PNEUMONIA DETECTION WITH GAME-THEORETIC ROUGH SETS

A Thesis

Submitted to the Faculty of Graduate Studies and Research

In Partial Fulfilment of the Requirements

For the Degree of

Masters of Science

In

Computer Science

University of Regina

By

Suby Singh

Regina, Saskatchewan

January, 2022

UNIVERSITY OF REGINA

FACULTY OF GRADUATE STUDIES AND RESEARCH

SUPERVISORY AND EXAMINING COMMITTEE

Suby Singh, candidate for the degree of Master of Science in Computer Science**,** has presented a thesis titled, ***Pneumonia Detection with Game-Theoretic Rough Sets,*** in an oral examination held on December 8, 2021. The following committee members have found the thesis acceptable in form and content, and that the candidate demonstrated satisfactory knowledge of the subject material.

External Examiner:	*Dr. Yin Choong Yow, Software Systems Engineering
Supervisor:	*Dr. JingTao Yao, Department of Computer Science
Committee Member:	*Dr. Yiyu Yao, Department of Computer Science
Committee Member:	*Dr. Shakil M. Khan, Department of Computer Science
Chair of Defense:	*Dr. Chun-Hua Guo, Department of Mathematics & Statistics

*via Zoom Conferencing

Abstract

Pneumonia causes severe repercussions if not acted upon at the right time. Machine learning has been applied to classify chest X-ray images into pneumonia-positive and pneumonia-negative classes to allow an early diagnosis and support medical experts' decisions about pneumonia. Nonetheless, the previous attempt focus on binary classification that may not consider the possibility of the presence of uncertain information in chest X-ray images. This method forces the system to have definite and unreliable decision on suspicious cases of pneumonia. The three-way classification can overcome the shortcoming of binary classification by classifying X-ray images into three classes, pneumonia-positive class to categorize instances that represent pneumonia; pneumonia-negative class to categorize instances that represent normal chest X-ray; pneumonia-indecisive class to categorize instances, for which making decision based on the available information is difficult.

In this study, chest X-ray images are used to experiment and we employ three-way classification, supported by rough sets, to interpret the doubtful X-ray images that lack complete information. The training set of X-ray images is divided into three distinct regions based on testing and test-treatment thresholds. This threshold pair determines the decision of no treatment, treatment, and delayed treatment. The three regions in three-way classification are termed as positive, negative, and boundary. Definite decisions, such as "treatment" or "no treatment" for pneumonia can be inferred from positive and negative regions, respectively. No decision is made for

the boundary region until supporting facts are collected. Conventional rough sets uses a stringent threshold pair that restricts the number of instances being classified for definite decision. Conversely, probabilistic rough sets mitigates this limitation by allowing the use of a modified value for the threshold pair. This theory can classify more instances into positive and negative regions. Hence, we apply probabilistic rough sets to derive three classes, but it is crucial to achieve an appropriate testing and test-treatment threshold pair.

In this study, we apply game-theoretic rough sets (GTRS) to determine a suitable threshold pair. The evaluation criteria, accuracy, and coverage of the three-way classification model are played as players in a competitive game. The game is reformulated continuously as the player modifies their strategy. Consequently, a suitable threshold pair can be learned that balances these aspects and achieves an optimal trade-off. By using the achieved threshold pair from GTRS, three classes are defined for the pneumonia classification system. From the first two classes, pneumonia-positive or pneumonia-negative decisions can be made. In the remaining class, we explore X-ray images that lack the crucial information for making inferences. With the third class, the healthcare system would be able to re-examine the indecisive chest X-ray images by conducting tests instead of taking definite decisions based on incomplete data.

Based on the experiment, we achieve an accuracy score of 96.25% while covering 64.01% of the test set for certain decisions. The test results are also compared with various other related work and rough sets models. The GTRS based probabilistic rough sets model obtains a greater coverage of data than the conventional rough sets and better accuracy as compared to 0.5-probabilistic rough sets. This thesis is expected to provide an understanding of GTRS from its application perspective.

Acknowledgement

I would like to express my sincere and most profound regards to my supervisor Dr. JingTao Yao for his time, guidance, encouragement, patience, and financial support when I carried out this research. His invaluable advice assisted me in staying focused on the research and writing of this thesis. It was my distinguished honor and pleasure to work under his immense supervision.

I am also grateful to Dr. Shakil M. Khan and Dr. Yiyu Yao for their valuable comments and suggestions about the thesis. They helped me in improvising the proof of concept and presentation of the thesis.

I would also like to take this opportunity to acknowledge my gratitude towards the financial support of the Department of Computer Science and the Faculty of Graduate Studies and Research. A special acknowledgement and big thank you to my parents, family, and friends for their constant support throughout my tenure at the University of Regina.

Post Defense Acknowledgement

At this point I thank my internal committee members, Dr. Shakil M. Khan and Dr. Yiyu Yao, and my supervisor, Dr. JingTao Yao for being available during my defense and providing their comprehensive and useful feedback. I would also like to extend my appreciation to my external examiner, Dr. Kin Choong Yow for his insightful comments and suggestions. I also thank Dr. Chun-Hua Guo for conducting and chairing the defense session.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

In this chapter, we analyze the shortcomings of previous work and introduce a method to overcome those limitations. Furthermore, the objective of the thesis and the remaining chapters of the thesis are described in Sections 1.3 and 1.4, respectively.

1.1 Problem Statement and Motivation

Pneumonia is an inflammatory lung disease that is caused by a bacterial and fungal infection. It is a significant cause of death in all age groups. According to the World Health Organization, pneumonia has a range of severity from mild to life-threatening. Approximately 4 million patients die from it every year [36]. The high death rate emphasizes the importance of early diagnosis.

Medical experts use chest X-ray images as the most traditional way to diagnose pneumonia. Quick recognition of pneumonia from chest X-ray images requires experts with adequate experience and good eyesight to identify pneumonia fluid simply from looking at the X-ray image. Furthermore, health care providers must have an adequate number of experts or a reliable classification system to effectively treat the growing number of pneumonia cases. Early detection of pneumonia is essential to its treatment [45]. Therefore, computer-based systems are being explored to assist

experts and speed up the process of pneumonia treatment [15, 45].

A developing application of classification techniques in the medical imaging field can assist experts in determining the disease from images [15, 28]. In these techniques, features are extracted from medical images and a classification algorithm is applied to classify them into disease-positive and disease-negative categories based on the extracted features [43].

Classifying medical images using a computer-based method helps making a swift decision on having the disease [15, 45]. On the basis of this concept, the thesis presents a method to examine chest X-ray images and delivers a classification system to determine the presence or absence of pneumonia. Medical experts can make prompt decisions on chest X-rays using the technique used in this study.

Pneumonia identification from chest X-ray images thus far has been considered as a binary classification problem. There have been research performed for detecting pneumonia using machine learning classification techniques [45], especially deep learning [12, 22, 44, 47, 50, 61]. In these experiments, the chest X-ray images were categorized into pneumonia-positive and pneumonia-negative classes based on the extracted statistical features. Using 3-layer convolutional neural networks (CNNs) and residual network architecture, Saul et al. developed a system that classifies each chest X-ray image into two definite classes [47]. In addition, Luka et al. proposed a 5-layer CNN architecture [44], Okeke et al. proposed a 4-layer CNN architecture [50], and Verma et al. proposed a 6-layer CNN architecture [12] to generate a binary classification system for pneumonia. In another experiment [45], Rajpurkar used 121-layer CNN to develop a classification system to predict the probability of pneumonia infection.

Previous work has employed a binary classification method that categorizes chest X-ray images into pneumonia-positive and pneumonia-negative classes, yielding a definite decision of either "treatment" or "no treatment" for pneumonia disease. Ibrahim

et al. conducted a similar experiment by using a pre-trained deep learning model, the AlexNet model, to develop various pneumonia classification systems that classify various types of chest X-ray images as belonging to pneumonia disease, including bacterial pneumonia, viral pneumonia, and COVID-19 [22]. In spite of this approach classifying chest X-ray images into multiple classes, it overlooks the uncertainty of data in chest X-ray images and forces the system to have definite opinions on each X-ray image.

Binary classification forces the chest X-ray image classification system to make a definite decision for every instance. A real-world object, however, necessarily contains uncertain and incomplete information [52]. This means that there may be chest X-ray images that lack certain information that would be required to make a firm decision. This inability to handle uncertain information and constraining the system to make definite decisions on doubtful cases could be the pitfall of binary classification.

Binary classification also suffers from an increased false prediction rate [52]. It is possible that reducing the misclassification of pneumonia-positive X-ray images into negative class will increase the misclassification of pneumonia-negative X-ray images into positive class, and vice versa. Consequently, binary classification may not be a good method for developing a high quality pneumonia classification system. It is important to consider the uncertainties associated with the data in its classification [52]. These limitations provide the basis for the proposed system in this study.

With computer-aided classification, the unsupported decision can be mitigated if a carry-over option is included in the algorithm. By not mandating every instance be classified only for definite decisions, this option can reduce the false prediction rate. Furthermore, although deep learning and machine learning concepts have been derived from a phenomenon that humans do every day in nature, such techniques often require high configuration machines to compute the complex algorithms [38]. Aspiring to overcome these limitations, this research proposes a game-theoretic rough

sets (GTRS)-based three-way classification approach for pneumonia. As a result, the method does not require a high configuration machine for computation of the algorithm and can probably be considered an optimal solution in the literature.

Three-way classification method, proposed by Yao, extends the widely applied binary classification technique by adding a third class to account for instances with uncertain information [70]. The third class defers making decisions on instances that lack necessary information to make a definite decision on them [46, 74]. All three classes are implemented based on the concept of acceptance, rejection, and deferment. The acceptance and rejection classes are treated with firm decisions, whereas the deferment classes are not, in fact, a decision is detained until supportive information is obtained.

By the same token, chest X-ray images can also be classified into three classes based on the presence of pneumonia. It follows that acceptance represents a class of pneumonia-positive X-ray images; rejection represents a class of pneumonia-negative or normal X-ray images; deferment is a category of indecisive or suspicious X-ray images that cannot be identified as either pneumonia-positive or pneumonia-negative. In comparison with binary classification, the three-way pneumonia classification system can confidently support the diagnosis of pneumonia-positive and pneumonia-negative cases (that is, experts can determine whether to treat or not treat pneumonia in these cases). By not forcing the system to make definite decisions on indecisive X-ray images, it also reduces false prediction rates. Images of this type can be re-examined with additional tests and information rather than risking making a mistake based on incomplete information available at the moment.

In the field of data analysis, rough sets, proposed by Pawlak [41], is a mathematical method to deal with incomplete and uncertain information about instances. It treats uncertainty with an information table. The information table is a normal table with instances arranged in the form of rows and columns. The basic element of

this theory is the equivalence class that is developed by grouping instances with the same values across all attributes (columns) in the information table. In rough sets, according to three-way classification, the equivalence classes are classified into three disjoint regions, namely, positive, negative, and boundary, which can correspond to pneumonia-positive, pneumonia-negative, and pneumonia-indecisive, respectively.

Conventional rough sets (Pawlak's rough sets) considers the qualitative relationship between a set and equivalence classes [66]. Nevertheless, qualitative relationships fail to consider the degree of overlap between the set and equivalence classes. In detail, a set of objects, in which most of the instances belong to one region (i.e., either positive or negative), regardless of this fact, all instances of the set are moved to the boundary region. Thus, Pawlak's rough sets has the shortcoming of requiring the positive and negative regions to be error-free [38, 63, 65]. As a consequence, this phenomenon states that pneumonia-positive and pneumonia-negative classes can tolerate no uncertainty, placing all uncertain data into the pneumonia-indecisive class. This limitation results in fewer instances being classified into positive and negative regions for the decision of treatment and no treatment, respectively.

Pawlak's rough sets can only classify a small portion of the data when most of it is uncertain, resulting in a lower usability of the system on uncertain data. A measure of the system's usability is the percentage of data that can be covered for making definite decisions. In other words, it shows the degree of usefulness or applicability of the classification model..

To overcome Pawalak's rough sets' stringent rules, probabilistic rough sets theory is proposed . It is a quantitative model that takes into account the degree of overlap between the equivalence class and set and tolerates a degree of uncertainty or error in positive and negative regions [71, 73]. In probabilistic rough sets, the degree of overlap between the equivalence class and set is coined as the conditional probability of the equivalence class. This is equivalent to the probability of pneumonia in the

chest X-ray. It contributes significantly to the definition of three regions in three-way classification [63, 66, 80]. In consequence, probabilistic rough sets enables a greater number of instances to be classified. As a result, the boundary region is reduced, which increases the usability of the system.

Several approaches have been proposed to interpret the three regions of three-way classification system, including decision-theoretic rough sets (DTRS) [60, 62], information-theoretic rough sets (ITRS) [59], game-theoretic rough sets or GTRS [59], decision-theoretic shadowed set model (DTSS) [63], and game-theoretic shadowed sets (GTSS) [64]. According to all of these approaches, the equivalence classes of an information table are divided into three regions based on their conditional probability and a threshold pair (α, β).

A threshold pair (α, β) is defined for the positive and negative regions, where α and β are the magnitudes of positive and negative regions, respectively. The threshold pair therefore plays an important role in controlling the three regions of three-way classification. This pair of thresholds may be referred to as testing and test-treatment thresholds [40, 59]. The development of a high quality classification system based on a three-way classification technique in probabilistic rough sets requires identifying an optimal threshold pair of test and test-treatment thresholds that allows more instances in positive and negative regions and as accurately as possible [21].

Game-theoretic rough sets, proposed by Yao and Herbert [21], applies game theory to tackle the challenging problem of finding the optimal threshold pair (α, β) in probabilistic rough sets. According to game theory, a competitive game is formulated between two or more players. The game is considered to have reached equilibrium when all players have obtained their maximum reward and no player is interested in updating their strategy further. Game-theoretic rough sets formulates the game by considering performance evaluation criteria of the three-way classification system as the game players and their rewards, and a set of possible threshold pairs as their

strategies.

