature in DHNN known as Content Addressable Memory (CAM), stores the generated synaptic weights in matrix form before they are retrieved in the retrieval phase.

$$H_{\Gamma_{r2\mathrm{SAT}}} = -\frac{1}{2}\sum_{i=1,i\neq j}^{N}\sum_{j=1,i\neq j}^{N} W_{ij}^{(2)} S_i S_j - \sum_{i=1}^{N} W_i^{(1)} S_j \tag{9}$$

During the retrieval phase, the neuron states are updated through local field computation by considering the synaptic weight values obtained during the learning phase. The local field computation of Γ_{r2SAT} can be formulated as follows:

$$h_i = \sum_{j=1,i\neq j}^{N} W_{ij}^{(2)} S_j - W_i^{(1)} \tag{10}$$

Then, the final neuron states, S_i^f can be retrieved by squashing the values obtained from Eq. (10) by using Hyperbolic Tangent Activation Function (HTAF) as shown in Eq. (11)- Eq. (12).

$$\tanh\left(h_i\right) = \frac{e^{h_i} - e^{-h_i}}{e^{h_i} + e^{-h_i}} \tag{11}$$

$$S_i^f = \begin{cases} 1, & \tanh\left(h_i\right) \geq 0 \\ -1, & \text{otherwise} \end{cases} \tag{12}$$

After all S_i^f are successfully retrieved, the quality of the states will be measured based on the final and expected $H_{\Gamma_{r2\mathrm{SAT}}}$ as follows:

$$\left| H_{\Gamma_{r2\mathrm{SAT}}} - H_{\Gamma_{r2\mathrm{SAT}}}^{\min} \right| \leq Tol \tag{13}$$

Referring to Eq. (13), if the difference between the energy value is less that predefined tolerance value Tol, the retrieved S_i^f is classified as global minimum energy or else the S_i^f is trapped in local minimum energy. According to Sathasivam [10], Tol is set to be 0.001 to reduce the errors within the neuron states. In this case, it is crucial for the final neuron states to obtain global minima energy because it will reflect the satisfiable property of Γ_{r2SAT}. To enhance the network's convergence toward final neuron states that achieve global minimum energy, a metaheuristic algorithm is required during the learning phase of DHNN. This is because the optimal synaptic weights obtained during learning increase the chances of retrieving final neuron states that reach global minimum energy. Several studies in the literature have implemented Election Algorithm (EA) in the learning phase of DHNN to improve the process of finding satisfied interpretations associated with logical rules and based on the findings, EA shown strong performance in obtaining optimal synaptic weights [11-13]. Therefore, this study implemented EA in the learning phase to minimize Φ_{r2SAT} by finding satisfied interpretation of Γ_{r2SAT}. All fundamental components of Γ_{r2SAT} have been discussed and will serve as the foundation for the proposed logic mining model.

2.3 Supervised Weighted Random 2 Satisfiability Reverse Analysis

The development of the proposed Supervised Weighted Random 2 Satisfiability Reverse Analysis (S-*r*2SATRA) involves several key components. This section explores how S-*r*2SATRA is applied to classify and extract knowledge on factors influencing airline passenger satisfaction behaviours. The process is divided into four main phases: pre-processing, logic phase, learning phase, and retrieval phase. The pre-processing phase includes data preparation, attribute selection, and data splitting. During data preparation, the dataset is transformed into a bipolar form of 1 and −1. This transformation is essential since real-world classification problems often involve qualitative, quantitative, or mixed data types. To facilitate this, the *k*-means clustering approach is used to convert the data into a bipolar from as shown in Eq. (14) as follows:

$$S_i = \begin{cases} 1, & \text{if } \; e_{ij} \geq \bar{E}_i, \\ -1, & \text{otherwise.} \end{cases} \tag{14}$$

where e_{ij} represent the values within the attributes, E_i while $\overline{E_i}$ referred to the mean value of each attribute in the dataset.The next step in data preparation involves handling missing values. In this study, the missing values of the dataset will be replaced with random value of 1 and -1 since the percentage of missing value in the dataset is very small.

The next step in the pre-processing phase is selecting the attributes to be learned by the network. Identifying significant attributes is crucial for enhancing the interpretability of the extracted knowledge. Inspired by the work of Kasihmuddin *et al.,* [14] and Rusdi *et al.,* [15], this study implements a correlation analysis test to systematically select representative attributes. This approach is more structured than random selection, as the correlation test serves as a benchmark for identifying the most relevant attributes. In this method, the correlation coefficient measures the association between each attribute and the output of the dataset. Attributes that meet or exceed the predefined threshold are considered significant and selected for further analysis. The selection process involves comparing significance value for each attribute obtained through the correlation test, with a predefined threshold value α. Attributes are selected only if their significance value is less than α. In this case, the learning logic, Γ_i can be generalized as in Eq. (15). This systematic approach enhances the performance of S-r2SATRA by improving interpretability.

$$\Gamma_i = \bigwedge_{i=1, i \neq j}^{m} \left(E_i^{\min|p_i|} \vee E_j^{\min|p_j|} \right) \bigwedge_{k=1}^{m} \left(E_k^{\min|p_k|} \right). \tag{15}$$

Referring to the above Eq. (15), the attributes will only be selected if it satisfies $0 \leq p_i \leq \alpha, 0 \leq p_j \leq \alpha, 0 \leq p_k \leq \alpha$, where $\alpha = 0.05$. In this case, 10 best attributes from the correlation tests will be selected to represent the literals of Γ_{r2SAT} with selected outcome of the dataset is denoted as Γ_d. Additionally, this study incorporates permutation operator during the pre-processing phase [16]. This operator aims to enhance the connectivity among attributes in the clause by rearranges the order of the attributes. The rearrangement of the attributes will generate new combinations and potentially identify new patterns that were not previously discovered. Consequently, this strategy indirectly expands the search space in finding optimal induced logic. The final step in this pre-processing phase involves splitting the datasets into learning data, Γ_{learn} and testing data, Γ_{test}. Therefore, this study considered splitting ratio of $60:40$ whereby 60% of the data will be used as learning data to learn the pattern of the dataset 40% of the data will be used to validate the retrieved induced logic during the retrieval phase. This ratio has good agreement with previous existing works. In addition, *k*-fold cross validation is considered in this study as an alternative way in addressing imbalance dataset. The average of all folds will be taken as the result for all metrics considered [17].

