Y-TYPE RANDOM 2-SATISFIABILITY IN DISCRETE HOPFIELD NEURAL NETWORK

GUO YUELING

UNIVERSITI SAINS MALAYSIA

2024

Y-TYPE RANDOM 2-SATISFIABILITY IN DISCRETE HOPFIELD NEURAL NETWORK

by

GUO YUELING

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

September 2024

ACKNOWLEDGEMENT

The exploration of this thesis has been one of the most significant academic challenges for me. I would like to express my sincere and respectful gratitude to my supervisor, Dr. Mohd Shareduwan Mohd Kasihmuddin who is a senior lecturer of Mathematics, Universiti Sains Malaysia. I am truly obliged for his mentorship, insightful comments, constructive feedback, and constant assistance throughout this adventure. He has the attitude and the substance of a genius where he continually and convincingly conveyed a spirit of adventure during the PHD journey. When I hesitate and cannot make a decision, he always guides me back on the right path. His high standards, strict requirements, and sense of responsibility toward students have had a profound impact on my Ph.D. journey and even my life, which will be role models for my lifelong learning. I am very lucky to be his student. I also would like to honor and respect my co-supervisor Dr. Mohd. Asyraf Mansor for guidance and support.

Secondly, I am using this opportunity to express my deepest gratitude to all of my family members. I express my earnest gratitude to my husband Li Chunbo for his unconditional support and blessings. His trust in me and his willingness to assist whenever needed have been instrumental in the completion of this thesis. Special thanks to my sons Li Zexi and Li Zekai for their kind support which makes me always strong and motivated. I am deeply grateful for their love, understanding, and untiring presence throughout this journey. Their support reduces my family challenges and gives me a huge mental support.

I would be remiss in not mentioning my circle of colleagues, collaborators, and friends in the Artificial Intelligence Research Development Group (AIRDG). As a member of this research group, I realized that my Ph.D. journey was not lonely. My experience as a mentor as well as a mentee in this group has helped me tremendously in my research journey. Their undying spirit and passion for research have kept me in my spirits and highly motivated during my Ph.D. journey. I also had the good fortune to have Nur Ezlin Zamri as my mentor who helps me selflessly. Sharing ideas with her has been very productive, rewarding, and a lot of fun. In addition, I would like to mention several members from whom I received emotional support, Alyaa Alway, Atiqah Romli, Chang Yunjie, Syatirah Sidik, Siti Zulaikha Jamaludin, Nur'Afifah Rusdi, Suad Abdeen, and Gaeithry Manoharam.

Finally, I would like to extend my thanks to the management of the library and coffee shop. You have provided me with a wealth of literature resources and a good environment for discussing issues. All available resources provided by Universiti Sains Malaysia (USM) in the completion of the thesis are much appreciated. Also, Sincere thanks to the anonymous reviewers for their valuable feedback and comments.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF SYMBOLS

P_2	2SAT
P_{2SAT}	2SAT formula
P_3	3SAT
P_{3SAT}	3SAT formula
F_i	A scoring mechanism
$\wedge$	AND (conjunction)
f_i	Current fitness
P_Y^k	Different cases of YRAN2SAT
FN	False Negative
FP	False Positive
$J_i^{(1)}$	First-order clause
$P_{YRAN2SAT}$	General formula of YRAN2SAT
$tanh(h_i)$	Hyperbolic Tangent Activation Function
$S_i^{\max}$	Ideal neuron state
$V^{'}$	Individuals after cloning
NH	Learning iterations
a_i^*	Literal
h_i	Local field
H_P	Lyapunov energy function of the DHNN-SAT
$H_{P_{YRAN2SAT}}$	Lyapunov energy function of the DHNN-YRAN2SAT
$f_{\max}$	Maximum fitness
f_{NC}	Maximum fitness achieved
$\tilde{\alpha}$	Mutation ratio for Swarm Mutation
$\neg$	NOT (negation)
H_0	Null hypothesis

l_1	Number of $\{S_i^{\max}=1,\ S_i=1\}$
m_1	Number of $\{S_i^{\max}=1,\ S_i=-1\}$
n_1	Number of $\{S_i^{\max}=-1,\ S_i=1\}$
o_1	Number of $\{S_i^{\max}=-1,\ S_i=-1\}$
m	Number of first-order logic clauses
$G(n)$	Number of global minimum energy
NL	Number of learning
$L(n)$	Number of local minimum energy
C	Number of neuron combination
NN	Number of neurons
$F_{\max-v}$	Number of non-repetitive solutions that reached maximum fitness
n	Number of second-order clauses
$F_{\max}$	Number of solutions that reached maximum fitness
V	Number of the total neuron variation
o	Number of third-order clauses
NT	Number of trials
$\vee$	OR (disjunction)
P_{R2}	RAN2SAT
$P_{RAN2SAT}$	RAN2SAT formula
P_{R3}	RAN3SAT
$P_{RAN3SAT}$	RAN3SAT formula
Z_m	Ratio of global minimum energy
$\Re$	Relaxation rate
P	SAT formula
$J_i^{(2)}$	Second-order clause
∂	Significance level
S_i	State of neuron i

Δn	Step size
W_{ij}	Synaptic weight between i and neuron j
W_i	Synaptic weight of neuron i
Aff_i	The affinity of each antibody
ASC_{clause}	The average satisfied clauses in the retrieval phase
$P_{induced}^{best}$	The best induced logic
P_{best}	The best learning logic
ΔH	The changes in the energy level
γ_2	The cloning rate in HDEA
W_i^{true}	The correct synaptic weight
E_P	The cost function of SAT formula
$E_{P_{YRAN2SAT}}$	The cost function of the DHNN-YRAN2SAT
CR	The crossover factor
ν_2	The crossover rate in GA
DI_{error}	The diversity error
$e_{ij}^{*}(k_{mean})$	The formulation of k-mean clustering
P_{learn}^{logic}	The formulation of learning logic in YR2SATRA
$H_P^{\min}$	The global minimum energy
Z_m	The global minimum solution ratio
γ_1	The greedy selection rate in HDEA
L_{ij}	The inconsistency between neuron i and neuron j
$P_{induced}$	The induced logic
X	The initial population set
X_i	The ith string
Z_i	The ith intermediate string during mutation stage in DEA
$h_i(t)$	The local field of the ith neuron at time t
MAE_{energy}	The mean error of the final energy