The final threshold pair is derived from the strategy of the player at the game equilibrium point [60]. When a player plays its strategy to maximize the reward, other players' rewards are affected too. Therefore, we need to establish a trade-off between them to get effective rewards for all players. The idea of a trade-off is that it is a compromise between players, in which all players receive their best rewards without compromising the rewards of other players.

The GTRS reformulates the game with an updated strategy until the players achieve their optimal trade-off. By analyzing the trade-off between the evaluation measures of the system [60], GTRS assists in finding the final threshold pair. The main benefit of using GTRS is that it allows the system to learn classification rules from the data itself because the rewards of players are the criteria used to evaluate the classification performance of the system. So, in every iteration, it learns from evaluating the system rigorously. Previously, GTRS has been applied to three-way email spam filtering [74], uncertainty analysis [5], and recommender systems [4].

This study employs GTRS to determine the final threshold pair based on the trade-off perspective of the players. In a three-way pneumonia classification system, the selected threshold pair can effectively control the three regions (classes). The three classes are: 1. pneumonia-positive, 2. pneumonia-negative, and 3. pneumonia-indecisive. A classification rule corresponding to each class is also interpreted in this study. The proposed system has been experimented with Mendeley's chest X-ray image dataset [26]. A comparison is then made between the results of the proposed system and those of the classification techniques based on 0.5-probabilistic rough sets, Pawlak's rough sets, and machine learning.

1.2 Research Objectives

In this section, we will discuss briefly the purpose of this research, which is to apply three-way classification techniques using GTRS to develop a pneumonia classification system.

1.2.1 A Framework to Handle Uncertainty in Three-way Pneumonia Classification System

In order to overcome the limitations of binary classification, a rough sets-based method of uncertainty analysis, three-way classification has been employed in this study. It provides a framework to develop a pneumonia classification system by partitioning the set of chest X-ray images into three pairwise disjoint classes using a threshold pair (α, β). Pneumonia-positive, pneumonia-negative, and pneumonia-indecisive are the classes.

The pneumonia-positive class is defined by a test-treatment (acceptance) threshold value (α) and contains X-ray images that are positive for pneumonia and require treatment without further testing. The pneumonia-negative class is defined by a testing (rejection) threshold value (β) and it consists of X-ray images that are without pneumonia infection, in other words, can be termed as normal and do not require treatment or further testing of pneumonia disease. X-ray images from this class are rejected as pneumonia-positive. On the basis of the X-ray images from these two classes, decisions (e.g., "treatment" or "no treatment") can be made. The pneumonia-indecisive class covers the suspicious cases of pneumonia, for which neither pneumonia-positive nor pneumonia-negative decision can be considered. In fact, decisions on X-ray images from this class are delayed until supporting details are gathered and analyzed.

1.2.2 Application of Game-theoretic Rough Sets

The proposed system in this thesis applies GTRS in a three-way classification algorithm to resolve the key issue of probabilistic rough sets. The final threshold pair is determined by the trade-off between the two evaluation criteria of the model, accuracy and coverage. Accuracy is the percentage of X-ray images classified correctly into pneumonia-positive and pneumonia-negative classes, whereas coverage is the fraction of X-ray images classified into pneumonia-positive and pneumonia-negative classes [4]. These two evaluation measures are incompatible, since an increase in classification accuracy can lead to a decrease in coverage, and vice versa. In this study, it is found that high accuracy of pneumonia classification may result in a lower coverage of X-ray images in definite regions (i.e., positive and negative), while high coverage may result in lower accuracy of pneumonia classification.

In this experiment, all three elements of GTRS are used, namely, the formulation of a competitive game between accuracy and coverage, a mechanism of repetitive learning, and the determination of the threshold pair from the trade-off. The first step involves defining GTRS artifacts like game players, their payoffs, and strategies. In the second step, the game formulation is repeated with the selected threshold pair from the previous iteration as the initial pair. In the third step, the game is stopped based on some criteria and the final threshold pair is selected. The optimal threshold pair is a threshold pair that accurately classifies as many objects as possible into pneumonia-positive and pneumonia-negative classes and is determined by the game equilibrium or by the player's trade-off.

1.2.3 Selection of Threshold Pair using Pareto Optimality Strategy

For three-way classification, the determination of a suitable threshold pair is crucial since it controls the derivation of three classes. A threshold pair that decreases the number of X-ray images in pneumonia-indecisive class may increase the system's usability while reducing its classification accuracy at the same time. Additionally, threshold pairs to improve classification accuracy may decrease the system's usability. With GTRS, we can determine an optimal threshold pair that targets both accuracy and usability factors of the model [60].

Since we want to achieve a trade-off between two conflicting evaluation measures (i.e., accuracy and coverage), the Pareto optimality strategy can be employed to find the game outcome by establishing a co-operation between these two criteria [11]. Additionally, this study experiments with Nash equilibrium in finding game solutions, but observes the best result with Pareto optimality. The results of the experiment are also compared with Pawlak's rough sets and other classification models in this literature that have been explored previously to develop a pneumonia classification system. In contrast to Pawlak's rough sets technique, three-way classification based on GTRS in probabilistic rough sets improves coverage of the data, but slightly compromises the accuracy of the system. Coverage of the data in positive and negative regions defines usability of the system from a three-way classification perspective [4].

1.3 Thesis Contributions

The objective of this section is to highlight the contribution of the thesis towards using GTRS to develop a pneumonia classification system based on three-way classification.

1.3.1 A Provision to Re-examine Chest X-ray Images with Incomplete Information

In the three-way pneumonia classification system, the three classes represent distinct levels of information uncertainty in X-ray images. This research contributes to the definition of pneumonia-indecisive class, which includes X-ray images with incomplete and uncertain information. By doing so, we have the opportunity to gather and analyze additional information about the patient before making a concrete decision on indecisive chest X-ray images. This would allow the system to align with manual approach of the medical expert to deal with doubtful cases of pneumonia and reduce the expected cost by avoiding to make incorrect decision on them.

The level of uncertainty and probability of pneumonia disease can be expressed by the conditional probability of the equivalence class. The definite equivalence classes will have the conditional probability of "1.0" for pneumonia-positive and conditional probability of "0.0" for pneumonia-negative classes. Remaining equivalence classes are considered to have a certain level of uncertainty between 0 and 1. In this study, we apply probabilistic rough sets so that the positive and negative classes can be defined with a certain degree of uncertainty and more instances can be classified for definite decisions.

1.3.2 Trade-off Analysis between Accuracy and Coverage using GTRS

In three-way classification systems, accuracy and coverage are two major contrasting evaluation criteria. To achieve optimal outcomes for accuracy and coverage criteria, this thesis applies GTRS to establish an agreement between these two criteria and to obtain a suitable threshold pair. In order to investigate trade-offs, a two-player game between accuracy and coverage is created. The game considers the different

configurations of the threshold pair as a set of strategies for the players. A repetitive game formulation and analysis of its results while modifying the threshold pair allows the system to learn a suitable threshold pair automatically from the data itself.

1.3.3 Applying Pareto Optimality Strategy to Obtain the Game Solution

The proposed research contributes to finding a game solution that is optimal. In this thesis, Nash equilibrium and Pareto optimality are used to search for the optimal threshold pair. We observe that changing a player's strategy unilaterally affects the payoff of another player using the Nash equilibrium approach. By applying Pareto optimality, the two players can achieve a better level of cooperation than in a Nash equilibrium. Almost all existing applications of game-theoretic rough sets are based on Nash equilibrium [74]. This thesis particularly focuses on finding an optimal threshold pair using the Pareto optimality strategy.

1.4 Structure of the Thesis

The thesis is divided into the following chapters:

- Chapter 2 focuses on understanding research related concepts, such as rough sets, three-way classification, probabilistic rough sets, and game-theoretic rough sets. Section 2.1 summarizes the general approach used in previous attempts of developing a pneumonia classification system. Section 2.2 states the basic notion of rough sets and elaborates on handling the uncertainty of the objects using probabilistic rough sets. Section 2.3 provides an overall idea about game theory and game-theoretic rough sets.

- Chapter 3 introduces three-way pneumonia classification system with game-

theoretic rough sets in Section 3.1. Section 3.2 details a step by step implementation of the proposed method. Section 3.3 visualizes the difference between conventional rough sets and probabilistic rough sets. Section 3.4 states the evaluation measures that can be used to evaluate classification performance of the system and formulate the game. Section 3.5 demonstrates on establishing a trade-off between accuracy and coverage to reach a suitable threshold pair.

- Chapter 4 focuses on the empirical implementation of the proposed model using a chest X-ray image dataset and evaluates its performance. Section 4.1 outlines the dataset used in this research and determines the attributes of the decision (information) table in rough sets analysis. Section 4.2 leverages GTRS in determining the suitable threshold pair from a trade-off perspective and implements the proposed model. Section 4.3 interprets the three classification rules, evaluates the classification performance of the proposed model, and compares its experimental results with Pawlak's rough sets and other classification models.

- Chapter 5 concludes the research work in Section **??** and summarizes the key contributions of the thesis towards developing a pneumonia classification system in Section 5.1. It also highlights the future research work in Section 5.2.

1.5 Summary

The chapter outlines the problem statement in the context of the pneumonia classification system highlighting the shortcoming of binary classification. This chapter offers a detailed explanation and motivation behind the application of game-theoretic rough sets in probabilistic rough sets in order to identify the threshold pair. A brief summary of this research is provided in which the contribution of this research to developing a high quality pneumonia classification system is explained, followed by an explanation of the thesis' organization.

Chapter 2

Related Work and Background Knowledge

In this chapter, we describe a generally followed approach to developing a pneumonia classification system in Section 2.1. Then, in Section 2.2, we present an insight into handling uncertainty of the data using conventional rough sets and probabilistic rough sets. Section 2.3 presents a detailed overview of game theory in rough sets (i.e., game-theoretic rough sets). The basic notion of rough sets and game-theoretic rough sets are also discussed in this chapter.

2.1 General Approach for Pneumonia Classification in Related Work

The related attempts to identify pneumonia rely mainly on extracting statistical features from chest X-ray images and feeding those features to a classification algorithm to categorize them [12, 47]. Figure 2.1 illustrates some of the steps of methodology used in previous experiments.

Chest X-ray images are initially transformed into a consistent format using noise

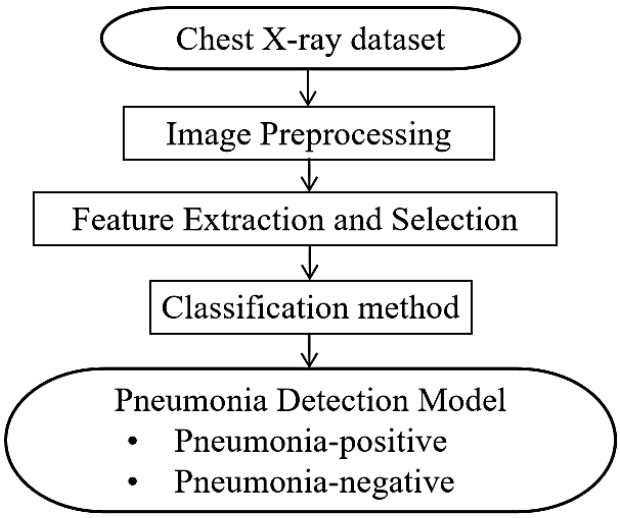

Figure 2.1: Existing approach (oval: input and output. rectangle: process)

removal and image enhancement preprocessing techniques. From preprocessed chest X-ray images, statistical features such as contrast and keypoint descriptors are extracted. Later, a set of discriminative features from the large pool of feature vectors is selected using principal component analysis and sequential forward search methods [53]. The selected features are then fed to a classification algorithm such as support vector machine or convolutional neural networks that generate a model to classify X-ray images [12, 22, 44, 47, 50, 61]. For a given chest X-ray image, the classification model predicts a label from two classes (i.e., pneumonia-positive and pneumonia-negative).

The following subsections discuss each step of the generally followed approach in existing experiments, such as image preprocessing, feature extraction and selection, and classification.

2.1.1 Image Preprocessing

The chest X-ray image dataset may consist of X-ray images that represent inconsistent and noisy information. There may also be parts of the X-ray images that are not helpful in the classification process, but could increase the computational and time

complexity of the classification algorithm. Moreover, images in the dataset can have different illumination scales; some may be brighter, while others may be darker.

By preprocessing the data before applying the classification algorithm, data analysis techniques can be significantly improved [68]. Hence, X-ray images are first preprocessed to transform them into a suitable format required for the classification algorithm. Some of the preprocessing techniques applied for images in existing work involve cropping of the image to an interesting area, widening of the image color space, and utilizing histogram equalization to enhance contrast or preserve brightness of the image [47, 61].

This step is necessary to ensure that meaningful features of the images are extracted before they are fed into a classification algorithm. The image augmentation technique is another technique used in image preprocessing to handle imbalanced datasets and avoid overfitting of the classification system (i.e., modeling error that is too closely aligned to a group of data points). In order to increase the diversity of data being classified, image augmentation techniques are used. Examples of these techniques include rotation, flipping, shifting, increasing brightness levels, etc [12].

Luke et al. resized the chest X-ray images to 200×200 pixels and normalized the value of pixel intensity by dividing it by 255 during the image preprocessing step. As a result, every image has floating point values for pixels instead of integer values between 0 and 255. This process should provide a positive impact in the performance of CNN. Additionally, in order to avoid overfitting caused by an imbalanced dataset, their experiment involves cropping, rotating, and zooming asymmetrically the input X-ray image [44].