The second phase in the development of the proposed S-r2SATRA is called logic phase. During this phase, *r*2SAT will be developed based on various *r* where $r=[0.1,0.9]$ with $\Delta r=0.1$. Notably, Artificial Bee Colony (ABC) efficiently optimizes the distribution of negative literals in r2SAT. Analysing all possible *r*2SAT structures is crucial to ensure the proposed logic mining model does not overlook any potential neuron connections, especially when negative linkages are placed adjacent to other literals. Each *r*2SAT structure, corresponding to different values of *r*, will proceed to a separate DHNN learning phase. The next stage is the learning phase, which consists of two main processes. Firstly, the best logic, Γ^{best} is determined by considering Eq. (16) with 60% learning data, Γ_{learn}. In this case, Γ_d based on Γ_i is denoted as predicted class and Γ_d based on Γ_{learn} is denoted as actual class. By comparing Γ_d based on predicted class and actual class, the confusion matrix of True Positive (*TP*) and True Negative (*TN*) will be obtained. Then, Γ^{best} is selected based on Γ_i that achieve maximum TP and TN such that:

$$\Gamma^{best}=\max[TP+TN]. \tag{16}$$

Secondly, DHNN will learn the generated Γ_i^{best} and minimize the cost function of the network by considering Eq. (7). To enhance the process of finding a satisfied interpretation that leads to a zero-cost function, EA is utilized in this phase. Next, the WA method is used to evaluate the synaptic weight values which represent the neuron connections between all attributes in the dataset. The optimal synaptic weight values retrieved through the learning phase with EA will then be stored in the CAM.

The final stage in the development of the proposed logic mining model is the retrieval phase, which aims to produce induced logic that represents the overall behaviour of the dataset. In this context, induced logic refers to the optimal final neuron states produced by the DHNN. This phase involves four consecutive steps, which follow from the previous steps. Firstly, the synaptic weight values stored in CAM will be retrieved and use to update the neuron states through local field computation as discussed in Eq. (10). Then, the values will be squash into bipolar form of 1 and -1 by considering Eq. (11) and the final neuron states are obtained based on Eq. (11). Secondly, the final neuron states, S_i^f will be transformed in form of logical rule, $S_i^{induced}$ based on Eq. (17) as follows:

$$S_i^{induced} = \begin{cases} X, & S_i^f = 1 \\ \neg X, & S_i^f = -1 \end{cases}. \tag{17}$$

Referring to Eq. (17), X consists of arbitrary literals of $\{A_i, B_i, C_i\}$. Then, the combination of $S_i^{induced}$ will produced induced logic, $\Gamma_i^{induced}$. Following this, the next step in this phase is to validate these induced logics by using 40% Γ_{test}. By comparing Γ_d based on predicted class, $\Gamma_i^{induced}$ and actual class, Γ_{test}, the confusion metric of TP, TN, False Positive (FP) and False Negative (FN) will be obtained. *TP* are the cases when both the actual and predicted Γ_d are positive, while *TN* refers to both being negative. *FP* denotes cases where the predicted Γ_d is positive but the actual Γ_d is negative whereas *FN* occurs when the predicted Γ_d is negative but the actual Γ_d is positive. All these components of the confusion matrix are utilized in the calculation of various performance metrics. In this study, several performance metrics will be measured including accuracy (*ACC*), precision (*PRE*), specificity (*SPE*), F1-Score (*F1*), and Matthew correlation coefficient (*MCC*) to assess the efficiency of the proposed logic mining model S-*r*2SATRA for different k fold. The final step in this phase is to select the best induced logic, $\Gamma_{best}^{induced}$. The $\Gamma_{best}^{induced}$ will be selected based on the repeated values of induced logic. The repeated values of induced logic show consistency of the induced logic representing the patterns of dataset. Figure 1 presents the flowchart of the overall process involved in S-*r*2SATRA.

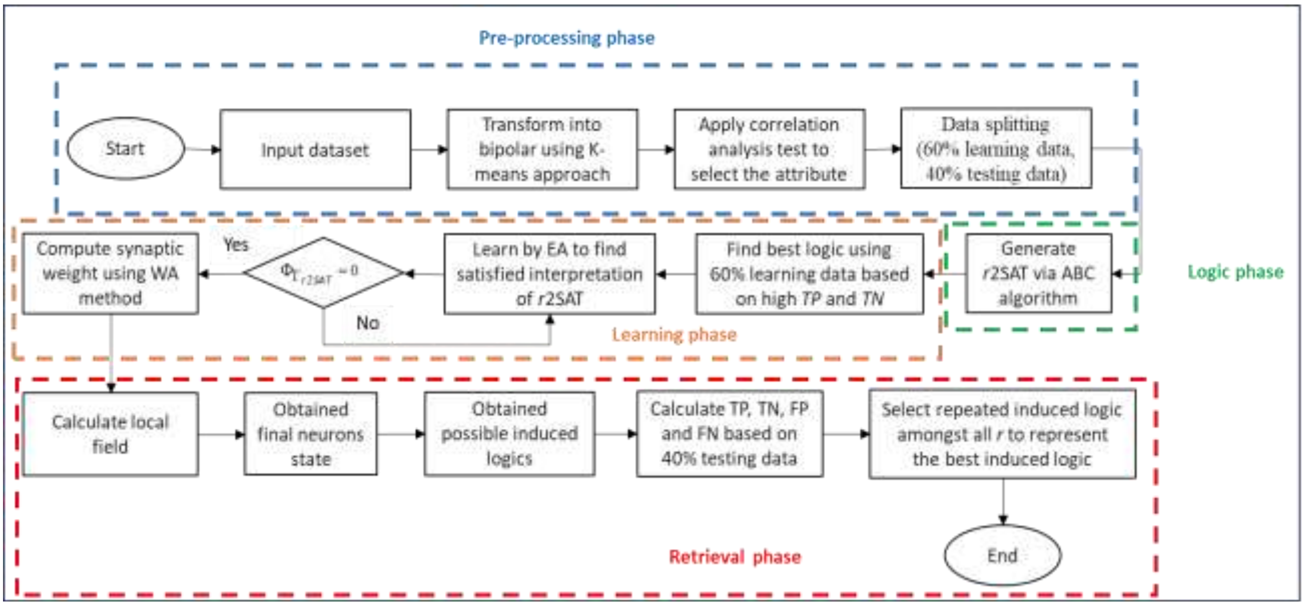


Fig. 1. Flowchart of overall process in S-r2SATRA

3. Experimental Setup

This section presents the experimental setup, including the experimental design for S-*r*2SATRA, data description, and performance metric measurement. Each aspect is detailed in the following subsections.