MAE_{weight}	The mean error of the synaptic weight
$H_{P_{YRAN2SAT}}^{\min}$	The minimum energy of the DHNN-YRAN2SAT
ν_2	The mutation rate in GA
$R_{mutation}$	The mutation ratio in CSA
$S_j(t)$	The neuron states of the *j*th neuron at time t
N	The number of antibodies after cloning
N_i	The number of cloning
$n\left(J_i^{(k)}\right)$	The number of *k*-order clauses
λ_{2SAT}	The number of literals in 2SAT
λ_{3SAT}	The number of literals in 3SAT
N_c	The number of logic
F_i	The number of satisfied clauses
N_w	The number of synaptic weights in a particular logic
λ	The number of variables
β	The percentage of negative neuron state
α	The percentage of positive neuron state
γ	The percentage of random neuron state
ρ	The probability that differences among ranks occurred by chance
W_i^{rand}	The random synaptic weight
FS	The scaling factor
sl	The selection rate in CSA
ν_1	The selection rate in GA
e^*	The set of selected attributes
PS	The size of initial strings
PS^2	The total number of strings to crossover
PS^3	The total number of strings to mutate
PS^1	The total number of strings to select

N_{wc}	The total number of synaptic weights
U_i	The *i*th string during the crossover stage in DEA
$J_i^{(3)}$	Third-order clause
ξ	Threshold value
τ	Tolerance value
TC	Total clauses
TN	True Negative
TP	True Positive
a_i	Variable
$P_{YRAN2SAT}$	YRAN2SAT formula
$P_{YRAN2SAT}^{(1,2)}$	YRAN2SAT formula (both first and second-order clauses)
$P_{YRAN2SAT}^{(1)}$	YRAN2SAT formula (only first-order clauses)
$P_{YRAN2SAT}^{(2)}$	YRAN2SAT formula (only second-order clauses)

LIST OF ABBREVIATIONS

2SAT	2 Satisfiability
2SATRA	2 Satisfiability Reverse Analysis
3SAT	3 Satisfiability
3SATRA	3 Satisfiability Reverse Analysis
ACC	Accuracy
AD	Alzheimer's Disease
ADNI	Alzheimer's Disease Neuroimaging Initiative
ABC	Artificial Bee Colony
AIS	Artificial Immune System
AI	Artificial intelligence
ANNs	Artificial Neural Networks
BM	Bookmark Informedness
CSA	Clonal Selection Algorithm
CNF	Conjunctive Normal Form
CNNs	Continuous Neural Networks
CAM	Control Addressable Memory
DEA	Differential Evolution Algorithm
DHNN	Discrete Hopfield Neural Network
LRR	DNA copy variation
EA	Election Algorithm
E2SATRA	Energy-based 2 Satisfiability Reverse Analysis
EDA	Estimation Distribution Algorithm
EVA	Evolutionary Algorithms

ES	Exhaustive Search
GA	Genetic Algorithm
GWO	Grey Wolf Optimization
HL	Hebbian Learning
HornSAT	Horn Satisfiability
HDEA	Hybrid Differential Evolution Algorithm
HEA	Hybrid Election Algorithm
HTAF	Hyperbolic Tangent Activation Function
ICA	Imperialistic Competitive Algorithm
L2SATRA	Log linear 2 Satisfiability Reverse Analysis
MAJ2SAT	Major 2 Satisfiability
MCC	Matthews's Correlation Coefficient
MAE	Mean Absolute Error
MAPE	Mean Absolute Percent Error
MCAM	Modified Correlation Analysis Method
G3SATRAμ	Multi-unit 3-satisfiability-based Reverse Analysis
PSO	Particle Swarm Optimization
P2SATRA	Permutation 2 Satisfiability Reverse Analysis
A2SATRA	Permutation and Log linear 2 Satisfiability Reverse Analysis
PRE	Precision
RBFNN	Radial Basis Function Neural Network
RAN2SAT	Random 2 Satisfiability
RAN3SAT	Random 3 Satisfiability
RAN*k*SAT	Random k Satisfiability
RNNs	Recurrent Neural Networks

RA	Reverse Analysis
SAT	Satisfiability
DHNN-SAT	Satisfiability in DHNN
SEN	Sensitivity
SCA	Sine Cosine Algorithm
SPE	Specificity
STD	Standard Deviation
S2SATRA	Supervised 2 Satisfiability Reverse Analysis
SHoRA	Supervised High Order Reverse Analysis
SVM	Support Vector Machine
UCI	University of California Irvine
WA	Wan Abdullah
*r*2SAT	Weighted Random 2 Satisfiability
WOA	Whale Optimization Algorithm
YR2SATRA	Y-Type Random 2-Satisfiability Reverse Analysis
YRAN2SAT	Y-Type Random 2-Satisfiability

LIST OF APPENDICES

2-SATISFIABILITI RAWAK JENIS-Y DALAM RANGKAIAN NEURAL HOPFIELD DISKRET

ABSTRAK

Dalam perkembangan semasa dalam Kecerdasan Buatan, Satisfiabiliti memainkan peranan penting sebagai bahasa simbolik Kecerdasan Buatan untuk ketelusan model kotak hitam. Walau bagaimanapun, masalah utama yang terkini dalam Satisfiabiliti ialah kekangan peraturan dalam logik gabungan yang masih belum disiasat. Maka, satu peraturan logik gabungan iaitu Jenis-Y 2- Satisfiabiliti Rawak telah dicadangkan dengan menggabungkan peraturan logik yang sistematik dan tidak sistematik. Peraturan logik yang dicadangkan akan menjadi arahan simbolik kepada neuron dalam Rangkaian Diskrit Neural Hopfield. Berdasarkan keputusan eksperimen, peraturan logik yang dicadangkan dilihat serasi dengan Rangkaian Diskrit Neural Hopfield. Selain itu, Algoritma Evolusi Perbezaan Hibrid yang dicadangkan telah dilaksanakan pada fasa latihan untuk memastikan fungsi kos Rangkaian Diskrit Neural Hopfield dapat diminimumkan. Pada fasa dapatan, fungsi pengaktifan yang baru dan mutasi kerumunan telah dicadangkan untuk meningkatkan kepelbagaian keadaan neuron. Algoritma dan mekanisma mutasi yang dicadangkan menunjukkan prestasi yang optimum berbanding dengan algoritma sedia ada. Akhirnya, model yang dicadangkan akan dilaksanakan dalam perlombongan logik baru iaitu Analisis Berbalik Jenis-Y 2-Satisfiabiliti Rawak. Model perlombongan logik yang baru dapat mengekstrak corak optimum set data dalam bentuk logik teraruh berbanding model perlombongan logik yang sedia ada. Model perlombongan logik yang dibangunkan telah digunakan untuk menganalisis dataset Inisiatif Neuroimaging Penyakit Alzheimer.