Saul et al. used three types of image processing techniques: contrast increment, X-ray image color space widening, and brightness increment. Using the phenomenon of medical experts examining the X-ray image under light, they followed to increase brightness of the image by parsing its every pixel to add a constant to Red Green

Blue (RGB) values. Similarly, they applied the second technique to increase the contrast of the image to make the edges solid and certain parts promptly visible. Their experiment also expands the color space by finding the average of the RGB value and multiplying RGB with its average value [47].

Similarly, Okeke et al. implemented image augmentation techniques such as rotating images by 40 degrees, shifting widths by 0.2%, shifting heights by 0.2%, clipping images in a counterclockwise direction by shear range of 0.2, zooming images randomly to 0.2%, and horizontally flipping images to solve the problem of overfitting [50]. In the same way, Garima and Shiva rescaled the raw chest X-ray images to 1/255, rotated the image by 45 degrees, performed height shifting and width shifting by 0.2%, zoomed the image with a shear range of 0.2, and flipped the image horizontally [12]. In another experiment, Gurmail and Kin-Choong used chest CT scans to detect COVID-19 from medical images. They resize the image to 224×224 dimension as expected by the model [14]. By using a radius of 3, Suganthi et al. modified the contrast of the image, and applied morphological operations to discard the irrelevant bone parts and dilation to smooth the lung boundary [51]. Later, these preprocessed chest X-ray images are used for the segmentation process as per region of interest using K-means clustering [51].

2.1.2 Feature Extraction and Selection

In this step, the two-dimensional chest X-ray image is converted to a feature vector that can be used as input to a classification algorithm. Some features that have been worked with in previous attempts include the entropy of the local binary pattern, as well as the texture information of the image, such as contrast, which describes the structural arrangement of the image surface [55].

In order to reduce the processing time and complexity of the algorithm, only the most relevant features are considered. Selection of features is important for opti-

mization of the classification algorithm. The subset of features that maximizes the criterion function is selected. In previous approaches, sequential forward selection, principal component analysis, or a combination of the two, such as intersection or union, have been used to select a subset of the effective features [47, 61]. An X-ray image can be decomposed according to different features. The aim is to extract useful information from the images.

Suganthi et al. extracted gray level co-occurrence matrix (GLCM) features from preprocessed X-ray images and selected dominant features using Analysis of variance (ANOVA) tests [51]. Similarly to the histogram calculation, pixel pair values are selected and the count is incremented by 1 at that pixel location in the GLCM matrix. The entire image is segmented into smaller chunks of 8×8, 16×, and 32× and a GLCM matrix is generated for each chunk. In this step, GLCM features such as correlation, homogeneity, energy, entropy, and contrast are calculated [51]. Shabnam and Mohamed suggested using the fuzzy C-means algorithm instead of discrete wavelet transforms, wavelet frame transforms, and wavelet packet transforms, etc. Clustering by fuzzy C-means stores the weight of each point associated with respect to a cluster, but it is slower than all other techniques due to exhausted calculations for every pixel [24].

For most experiments using CNN, which has an in-built feature extraction function, X-ray images are directly fed to CNN after preprocessing and augmentation [12, 22, 44, 47, 50, 61]. The training input dataset is first passed to a neural network called a feature extraction network. As a result of this network, features have been extracted that are passed onto a classifier network [35]. The feature network is composed of convolutional and pooling layers. The convolutional layer comprises a set of digital filters that execute the convolution operation on the input image dataset and the pooling layer performs dimensionality reduction task [35].

The convolution layer transforms the input image by applying the convolution

network. With an image classification system, for example pneumonia classification using chest X-ray images, the convolution network determines which filters will extract the most discriminating features to differentiate the classes. Using aggregation on the pixel values, a pooling layer samples pixel values from neighbors to get a single pixel value. One method of aggregation is to take the maximum value of a pixel from the group (max pooling) and another is to take the average of all pixels in the group (average pooling) [35].

In the experiment performed by Okeke et al., the feature extraction network consists of 3×3, 32; 3×3, 64; 3×3, 128; and 3×3, 128 convolutional layers, 2×2 max-pooling layer, and a Rectified Linear Unit (ReLU) activation function [50]. Luka et al. applied five convolutional layer of 3×3, 64; 3×3, 128; 3×3, 128; 3×3, 256; and 3×3, 512 sizes[44]. Similar approaches, 3-layer convolutional neural networks (CNN) with residual network architecture [47] and 6-layer CNN [12] were used on the training set. Rajpurkar used 121-layer CNN on the training set of 100,000 chest X-ray images to develop a classification model that classifies the chest X-ray images according to 14 diseases [45].

2.1.3 Classification Limitations of Existing Work

In general, existing approaches have focused mainly on applying deep learning to develop a classification system that can classify chest X-ray images for pneumonia. Deep learning is a branch of machine learning that is based on human cognitive processes and neural network architectures. A number of features of the images are extracted at multiple levels in which some lower levels define a higher level [34].

The most commonly used neural network for image classification is a convolutional neural network, which implements convolutional filters. Deep learning has the advantage that researchers don't have to work on feature extraction. The set of X-ray images are fed to convolutional neural networks that convert the two-dimensional

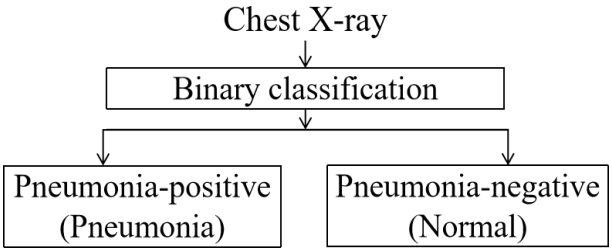

Figure 2.2: The two classes for pneumonia classification

feature vectors of all layers to a one-dimensional output layer. However, neural networks require a large, labeled, balanced, and diverse dataset to train the model effectively [34], whereas the dataset used in this thesis is imbalanced and inconsistent in terms of the number of X-ray images of two classes. Apart from the use of convolutional neural networks, Suganthi et al. utilized K-Nearest Neighbor (K-NN) and Support Vector Machine (SVM) classifiers in order to detect pneumonia [51].

Once the model is trained, it can be used to classify chest X-ray images into pneumonia-positive and pneumonia-negative categories. As shown in Figure 2.2, existing approaches have used binary classification of chest X-ray images. The majority of existing approaches provide good accuracy scores between 85 and 94 percent when evaluating the performance of the classification system, but do not address the uncertainty associated with chest X-ray images.

We encounter two fundamental issues when dealing with binary classification. In the first place, the system needs to make a definite decision for every case. The reality is, however, that imprecision and uncertainty are integral to the real-world object [52] and could prevent us from making definitive decisions. Making a solid decision based on X-ray images that are insufficient to support the decision and forcing the system to conclude a groundless decision. In this thesis, such images are referred to as indecisive X-ray images. In order to deliver an effective pneumonia classification algorithm, we must account for the uncertainty of the data. Another problem with binary classification is that it fails to improve the actual prediction rate since it may

misclassify indecisive X-ray images for definite decisions. Three-way classification based on rough sets provides a solution to the binary classification issue [70]. In a pneumonia classification system, convolutional networks can be used for multiclass classification, but it requires labeled data for each class. Available datasets have chest X-ray images only labeled as pneumonia-positive or pneumonia-negative. There are no X-ray images that have been labeled pneumonia-indecisive. The rough sets theory can be used to identify uncertain X-ray images from a set [41].

Rough sets is a mathematical data analysis approach to analyze incomplete and uncertain information of the instances [41]. The fundamental element of rough sets is an information table that is partitioned based on the same values on a set of attributes. A table with information is a regular table whose instances are arranged in the form of rows and columns. Three-way classification in rough sets divides the set of partitions from the information table into three pairwise disjoint regions, namely, positive, negative, and boundary. Those instances representing uncertain or incomplete information fall into the boundary region. The theory has been applied in many different fields, including engineering, banking, medicine, decision support, and others [41]. Examples include the three-way spam filtering system [78], recommender system [5], and the web-based medical decision support system [59].

In order to handle indecisive chest X-ray images and develop an effective decision support system for medical experts, rough sets theory can be applied in the field of pneumonia classification systems. Instead of making a wrong decision based on incomplete information in identified indecisive chest X-ray images, the system allows medical experts to gather additional details about such patients and analyze them more thoroughly before making a judgment. A wrong decision about the patient can result in more expenses or endanger the patient's life. Hence, the outcome of pneumonia-positive and pneumonia-negative from the three-way pneumonia classification system aligns with manual analysis of the medical experts that would

confidently lead the healthcare system to proceed with the further procedure. Alternatively, the pneumonia-indecisive outcome of three-way pneumonia classification system provides a provision to delay certain decisions on uncertain instances and a chance to look at the details again before concluding a decision. To my knowledge, this is the first application of three-way classification in pneumonia classification domain.

2.2 An Overview of Rough Sets

Pawlak proposed rough sets as a mathematics technique for analyzing instances that contain incomplete and uncertain information [41]. It deals with uncertainty in classification by constructing three regions of the universal set (U). In rough sets, U is represented by a decision table. Each row of the decision table contains a set of instances. All instances in the table are described by a finite set of attributes (column). A decision table can be represented by a mathematical tuple as follows [41],

$$DT = (U, At, \{V_a | a \in At\}, \{I_a | a \in At\}) \tag{2.1}$$

Where U is the finite set of instances, $At = \{C \cup D\}$ is the finite set of conditional C and decisional D attributes, V_a is the set of values for $\forall a \in At$ (domain of each attribute in At), and I_a is the information function ($I_a : U \rightarrow V_a$) to map an instance of U to a value from V_a. The Decisional attribute indicates which class an instance belongs to; the remaining attributes in the table are conditional attributes that represent features of the instance.

Equivalence classes are formed by grouping instances of U that have the same value for each conditional attribute. The equivalence relation on a set of instances

Table 2.1: An example of a decision table

Instance	A_1	A_2	A_3	Class
I_1	1	2	1	0
I_2	1	1	1	1
I_3	4	2	3	0
I_4	1	2	3	1
I_5	1	2	1	1
I_6	2	2	3	0
I_7	1	1	1	0
I_8	1	1	1	1
I_9	4	2	3	1
I_{10}	2	2	1	1
I_{11}	1	2	3	1
I_{12}	2	2	3	1

can be given based on a set of attributes $A \subseteq At$ such as,

$$E_A = \{(X_1,\ X_2) \in U \times U | \forall a \in A, I_a(X_1) = I_a(X_2)\} \tag{2.2}$$

where E_A is the equivalence relation according to the set of attributes $A \subseteq At$. Equation (2.2) suggests that two instances, X_1 and X_2 are indiscernible or equivalent if they have the same value for every attribute in A.

Equivalence relation E_A can be used to partition the universal set U based on the set of attributes A and can be denoted by U/E_A. Every subset from the partition (U/E_A) is referred to an equivalence class [66]. The equivalence class containing instance X_1 can be represented by $[X_1] = \{X_2 \in U | X_1\ E_A\ X_2\}$.

For example, consider the decision table (shown as Table 2.1), where the set of conditional attributes is $C = \{A_1, A_2, A_3\}$, and attribute "Class" is the decisional attribute D. Given an equivalence relation E_A, where $A \subseteq C$ and $A = \{A_1, A_2\}$, the partition of the table can be given by, $U/E_A = \{\{I_1, I_4, I_5, I_{11}\}, \{I_2, I_7, I_8\}, \{I_3, I_9\}, \{I_6, I_{10}, I_{12}\}\}$. Each subset in this partition is termed an equivalence class. For another equivalence relation E_A, where $A = \{A_1, A_2.A_3\}$, the partition of the table can be given by, $U/E_A = \{\{I_1, I_5\}, \{I_4, I_{11}\}, \{I_2, I_7, I_8\}, \{I_3, I_9\}, \{I_6, I_{12}\}, \{I_{10}\}\}$.

In rough sets, a set of instances can be described by two approximations, lower and upper approximations. These approximations are developed around the equivalence classes. For a given subset $P \subseteq U$ (i.e., some target concept or set), the lower approximation can be provided by the set of equivalence classes contained in P, and the upper approximation can be given by the set of equivalence classes containing P. The lower and upper approximations can be mathematically represented as [41],

$$\begin{aligned} \underline{apr}(P) &= \{a \in U | [a] \subseteq P\}, \\ \overline{apr}(P) &= \{a \in U | [a] \cap P \neq \phi\} \end{aligned} \tag{2.3}$$

From Table 2.1 and equivalence classes derived based on a set of attributes $A = \{A_1, A_2, A_3\}$, a set on instances, $P = \{I_2, I_3, I_6, I_9, I_7, I_{10}\}$ can be described by the lower approximation, $\underline{apr}(P) = \{\{I_3, I_9\} \cup \{I_{10}\}\}$ and upper approximation, $\overline{apr}(P) = \{\{I_3, I_9\} \cup \{I_{10}\} \cup \{I_2, I_7, I_8\} \cup \{I_6, I_{12}\}\}$, where $\cup$ is the union of equivalence classes in both approximations.

Positive, negative, and boundary (indecisive) regions are the results of lower and upper approximations. The positive region is equivalent to the lower approximation. The negative region is the complement of the upper approximation (i.e., objects from neither P nor equivalence classes containing P). The boundary or indecisive region can be given by the difference between lower and upper approximations.

$$\begin{aligned} POS(P) &= \underline{apr}(P) = \{a \in U | [a] \subseteq P\}, \\ NEG(P) &= \overline{apr}(P)^c = \{a \in U | [a] \cap P = \phi\}, \\ BND(P) &= \overline{apr}(P) - \underline{apr}(P) = \{a \in U | [a] \not\subseteq P, [a] \cap P \neq \phi\}. \end{aligned} \tag{2.4}$$

Equation (2.4) shows the degree of quantitative relationship between the target set P and equivalence class $[a]$. POS(P), NEG(P), and BND(P) are the positive, negative, and boundary regions, respectively. Condition $[a] \subseteq P$ represents that $[a]$

is completely contained in P. On flip, condition $[a] \not\subseteq P$ shows that $[a]$ should not be completely in P. Condition $[a] \cap P \neq \phi$ shows if $[a]$ intersect with the concept P. So, there can be some degree of overlap between $[a]$ and P in the boundary region, but no overlap between $[a]$ and P in case of the negative region.