3.1 Experimental Design

The experimental design focuses on conducting experiments from a logic mining perspective. Simulations were executed using Microsoft Dev C++ Version 5.11, chosen for its user-friendly interface and open-source accessibility. To ensure unbiased result interpretation, all simulations were performed on a single personal computer. For consistency, experiments should be conducted with the same compiler settings and on devices with comparable processing capabilities. Following this, the parameter setting of S-*r*2SATRA througout this study is shown as in Table 1.

Table 1
Parameter settings involved in S-*r*2SATRA

Parameter/Notation	Value/Remark
Range of *r*	[0.1, 0.9] [12]
Logic phase algorithm	ABC [17]
Attribute selection	Correlation analysis test
Number of learning	100 [23]
Learning algorithm	EA
K-fold cross validation	5
Number of trials	100 [25]
$\Gamma_{learn} : \Gamma_{test}$	60: 40 [5]
Logical condition	Γ_{r2SAT} [12]
Best logic	Max [*TP* + *TN*] [5]
Best induced logic	Highest repeated logic

3.2 Data Description

The Airline Passenger Satisfaction dataset was obtained from a reputable machine learning repository, Kaggle.com (https://www.kaggle.com/datasets), specifically from the airline service category. This dataset comprises an airline passenger satisfaction survey with a total of 25,976 entries. For this study, 15,586 entries are used as Γ_{learn} to learn patterns, while 10,390 entries are used as Γ_{test} for validation. The dataset consists of 23 attributes, where 22 represent various passenger-related factors, and one serves as the outcome variable. In this study, inflight service satisfaction is selected as Γ_d. The class outcomes are denoted as {1, -1}. The details of the attributes are presented in Table 2 until Table 4.

Table 2
Attributes of airline passenger satisfaction dataset

Attribute label	Name of attributes	Description	Bipolar class
E_1	Gender	Female and male passengers of airline.	$E_1 = \begin{cases} 1, & \text{if } e_{1_{ij}} = \text{"Female"} \\ -1, & \text{otherwise} \end{cases}$
E_2	Age	The different age groups between younger and older passenger.	$E_2 = \begin{cases} 1, & \text{if } e_{2_{ij}} \geq 39 \\ -1, & \text{otherwise} \end{cases}$

Table 3
Attributes of airline passenger satisfaction dataset

Attribute label	Name of attributes	Description	Bipolar class
E_3	Class	Two types of class for the passengers (business class and economy class)	$E_3 = \begin{cases} 1, & \text{if } e_{3_{ij}} = \text{"Business"} \\ -1, & \text{otherwise} \end{cases}$
E_4	Inflight Wi-Fi service	The satisfaction of passengers over the availability of Wi-Fi services.	$E_4 = \begin{cases} 1, & \text{if } e_{4_{ij}} = 0 \text{ or } e_{4_{ij}} = 1 \text{ or } e_{4_{ij}} = 2 \\ -1, & \text{otherwise} \end{cases}$
E_5	Inflight Entertainment	In-flight entertainment refers to the various entertainment provided to passengers.	$E_5 = \begin{cases} 1, & \text{if } e_{5_{ij}} = 0 \text{ or } e_{5_{ij}} = 1 \text{ or } e_{5_{ij}} = 2 \\ -1, & \text{otherwise} \end{cases}$
E_6	On-board service	Various amenities provided to passengers during a flight.	$E_6 = \begin{cases} 1, & \text{if } e_{6_{ij}} = 0 \text{ or } e_{6_{ij}} = 1 \text{ or } e_{6_{ij}} = 2 \\ -1, & \text{otherwise} \end{cases}$
E_7	Leg room service	The amount of space provided for passenger legs between seats.	$E_7 = \begin{cases} 1, & \text{if } e_{7_{ij}} = 0 \text{ or } e_{7_{ij}} = 1 \text{ or } e_{7_{ij}} = 2 \\ -1, & \text{otherwise} \end{cases}$
E_8	Baggage handling	Baggage handling involves the process of transporting passenger luggage.	$E_8 = \begin{cases} 1, & \text{if } e_{8_{ij}} = 1 \text{ or } e_{8_{ij}} = 2 \text{ or } e_{8_{ij}} = 3 \\ -1, & \text{otherwise} \end{cases}$
E_9	Satisfaction	Overall satisfactions based on the rating received.	$E_9 = \begin{cases} 1, & \text{if } e_{9_{ij}} = \text{"Satisfied"} \\ -1, & \text{otherwise} \end{cases}$
E_{10}	Customer type	Type of customer that repeated the service or not.	$E_{10} = \begin{cases} 1, & \text{if } e_{10_{ij}} = \text{"Loyal"} \\ -1, & \text{otherwise} \end{cases}$
E_{11}	Inflight service	The services that customers get from the cabin crew.	$E_{11} = \begin{cases} 1, & \text{if } e_{11_{ij}} = 0 \text{ or } e_{11_{ij}} = 1 \text{ or } e_{11_{ij}} = 2 \\ -1, & \text{otherwise} \end{cases}$
E_{12}	Type of class trip	Two types of class trip which is for business and personal trips	$E_{12} = \begin{cases} 1, & \text{if } e_{12_{ij}} = \text{"Business trip"} \\ -1, & \text{otherwise} \end{cases}$
E_{13}	Departure/Arrival time convenient	Favourable suitable of scheduled times for flight departure and arrival.	$E_{13} = \begin{cases} 1, & \text{if } e_{13_{ij}} = 0 \text{ or } e_{13_{ij}} = 1 \text{ or } e_{13_{ij}} = 2 \\ -1, & \text{otherwise} \end{cases}$
E_{14}	Ease of online booking	The straightforward and user-friendly of online flight booking process.	$E_{14} = \begin{cases} 1, & \text{if } e_{14_{ij}} = 0 \text{ or } e_{14_{ij}} = 1 \text{ or } e_{14_{ij}} = 2 \\ -1, & \text{otherwise} \end{cases}$

Table 4
Attributes of airline passenger satisfaction dataset

Attribute label	Name of attributes	Description	Bipolar class
E_{15}	Gate location	Service for passengers while waiting for their flights to departure and arrives.	$E_{15} = \begin{cases} 1, & \text{if } e_{15_{ij}} = 1 \text{ or } e_{15_{ij}} = 2 \text{ or } e_{15_{ij}} = 3 \\ -1, & \text{otherwise} \end{cases}$
E_{16}	Food and beverages	Service food and drinks for long flight or for those buy the package during flight.	$E_{16} = \begin{cases} 1, & \text{if } e_{16_{ij}} = 0 \text{ or } e_{16_{ij}} = 1 \text{ or } e_{16_{ij}} = 2 \\ -1, & \text{otherwise} \end{cases}$
E_{17}	Online boarding	Online check-in, boarding pass and ease of selecting seat before flight.	$E_{17} = \begin{cases} 1, & \text{if } e_{17_{ij}} = 0 \text{ or } e_{17_{ij}} = 1 \text{ or } e_{17_{ij}} = 2 \\ -1, & \text{otherwise} \end{cases}$