Y-TYPE RANDOM 2-SATISFIABILITY IN DISCRETE HOPFIELD NEURAL NETWORK

ABSTRACT

In the current development of Artificial Intelligence, Satisfiability plays a crucial role as a symbolic language of Artificial Intelligence for the transparency of black box models. However, the main problem of existing Satisfiability is the lack of combined logical rule, so the benefits of combined logical rule have not yet been investigated. The rule namely Y-Type Random 2-Satisfiability is proposed by combining the systematic and non-systematic logical rule. Next, the newly proposed logical rule as the symbolic instruction was implemented into the Discrete Hopfield Neural Network to govern the neurons of the network. Experimental results demonstrated the compatibility of the proposed logical rule and the Discrete Hopfield Neural Network. Additionally, the proposed Hybrid Differential Evolution Algorithm was implemented into the training phase to ensure that the cost function of the Discrete Hopfield Neural Network is minimized. During the retrieval phase, a new activation function and Swarm Mutation were proposed to ensure the diversity of the neuron states. The proposed algorithm and mutation mechanism showed optimal performances as compared to the existing algorithms. Finally, a new logic mining model namely Y-Type Random 2-Satisfiability Reverse Analysis was proposed, which showed optimal performances in terms of several metrics as compared to the existing classification models. The developed logic mining will be used to analyze the Alzheimer's Disease Neuroimaging Initiative dataset.

CHAPTER 1

INTRODUCTION

1.1 Overview

Artificial intelligence (AI) is a field of study that gradually changes the way we work and live. With the continuous development of technologies, more and more fields have begun to apply AI, bringing huge innovations and changes. In this chapter, the explanation of the research background, research direction, and research philosophy are presented to align with the context of this thesis. Subsequently, this thesis will also discuss the research problem statements, followed by the presentations of research questions, research objectives, methodology, and scopes. Lastly, the organization of this thesis and summary are presented.

1.2 Research Background

AI possesses the cognitive abilities exhibited by machines or software, which can carry out tasks that typically require human intelligence, including learning, logical reasoning, problem-solving, and language. AI technology is extensively employed across various sectors, including industry, government, and scientific domains. A few prominent applications include advanced web search engines (i.e., Google Search), recommendation systems (implemented by YouTube, Amazon, and Netflix), comprehension of human speech (i.e., Google Assistant, Siri, and Alexa), autonomous vehicles (such as Waymo), generative and innovative tools (like ChatGPT and AI art), and surpassing human capabilities in strategic games, such as chess and go. However, many AI systems possess such complexity that the designers themselves are unable to explain the decision-making process, which is commonly known as the black box model (Sample, 2017). Within the field of AI, a black box model refers to a system that generates outputs based on inputs, yet fails to disclose the underlying mechanisms responsible for such outputs. These models primarily rely on machine learning or deep learning techniques, which acquire knowledge from data without the need for human intervention. The presence of black box models in the field of AI presents challenges in terms of explicability, transparency, and accountability.

Numerous instances have arisen in which a machine learning program successfully passed rigorous assessments, only to unexpectedly acquire knowledge that diverged from the original intentions of the programmers (Amodei *et al.*,2016). For instance, a machine learning system specifically designed to allocate medical resources efficiently was discovered to classify asthma patients as “low-risk” patients who die from pneumonia, which means asthma patients are considered less likely to die from pneumonia under this classification (Odeyemi *et al.*, 2024). Despite asthma being an important health risk factor, the training data used by the machine learning system demonstrates that asthma patients often receive heightened medical care, resulting in relatively lower death rates. Although the correlation between asthma and reduced risk of pneumonia-related mortality is true in the data, this correlation is misleading. This is because it may lead to the system mistakenly categorizing asthma patients as “low-risk” patients dying from pneumonia, potentially resulting in the neglect or underestimation of actual health risks during resource allocation. Therefore, individuals who are affected by machine learning system decisions and suffer harm have the right to receive comprehensive explanations to understand the decision-making process of the system. The interpretability of black box models enables users to understand the knowledge acquired by AI and generate outputs that reflect the learning process of AI. An emerging field known as the neuro-symbolic method strives to tackle the issue of interpretability within the field of AI. More precisely, the neuro-symbolic method refers to combining the data-driven learning of neural networks with the logic-based reasoning of symbolic systems, which leverages the strengths of both approaches to enhance AI capabilities in learning and reasoning tasks.

1.3 Research Direction

Although both systematic and non-systematic logical rules can serve as languages of AI, they are limited in representing data across a broad range of datasets. To enhance the transparency of AI, the research direction in this thesis is as follows:

Expanding the symbolic language of AI.

The formal language utilized for expressing, storing, and manipulating knowledge within AI is known as the symbolic language. This language enables computers to possess the ability to understand, reason, and solve problems in a manner like humans. Expanding the symbolic language of AI involves enhancing the capabilities for representation, reasoning, and application to effectively handle complex knowledge and problems. One approach to expanding symbolic languages is to enhance the ability of knowledge representation to describe complex concepts and relationships more accurately. This may involve introducing more symbol types, relationship types, and attribute types. By introducing a new symbol in this thesis, the symbolic language of AI can be effectively expanded. Another approach to expanding symbolic languages is to combine Artificial Neural Networks (ANNs) with the new symbol. This combination can compensate for the lack of logical reasoning and interpretability of ANNs in processing large-scale data and complex pattern recognition, aiming to achieve improved performance and interpretability. In summary, the expansion of the symbolic language of AI includes various aspects of work, including enhancing knowledge representation and integrating ANNs. These methods effectively expand the symbolic language of AI, enabling AI to accurately represent and understand a wider range of information contained in various datasets.