For the given target concept or set in previous example, $P = \{I_2, I_3, I_6, I_9, I_7, I_{10}\}$, the Table 2.1 can be used to derive following three regions,

- $POS(P) = \{\{I_3, I_9\} \cup \{I_{10}\}\}$
- $NEG(P) = \{\{I_1, I_5\} \cup \{I_4, I_{11}\}\}$
- $BND(P) = \{\{I_2, I_7, I_8\} \cup \{I_6, I_{12}\}\}$

2.2.1 Three-way Decision

Three-way decision or three-way classification in rough sets was proposed by Yao [67, 69]. Three-way decision divides the universal set U into three pairwise disjoint regions and leads to the three classification rules. It can be pictorially represented as shown in Figure 2.3. The three decisions are acceptance, indecisive, and rejection corresponding to $POS(P)$, $BND(P)$, and $NEG(P)$ regions, respectively (Equation (2.4)).

Three decisions are controlled by a threshold pair (α, β). The threshold pair is defined for the positive and negative regions, where α and β represent the magnitude of positive and negative regions, respectively. Given a threshold pair (α, β) and based

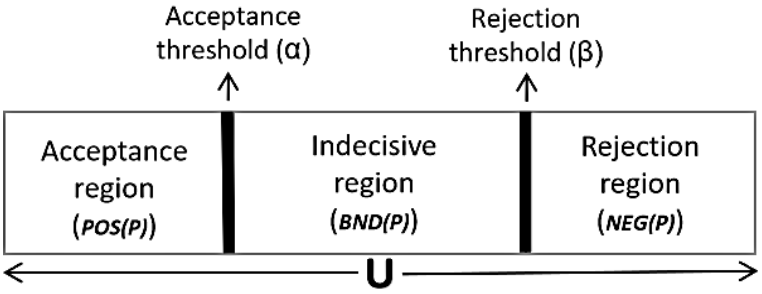

Figure 2.3: Three-way decision in rough sets

on the number of instances of an equivalence class $[a]$ that are in target concept P (i.e., degree of overlap between $[a]$ and P), one of the three decisions is determined for all the instances in the equivalence class $[a]$. We can formulate the degree of overlap (satisfiability level) as the conditional probability of an equivalence class $[a]$ with respect to the target concept P as [76],

$$Pr([a]) = \frac{|[a] \cap P|}{|[a]|}, \ where \ |.| \ denotes \ the \ cardinality \ of \ a \ set \tag{2.5}$$

- Acceptance of instances from an equivalence class $[a]$ in the target concept P is deduced if satisfiability level $Pr([a])$ of $[a]$ is greater than or equal to the acceptance threshold (α). Such instances become a part of the positive region $POS(P)$.

- Rejection of instances of $[a]$ as $POS(P)$ is determined if its $Pr([a])$ is less than or equal to the rejection threshold (β). Such instances become a part of the rejection region $NEG(P)$.

- Neither rejection nor acceptance, but a deferred decision on the instances of $[a]$ is opted if its $Pr([a])$ is greater than the rejection threshold (β) but less than the acceptance threshold (α). Such instances become a part of the boundary (indecisive) region $BND(P)$.

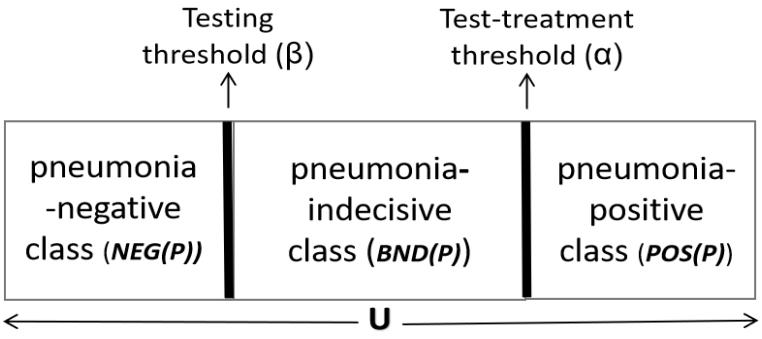

Figure 2.4: Threshold approach for pneumonia classification

Threshold approach for a disease treatment decision was proposed by Paukar and Kassirer [40]. A pair of test-treatment (acceptance) and testing (rejection) thresholds were defined with the constraint of the test-treatment threshold being greater than the testing threshold (Figure 2.4). In this research, we use conditional probability to represent the probability of pneumonia disease and is used to evaluate a particular decision of treatment, no treatment, and further tests. Conceptually, if the probability of pneumonia is lower than the testing threshold (β), it suggests the absence of pneumonia disease and additional diagnosis is not required. It defines the magnitude of the pneumonia-negative class and rejects X-ray images of this class to be pneumonia-positive. If the probability of pneumonia disease is greater than the test-treatment threshold (α), it suggests the presence of pneumonia disease and immediate treatment of pneumonia is concluded. It defines the magnitude of the pneumonia-positive class and accepts X-ray images of this class as pneumonia-positive. If the probability of pneumonia is between these two thresholds, the absence or presence of pneumonia is inconclusive and additional medical tests are needed while the treatment or no treatment decision is pending on the additional test results.

2.2.2 Limitation of Conventional Rough Sets

In Conventional (Pawlak's) rough sets, the three-way decision is formed based on the following notion, 1. equivalence classes that are fully contained in P or satisfy $[a] \subseteq P$ (i.e., $Pr([a]) = 1$) condition, are described in the positive region; 2. equivalence classes that satisfy $[a] \cap P = \phi$ condition such that zero overlap between $[a]$ and P (i.e., $Pr([a]) = 0$) are considered into the negative region; 3. the remaining equivalence classes that do not satisfy any of these conditions form the boundary or indecisive region that leads to account for uncertainty in data. This formation of three-way classification suggests either complete or no overlap between $[a]$ and P for positive and negative regions, respectively. This also suggests that α and β have to be 1 and

0, respectively.

A drawback of threshold pair (1, 0) is that it limits the number of instances being classified into positive and negative regions, and thus limited coverage of the data is obtained for definite decisions. As a consequence, it reduces usability of the system [66]. Usability or applicability of the system depends on the coverage of the data for making certain decisions about them, the less data coverage by the system, the lower usability of the system. However, this is not the effective exploitation of rough sets in uncertainty analysis [57, 76, 78]. A strategy to overcome this limitation would be to allow more instances to be classified into positive and negative regions while shrinking the size of the boundary region.

In rough sets, reducing the boundary region size is critical to increasing usability of the model. Conversely, relieving the stringent requirement of value for thresholds of positive and negative regions can allow the inclusion of more instances into these regions. By considering the degree of overlap and modifying the threshold values such as decreasing α and increasing β can overcome the shortcomings of conventional rough sets. The modified threshold pair can be referred to as the probabilistic threshold pair. The probabilistic rough sets leverages the advantage of the probabilistic threshold pair in developing a high quality pneumonia classification model [57, 64].

2.2.3 Probabilistic Rough Sets

Probabilistic rough sets eases upon the rigid requirement of conventional rough sets by considering the conditional probability of equivalence classes [63, 64]. Conditional probability represents the degree of overlap between a target concept P and equivalence class $[a]$. As a result, the positive and negative regions can be inferred with a certain level of misclassification error tolerance [66]. Conditional probability $Pr(P|[a])$ can be defined as the fraction of instances in $[a]$ that satisfy the target

concept P or the degree of intersection between $[a]$ and P [64],

$$Pr(P|[a]) = \frac{|P \cap [a]|}{|[a]|}, \quad where\ P \subseteq U \quad and\ where\ |.|\ is\ the\ cardinality\ of\ a\ set \tag{2.6}$$

The lower and upper approximations of P can also be given by the conditional probability $Pr(P|[a])$ of equivalence classes $[a]$ and a threshold pair (α, β) as follows [63, 66]:

$$\begin{aligned} \underline{apr}_{(\alpha,\beta)}(P) &= \{a \in U | Pr(P|[a]) \geq \alpha\}, \\ \overline{apr}_{(\alpha,\beta)}(P) &= \{a \in U | Pr(P|[a]) > \beta\} \\ & where\ 0 \leq \beta < \alpha \leq 1 \end{aligned} \tag{2.7}$$

Given the conditional probability $Pr(P|[a])$ of each equivalence class and an optimal threshold pair (α, β), each equivalence class can be classified according to three probabilistic disjoint regions. This procedure is also referred to as the probabilistic three-way decision [64].

$$\pi_{(\alpha,\beta)}(P) = \{POS_{(\alpha,\beta)}(P),\ BND_{(\alpha,\beta)}(P),\ NEG_{(\alpha,\beta)}(P)\} \tag{2.8}$$

The three regions of probabilistic three-way decision can be defined using a threshold pair (α, β) and conditional probability $Pr(P|[a])$ of the equivalence classes as

follows,

$$
\begin{aligned}
POS_{(\alpha,\beta)}(P) &= \underline{apr}_{(\alpha,\beta)}(P) \\
&= \{a \in U | Pr(P|[a]) \geq \alpha\}, \\
NEG_{(\alpha,\beta)}(P) &= (\overline{apr}_{(\alpha,\beta)}(P))^c = U - \overline{apr}_{(\alpha,\beta)}(P) \\
&= \{a \in U | Pr(P|[a]) \leq \beta\}, \\
BND_{(\alpha,\beta)}(P) &= \overline{apr}_{(\alpha,\beta)}(P) - \underline{apr}_{(\alpha,\beta)}(P) \\
&= \{a \in U | \beta < Pr(P|[a]) < \alpha\}, \\
&where\ 0 \leq \beta < \alpha \leq 1
\end{aligned}
\tag{2.9}
$$

We scrutinize the conditional probability as the level of certainty in probabilistic rough sets model. According to Equation (2.9), we accept that instances from an equivalence class $[a]$ have same description and are the part of P if its certainty level $Pr(P|[a])$ is greater than or equal to the acceptance level α (i.e., $Pr(P|[a]) \geq \alpha$). We reject the instances in $[a]$ from being the part of P if its $Pr(P|[a])$ is less than or equal to the rejection level β (i.e., $Pr(P|[a]) \leq \beta$). We defer to have decision on the instances of $[a]$ if its $Pr(P|[a])$ lies between β and α (i.e., $\beta < Pr(P|[a]) < \alpha$).

In general, the limit [0,1] is divided into three regions, namely, positive [α,1], negative [0,β], and boundary (indecisive) (α, β). Positive and negative regions provide definite responses to instances, either "yes" or "no". Alternatively, indecisive region does not provide a definitive decision regarding its instances due to inconsistency and uncertainty in the information. It is delayed to allow for additional exploration and collection of certain information in such cases.

Pawlak's rough sets can also be reformulated with respect to probabilistic rough sets by using the threshold pair, $(\alpha, \beta) = (1, 0)$. Condition $[a] \subseteq P$ can be equivalently considered as $Pr(P|[a]) = 1$ and condition $[a] \cap P \neq \phi$ as equivalent to $Pr(P|[a]) > 0$. Hence, the conventional three-way decision can be given by,

$$
\begin{aligned}
POS_{(1,\ 0)}(P) &= \{a \in U | Pr(P|[a]) \geq 1\}, \\
NEG_{(1,\ 0)}(P) &= \{a \in U | Pr(P|[a]) \leq 0\}, \\
BND_{(1,\ 0)}(P) &= \{a \in U | 0 < Pr(P|[a]) < 1\}
\end{aligned}
\tag{2.10}
$$

Specifically, Pawlak's rough sets is a special case of probabilistic rough sets, in which we must strictly consider $Pr(P|[a])$ equal to either 0 or 1 while making a definite decision of negative or positive. Likewise, another special case of the probabilistic model is 0.5-probabilistic rough sets that uses the threshold pair, $(\alpha, \beta) = (0.5, 0.5)$ such that [77],

$$
\begin{aligned}
POS_{(0.5,\ 0.5)}(P) &= \{a \in U | Pr(P|[a]) > 0.5\}, \\
NEG_{(0.5,\ 0.5)}(P) &= \{a \in U | Pr(P|[a]) < 0.5\}, \\
BND_{(0.5,\ 0.5)}(P) &= \{a \in U | Pr(P|[a]) = 0.5\}
\end{aligned}
\tag{2.11}
$$

The most uncertain equivalence classes (i.e., $Pr(P|[a]) = 0.5$) are moved to the boundary region in 0.5-probabilistic rough sets. It almost always classifies all instances for definite decisions, making it a binary classification model.

The optimal threshold pair can control the three regions effectively. Different threshold pairs, however, provide different interpretations of the three-way decision. It is therefore crucial to select a suitable threshold pair when using probabilistic rough sets in a three-way classification model.

2.2.4 Evaluation Criteria of Three-way Model

Provided a probabilistic threshold pair (α, β) and the target concept P, we can obtain the three-way classification in probabilistic rough sets according to Equation (2.9). The most commonly used evaluation measures to evaluate the classification efficiency of the model are accuracy and coverage [4].