E_{18}	Cleanliness	Cleanliness of airplane interior including seats, restrooms and common area.	$E_{18}=\begin{cases}1, & \text{if } e_{18_{ij}}=0 \text{ or } e_{18_{ij}}=1 \text{ or } e_{18_{ij}}=2\\ -1, & \text{otherwise}\end{cases}$
E_{19}	Departure delayed in minutes	The number of minutes the flight departure was delayed beyond its scheduled time.	$E_{19}=\begin{cases}1, & \text{if } e_{19_{ij}}\geq 14.30\text{min}\\ -1, & \text{otherwise}\end{cases}$
E_{20}	Arrival delayed in minutes	The number of minutes the flight arrival was delayed beyond its scheduled time.	$E_{20}=\begin{cases}1, & \text{if } e_{20_{ij}}\geq 14.74\text{min}\\ -1, & \text{otherwise}\end{cases}$
E_{21}	Flight distance	The distance of the flights to the destinations.	$E_{21}=\begin{cases}1, & \text{if } e_{21_{ij}}\geq 1194\text{km}\\ -1, & \text{otherwise}\end{cases}$
E_{22}	Check In services	The check-in process at the airport includes the efficiency and ease of navigating check-in procedures.	$E_{22}=\begin{cases}1, & \text{if } e_{22_{ij}}=1 \text{ or } e_{22_{ij}}=2 \text{ or } e_{22_{ij}}=3\\ -1, & \text{otherwise}\end{cases}$
E_{23}	Seat comfort	Factors like cushioning, leg room and overall support.	$E_{23}=\begin{cases}1, & \text{if } e_{23_{ij}}=1 \text{ or } e_{23_{ij}}=2 \text{ or } e_{23_{ij}}=3\\ -1, & \text{otherwise}\end{cases}$

3.2 Evaluation Metrics

In terms of metric evaluation performance, five standard classification metrics are used to measure the performance of the proposed model. This includes accuracy (*ACC*), precision (*PRE*), specificity (*SPE*), Matthew Correlation Coefficient (*MCC*), and F1 score (*F1*). *ACC* measures the ability of the model to correctly classify *TP* and *TN* with optimal value of 1 [18]. The *ACC* value may be measured using Eq. (18):

$$ACC=\frac{(TP+TN)}{(TP+TN+FP+FN)} \tag{18}$$

PRE is used to assess the the accuracy of positive predictions made by the model [19]. High PRE values indicate a lower rate of FP which ensure that the model's predictions are reliable. The formulation of PRE is as follows:

$$PRE=\frac{TP}{(TP+FP)} \tag{19}$$

Next, the ratio of *TN* to total *TN* and *FN* is displayed using *SPE* values. As a result, *SPE* evaluates how well the model can predict the results [20]. Here is how *SPE* is formulated:

$$SPE=\frac{TN}{(TN+FP)} \tag{20}$$

F1 is also a statistic used to assess accuracy. F1 represents the modulation index for the sensitivity and precision parameters [21]. Formula of F1 is shown as follows:

$$F1=\frac{2(PRE\times SEN)}{(PRE+SEN)} \tag{21}$$

Finally, the Matthews Correlation Coefficient (*MCC*) which was created by Matthew in 1975 was used in this experiment [22]. *MCC* includes all four confusion matrix forms. Formula of *MCC*:

$$MCC=\frac{((TP\times TN)-(FP\times FN))}{\sqrt{((TP+FP)(TP+FN)(TN+FP)(TN+FN))}} \tag{22}$$

4. Results and Discussion

The performance of S-r2SATRA on handling Airline passanger satisfaction behaviours will be discussed in this section. Table 5 presented the significant values obtained for all attributes based on the correlation test analysis. Typically, an attribute is considered highly correlated with the outcomes if the significant value obtained is less than 0.05. From all attributes analysed, ten will be selected for this study as *r*2SAT consists of ten literals. Thus, attributes with significant values below 0.05 will be chosen. Each attribute will correspond to each literal in the logic of *r*2SAT. Subsequently, the results of five performance metrics such as values of *ACC*, *PRE*, *SEN*, *F1* and *MCC* are summarized based in Table 6. Additionally, the classification of the confusion matrix based on actual and predicted Γ_d for all 5 folds were displayed in Figure 2 until Figure 6.

Table 5
Attributes of airline passenger satisfaction dataset

Attributes	Labelled	Significant value
Gender	A_1	5.0436×10^{-13}
Age	B_1	5.6889×10^{-24}
Class	A_2	6.4598×10^{-51}
Inflight Wi-Fi service	$\neg B_2$	8.1134×10^{-68}
Inflight Entertainment	A_3	0.000
On-board service	B_3	0.000
Leg room service	A_4	0.000
Baggage handling	B_4	0.000
Satisfaction	$\neg C_1$	2.3977×10^{-99}
Customer Type	C_2	6.3376×10^{-15}
Inflight Service	$\Gamma_{best}^{induced}$	

Table 6
The results of *ACC, PRE, SEN*, F1 and *MCC* by S-r2SATRA in analysing the airline passenger satisfaction

Fold	1	2	3	4	5
ACC	0.84706	0.84851	0.84851	0.84783	0.84783
PRE	0.63041	0.64279	0.62356	0.63112	0.63112
SPE	0.95704	0.95730	0.95393	0.95501	0.95501
F1	0.44069	0.45423	0.44835	0.45389	0.45389
MCC	0.38471	0.39808	0.38756	0.39452	0.39452