The initial perspective considers that the language employed in the field of AI can be characterized as a type of formal language rooted in symbolic logic. In the development process of AI, ANNs have played an important role and the design inspiration comes from imitating the biological brain. By adjusting connections between neurons, ANNs simulate the way humans learn and process information. Logic rules within ANNs play a pivotal role in serving as governing principles that represent the neurons in the network. A previous study conducted by Pinkas (1991) initially asserted that connectionist ANNs should utilize Satisfiability (SAT) to showcase the energy minimization capabilities of the network. In the field of ANNs, minimizing network energy is a key objective of optimization and helps the network converge to a suitable state. The concept of SAT involves determining whether there exists an assignment of truth values to variables that makes a given logical formula true. Employing SAT to demonstrate the energy minimization capability of ANNs can help evaluate the performance of networks on specific tasks such as classification, regression, and clustering. This approach is particularly applicable to networks with symmetric weights, such as Discrete Hopfield Neural Network (DHNN). The incorporation of SAT into DHNN has provided a more comprehensive understanding of the hidden units within the network. Subsequently, Abdullah (1992) proposed an approach that utilized Horn Satisfiability (HornSAT) as a language to govern DHNN, which consists of at most one negative literal in each clause. The study by Abdullah (1992) emphasized the development of a language through the minimization of logical inconsistencies in DHNN using HornSAT.

According to Abdullah (1992), DHNN can make logical inferences by minimizing the cost function and generating an optimal solution based on the learned information. To be more concise, the optimal solution explains the behavior of HornSAT as a neuron representation in DHNN. As a language, HornSAT facilitates an understanding of the overall dynamics of neuron states in the network. The robustness of SAT as a language in DHNN enables a deeper comprehension and transparency of neuron connections compared to other ANN models that function as black boxes. Despite the effectiveness of HornSAT as a language embedded into DHNN, it is crucial to highlight that using a single AI language can only address a limited number of real-world problems. As a result, it is critical to develop new AI languages for managing neurons in networks.

1.4 Research Philosophy

The previous research conducted by Abdullah (1992) demonstrated the importance of employing the logical rule as the neuro-symbolic model. Nevertheless, the presence of repetitive information in HornSAT was evident in the form of redundant literals. Therefore, it becomes necessary to establish a new logical rule to govern the behavior of neurons in DHNN. This gives rise to a more comprehensive philosophical perspective of this thesis:

Creating a flexible neuro-symbolic model.

The neuro-symbolic model integrates symbolic AI and connectionist AI. Symbolic AI employs logical rules for reasoning but struggles with uncertainty and complexity. Conversely, connectionist AI captures complex nonlinear relationships through ANNs but lacks symbolic reasoning. The neuro-symbolic model effectively combines these strengths, achieving more robust and interpretable AI systems. This thesis explores various perspectives to create a flexible neuro-symbolic model, each contributing uniquely to the framework.

The first perspective involves a flexible logical rule, which is essential for neuro-symbolic model development. Existing logical rules are inflexible, limiting their ability to handle large-scale datasets. In contrast, flexible logical rules enhance the capacity to adaptively process diverse datasets. However, logical rules alone cannot optimize the information in datasets, necessitating a supplementary medium for efficient optimization. The second perspective integrates the flexible logical rule with an ANN, highlighting their interdependence. This relationship acknowledges that the logical rule needs the ANN for optimization, while the ANN requires the logical rule for neuron management. One noteworthy approach is the DHNN-SAT models, such as DHNN-2SAT (Kho *et al.*, 2022), DHNN-MAJ2SAT (Alway *et al.*, 2022) and DHNN-RAN2SAT (Sathasivam *et al.*, 2020a). Nonetheless, ensuring compatibility between the proposed logical rule and the DHNN remains a critical concern. Therefore, the third perspective focuses on utilizing the proposed algorithm to optimize the proposed DHNN-SAT model. The core idea of this method is to adjust the parameters and structure of DHNN through the proposed algorithm (Zamri *et al.*, 2020). In this process, the selection and design of algorithms are crucial as they directly affect the performance of ANN systems. For example, optimization algorithms can help adjust the parameters of the network to minimize the loss function and improve the accuracy and generalization ability of the ANN systems. In addition, algorithms can also help optimize the structure of the network, including the number of neurons and connection methods, thereby making the network more efficient and robust. In summary, the first three perspectives have theoretically validated the effectiveness of the proposed neuro-symbolic model but have not been validated in practical applications. Hence, the fourth perspective refers to creating and implementing a classification model as a means of validating the effectiveness and efficacy of the proposed neuro-symbolic model (Kasihmuddin *et al.*, 2022). This model should fully utilize the characteristics of the flexible neuro-symbolic model and undergo thorough testing and assessment in practical scenarios. Each of these perspectives is crucial and will be thoroughly examined, discussed, and analyzed in the subsequent problem statement section of this thesis.

1.5 Problem Statements

The main problem addressed in the thesis is the absence of a flexible neuro-symbolic model. The existing neuro-symbolic models face restrictions in combining systematic and non-systematic logical rules and optimizing information during the training and retrieval phases, which leads to inefficiencies and constraints in logic mining. The following content will provide a detailed explanation of five problem statements.

Problem Statement 1: Recently researchers have faced a series of challenges regarding the development of Satisfiability logic representations, which involve two main perspectives: systematic logical rules and non-systematic logical rules. Firstly, the concept of systematic logical rules was initially proposed by Kasihmuddin *et al.* (2017a) through the 2 Satisfiability (2SAT), where each clause only consists of two literals that are connected through logic operator disjunction. Following closely, Mansor *et al.* (2017) extended the concept by introducing the 3 Satisfiability (3SAT), where each clause is only composed of three literals. On the other hand, Sathasivam *et al.* (2020a) proposed the concept of a non-systematic logical rule namely Random Satisfiability (RAN*k*SAT), where *k* is less than or equal to 2. Bazuhair *et al.* (2021)

further proposed a higher-order non-systematic logical rule namely Random 3 Satisfiability (RAN3SAT). Systematic logical rules provide consistency and predictability, while non-systematic logical rules provide diversity and unpredictability of representation. However, the main problem with the current research is a lack of flexibility in representing the combination of features between systematic and non-systematic logical rules. This constraint has resulted in a significant obstacle to fully utilize the potential advantages of these two types of logical rules. Therefore, it is imperative to incorporate more random behaviors into existing logical rules and design a new logical rule with greater flexibility in clause structure by combining systematic and non-systematic logical rules. Through this approach, the advantages of both types of logical rules can be demonstrated simultaneously, and the limitations of current logical rules can be addressed.