The degree of accuracy can be assessed by the proportion of instances classified

correctly into positive and negative regions of a three-way classification to the total number of instances classified into these two regions. This captures the degree of accurate classification rate of the prediction model. The degree of correctness lies between 0% and 100%. The accuracy efficiency of $\pi_{(\alpha,\beta)}(P)$ can be given as follows,

$$Acc_{(\alpha,\beta)}(P) = \frac{|P \cap POS_{(\alpha,\beta)}(P)| + |P^c \cap NEG_{(\alpha,\beta)}(P)|}{|POS_{(\alpha,\beta)}(P)| + |NEG_{(\alpha,\beta)}(P)|} \times 100, \\ where\ |.|\ denotes\ the\ cardinality\ of\ a\ set \tag{2.12}$$

In Equation (2.12), term $|P \cap POS_{(\alpha,\beta)}(P)|$ is the number of instances from positive region that actually belongs to the target concept P (the set of pneumonia-positive X-ray images), and term $|P^c \cap NEG_{(\alpha,\beta)}(P)|$, where P^c is the complement of P (the set of pneumonia-negative X-ray images), is the number of instances from the negative region which are actually not in the target concept P. Similarly, terms $|POS_{(\alpha,\beta)}(P)|$ and $|NEG_{(\alpha,\beta)}(P)|$ are the number of instances classified into positive and negative regions, respectively.

Similarly, coverage can be defined by the proportion of instances that have been covered in positive and negative regions of the three-way classification model to the total number of instances in the universal set (U). It reflects the coverage of data for definite decisions and the usability of the three-way classification model. The degree of usability lies between 0% and 100%. The coverage efficiency of classification $\pi_{(\alpha,\beta)}(P)$ can be given as follows,

$$Cov_{(\alpha,\beta)}(P) = \vartheta_{cov} = \frac{|POS_{(\alpha,\beta)}(P)| + |NEG_{(\alpha,\beta)}(P)|}{|U|} \times 100 \tag{2.13}$$

2.2.5 Impact of Threshold Pair on Accuracy and Coverage

Different pairs of threshold values for α and β can result in a different level of classification efficiency of the three-way model. A threshold pair to allow for more instances

to be classified (greater coverage) may negatively impact the accuracy of the classification model as it increases the chances of misclassifying more uncertain instances for definite decisions. Likewise, a threshold pair for accurately identifying instances may result in fewer instances being classified for definite decisions. These threshold pairs lead to a lower usability of the model. Therefore, an arbitrary choice of threshold pair may not be the best approach.

A key issue in three-way classification based on probabilistic rough sets is the selection of an optimal threshold pair to balance accuracy and coverage. As part of the Bayesian rough sets, the certainty gain measure has been used to evaluate the performance of the model [49]. In Parameterized rough set, the degree of instance's membership in an equivalence class has been used as a basis of providing the set approximations [13].

In this research, game-theoretic rough sets or GTRS is used to obtain an optimal threshold pair. The method evaluates multiple evaluation measures simultaneously while determining a suitable threshold pair for three-way classification [30]. It utilizes game theory in rough sets to create a competitive game between a number of players. GTRS helps to reach a suitable threshold pair gradually while continuously learning from the impact of modified threshold pairs on accuracy and coverage.

Using GTRS, the model can learn classification rules from a data perspective. This research considers accuracy and coverage as players in the game and their payoff functions. In detail, since players and their rewards in GTRS are the evaluation criteria themselves that we use to evaluate classification performance of the model, we learn which strategy is most beneficial for both, accuracy and coverage. In addition, reformulation of the game allows the model to learn continuously from the data at every iteration by continuously evaluating the model [30]. Hence, to improve the classification efficiency of the three-way pneumonia classification model, this thesis experiments with game-theoretic rough sets.

2.3 Game-theoretic Rough Sets

The game-theoretic rough sets or GTRS uses game theory in probabilistic rough sets to determine the value of the threshold pair (α, β). To determine the threshold pair using GTRS, three steps are required, namely, the formulation of competitive games between players, the repetitive learning mechanisms within the game, and the determination of threshold pairs based on trade-offs. First, GTRS-related artifacts such as game players, their payoffs, and strategies are defined. In the second step, the game formulation is repeated using an initial pair based on the previously selected threshold pair. The third step is to decide whether to stop reformulating the game and choose the final threshold pair.

The optimal threshold pair that accurately categorizes the maximum number of instances into positive and negative regions is determined by the game equilibrium or the player's trade-offs. Game theory helps to analyze the decision-making problem by formulating a game in the form of mathematical formulae [30].

2.3.1 Game Theory

In decision science, game theory is a mathematical technique for analyzing the decision-making process in an interactive environment [30]. A number of fields have applied its problem-solving technique, including cryptography [7], economics [16], and political science [10], etc. Its application can be described as a game played by two or more people. Players in the game are considered to be rational such that their actions are more predictable than those of irrational players. As players can determine another's strategy, they may not intentionally wish to make their payoffs worse than they were before [30].

During the game, the players execute their possible actions or make decisions that maximize their rewards. It may also affect other players' rewards at the same time.

As soon as all players have received their maximum reward and no player wishes to continue to update their strategy, we say that the game has reached an equilibrium state. By determining a trade-off between players, the game theory ensures that players deviating from the plan will not profit or may negatively impact their fellow players. Computer science researchers have utilized game theory to formulate the decision-making problem of distributed computing [2], machine learning [18, 48], network [6, 27], computational complexity theory [58], and rough sets [19, 60].

Game theory can be modeled in two types, 1. conflicting (non-cooperative) games, and 2. cooperative games (coalition game theory) [3, 5]. Depending on the formulation of the game, these two types can be distinguished. Conflicting or non-cooperative games require a single player with its own preferences without taking into account other players' actions. As a contrast, in cooperative games, the primary modeling unit is the group of players, and it considers the payoff of all players at the same time [3].

2.3.2 Game Representation

A game can be mathematically described by a tuple, $G = \{P, S, \vartheta\}$ [60]. It contains the following three entities:

- P is a set of n players, $P = \{p_1, p_2, ..., p_n\}$. Players in GTRS can be the evaluation measures that are used to evaluate classification performance of the model, such as accuracy and coverage [4, 20, 75].

- S is a set of strategies available to players, $S = S_1 \times S_2 \times ... \times S_n$, where S_i is the strategy profile of player p_i such as $S_i = \{a_1, a_2, \ldots, a_n\}$. Strategies in GTRS can be used to define different sets of values for acceptance α and rejection β thresholds. The goal of the game is to find an optimal threshold pair.

- ϑ is the reward or gain obtained by performing a possible action. It can be

termed as payoff function or real value utility for players, $\vartheta = \{\vartheta_1, \vartheta_2, ..., \vartheta_n\}$, where u_i is the payoff function for player p_i. Payoff functions in GTRS can be the same evaluation measures that we consider for players. By considering evaluation measures as the player's reward function, data itself can be used to determine the model's decision or classification rules as discussed in Section 1.1.

Each player performs an action from S as per its profile and obtains an expected payoff from ϑ. Some a_i from S can maximize the payoff of one player $\vartheta_i(a_i)$ while minimizing the payoff of remaining players.

In game theory, all information about a game can be represented in a table, which is referred to as a payoff table. This table provides a visual representation of each player's possible actions and their corresponding payoffs or results. Each cell in the table represents the payoff of all players according to their action profiles. For instance, Table 2.2 shows the payoff matrix of a two-player game. The two players are $P = \{p_1,\ p_2\}$. The strategy profiles of both players are $PS_1 = \{a_1,\ a_2,\ \dots\}$ and $PS_2 = \{c_1,\ c_2,\ \dots\}$, respectively. Columns indicate a possible strategy of player p_2 and rows indicate a possible strategy of player p_1.

Rewards for players p_1 and p_2 can be given by ϑ_1 and ϑ_2, respectively. For a strategy profile (a_1, c_1), where player p_1 performing action a_1 and player p_2 executing strategy c_1, the payoff of players $p1$ and p_2 are $\vartheta_1(a_1, c_1)$ and $\vartheta_2(a_1, c_1)$, respectively. Hence, the first cell in payoff table (Table 2.2) contains $(\vartheta_1(a_1, c_1),\ \vartheta_2(a_1, c_1))$, where $a_1 \in PS_1$ and $c_1 \in PS_2$.

Table 2.2: An example of payoff table

		p_2		
		c_1	c_2	...
p_1	a_1	$(\vartheta_1(a_1, c_1),\ \vartheta_2(a_1, c_1))$	$(\vartheta_1(a_1, c_2),\ \vartheta_2(a_1, c_2))$	...
	a_2	$(\vartheta_1(a_2, c_1),\ \vartheta_2(a_2, c_1))$	$(\vartheta_1(a_2, c_2),\ \vartheta_2(a_2, c_2))$	...
	...	...	...	...

2.3.3 Game Solution

Players in an interactive game environment expect to increase their payoff. A player's payoff depends on its own and every other player's actions. Therefore, the payoff of a player is affected by the strategy of other players. The best action profile is decided based upon the payoffs of all involved players. In game theory, there are some approaches to finding a rational choice. Examples include dominant strategy, best response, Nash equilibrium, and Pareto optimality. We apply Pareto optimality to find a solution to a game in this research.

Dominant Strategy: Let $\{a_i, a_i^{'}\} \in A_i$ be the action profile of player p_i and A_{-i} be the set of all possible action profiles of other $n-1$ players, a_i strictly dominates player p_i's strategy $a_i^{'}$ if and only if [11, 30],

$$\forall a_{-i} \in A_{-i}, \vartheta_i(a_i, a_{-i}) > \vartheta_i(a_i^{'}, a_{-i}),\ where\ (a_i^{'} \neq a_i) \tag{2.14}$$

It means that strategy a_i provides a better payoff than strategy $a_i^{'}$ irrespective of the action of other players. This is called the dominant strategy.

Weakly dominant strategy: Let $\{a_i, a_i^{'}\} \in A_i$ be the action profile of player p_i and A_{-i} be the set of all possible action profiles of other $n-1$ players, a_i weakly dominates player p_i's strategy $a_i^{'}$ if and only if [11, 30],

$$\begin{aligned} \forall a_{-i} \in A_{-i}, \vartheta_i(a_i, a_{-i}) \geq \vartheta_i(a^{'}{}_i, a_{-i}) \\ \exists a_{-i} \in A_{-i},\ such\ that\ \vartheta_i(a_i, a_{-i}) > \vartheta_i(a^{'}{}_i, a_{-i}) \end{aligned} \tag{2.15}$$

It states that strategy a_i highlights at least as good payoff as any other strategy irrespective of the action of other players, also strictly dominating some actions of other players at the same time. This is called weakly dominant strategy.

Nash Equilibrium: In most applications of game theory [5, 76], Nash equilibrium strategy has been considered as the best solution. Considering a situation, where all

opponents of player p_i are to commit a_{-i}, the best response of player p_i to the action profile $a_{-i} \in A_{-i}$ can be given by $a_i \in A_i$ if [30, 75],

$$\forall a_i^{'} \in A_i,\ we\ have\ \vartheta_i(a_i, a_{-i}) \geq \vartheta_i(a_i^{'}, a_{-i}),\ and\ (a_i^{'} \neq a_i) \tag{2.16}$$

Given a player's opponent's strategy equilibrium, the best reward for the player is the maximum of its reward. An action profile $A = (a_1, a_2, \ldots, a_n)$ is a Nash equilibrium, if for all players $p_i \in P$, a_i is the best response to $a{-}i$. This strategy can be mathematically formulated as [74],

$$\forall p_i \in P,\ \forall a^{'}{}_i \in A, \vartheta_i(a_i, a_{-i}) \geq \vartheta_i(a^{'}{}_i, a_{-i}),\ where\ (a_i^{'} \neq a_i) \tag{2.17}$$

If opponents are to play strategy a_{-i}, player p_i has to determine its best response when others are playing $a{-}i$. It depicts that a player's payoff does not get any better by updating its strategy unilaterally unless the player knows what actions other players are taking. It also suggests that no player has an intention to deviate from its strategy. This strategy is called Nash equilibrium and deviation from the equilibrium is a loss for any player [30].

Pareto dominant strategy: Action profile $a \in A$ Pareto dominates action profile $a^{'} \in A$ if for all players $p_i \in P, \vartheta_i(a) \geq \vartheta_i(a^{'})$, and there also exist some $p_k \in P\ for\ which\ \vartheta_k(a) > \vartheta_k(a^{'})$. It means that [30, 75],

$$\forall p_i \in P,\ \vartheta_i(a) \geq \vartheta_i(a^{'})),\ and\ \exists p_k \in P,\ \vartheta_l(a) > \vartheta_k(a^{'}) \tag{2.18}$$

Here the strategy profile $a \in A$ is the Pareto dominant strategy profile. According to Pareto dominant strategy profile, a player achieves a better payoff without sacrificing other players' payoffs.

Pareto optimality strategy: An action profile $a \in A$ is Pareto optimal if there does

not exist an action profile that Pareto dominants action a [11, 30, 75]. There may be more than one Pareto optimal solution in some games. With Pareto dominant strategy, no player gets a worse payoff or at least one player gets a better payoff than any other strategy profile.

2.3.4 Repetitive Learning Mechanism

In cooperative games, the game solution can be determined by the Pareto optimality strategy, and in non-cooperative games, by the Nash equilibrium strategy. The outcome corresponding to a particular game solution strategy for the current formulated game is the best outcome within the current strategy profile of the player [60]. In the case where the current selection does not provide a satisfactory classification result, we rebuild the game with the currently selected solution to see if any solution near it provides a better classification.

For the game to stop reformulating at a certain point, we must also define stop criteria. In GTRS, we define stop criteria as players obtain their payoffs greater than some specified values. Another stop criteria can be the constraint on the threshold pair, such as $0 \leq \beta < \alpha \leq 1$. Once the stop criteria are accomplished, we stop reformulating the game and select the final threshold pair from that iteration of the game.

2.4 Summary

As a result of this chapter, we are made aware of the related literature on pneumonia classification. It briefly discusses the existing experiment and highlights the related issues. It explains how three-way classification can be used to overcome the existing research shortcoming. Moreover, it provides a detailed overview of the use of game theory rough sets to achieve a tradeoff between evaluation criteria.