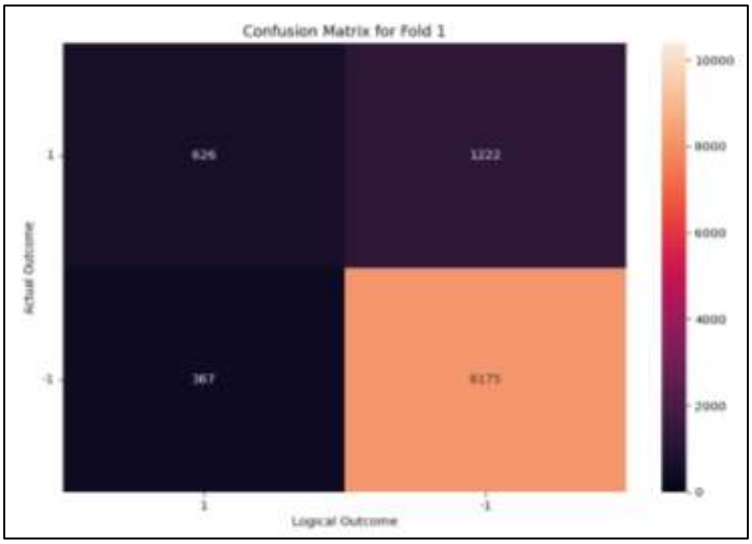


Fig. 2. Confusion matrix classification of S-r2SATRA- fold 1

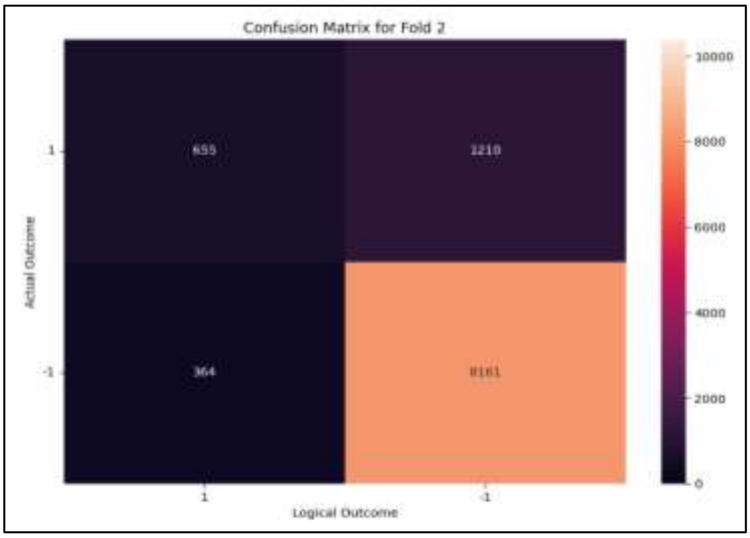


Fig. 3. Confusion matrix classification of S-r2SATRA- fold 2

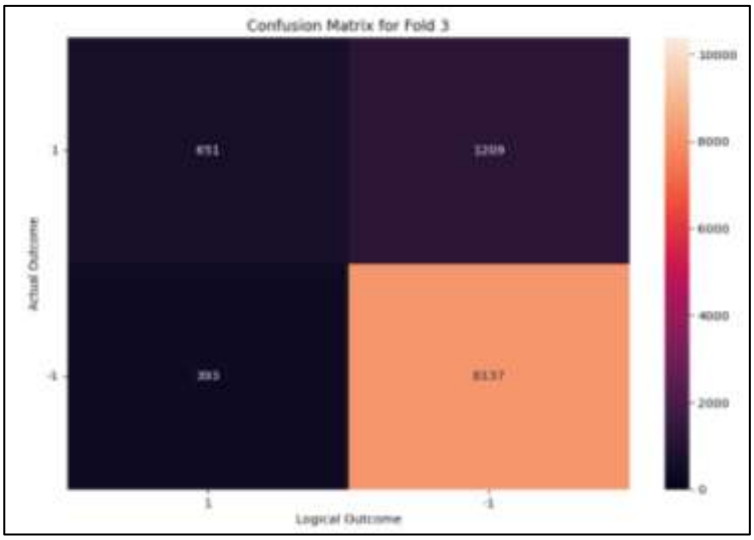


Fig. 4. Confusion matrix classification of S-r2SATRA- fold 3

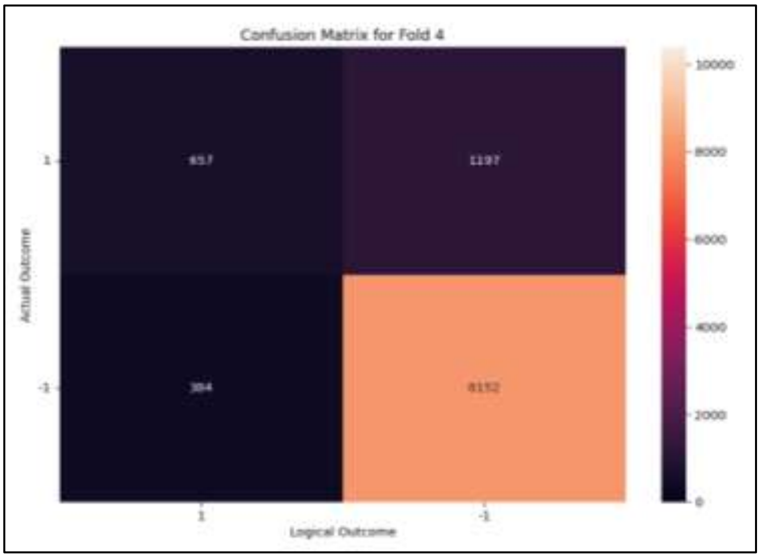


Fig. 5. Confusion matrix classification of S-r2SATRA- fold 4

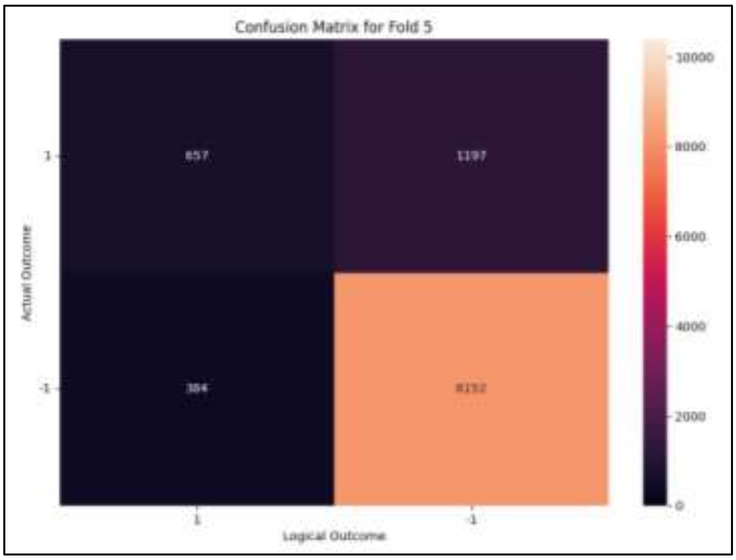


Fig. 6. Confusion matrix classification of S-r2SATRA- fold 5

Based on the illustrated figures, the logical outcome aligns with the predicted outcome generated by the proposed logic mining model, S-*r*2SATRA. In these figures, darker-colored boxes represent lower values in the confusion matrix. Overall, S-*r*2SATRA captures the highest number of TN. This pattern indicates that the model is more inclined to classify samples as negative. This tendency may be due to the dynamic structure of *r*2SAT, particularly its focus on the distribution of negative literals. With the highest *TN* across all folds, S-*r*2SATRA achieves *ACC* and *SPE* above 80% for all folds. However, this emphasis on capturing *TN* leads to an increase in *FP*, as observed in the figures. The higher FP count negatively impacts PRE, resulting in lower *PRE* values. Consequently, the reduced precision lowers the F1. Besides that, *MCC* considers all elements of the confusion matrix, including *TP*, *TN*, *FP*, and *FN*. Achieving a balanced distribution of these elements is crucial for obtaining an optimal *MCC* value. This metric is important for assessing the efficiency of a logic mining model in classification, as it helps determine whether the proposed S-r2SATRA model performs similarly, worse, or better than a random classifier [15]. Based on the *MCC* values, the lower values obtained for all folds indicate that the proposed S-*r*2SATRA performs better than a random classifier.

The retrieved $\Gamma_i^{induced}$ for all 5-folds are presented as follows:

$$\Gamma_{fold1}^{induced} = (\neg A_1 \vee B_1) \wedge (A_2 \vee \neg B_2) \wedge (A_3 \vee B_3) \wedge (A_4 \vee B_4) \wedge \neg C_1 \wedge C_2 \quad (23)$$

$$\Gamma_{fold2}^{induced} = (\neg A_1 \vee B_1) \wedge (A_2 \vee \neg B_2) \wedge (A_3 \vee B_3) \wedge (A_4 \vee B_4) \wedge \neg C_1 \wedge C_2 \quad (24)$$

$$\Gamma_{fold3}^{induced} = (A_1 \vee B_1) \wedge (A_2 \vee \neg B_2) \wedge (A_3 \vee B_3) \wedge (A_4 \vee B_4) \wedge \neg C_1 \wedge C_2 \quad (25)$$

$$\Gamma_{fold4}^{induced} = (A_1 \vee B_1) \wedge (A_2 \vee \neg B_2) \wedge (A_3 \vee B_3) \wedge (A_4 \vee B_4) \wedge \neg C_1 \wedge C_2 \quad (26)$$

$$\Gamma_{fold5}^{induced} = (A_1 \vee B_1) \wedge (A_2 \vee \neg B_2) \wedge (A_3 \vee B_3) \wedge (A_4 \vee B_4) \wedge \neg C_1 \wedge C_2 \quad (27)$$

Fold 3 have repeated value of induced logic between others fold. Hence, fold 3 had been chosen as the best induced logic, $\Gamma_{best}^{induced}$ from the 5-fold cross validation.

$$\Gamma_{best}^{induced} = (A_1 \vee B_1) \wedge (A_2 \vee \neg B_2) \wedge (A_3 \vee B_3) \wedge (A_4 \vee B_4) \wedge \neg C_1 \wedge C_2 \quad (28)$$

Based on Eq. (28), the possible deduction for airline passenger satisfaction behaviour can be presented as follows:

Case 1: Based on A_1 or B_1 and $\neg C_1$