Problem Statement 2: Logic rules as the cornerstone of AI language can represent various types of information, but they cannot independently optimize information. This is where the DHNN becomes crucial, serving as a medium to optimize information. However, the DHNN relies heavily on logical rules that define neuron connections and synaptic weight allocations, which impact the behavior and performance of the network. A significant challenge arises as the number of neurons increases, causing the storage capacity of DHNN to grow exponentially. If neurons fail to converge to a global minimum energy state, the final neuron states will be suboptimal, which does not contribute to solving any specific optimization problem. Many existing DHNN-SAT models such as DHNN-3SAT (Zamri *et al.*, 2020), DHNN-MAJ2SAT (Alway *et al.*, 2022), and DHNN-RAN2SAT (Sathasivam *et al.*, 2020a) attempt to address this issue using the WA method (Abdullah, 1992), which minimizes the cost function during the training phase by maximizing the number of satisfactory clauses. While this method can transform the cost-minimizing problem into a fitness-maximizing problem, it has a major drawback: rigid logical rules can lead to suboptimal synaptic weights and neuron states, reducing the effectiveness of DHNN-SAT models. To overcome this, a flexible logical rule is proposed, combining systematic and non-systematic rules through random clause generation. This approach is expected to help DHNN achieve optimal synaptic weights and neuron states, improving the interpretability of black-box models.

Problem Statement 3: The proposed logical rule primarily uses lower-order clauses, including first-order clauses, which reduces the likelihood of achieving Satisfiability compared to higher-order clauses. To ensure that all clauses are satisfiable during the training phase, a robust algorithm is necessary. The training phase is critical because if DHNN achieves a zero-cost function, it indicates that the derived cost function aligns with the Lyapunov energy function, leading to optimal synaptic weights. Conversely, a non-zero cost function suggests that DHNN cannot satisfy the logical rule, negatively affecting the synaptic weights of the network (Kasihmuddin *et al.*, 2017). To address this, many studies have incorporated metaheuristics into the DHNN as a training algorithm to achieve a zero-cost function. For example, Zamri *et al.* (2020) used the Clonal Selection Algorithm (CSA) to optimize 3SAT in the training phase, effectively finding near-optimal solutions with fewer iterations than Genetic Algorithms (GA) and Evolution Strategies (ES). However, CSA and similar algorithms are limited by their reliance on a single global search operator, which restricts their ability to perform comprehensive global searches. These algorithms often stop learning after they have discovered a satisfied interpretation, without fully exploring the entire solution space. To overcome these challenges, the proposed training algorithm in this thesis seeks to overcome these limitations by incorporating two additional exploration operators into the conventional Differential Evolution Algorithm (DEA). These operators allow the algorithm to explore other promising regions of the search space, increasing the chances of finding optimal solutions for the proposed DHNN-SAT model. The introduction of exploration operators helps the algorithm expand the search range and thoroughly consider the solution space, leading to more effective optimization of DHNN-SAT models.

Problem Statement 4: Despite using various algorithms to improve the training phase of DHNN-SAT models, an efficient training phase does not ensure an efficient retrieval phase. Many DHNN-SAT models produce repetitive final neuron states, meaning the same output is generated for different inputs, leading to a lack of diversity. This repetition limits the flexibility and adaptability of the models, making it difficult to handle real-life scenarios where different inputs should produce different outputs (Zamri *et al.*, 2022). The optimal retrieval capability of DHNN is crucial for ensuring that final neuron states converge to the global minimum energy. However, the retrieval phase is often influenced by various initial states, causing the network to get stuck in local minima and overfit. To address this problem, Kasihmuddin *et al.* (2019) enhanced the DHNN retrieval phase by implementing a mutation-based approach using the Estimation Distribution Algorithm (EDA). This method provided a global search capability, allowing the retrieval of various final neuron states with global minimum energy. However, as the number of neurons increases, the performance of this mutation-based approach deteriorates, leading to repetitive neuron states. Additionally, many existing works use the Hyperbolic Tangent Activation Function (HTAF) for output scaling after local field computation. While HTAF has shown compatibility with DHNN retrieval (Karim *et al.*, 2021), HTAF tends to favor one extreme (The neuron state is highly likely to be updated to 1), increasing the likelihood of overfitting. To overcome these challenges, this thesis proposes two modifications to enhance the retrieval phase of DHNN. The first modification ensures that the output after local field computation achieves the global minimum energy, while the second modification ensures diversity in the global solutions. These enhancements aim to improve the adaptability and flexibility of DHNN-SAT models in practical applications.

Problem Statement 5: The performance of a model needs to be evaluated not only with simulated datasets but also with real-life datasets. Logic mining is an important branch of data mining, and the main goal is to discover the behavioral patterns of datasets through logical rules. Unlike traditional machine learning models, logic mining not only provides final classification outcomes but also the specific logical rules behind these outcomes, making it a powerful tool for decision support, knowledge extraction, and inference. The first logic mining model proposed by Sathasivam and Abdullah (2011) used Reverse Analysis (RA) to represent real-life datasets with HornSAT. Although effective, this model struggled with generalizing the induced logic. To address this, Jamaludin *et al.* (2021a) introduced an Energy-based logic mining model namely E2SATRA, which considered only the final neuron states corresponding to the global minimum energy. E2SATRA performed well in classifying the e-recruitment dataset but had limitations due to the fixed attribute arrangements. To improve E2SATRA, Jamaludin *et al.* (2023) added a permutation operator to 2SAT namely P2SATRA, which allowed different combinations of attributes, resulting in better performance across various metrics. Despite these advances, these models commonly generate learning logic based on the frequency of clauses. Rusdi *et al.* (2023) improved RA by using 3SAT logical rules in DHNN, developing a new model namely Supervised High Order Reverse Analysis (SHoRA). SHoRA used the highest true values to generate the best learning logic. However, all these models focused only on systematic SAT formulations, neglecting non-systematic SAT logical rules. To overcome these limitations, the proposed logic mining model must be able to represent both systematic and non-systematic logical rules and generate the best learning logic based on both frequency and highest true values. This approach aims to extract the most accurate induced logic, reflecting the overall behavior of the dataset more effectively.

1.6 Research Questions

Aligned with the biggest problem statement in this thesis, the biggest research question of the thesis is how can a flexible neuro-symbolic model be developed effectively. To clarify the consistency with the problem statements given in Section 1.5, this section lays out several important research questions. Notably, these research questions in Section 1.6 are based on all problems mentioned in Section 1.5. Therefore, the detailed research questions involved in this thesis are listed as follows:

(a) What is the alternative formulation of a flexible logical rule that consists of lower-order clauses to represent the features of both systematic and non-systematic logical rules?