Chapter 3

Pneumonia Identification with Game-theoretic Rough Sets

The chapter introduces the proposed model (i.e., pneumonia classification using game-theoretic rough sets) with three-way classification in probabilistic rough sets. It explains the approach used in this study to deliver a high-quality pneumonia classification system. In Section 3.3, it illustrates the impact of different pairs of test-treatment and testing (α, β) thresholds on classification efficiency of various rough sets models described in this thesis.

3.1 Introduction

A lot of attention has been paid to rough sets in the field of data analysis to deal with uncertainty of the data in classification tasks. It has been applied in the recommender system [4], web-based decision system [74], spam filtering system, and uncertainty analysis [5]. The purpose of this thesis is to develop a pneumonia classification model that accounts for uncertainty in chest X-ray images by using game theory in rough sets (i.e., game-theoretic rough sets). A pneumonia classification system can be broken down into three classes, provided that the third class, pneumonia-indecisive, accounts

for uncertainty in chest X-ray images.

In order to mitigate the stringent requirements of conventional (Pawlak's) rough sets, the thesis employs probabilistic rough sets in the experiment to implement a three-way pneumonia classification model. Positive, negative, and boundary regions of three-way decisions correspond to pneumonia-positive, pneumonia-negative, and pneumonia-indecisive classes, respectively.

Given a probabilistic threshold pair (α, β), the three decisional regions can be derived using the test-treatment (α) and testing (β) thresholds for positive and negative regions, respectively [72] (Figure 2.4). According to medical terminology, in this study, we use testing (β) and test-treatment (α) threshold terms instead of rejection and acceptance thresholds [40]. The test-treatment threshold represents the probability of pneumonia above which the system directs treatment of the patient without further testing whilst the testing threshold represents the probability of pneumonia below which further testing is not likely to be necessary [40]. If the probability of pneumonia is between test-treatment and testing thresholds, it is advisable to conduct more tests before making a definite decision.

A higher number of instances in positive and negative regions leads to making definitive decisions on some or more possible instances and increases the coverage of the data by the system. However, adding more uncertain instances into pneumonia-positive and pneumonia-negative decisions may decrease the model's classification accuracy, and vice versa. The probabilistic threshold pair (α, β) can be adjusted to control the level of uncertainty in positive and negative regions.

A level of uncertainty that does not significantly compromise classification accuracy and number of instances into definite regions at the same time requires careful changing of the test-treatment and testing thresholds (α, β). Currently, our goal is to adjust threshold pair value until we reach a pair that provides a better outcome for accuracy and coverage. It suggests that we have two competing measures, and

we need to find a way to balance between them in order to implement an effective pneumonia classification system.

For establishing the best possible compromise between accuracy and coverage, game-theoretic rough sets (GTRS) seems to be the best option. Specifically, it provides a way to formulate and resolve the problem of competing decisions between two or more players [30]. This thesis uses GTRS to formulate a competitive environment between accuracy and coverage criteria, and then determine an optimal trade-off between them. The result is a suitable threshold pair for a three-way classification system.

3.2 The Proposed Pneumonia Classification Model

The architecture of the three-way chest X-ray classification model for pneumonia is described in Subsection 3.2.1. It also discusses an algorithm and shows step-wise development of the proposed model in Subsection 3.2.2.

3.2.1 Architecture of Three-way Pneumonia Classification

The classification algorithm in the existing approach can be replaced with a three-way classification method. The proposed model classifies X-ray images into three classes as represented in Figure 3.1. Instead of two classes, three classes are derived: 1. pneumonia-positive, 2. pneumonia-negative, and 3. pneumonia-indecisive.

X-ray images from the pneumonia-positive class represent positive pneumonia

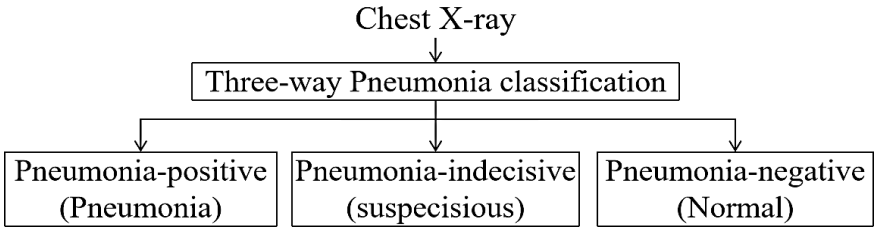

Figure 3.1: Three-way pneumonia classification with chest X-ray

cases, whereas images from the pneumonia-negative class represent negative pneumonia cases. We yield the decision on the X-ray image of pneumonia-indecisive class. Such images represent incomplete or uncertain information and can be suspected to be a positive or negative case of pneumonia, but can not be processed as either of them until certain information is gathered. A benefit of the three-class approach is that the system does not have to make a definite decision, it is aligned with the manual interpretation of suspected pneumonia cases, and it reduces misclassification of such X-ray images. By doing so, we will be able to deliver an efficient pneumonia classification model.

3.2.2 Pneumonia Classification with Game-theoretic Rough Sets

A two-phased approach is used to implement the proposed pneumonia classification system (Figure 3.2).

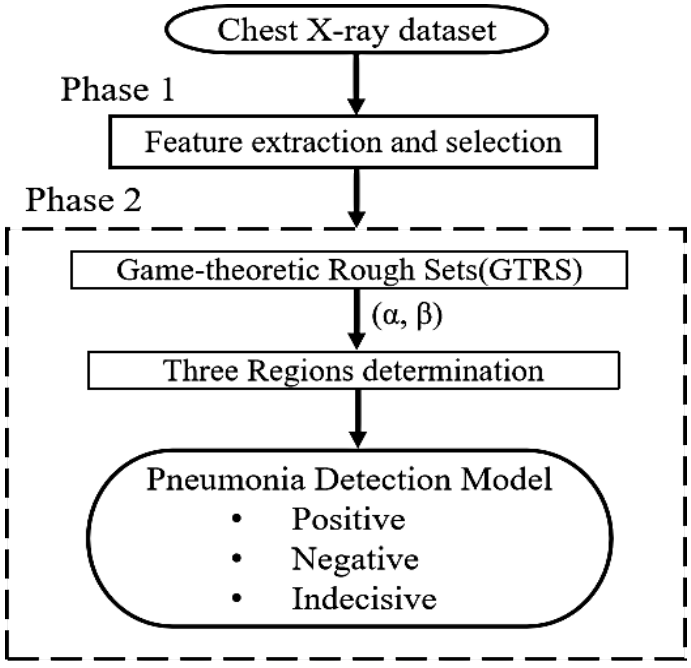

Figure 3.2: Classification approach for pneumonia using GTRS (oval: input and output. rectangle: process)

Phase 1: Feature Extraction and Selection from Chest X-ray Images

In the first phase, image preprocessing techniques are applied to enhance the chest X-ray image, obtain features from the enhanced image, and select effective features from the large pool of extracted features while disregarding irrelevant ones. By converting the two-dimensional images into a feature vector, we can illustrate the decision table (DT) (defined by Equation (2.1)) in rough sets theory.

Image Augmentation: The chest X-ray image dataset [26] used in this study contains a total of 5,878 chest X-ray images with pneumonia-positive instances more than twice the number of pneumonia-negative images. As a result, image augmentation is applied to create a balanced dataset. The Python library ($cv2$ (OpenCV)) is imported in order to create replica of pneumonia-negative chest X-rays by changing the contrast and brightness levels. The contrast and brightness of an image are affected by the gain α and bias β parameters [37]. The brightness of the copied images is altered by adding bias factors 10, 20, and 25, and contrast is adjusted by multiplying the pixel with gain factors 1.1, 1.2, and 1.3. This method is based on the observation that medical experts examine chest X-ray images under bright light, and contrast increases the visibility of edges and specific regions [47]. Additionally, various versions of pneumonia-positive X-ray images are generated by randomly selecting a set of images to approximate the count of both types.

Image Preprocessing: The selected dataset contains chest X-ray images with different sizes, such as 1744×1224, 1576×1288, 1264×1016, etc. In this experiment, the image size is rescaled to 150×150. We then apply a non-linear bilateral filter to reduce noise from the image without affecting its edges. This dataset contains chest X-ray images of differing illumination scales. Therefore, we also apply histogram equalization to improve and enhance contrast of chest X-ray images by extending their pixel intensity range [47, 61].

Feature Extraction: We extract a set of statistical and texture features from

chest X-ray images by computing a gray-level co-occurrence matrix with Python libraries, *cv*2 *and skimage*. The resulting features include contrast, dissimilarity, homogeneity, energy, correlation [39]. Furthermore, we extracted texture information such as Local Binary Pattern (LBP) [25]. In this method, images are decomposed into regions of interest that can be exploited for their local appearance properties [17, 32]. Keypoint descriptors such as oriented fast and rotated binary robust independent elementary features (*Orb*) [39], binary robust invariant scalable keypoints (*Brisk*) [29], and accelerated-kaze (*Akaze*) [31] are the local detector of the image. keypoint descriptor features simply describe the local regions of an image by their characteristics and play a key role in applications, including image matching and identification [17]. The value of keypoint descriptors is also extracted as part of this experiment.

Feature Selection: To reduce complexity of the classification algorithm, this step is necessary to remove irrelevant and redundant attributes. Using chi-square measure from Python library *sklearn* [42], the relevance of each extracted feature is

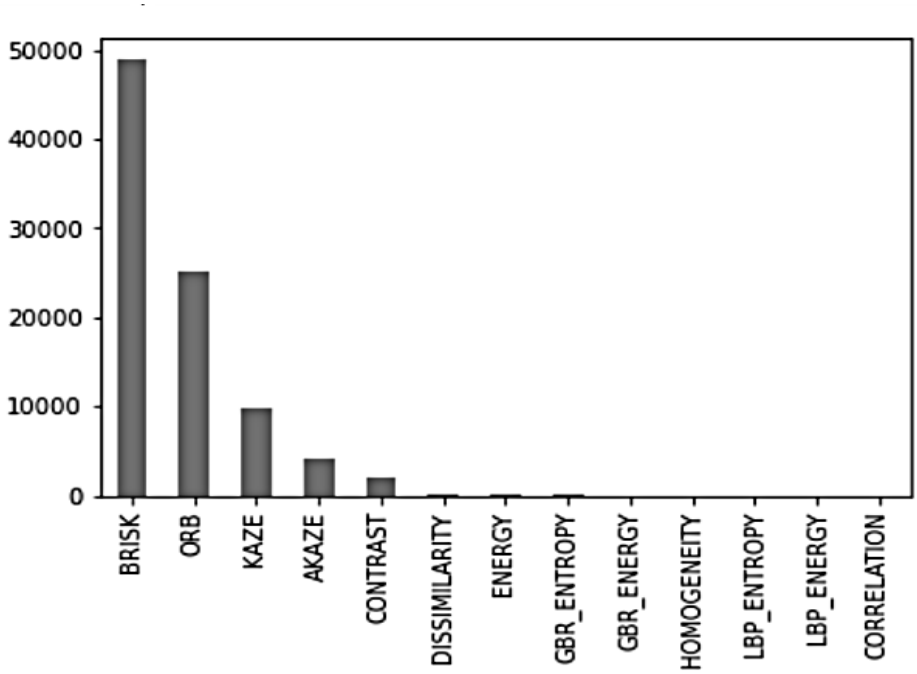

Figure 3.3: Relevancy between input features and output feature

calculated with respect to the output labels 1, 0, where "1" represents pneumonia-positive, and "0" represents pneumonia-negative. In Figure 3.3, the relevancy between the extracted features and output attribute is shown. For rough sets analysis, the most relevant attributes are selected, such as Orb, Kaze, and Brisk.

Phase 2: Application of GTRS

In the second phase, the selected set of features is used to learn to distinguish pneumonia-positive and pneumonia-negative X-ray images. Algorithm 1 introduces the related steps. Every X-ray image represents a single instance of the decision table DT. A set of conditional attributes is constructed from the features of the image. The decisional attribute is the label for each X-ray image, such as pneumonia-positive

Algorithm 1 : Phase 2 (employing probabilistic rough sets with GTRS)

Input: Feature vector from images
Output: Three classes of pneumonia classification

$DT \Leftarrow formDT(features)$
$E \Leftarrow formEquivalneceClass(DT)$
for e_i *in* E : **do**
  $pr[i] \Leftarrow calConditionalProbability(e_i)$
end for
$p \Leftarrow \{acc, cov\}$ $\quad \#define\ players$
$s_{acc} \Leftarrow \{update\ \alpha\}$ $\quad \#define\ strategy\ for\ acc$
$s_{cov} \Leftarrow \{update\ \beta\}$ $\quad \#define\ strategy\ for\ cov$
$u \Leftarrow \{acc(\alpha, \beta),\ cov(\alpha, \beta)\}$ $\quad \#define\ payoff$
while !$stopCriteria$ **do**
  $resetThresholdPair(\alpha, \beta)$
  $reformulateGameWith(\alpha,\ \beta)$
  for a *in* s_{acc} **do**
    for c *in* s_{cov} **do**
      $pt \Leftarrow calculatePayoff(a,\ c)$
    end for
  end for
  $(\alpha, \beta) \Leftarrow findParetoOptimal(pt)$
  $stopCriteria \Leftarrow isStopCriteriaReached()$
end while
$\pi_{(\alpha,\beta)}(Pneumonia) \Leftarrow formThreeClasses(pr, E, \alpha, \beta)$

or pneumonia-negative. The X-ray images with the same value on all conditional attributes are grouped into an equivalence class (Line 2 in Algorithm 1).