The higher airplane satisfactions ratings for in-flight services $\left(\Gamma_{best}^{induced} = -1\right)$ are significantly influenced by factors such as male passengers $(A_1 = -1)$. Particularly, those over the age of 39 $(B_1 = 1)$ prefer luxury facilities for comfort and convenience. Older male travellers may have greater travel experience giving them a better knowledge of what makes good service. They are more inclined to value and recognise high-quality services and facilities. Men passengers might have a high tolerance for service inconsistencies or be less likely to report dissatisfactions. Additionally, they could have more money available to them which would allow them to travel more frequently and in higher classes, which usually offer better service. Airlines that match these expectations through timely flights and professional crew are more likely to obtain high service satisfaction ratings $(\neg C_1 = 1)$ from male passengers. As a result, male passengers may be more inclined to evaluate their trip and positive services satisfaction rating $(\neg C_1 = 1)$.

Therefore, airlines should focus on customising their offerings to fulfil and surpass the anticipations of this particular consumer base. Indeed, excellent service quality can increase levels of customer satisfactions leading to consumer retention [23]. Airlines should undertake frequent feedback surveys focusing on male customers over the age of 39 to learn about their preferences and potential areas for economic growth. Additionally, recommendation for airlines company such as Malaysian Airlines is that the company can offer exclusive loyalty programs, special promotions and targeted marketing campaigns to appeal to older male passengers. Staff training programmes should be improved to emphasise personalised service and attention to detail in order to meet the demands of older passengers. Aside from that, upgrading premium service packages with perks such as high-quality meals and complementary services will help to increase customer satisfaction. Airlines may also look at collaborating with high-end companies to offer exclusive in-flight amenities and produce a distinctive and engaging experience. By focusing on these strategies, airlines may improve the entire experience for this crucial demographic potentially increasing their reputation and customer loyalty. Therefore, it will lead to higher sales with excellent service quality.

Case 2: Based on B_1, A_2 and A_3.

Younger passengers within the range of lower than 39 years old $(B_1 = -1)$ are often more adaptive and technologically aware can significantly enhance their happiness with economy class $(A_2 = -1)$ services when their preferences are matched resulting in higher rating in inflight services satisfaction $(\Gamma_{best}^{induced} = -1)$. Younger travellers frequently prefer economy since more affordable. This age range has grown accustomed to a digital lifestyle and places a high importance on connection and in-flight entertainment $(A_3 = -1)$. They are more inclined to appreciate and make full use of advanced in-flight entertainment systems like watching movies, Wi-Fi connectivity, and mobile applications. An airline ability to provide these technological soothes and make travellers flights more productive and enjoyable can have a substantial impact on younger passenger satisfaction levels. Younger passengers especially millennials and Gen Z always prioritize value for money. They are more likely to be happy if they believe they are getting excellent value from their flight experience, even if they are in economy class. The offering of essential amenities like spotless and comfortable sitting, delicious meals, dependable and fast service are all included in the affordable ticket prices. When these basic needs are well-addressed, passengers feel they are receiving good value for money which significantly enhances their overall satisfactions.