(b) What kind of logical rule can be embedded into the Discrete Hopfield Neural Network to create a flexible neuro-symbolic model that can govern the neurons?

(c) What kind of improvements can be considered to the existing Differential Evolutionary Algorithm to ensure an effective training phase that contributes to the optimal synaptic weights?

(d) What kind of mutation can be applied to the current retrieval phase of the Discrete Hopfield Neural Network to ensure an effective retrieval phase that can produce diverse final neuron states?

(e) What modifications can be implemented to the current pre-processing and training phases of logic mining models to ensure an optimal classification model with a high generalization ability that can extract the information from real-life datasets?

1.7 Research Objectives

The research objectives of this thesis concentrate on the modeling of DHNN using a flexible logical rule with a random generation of first and second-order clauses. The training phase in DHNN is centered around the minimization of the cost function which corresponds to the management of optimal synaptic weights. Due to the constraints in the training phase, the proposed model will embed a metaheuristic algorithm to achieve the minimization of the cost function. In the retrieval phase, a mutation-based mechanism is utilized to expand the solution space. Eventually, the proposed model will extract the empirical patterns from real-life datasets for classification tasks. The objectives of this thesis are presented as follows:

(a) To formulate a logical rule namely Y-Type Random 2-Satisfiability that has the property of both systematic and non-systematic logic. The proposed Y-Type Random 2-Satisfiability logical rule will generate different orders of clauses randomly. The proposed logical rule enables better structural flexibility to represent information.

(b) To implement the formulated Y-Type Random 2-Satisfiability into the Discrete Hopfield Neural Network as the symbolic rule to govern the neurons. The proposed Y-Type Random 2-Satisfiability logical rule will be trained and the neuron connections will be stored as the Content Addressable Memory of the network. In this context, each literal in the proposed logical rule effectively represents one neuron in the Discrete Hopfield Neural Network.

(c) To propose a new training algorithm namely the Hybrid Differential Evolution Algorithm to minimize the cost function in the training phase of the Discrete Hopfield Neural Network. The proposed algorithm includes two hybrid operators namely greedy selection and cloning that originated from the state-of-art metaheuristics. The proposed algorithm will locate the satisfied interpretation of the logical rule for an arbitrary number of neurons.

(d) To develop a mutation-based retrieval phase of the Discrete Hopfield Neural Network by embedding a new mutation scheme namely Swarm Mutation. The proposed Swarm Mutation will optimize the neuron states of second-order clauses by comparing them with the benchmark neuron states. In this context, the newly developed retrieval phase is responsible for the diversification of final neuron states.

(e) To construct a logic mining model namely Y-Type Random 2-Satisfiability Reverse Analysis in doing classification tasks for various real-life data. The proposed logic mining model employs a new method to generate the best learning logic by considering the highest frequency of clauses and the accuracy of logical rules during the pre-processing phase. In this context, the best learning logic will represent the highest accumulation of both True Positive and True Negative classifications for the training data. Capitalizing on the proposed retrieval phase, the best induced logic will be selected based on the highest accuracy of the testing data. Note that, the extracted best induced logic represents the overall behavior of real-life data.

1.8 Methodologies and Scopes of the Thesis

To address these 5 research questions in Section 1.5, this section will introduce the corresponding 5 methodologies and scopes. Firstly, a flexible logical rule is proposed, which combines systematic and non-systematic logical rules. Second, this flexible logical rule is embedded into DHNN to manage the behavior of neurons in the network. Third, a Hybrid Differential Evolution Algorithm will be adopted to optimize the training phase of DHNN, ensuring the compatibility of the proposed logical rule with DHNN and obtaining the optimal synaptic weights. Fourth, introducing Swarm Mutation to optimize the retrieval phase of DHNN to ensure the diversity of final neuron states. Finally, the modified Y-Type Random 2-Satisfiability Reverse Analysis method will be combined with the Hybrid Differential Evolution Algorithm and Swarm Mutation for logic mining research. The specific methodologies and limitations in this thesis will be explained in detail as follows:

Methodology for Objective 1: Recently, there has been a substantial increase in the number of SAT formulations. Within this set of SAT formulas, there exist both systematic SAT and non-systematic SAT. However, thus far, there is a lack of a flexible SAT formula that effectively combines both systematic SAT and non-systematic SAT. Hence, a flexible SAT representation is proposed that adequately includes both systematic SAT and non-systematic SAT. With regards to this new SAT formulation, this thesis presents the Y-Type Random 2-Satisfiability by capitalizing on the systematic and non-systematic SAT structure through the generation of the first and second-order clauses. The foundation of the proposed SAT formulation draws inspiration from the following works. First, the systematic approach adopted in the proposed Y-Type Random 2-Satisfiability is inspired by the work of Kasihmuddin *et al.* (2017b) where a systematic SAT structure was introduced by strictly defining 2 literals per clause, which is commonly referred to as 2SAT. Second, the non-systematic approach in the proposed Y-Type Random 2-Satisfiability is influenced by the work of Sathasivam *et al.* (2020a), which proposed the incorporation of different order clauses in a single SAT formulation. A flexible logical rule that combines the characteristics of both systematic and non-systematic SAT is proposed using a random clause generator, where the number of first and second-order clauses is randomly generated. The decision of which literal to negate (negation) is also randomly determined. The systematic component in the proposed logical rule offers a higher storage capacity in terms of clauses, as it allows for the inclusion of more variables. On the other hand, the non-systematic component in the proposed logical rule enhances the connectivity in the first-order clause, contributing to a more diverse range of connections. By combining both features, the proposed logical rule maximizes the benefits of both systematic SAT and non-systematic SAT, ultimately creating a flexible logical rule.

Scope of Methodology 1: While this methodology may present an optimal strategy for producing the proposed Y-Type Random 2-Satisfiability, certain scopes are implemented to guarantee the replicability of the first objective in this thesis. First, it should be noted that the literals found within all clauses do not exhibit any instances of repetitiveness or redundancy. The presence of repetitive literals signifies the representation of repetitive information. Moreover, it is important to acknowledge the drawbacks associated with redundant literals, which can lead to increased complexity and reduced efficiency. Secondly, it is crucial to highlight that this thesis solely focuses on the consideration of lower-order clauses, specifically those of the first and second-order clauses. Conversely, higher-order clauses (such as third-order clauses) are omitted from this thesis. The rationale behind this exclusion lies in the fact that higher-order clauses introduce additional complexities when compared to lower-order clauses.