For each equivalence class (E_i), we calculate the conditional probability ($Pr(Pneumonia|E_i)$) according to the target concept (Line 4 in Algorithm 1). In this research, we define target concept ($Pneumonia$) as *the proportion of pneumonia-positive instances in equivalence class such that the decisional attribute, "Label = 1"*; therefore, $Pr(Pneumonia|E_i)$ of an E_i can be measured as the proportion of chest X-ray images in E_i labeled as pneumonia-positive to the total number of chest X-ray images in E_i.

$$\begin{aligned} Pr(Pneumonia|E_i) &= \frac{number\ of\ pneumonia\_positive\ image\ instances}{total\ number\ of\ image\ instances\ in\ E_i} \\ &= \frac{|Pneumonia \cap [E_i]|}{|E_i|} \end{aligned} \tag{3.1}$$

where $|.|$ is the cardinality of a set. As shown in Equation (3.1), $Pneumonia \cap [E_i]$ is a rule (p) such as p: $label = 1$. The meaning set, $m(p)$ consists of all the X-ray images that satisfy rule p. Hence, the meaning set $m(p)$ can be given as [64]:

$$m(p) = \{x \in E_i \mid x \models p\} \tag{3.2}$$

After obtaining the conditional probability for each equivalence class, we use GTRS to determine the optimal test-treatment and testing threshold pair (from Line 6-19 in Algorithm 1). Employing GTRS for a three-way classification model includes defining a set of players, strategy profile, and reward function for each player. The reward function displays the gain achieved by performing an action from the strategy profile. Based on the achieved threshold pair, all equivalence classes are grouped into three sets based on their conditional probability. These sets are analogous to the three classes that contain instances from corresponding equivalence classes. Given a test-treatment and testing threshold pair (α, β) and referring to Equation (2.8), the

three classes for pneumonia classification system can be given by:

$$\begin{aligned}\pi_{(\alpha,\beta)}(Pneumonia) =&\{POS_{(\alpha,\beta)}(Pneumonia),\ BND_{(\alpha,\beta)}(Pneumonia),\\ &NEG_{(\alpha,\beta)}(Pneumonia)\}\end{aligned} \tag{3.3}$$

In probabilistic rough sets, the training set of chest X-ray images is partitioned into equivalence classes. Each equivalence class $[a]$ contains X-ray images that have the same conditional attribute values.

Given an equivalence class $[a]$, if most or all of the X-ray images in $[a]$ are pneumonia-positive such that $Pr(Pneumonia|[a]) \geq \alpha$, system keeps all images of $[a]$ into pneumonia-positive class and treatment is conducted without further testing; if most or all of the X-ray images in $[a]$ are pneumonia-negative such that $Pr(Pneumonia|[a]) \leq \beta$, system places all images of $[a]$ into pneumonia-negative class and X-ray images are considered not having pneumonia infection; otherwise, system classifies all images of $[a]$ into a suspected folder such as pneumonia-indecisive class if $\beta < Pr(Pneumonia|[a]) < \alpha$. It cannot determine pneumonia-negative or pneumonia-positive based on the available information. Additional diagnostic tests are therefore necessary. The three classes of pneumonia classification system using probabilistic test-treatment and testing threshold pair (α, β) can be given as follows,

$$\begin{aligned}Pneumonia - positive &= POS_{\alpha,\ \beta)}(Pneumonia)\\ &= \{a \in U | Pr(Pneumonia|[a]) \geq \alpha\},\\ Pneumonia - negative &= NEG_{\alpha,\ \beta)}(Pneumonia)\\ &= \{a \in U | Pr(Pneumonia|[a]) \leq \beta\},\\ Pneumonia - indecisive &= BND_{(\alpha,\ \beta)}(Pneumonia)\\ &= \{a \in U | \beta < Pr(Pneumonia|[a]) < \alpha\}\end{aligned} \tag{3.4}$$

3.3 Evaluating Three-way Pneumonia Classification in Rough Sets

A pair of test-treatment and testing thresholds (α, β) determines and affects the decision on X-ray images in probabilistic rough sets. With different threshold values, the classification technique can produce different decisions for the same X-ray image. An X-ray image detected as pneumonia-positive within a certain pair of threshold values may also be identified as pneumonia-negative or deferred to a decision for a different pair of thresholds. Different threshold pairs can also affect the accuracy and coverage evaluation criteria of the model.

The accuracy criterion of the model captures the degree of correct classification performed by the model. The coverage criterion varies depending on the context [65]. In this thesis, coverage refers to the proportion of X-ray images classified for concrete decisions (i.e., positive and negative regions). It represents the usefulness of the model. Both criteria reflect different aspects of evaluating the quality and performance of rough sets models.

Figure 3.4 (according to Pawlak's rough sets) and Figure 3.5 (according to prob-

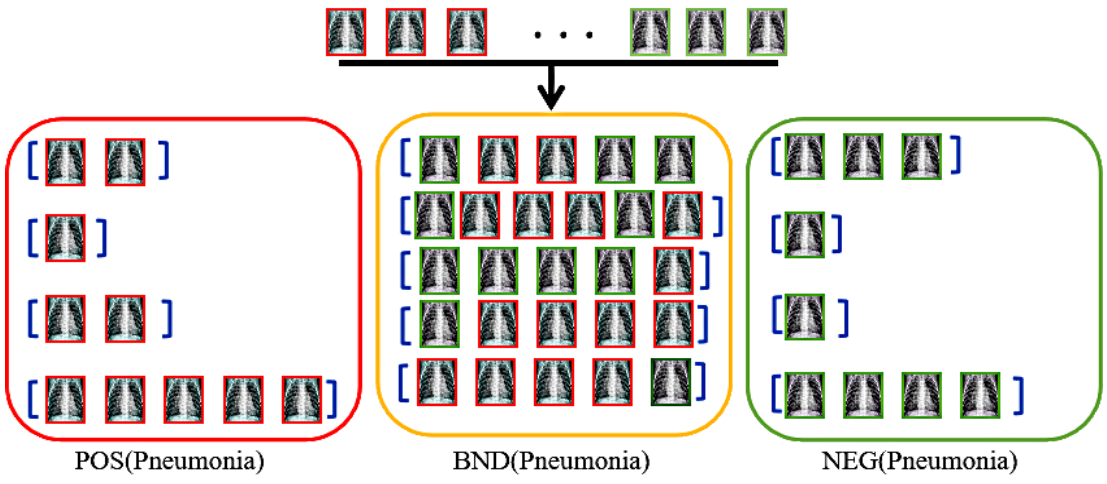

Figure 3.4: Evaluation criteria in Pawlak's rough sets

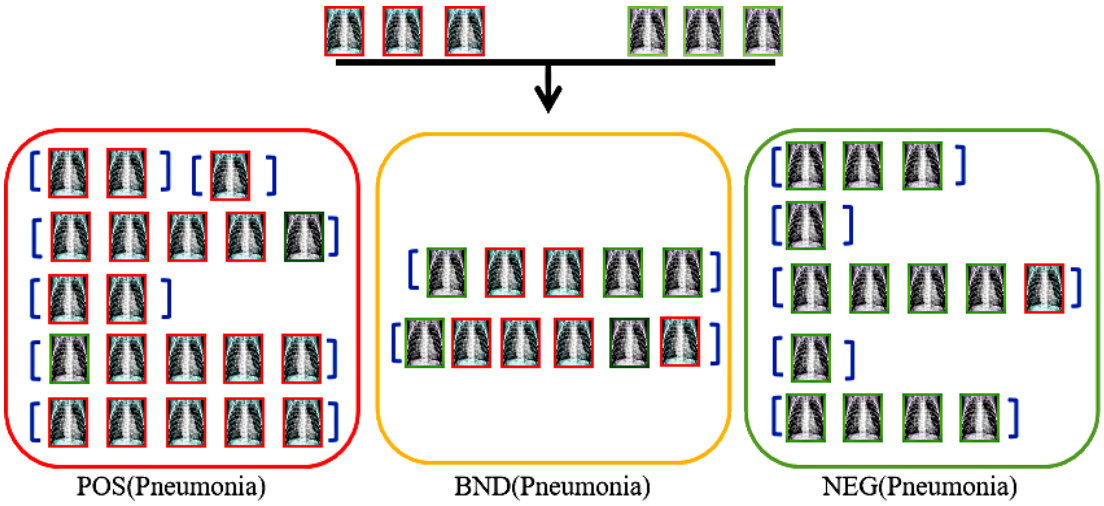

Figure 3.5: Evaluation criteria in probabilistic rough sets

abilistic rough sets) illustrate the interrelationship between the evaluation criteria (accuracy and coverage) and threshold pair. A red bordered X-ray image indicates pneumonia is present, whereas a green bordered X-ray image indicates pneumonia is not present. The rectangular brackets (blue color) around the X-ray images indicate the equivalence class. Figures 3.4 and 3.5 show the concept of approximation.

The rough sets of Pawlak are concluded based on the full inclusion and exclusion of the equivalence class as shown in Figure 3.4. Positive and negative regions in probabilistic rough sets are defined based on the number of objects in an equivalence class that satisfy the target concept (i.e., according to the degree of inclusion), as shown in Figure 3.5. A target concept is the proportion of chest X-ray images that show pneumonia in an equivalence class. In comparison to Pawlak's rough sets, the probabilistic rough sets cover more instances of X-ray images into positive and negative regions as shown in Figures 3.4 and 3.5.

The threshold pair (α, β) determines the region of an equivalence class in probabilistic rough sets. Accuracy and coverage criteria can also be observed from figures 3.4 and 3.5. The accuracy criterion represents the proportion of X-ray images correctly classified as pneumonia-positive and pneumonia-negative for treat-

ment and no treatment decisions, respectively. The coverage criterion corresponds to the proportion of X-ray images that have been considered for treatment and no treatment decisions (i.e., the total number of X-ray images in $(POS(Pneumonia) \cup NEG(Pneumonia))$ to the total number of X-ray images in $(POS(Pneumonia) \cup BND(Pneumonia) \cup NEG(Pneumonia))$, such as the universal set U).

Here is a decision table corresponding to the X-ray images in Figures 3.4 and 3.5. As shown in Table 3.1, X-ray images in the decision table are described by a set of conditional and decisional attributes. The set of attributes $\{A_1, A_2, A_3, A_4\}$ corresponds to the conditional attributes that have domain {1, 2, 3, 4}, where "4" depicts the highest value of any feature in the image and "1" represents its lowest value. "Label" is the decisional attribute that shows the actual class of an X-ray image. Its domain is {0, 1}, where"0" represents pneumonia-negative, and "1" represents pneumonia-positive. Now, let's examine the calculation of accuracy and coverage criteria.

Let's consider, E_i is the equivalence class which is basically a set of X-ray images that have the same value for all conditional attributes (i.e., except for "*Label*"). The decisional attribute can have an actual label from {0, 1} for each instance in the equivalence class. Following equivalence classes result from Table 3.1,

- $E_1 = \{I_1, I_7, I_{13}\}$
- $E_2 = \{I_{31}, I_{35}\}$
- $E_3 = \{I_6, I_{12}, I_{18}, I_{45}\}$
- $E_4 = \{I_{24}, I_{30}, I_{36}, I_{42}, I_{43}\}$
- $E_5 = \{I_3, I_9, I_{15}, I_{21}, I_{27}\}$
- $E_6 = \{I_2, I_8\}$
- $E_7 = \{I_{26}\}$
- $E_8 = \{I_4, I_{10}, I_{16}, I_{22}, I_{28}\}$
- $E_9 = \{I_{34}, I_{40}, I_5, I_{11}, I_{17}\}$
- $E_{10} = \{I_{23}, I_{29}, I_{41}, I_{33}, I_{38}\}$
- $E_{11} = \{I_{37}, I_{39}, I_{44}, I_{32}, I_{14}, I_{20}\}$
- $E_{12} = \{I_{25}\}$
- $E_{13} = \{I_{19}\}$

Table 3.1: The decision table based on X-ray images in Figures 3.4 and 3.5

Img_Name	A_1	A_2	A_3	A_4	Label
I_1	1	2	1	4	0
I_2	1	2	1	1	1
I_3	4	2	3	1	0
I_4	2	2	3	0	1
I_5	3	2	3	0	1
I_6	2	2	3	1	0
I_7	1	2	1	4	0
I_8	1	2	1	1	1
I_9	4	2	3	1	1
I_{10}	2	2	3	0	1
I_{11}	3	2	3	0	1
I_{12}	2	2	3	1	0
I_{13}	1	2	1	4	0
I_{14}	2	2	1	4	1
I_{15}	4	2	3	1	1
I_{16}	2	2	3	0	1
I_{17}	3	2	3	0	1
I_{18}	2	2	3	1	0
I_{19}	4	2	1	4	0
I_{20}	2	2	1	4	1
I_{21}	4	2	3	1	1
I_{22}	2	2	3	0	1
I_{23}	4	1	4	3	1
I_{24}	3	2	3	1	1
I_{25}	3	2	1	4	0
I_{26}	4	2	1	1	1
I_{27}	4	2	3	1	1
I_{28}	2	2	3	0	1
I_{29}	4	1	4	3	0
I_{30}	3	2	3	1	1
I_{31}	1	2	2	4	1
I_{32}	2	2	1	4	1
I_{33}	4	1	4	3	0
I_{34}	3	2	3	0	1
I_{35}	1	2	2	4	1
I_{36}	3	2	3	1	0
I_{37}	2	2	1	4	0
I_{38}	4	1	4	3	0
I_{39}	2	2	1	4	0
I_{40}	3	2	3	0	0
I_{41}	4	1	4	3	0
I_{42}	3	2	3	1	0
I_{43}	3	2	3	1	0
I_{44}	2	2	1	4	1
I_{45}	2	2	3	1	0

Table 3.2: Summary of example data in Table 3.1

X_i	E_1	E_2	E_3	E_4	E_5	E_6	E_7	E_8	E_9	E_{10}	E_{11}	E_{12}	E_{13}
$Pr(E_i)$	0.067	0.045	0.089	0.11	0.11	0.045	0.022	0.11	0.11	0.11	0.134	0.022	0.022
$Pr(P\|E_i)$	0	1	0	0.4	0.8	1	1	1	0.8	0.2	0.67	0	0

There are thirteen equivalence classes formed from Table 3.1. Equivalence class E_1 contains X-ray instances I_1, I_7, and I_3 that have same value on each conditional attribute. The probability of an equivalence class is the frequency of that equivalence class from Table 3.1. It can be denoted by $Pr(E_i)$ where $i = 1, 2, 3, \ldots, 13$. The sum of $Pr(E_i)$ is equal to 1. $Pr(P|E_i)$ is the conditional probability of equivalence class E_i that represents the proportion of pneumonia-positive instances in E_i.