Additionally, recommendations for the airline company should modern their inflight entertainment systems by provide a diverse selection of movies, music and interactive games. This can accommodate younger passenger diverse interests. Touchscreen interfaces with user-friendly navigation are particularly appealing. Furthermore, airlines should provide USB charging connections and power outlets at each seat allowing customers to charge their gadgets during the trip. Next, adding tiered economy services such as premium economy with extra amenities like priority boarding and expanded food selections might attract customers ready to spend a little more for increased comfort. Long flights might be more comfortable if the airline provides ergonomic chairs with adjustable headrests and more legroom in the economy class. Well-trained workers who are attentive to customer demands could significantly enhance the travel experience. Younger passengers are frequent users of social media sites where they passionately share their experiences. Good evaluations have the power to boost a brand's reputation and draw in additional clients. As a result, airlines who continuously provide great services and facilities to passengers can profit from good word-of-mouth produced by viral content, improving brand awareness and revenue by attracting new consumers.

Case 3: Based on A_2 and A_4 or B_4.

Business class $(A_2 = 1)$ provides a wide variety of luxuries and services that help customers be more satisfied with their in-flight services $(\Gamma_{best}^{induced} = -1)$. Adequate legroom services $(A_4 = -1)$ adds to a more comfortable travel by allowing passengers to stretch and relax especially on long-haul flights. This comfort is especially appreciated by business travellers, who frequently seek to maximise productivity or relax during flights. Airlines show their dedication to providing comfortable and well-received traveller experiences and cultivating favourable brand impressions by providing wide seating layouts. Next, business class travellers frequently have preferential access to amenities including priority baggage handling $(B_4 = -1)$ and designated check-in desks. Passengers appreciate airlines that prioritise safe, timely delivery of their luggage since it reduces the stress and frustration associated with missing or delayed bags. Business class passengers, in particular frequently travel with precious things or critical documents that they must have for professional or personal reasons. These benefits not only reduce waiting times and inconvenience but also make travelling easier and more pleasurable which raises customer satisfaction with the airline company offerings. This reflects the airline cabin crews professionalism and attention to detail. Airlines could examine the following tips to improve in-flight service and increase customer happiness. First and foremost, investing in the design and layout of business class cabins to maximise legroom and comfort may considerably improve the entire passenger experience. Rearranging the seats or adding cutting-edge features like lie-flat mattresses or adjustable footrests may be necessary to achieve this. Furthermore, airlines should prioritise employee training and operational efficiency to guarantee that baggage handling systems run smoothly. Implementing modern tracking technologies and communication protocols can help reduce the risk of luggage maltreatment while also increasing transparency for travellers. Prioritising these characteristics allows airlines to stand out in a competitive market and build long-term loyalty among business travellers and premium passengers. As a conclusion, airlines should prioritise enhancing priority services and providing high-quality services to satisfy specific requirements of various customer demographics to enhance overall passenger satisfaction. Airlines can greatly improve the travel experience and generate more income by implementing these focused methods into practice for a variety of customer categories.

5. Conclusion

In conclusion, this study successfully achieved the three main objectives outlined in the research. First, the study proposed the Weighted Random 2 Satisfiability (*r*2SAT) in DHNN as a neuron representation. The introduction of the *r*2SAT rule, incorporating a weighted feature and using the ABC algorithm to generate a dynamic distribution of negative literals, improved the structural organization of the neurons in DHNN, thereby enhancing performance in non-systematic logic. The optimized SAT was then incorporated as a symbolic rule to govern neuron behaviour in DHNN. The second objective focused on proposing the Supervised Weighted Random 2 Satisfiability Reverse Analysis (S-r2SATRA) as a logic mining model for classification tasks. The formulation of S-r2SATRA involved four main phases including pre-processing phase, logic phase, learning phase and retrieval phase. The implementation of the EA in the learning phase ensured that optimal synaptic weights were obtained. For the third objective, the study assessed the performance of S-r2SATRA by evaluating the classification outcomes using a confusion matrix. This process demonstrated the capability of S-r2SATRA in accurately classifying and extracting meaningful insights from the dataset.

The extracted logic not only aids in understanding general dataset patterns but also provides concrete evidence of the factors influencing airline passenger satisfaction. By analysing these factors, airline companies can improve inflight service satisfaction and potentially increase sales. Ultimately, the findings of this study contribute to a better understanding of what makes a computational model intelligent. In summary, this study contributes to the development of an advanced logic mining model that enhances classification accuracy and knowledge extraction. The findings not only provide insights into the key factors influencing airline passenger satisfaction but also demonstrate the effectiveness of incorporating non-systematic SAT and optimization algorithms in neural network models.

Despite promising results shown by the proposed S-r2SATRA, there are several limitations of this study which can be the basis consideration for the future works. First, this study only considered first- and second-order logic in capturing dataset behaviour. Future research could explore higher-order logic, which has a greater potential for obtaining satisfied interpretations [4]. Second, future studies could prioritize logic with the highest performance metric values, aligning with previous work that proposed different performance metrics for this purpose [17]. Additionally, future research could explore alternative attribute selection methods based on unsupervised learning approaches, such as the Jaccard similarity index [1]. Furthermore, future works can implement this model on different datasets to assess its capability in classification and knowledge extraction. This would provide a more comprehensive understanding of the model's generalizability and effectiveness. Additionally, future research should emphasize how this model can be made accessible and beneficial to practitioners outside the field, providing guidance on its practical application.

Acknowledgement
This research was not funded by any grant.

References

[1] Zamri, Nur Ezlin, Mohd Asyraf Mansor, Mohd Shareduwan Mohd Kasihmuddin, Siti Syatirah Sidik, Alyaa Alway, Nurul Atiqah Romli, Yueling Guo, and Siti Zulaikha Mohd Jamaludin. "A modified reverse-based analysis logic mining model with weighted random 2 satisfiability logic in discrete hopfield neural network and multi-objective training of modified niched genetic algorithm." *Expert Systems with Applications* 240 (2024): 122307. https://doi.org/10.1016/j.eswa.2023.122307

[2] Abdullah, Wan Ahmad Tajuddin Wan. "Logic programming on a neural network." *International journal of intelligent systems* 7, no. 6 (1992): 513-519. https://doi.org/10.1002/int.4550070604

[3] Kasihmuddin, Mohd Shareduwan Mohd, Mohd Asyraf Mansor, and Saratha Sathasivam. "Hybrid Genetic Algorithm in the Hopfield Network for Logic Satisfiability Problem." *Pertanika Journal of Science & Technology* 25, no. 1 (2017).