Methodology for Objective 2: DHNN is a type of ANNs that operates with feedback and can be utilized as either an associative memory or an optimization model. The inspiration behind embedding the logical rule into DHNN originates from the work of Abdullah (1992), where the first neuro-symbolic of Horn Satisfiability (HornSAT) was proposed by assigning each literal to represent one neuron. By observing the behavior of the neurons, Abdullah (1992) suggested that the formulation of the cost function in DHNN can be derived from the inconsistency of the logical rule. Consequently, minimizing the cost function can be achieved by obtaining satisfied interpretations of the logical rule. To assess the compatibility of the proposed logical rule with DHNN, two perspectives were introduced. Firstly, the training phase of DHNN is necessary to ensure the complete satisfaction of the Y-Type Random 2-Satisfiability. In other words, the primary objective of the training phase is to minimize the cost function by Exhaustive Search (ES). Achieving interpretation that led to the minimization of the cost function ensures that the synaptic weight of DHNN can be effectively obtained and stored in the Control Addressable Memory (CAM). The behavior of the proposed logical rule is then compared to existing Satisfiability logic using various error metrics. Secondly, the quality of the final neuron states in the retrieval phase of DHNN is conducted with regard to the energy profile and the variation of the neuron states. The assessment of the energy profile can be accomplished through an examination of the difference between the global minimum energy and the local minimum energy. Additionally, the evaluation of the variation of neuron states is carried out by utilizing the formulation of the similarity index.

Scope of Methodology 2: While the selection of this particular methodology can be viewed as an ideal approach to evaluate the performance of the proposed DHNN-SAT model, several scopes have been considered to ensure the reproducibility of Objective 2 within this thesis. Firstly, the states of the neurons will be represented in a bipolar form, taking on values of 1 and -1. According to Stern and Shea-Brown (2020), the dynamics of ANNs that rely on the Lyapunov energy function are unsuitable for representing information in a binary form (i.e., 1 and 0). This is primarily due to the potential elimination of crucial coefficients within the energy function. In situations where the states of the neurons are zero, the final energy of the network will always be zero. Consequently, it becomes impossible to determine the actual minimum energy of the network. Secondly, the method employed to obtain the optimal synaptic weights is the WA method, as opposed to the conventional Hebbian learning approach. This choice is made to mitigate the generation of spurious memories and to reduce neuron oscillations (Gülcü, 2022).

Methodology for Objective 3: The proposed YRAN2SAT into the DHNN model requires an effective and efficient training phase to ensure the generation of optimal synaptic weights. Nevertheless, due to the presence of first-order clauses in Y-Type Random 2-Satisfiability, it becomes challenging for DHNN to retrieve the neuron states that achieve the zero-cost function through the utilization of ES particularly as the number of first-order clauses progressively increases. To solve this issue, the thesis proposed a modified metaheuristic namely Hybrid Differential Evolution Algorithm as the training algorithm to minimize the cost function in the training phase of DHNN. This algorithm is built upon the fundamental DEA (He *et al.*, 2022) and incorporates two additional optimization operators. The primary objective of the proposed training algorithm is to optimize the objective function stated in

Objective 2. In other words, the focus of the proposed algorithm lies in learning the configuration of the Y-Type Random 2-Satisfiability logical rule, which possesses both systematic and non-systematic SAT qualities. Moreover, the ES in Objective 2 only needs to retrieve one neuron state that achieves the zero-cost function. However, the proposed Hybrid Differential Evolution Algorithm in Objective 3 requires identifying the maximum number of neuron states that achieve the zero-cost function. In this context, the Hybrid Differential Evolution Algorithm is required to locate as many neuron states as possible that satisfy all the clauses in the Y-Type Random 2-Satisfiability. The Hybrid Differential Evolution Algorithm is composed of five operators: greedy selection, cloning, mutation, crossover, and selection. There are several distinctions between the traditional operators in DEA and the operators in the proposed training algorithm. Firstly, there is an additional operator namely greedy selection, which draws inspiration from the selection operator of GA (Kasihmuddin *et al.*, 2017a). This operator plays a crucial role in segregating individuals based on the fitness function of Y-Type Random 2-Satisfiability. Individuals with the highest fitness will be selected as the fittest individuals, while those with lower fitness will be abandoned. This operator serves as a preservation mechanism to maintain optimal neuron states for the Y-Type Random 2-Satisfiability. Additionally, the second additional operator namely cloning draws inspiration from the cloning operator of CSA (Pantourakis *et al.*, 2022). The proposed cloning operator is responsible for proliferating the fittest individuals obtained from greedy selection. With an ideal cloning rate, this operator employs a double cloning exploration technique. Notably, both greedy selection and cloning are global search operators that aim to explore a wider range of satisfied interpretations of Y-Type Random 2-Satisfiability for any given number of neurons. To evaluate the performance of the training algorithm, various performance metrics will be assessed during the training phase of DHNN. Furthermore, to provide a clearer comparison between the proposed algorithm and previous works, the Hybrid Differential Evolution Algorithm will be compared to existing training algorithms such as ES, GA, ABC, EA, CSA, DEA, and so on.

Scope of Methodology 3: While the training algorithm that has been proposed can potentially serve as an optimal approach for minimizing the objective function in Objective 2, it is crucial to acknowledge that there exist a few scopes that must be addressed to ensure the reproducibility of Objective 3 within the context of this thesis. Primarily, it is essential to note that all training algorithms utilized in Objective 3 must adhere to the objective function that has been defined in Objective 2. Second, it is worth mentioning that the parameter settings for all existing training algorithms will be determined based on the successful experimentation conducted by previous works (i.e. Kasihmuddin *et al.*, 2019; Zamri *et al.*, 2020; and Karim *et al.*, 2022). More specifically, parameter settings are related to the training algorithms employed in the DHNN-SAT models.