[4] Mansor, M. A., M. S. M. Kasihmuddin, and S. Sathasivam. "Artificial Immune System Paradigm in the Hopfield Network for 3-Satisfiability Problem." *Pertanika Journal of Science & Technology* 25, no. 4 (2017). https://doi.org/10.9781/ijimai.2017.448

[5] Sathasivam, Saratha, Mohd Asyraf Mansor, Ahmad Izani Md Ismail, Siti Zulaikha Mohd Jamaludin, Mohd Shareduwan Mohd Kasihmuddin, and Mustafa Mamat. "Novel random k satisfiability for k≤ 2 in hopfield neural network." *Sains Malays* 49, no. 11 (2020): 2847-2857. https://doi.org/10.17576/jsm-2020-4911-23

[6] Saribatur, Zeynep G., and Thomas Eiter. "Omission-based abstraction for answer set programs." *Theory and Practice of Logic Programming* 21, no. 2 (2021): 145-195. https://doi.org/10.1017/S1471068420000095

[7] Alway, Alyaa, Nur Ezlin Zamri, Mohd Shareduwan Mohd Kasihmuddin, Mohd Asyraf Mansor, and Saratha Sathasivam. "Palm Oil Trend Analysis via Logic Mining with Discrete Hopfield Neural Network." *Pertanika Journal of Science & Technology* 28, no. 3 (2020).

[8] Jamaludin, Siti Zulaikha Mohd, Mohd Shareduwan Mohd Kasihmuddin, Ahmad Izani Md Ismail, Mohd Asyraf Mansor, and Md Faisal Md Basir. "Energy based logic mining analysis with hopfield neural network for recruitment evaluation." *Entropy* 23, no. 1 (2020): 40. https://doi.org/10.3390/e23010040

[9] Zamri, Nur Ezlin, Siti Aishah Azhar, Mohd Asyraf Mansor, Alyaa Alway, and Mohd Shareduwan Mohd Kasihmuddin. "Weighted random k satisfiability for k= 1, 2 (r2SAT) in discrete hopfield neural network." *Applied Soft Computing* 126 (2022): 109312. https://doi.org/10.1016/j.asoc.2022.109312

[10] Sathasivam, Saratha. "Upgrading logic programming in Hopfield network." *Sains Malaysiana* 39, no. 1 (2010): 115-118.

[11] Roslan, Muhammad Aqmar Fiqhi, Nur Ezlin Zamri, Mohd Asyraf Mansor, and Mohd Shareduwan Mohd Kasihmuddin. "Major 3 Satisfiability logic in Discrete Hopfield Neural Network integrated with multi-objective Election Algorithm." *AIMS Mathematics* 8, no. 9 (2023): 22447-22482. https://doi.org/10.3934/math.20231145

[12] Someetheram, Vikneswari, Muhammad Fadhil Marsani, Mohd Shareduwan Mohd Kasihmuddin, Nur Ezlin Zamri, Siti Syatirah Muhammad Sidik, Siti Zulaikha Mohd Jamaludin, and Mohd Asyraf Mansor. "Random maximum 2 satisfiability logic in discrete hopfield neural network incorporating improved election algorithm." *Mathematics* 10, no. 24 (2022): 4734. https://doi.org/10.3390/math10244734

[13] Sathasivam, Saratha, Mohd Asyraf Mansor, Mohd Shareduwan Mohd Kasihmuddin, and Hamza Abubakar. "Election algorithm for random k satisfiability in the Hopfield neural network." *Processes* 8, no. 5 (2020): 568. https://doi.org/10.3390/pr8050568

[14] Kasihmuddin, Mohd Shareduwan Mohd, Siti Zulaikha Mohd Jamaludin, Mohd Asyraf Mansor, Habibah A. Wahab, and Siti Maisharah Sheikh Ghadzi. "Supervised learning perspective in logic mining." *Mathematics* 10, no. 6 (2022): 915. https://doi.org/10.3390/math10060915

[15] Rusdi, Nur'Afifah, Mohd Shareduwan Mohd Kasihmuddin, Nurul Atiqah Romli, Gaeithry Manoharam, and Mohd Asyraf Mansor. "Multi-unit discrete hopfield neural network for higher order supervised learning through logic mining: Optimal performance design and attribute selection." *Journal of King Saud University-Computer and Information Sciences* 35, no. 5 (2023): 101554. https://doi.org/10.1016/j.jksuci.2023.101554

[16] Jamaludin, Siti Zulaikha Mohd, Mohd Asyraf Mansor, Aslina Baharum, Mohd Shareduwan Mohd Kasihmuddin, Habibah A. Wahab, and Muhammad Fadhil Marsani. "Modified 2 satisfiability reverse analysis method via logical permutation operator." *Computers, Materials & Continua* 74, no. 2 (2023). https://doi.org/10.32604/cmc.2023.032654

[17] Romli, Nurul Atiqah, Nur Fariha Syaqina Zulkepli, Mohd Shareduwan Mohd Kasihmuddin, Nur Ezlin Zamri, Nur'Afifah Rusdi, Gaeithry Manoharam, Mohd Asyraf Mansor, Siti Zulaikha Mohd Jamaludin, and Amierah Abdul Malik. "Unsupervised logic mining with a binary clonal selection algorithm in multi-unit discrete Hopfield neural networks via weighted systematic 2 satisfiability." *AIMS Mathematics* 9, no. 8 (2024): 22321-22365. https://doi.org/10.3934/math.20241087

[18] Sokolova, Marina, and Guy Lapalme. "A systematic analysis of performance measures for classification tasks." *Information processing & management* 45, no. 4 (2009): 427-437. https://doi.org/10.1016/j.ipm.2009.03.002

[19] Shieh, Chin-Shiuh, Thanh-Tuan Nguyen, and Mong-Fong Horng. "Detection of unknown ddos attack using convolutional neural networks featuring geometrical metric." *Mathematics* 11, no. 9 (2023): 2145. https://doi.org/10.3390/math11092145

[20] Huang, Changwu, Yuanxiang Li, and Xin Yao. "A survey of automatic parameter tuning methods for metaheuristics." *IEEE transactions on evolutionary computation* 24, no. 2 (2019): 201-216. https://doi.org/10.1109/TEVC.2019.2921598

[21] Sunil, G. C., Yu Zhang, Cengiz Koparan, Mohammed Raju Ahmed, Kirk Howatt, and Xin Sun. "Weed and crop species classification using computer vision and deep learning technologies in greenhouse conditions." *Journal of Agriculture and Food Research* 9 (2022): 100325. https://doi.org/10.1016/j.jafr.2022.100325

[22] Chicco, Davide, and Giuseppe Jurman. "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation." *BMC genomics* 21 (2020): 1-13. https://doi.org/10.1186/s12864-019-6413-7

[23] Hu, Hsin-Hui, Jay Kandampully, and Thanika Devi Juwaheer. "Relationships and impacts of service quality, perceived value, customer satisfaction, and image: an empirical study." *The service industries journal* 29, no. 2 (2009): 111-125. https://doi.org/10.1080/02642060802292932