Methodology for Objective 4: A robust training algorithm can ensure the optimal training phase of DHNN-SAT, contributing to the optimal synaptic weights (Bazuhair *et al.*, 2021). However, it is important to note that the attainment of an optimal training phase does not necessarily ensure an optimal retrieval phase. It is important to highlight that the term "optimal" in optimal final neuron states refers to the requirement of the neuron state to reach both the global minima and diversity. To achieve optimal final neuron states, the retrieval capability of DHNN-SAT can be enhanced by introducing a new activation function as well as a mutation-based scheme. The new activation function is a three-segment function that modifies the update rule when the local field is equal to 0, while keeping the update rule unchanged when the local fields are greater than or less than 0. By implementing this new activation function, the neuron state in the subsequent time step remains consistent with the state in the previous time step when the local field is equal to 0. Another important consideration in the retrieval phase is the diversity of solutions, which is particularly significant for logical mining purposes. To address the issue of high similarity, a mutation-based mechanism called Swarm Mutation has been proposed to enhance the diversity of the final neuron states. Swarm Mutation includes two processes: the detection process and the mutation process. In the detection process, the Swarm Mutation mechanism identifies which neuron states should be mutated and which should remain unaltered. Subsequently, the neuron states of 2SAT that match the benchmark neuron states are mutated to other optimal neuron states based on a specific probability. It is crucial to note that even after Swarm Mutation, the solutions still represent global minima. The outcomes and discussion throughout the retrieval phase are grounded on the utilization of a potent training algorithm HDEA during the training phase. Through HDEA, the proposed DHNN-YRAN2SAT model obtained the optimal synaptic weights. The outcomes and discussion of the retrieval phase are classified into two parts: (1) When using the proposed activation function but the absence of Swarm Mutation, performance metrics concentrate on the global minimum ratio, energy error, the total number of variations, and the similarity index. It is noteworthy that Swarm Mutation does not affect these parameters. (2) While utilizing both the proposed activation function and Swarm Mutation, diverse performance metrics focus on the mutation of the final neuron states. The total number of mutated neuron states and various similarity metrics are employed to analyze the transformations in neuron states before and after Swarm Mutation.

Scope of Methodology 4: Even though this particular methodology can be considered as an optimal approach for ensuring the optimality of the retrieval phase of the DHNN, several scopes need to be considered to ensure the reproducibility of the objective in this thesis. Firstly, it is important to note that the diversity function is solely applied to the second-order clauses. The rationale behind this is that second-order clauses tend to have a higher probability of satisfying interpretation when compared to first-order clauses (Alway *et al.*, 2022). To illustrate this further, let us consider the fact that for one second-order clause to be true, there are three out of four possibilities that can be realized. On the other hand, first-order clauses can only have one interpretation that can make them true. Consequently, it becomes clear that the diversity evaluation for the first-order clause becomes negligible. Secondly, it is worth mentioning that the diversity measure of the final solutions is exclusively based on the benchmark neuron states. The reason behind this lies in the fact that once the proposed logical rule is randomly generated, the benchmark neuron states become fixed. Moreover, it is crucial to highlight the fact that the benchmark neuron states correspond to the global minimum solution. At this particular point, the final neuron states can utilize the benchmark state as a reference point to explore other potential global and diverse solutions through Swarm Mutation.

Methodology for Objective 5: The existing logic mining models have demonstrated acceptable classification accuracy across various fields of datasets (i.e. Kasihmuddin *et al.*, 2022; Zamri *et al.*, 2024; and Jamaludin *et al.*, 2022). However, the main problems in most of these models lie in the absence of a flexible logical rule and suboptimal best learning logic. Consequently, existing models face difficulties in reaching high classification accuracy. To address this issue, this thesis proposes an improved logic mining model known as the Y-Type Random 2-Satisfiability Reverse

Analysis. Firstly, the proposed model will formalize a pre-processing phase to carry out data cleaning, supervised attribute selection processes, and data preparation. The data cleaning process involves three steps: data imputation for missing values, the frequency method for handling categorical entries, and the *k*-mean clustering method for discrete entries. Then, the supervised attribute selection method referred to as the Modified Correlation Analysis Method is employed to select the top *n* attributes with the highest correlation coefficient to the dependent attribute of the dataset. Subsequently, the selected attributes are transformed into bipolar forms of {1, -1} to ensure compatibility with the datasets in DHNN. Following this, the data preparation process proceeds with train test splitting and *k*-fold cross validation of all entries using a predefined ratio that aligns with existing works. Secondly, the proposed logic mining model introduced a novel method for generating the best learning logic that considered both frequency and the highest true values. The best learning logic was then trained using the proposed Hybrid Differential Evolution Algorithm as outlined in Objective 3 and aligned with the WA method. This approach was adopted to ensure that the generated best learning logic can represent the majority patterns of the dataset. Thirdly, the proposed Swarm Mutation in the retrieval phase of DHNN aids the logic mining model in widening the search space and identifying non-overfitting induced logic. Subsequently, the induced logic with the highest accuracy is considered as the best induced logic. The proposed logic mining model will be evaluated and investigated using 20 repository datasets obtained from reputable databases (UCI and Kaggle), as well as one case study on the Alzheimer's Disease Neuroimaging Initiative on the 19th of December 2022. The performance of the proposed logic mining model will be compared against all existing logic mining models based on accuracy, precision, sensitivity, specificity, and Matthew's correlation coefficient metrics.

Scope of Methodology 5: While this methodology may serve as an optimal approach for extracting information from real-life datasets, there exist several limitations that must be addressed to ensure the reproducibility of Objective 5 in this thesis. Firstly, the transformation to bipolar forms utilizing the frequency method will be implemented where the data is categorical. To facilitate fair comparison and reproducibility, the same train-test splitting and *k*-fold cross validation ratio will be employed across all models and held as constants (i.e. Rusdi *et al.*, 2023; Zamri *et al.*, 2024; Manoharam *et al.*, 2023). Lastly, the datasets obtained for this thesis were sourced from the University of California Irvine (UCI) Machine Learning Repository and Kaggle.com. By utilizing these acquired datasets, the proposed logic mining model can be compared against existing logic mining models in various fields.

1.9 The Organization of the Thesis

The thesis is structured in the following manner. Chapter 1 initiates the thesis by providing a comprehensive introduction, starting with an overview that establishes the research context, followed by the direction, philosophy, and problem statements that guide the study. The chapter also outlines the research questions, objectives, methodologies, scopes, a brief thesis organization, and the summary. In Chapter 2, a comprehensive review of various significant studies will be carried out on a series of follow-up studies regarding DHNN, the Satisfiability logical rule in DHNN, Evolutionary Algorithms, Optimization Algorithms in DHNN, and Logic mining. Moving forward to Chapter 3, a comprehensive explanation will be provided on the background study and foundational knowledge of all the previously mentioned domains. Transitioning to Chapter 4, a detailed explanation will be given on the specific strategies of the proposed model in this thesis. Firstly, the formulation