Table 3.2 summarizes the probabilistic details of all thirteen equivalence classes. The set of pneumonia-positive chest X-ray images is approximated in probabilistic rough sets theory by considering the conditional probability of the equivalence class. Conditional probability of E_i can be given by,

$$Pr(P|E_i) = Pr(label = 1|E_i) = \frac{|P \cap E_i|}{|E_i|} \tag{3.5}$$

3.4 Evaluation Measures

In three-way classification, accuracy and coverage are the most commonly used evaluation measures. In this thesis, the classification performance of the model is evaluated using these two measures. Contextually, accuracy is the ratio of chest X-ray images that can be classified correctly into pneumonia-positive ($POS_{(\alpha,\ \beta)}(P)$) and pneumonia-negative ($NEG_{(\alpha,\ \beta)}(P)$) classes to the total number of chest X-ray images classified into these two classes of the three-way classification model for pneumonia [4].

It can be given based on Equation (2.12), such as follows,

$$Acc_{(\alpha,\ \beta)}(P) = \frac{Correct\ classification\ of\ images\ in\ POS_{(\alpha,\ \beta)}\ and\ NEG_{(\alpha,\ \beta)}}{Total\ number\ of\ images\ in\ POS_{(\alpha,\ \beta)}\ and\ NEG_{(\alpha,\beta)}} \times 100$$

$$= \frac{|P \cap POS_{(\alpha,\ \beta)}(P)| + |P^c \cap NEG_{(\alpha,\ \beta)}(P)|}{|POS_{(\alpha,\ \beta)}(P)| + |NEG_{(\alpha,\ \beta)}(P)|} \times 100 \tag{3.6}$$

Similarly, coverage can be defined by the proportion of chest X-ray image instances from the universal set that are classified into pneumonia-positive ($POS_{(\alpha,\ \beta)}(P)$) and pneumonia-negative ($NEG_{(\alpha,\ \beta)}(P)$) classes of the three-way classification model for pneumonia [4]. It can be formulated based on Equation (2.13), such as follows,

$$Cov_{(\alpha,\beta)}(P) = \frac{Total\ number\ of\ images\ in\ POS_{(\alpha,\beta)}\ and\ NEG_{(\alpha,\beta)}}{Total\ number\ of\ images\ in\ U} \times 100$$

$$= \frac{|POS_{(\alpha,\ \beta)}(P)| + |NEG_{(\alpha,\ \beta)}(P)|}{|U|} \times 100 \tag{3.7}$$

In above formulation of accuracy and coverage criteria, terms $|P \cap POS_{(\alpha,\ \beta)}\ (P)|$ and $|P^c \cap NEG_{(\alpha,\ \beta)}(P)|$ are the accurate classification of pneumonia-positive and pneumonia-negative X-ray images into positive and negative regions, whereas terms $|POS_{(\alpha,\ \beta)}(P)|$ and $|NEG_{(\alpha,\ \beta)}(P)|$ are the total number of X-ray images classified into positive and negative regions, respectively.

Evaluating Pawlak's Rough Sets Model

In Pawlak's rough sets theory, according to the complete inclusion and exclusion, (1, 0) is considered as the value for threshold pair (α, β). Positive, negative, and boundary regions from Table 3.1, can be given by $POS_{(1,0)}(P) = \{E_2, E_6, E_7, E_8\}$, $NEG_{(1,0)}(P) = \{E_1, E_3, E_{12}, E_{13}\}$, and $BND_{(1,0)}(P) = \{E_4, E_5, E_9, E_{10}, E_{11}\}$, respectively. The accuracy and coverage criteria of this type of three-way classification are

as follows,

$$
\begin{aligned}
Acc_{\pi_{(1,0)}}(P) &= \frac{|P \cap POS_{(\alpha,\beta)}(P)| + |P^c \cap NEG_{(\alpha,\beta)}(P)|}{|POS_{(\alpha,\beta)}(P)| + |NEG_{(\alpha,\beta)}(P)|} \times 100 \\
&= \frac{|P \cap \{E_2 \cup E_6 \cup E_7 \cup E_8\}| + |P^c \cap \{E_1 \cup E_3 \cup E_{12} \cup E_{13}\}|}{|E_1 \cup E_2 \cup E_3 \cup E_6 \cup E_7 \cup E_8 \cup E_{12} \cup E_{13}|} \times 100 \\
&= \frac{19}{19} \times 100 \\
&= 100\%
\end{aligned}
$$

$$
\begin{aligned}
cov_{\pi_{(1,0)}}(P) &= \frac{|POS_{(\alpha,\beta)}(P)| + |NEG_{(\alpha,\beta)}(P)|}{|U|} \times 100 \\
&= \frac{|E_1 \cup E_2 \cup E_3 \cup E_6 \cup E_7 \cup E_8 \cup E_{12} \cup E_{13}|}{|E_1 \cup E_2 \cup E_3 \cup \cdots \cup E_{13}|} \times 100 \\
&= \frac{|E_1 \cup E_2 \cup E_3 \cup E_4 \cup E_5 \cup E_6 \cup E_7 \cup E_8|}{|\{I_1, I_2, I_3, \ldots I_5 2\}|} \times 100 \\
&= \frac{19}{45} \times 100 \\
&= 40\%
\end{aligned}
$$

Based on Pawlak's rough sets theory, the accuracy of three-way chest X-ray classification reaches 100%, however only 40% of the X-ray images are classified. Although the pneumonia classification system in this theory has a high level of accuracy, it does not reflect the effective use of the system. The conventional three-way classification provides decisions on a few instances of X-ray images.

It is expected that an ideal pneumonia classification system will achieve a high level of accuracy and coverage simultaneously in its actual application. A high level of accuracy suggests that the model can classify X-ray images mostly correctly. In contrast, a high coverage represents the system's ability to classify more X-ray instances into definite categories (i.e., pneumonia-positive and pneumonia-negative) and provide a definite decision on most instances. Generally, a highly accurate pneumonia classification system tends to achieve a low coverage, and a system with high coverage

may reflect a low classification accuracy.

There could be a tendency here to classify more X-ray images into pneumonia-positive and pneumonia-negative classes by compromising accuracy slightly, while at the same time being careful not to misclassify pneumonia-positive cases as negative, and vice versa. Even though pneumonia-negative into pneumonia-positive class may be acceptable with a certain level of classification error rate, vice versa may not be accepted. This concept requires updating the test-treatment and testing thresholds (α, β) to a certain level of misclassification error. Now, let's also observe the result of another extreme case of the probabilistic rough sets model (i.e., 0.5-probabilistic rough sets).

Evaluating 0.5-probabilistic Rough Sets Model

In 0.5-probabilistic rough sets, (0.5, 0.5) is considered as the value for threshold pair (α, β). The corresponding threshold pair provides following regions, captured from Table 3.1 and Equation (2.11). Positive region consists of $POS_{(0.5,0.5)}(P) = \{E_2, E_5, E_6, E_7, E_8, E_9, E_{11}\}$ and negative regions consist of $NEG_{(0.5,0.5)}(P) = \{E_1, E_3, E_4, E_{10}, E_{12}, E_{13}\}$. Boundary region contains equivalence classes with conditional probability exacting 0.5. From the example of Table 3.1, we can say that it does not contain any equivalence class with conditional probability as 0.5). Hence, $BND_{(0.5,0.5}(P) = \{\phi\}$. Accuracy and coverage criteria of this type of three-way classification (most likely binary classification) are as follows,

$$\begin{aligned} Acc_{\pi_{(0.5,0.5)}}(P) &= \frac{|P \cap \{E_2 \cup E_5 \cup \cdots \cup E_9 \cup E_{11}\}| + |P^c \cap \{E_1 \cup E_3 \cup E_4 \cup E_{10} \cup E_{12} \cup E_{13}\}|}{|E_1 \cup E_2 \cup E_3 \cup \cdots \cup E_{13}|} \times 100 \\ &= \frac{38}{45} \times 100 \\ &= 84\% \end{aligned}$$

$$cov_{\pi_{(0.5,0.5)}}(P) = \frac{|E_1 \cup E_2 \cup E_3 \cup E_4 \cup E_5 \cup E_6 \cup E_7 \cup E_8 \cup E_9 \cup E_{10} \cup E_{12} \cup E_{11} \cup E_{13}|}{|E_1 \cup E_2 \cup E_3 \cup \cdots \cup E_{13}|} \times 100$$
$$= \frac{45}{45} \times 100$$
$$= 100\%$$

Based on the above outcome, 0.5-probabilistic rough sets result in a coverage of 100% and an accuracy of 84%. In this case, a slightly low coverage of X-ray images in positive and negative regions may reduce the level of uncertainty in definite classes and provide a better level of accuracy.

Let's consider another pair of thresholds (0.8, 0.2) for classification, where test-treatment threshold (α) is decreased to 0.8 and testing threshold (β) is increased to 0.2. Positive, negative, and boundary regions can be given by $POS_{(0.8,0.2)}(P) = \{E_2, E_5, E_6, E_7, E_8, E_9\}$, $NEG_{(0.8,0.2)}(P) = \{E_1, E_3, E_{10}, E_{12}, E_{13}\}$, and $BND_{(0.8,0.2)}(P) = \{E_4, E_{11}\}$, respectively. Equivalence classes E_5 and E_9 are moved to positive region and E_{10} is moved to negative region. This threshold pair obtains an accuracy score of 92% and coverage score of 76% as interpreted below,

$$Acc_{\pi_{(0.8,0.2)}}(P) = \frac{|P \cap POS_{(\alpha,\beta)}(P)| + |P^c \cap NEG_{(\alpha,\beta)}(P)|}{|POS_{(\alpha,\beta)}(P)| + |NEG_{(\alpha,\beta)}(P)|} \times 100$$
$$= \frac{|P \cap \{E_2 \cup E_5 \cup E_6 \cup E_7 \cup E_8 \cup E_9\}| + |P^c \cap \{E_1 \cup E_3 \cup E_{10} \cup E_{12} \cup E_{13}\}|}{|E_1 \cup E_2 \cup E_3 \cup E_5 \cup E_6 \cup E_7 \cup E_8 \cup E_9 \cup E_{10} \cup E_{12} \cup E_{13}|} \times 100$$
$$= \frac{31}{34} \times 100$$
$$= 92\%$$

$$cov_{\pi_{(0.8,0.2)}}(P) = \frac{|POS_{(\alpha,\beta)}(P)| + |NEG_{(\alpha,\beta)}(P)|}{|U|} \times 100$$
$$= \frac{|E_1 \cup E_2 \cup E_3 \cup E_5 \cup E_6 \cup E_7 \cup E_8 \cup E_9 \cup E_{10} \cup E_{12} \cup E_{13}|}{|E_1 \cup E_2 \cup E_3 \cup \cdots \cup E_{13}|} \times 100$$
$$= \frac{34}{45} \times 100$$
$$= 76\%$$

The results indicate that thresholds can be adjusted to correctly classify more X-ray images. Pawlak's rough sets and 0.5-probabilistic rough sets classifications are special cases of probabilistic rough sets theory in which thresholds are set to (1, 0) and (0.5, 0.5). Adjusting the threshold pair to a suitable value allows for balancing accuracy and coverage as we observed in the third case.

It is likely that we will be able to achieve a better coverage of data than Pawlak's rough sets at the cost of a slight decrease in accuracy and a better accuracy than 0.5-probabilistic rough sets by compromising coverage. Nevertheless, it is crucial that the threshold pair avoid misclassification of pneumonia-positive X-ray images as pneumonia-negative. This thesis aims for the same objective when obtaining the final threshold pair. What values the thresholds α and β should be adjusted to, this analogy requires a trade-off analysis between accuracy and coverage criteria. Therefore, GTRS has been utilized in this thesis to determine a suitable value for a threshold pair from the accuracy and coverage trade-offs.

3.4.1 Other Measures to Evaluate Classification Efficiency of Three-way Pneumonia Classification Model

This study also evaluates classification performance of the model using True Positive Rate (TPR) or recall, True Negative Rate (TNR), and precision. TPR, also coined as recall, represents the number of actual pneumonia-positive X-ray images that the system correctly identifies (Equation (3.8)) [54].

On the other hand, TNR indicates the proportion of pneumonia-negative X-ray images predicted accurately as pneumonia-negative by three-way classification system (Equation (3.9)) [54], and precision expresses its positive predictive rate of the system (Equation (3.10)). Balanced accuracy represents the overall performance of the system in accurate classification of chest X-ray images (Equation (3.11)). The

mathematical formulation of all these measures can be provided as follows [79],

$$\begin{aligned} Recall_{(\alpha,\beta)}(P) &= TPR_{(\alpha,\beta)}(P) \\ &= \frac{|P \cap POS_{(\alpha,\beta)}(P)|}{|P \cap POS_{(\alpha,\beta)}(P)| + |P \cap NEG_{(\alpha,\beta)}(P)|} \times 100 \end{aligned} \tag{3.8}$$

$$Precision_{(\alpha,\beta)}(P) = \frac{|P \cap POS_{(\alpha,\beta)}(P)|}{|P \cap POS_{(\alpha,\beta)}(P)| + |P^c \cap POS_{(\alpha,\beta)}(P)|} \times 100 \tag{3.9}$$

$$TNR_{(\alpha,\beta)}(P) = \frac{|P^c \cap NEG_{(\alpha,\beta)}(P)|}{|P^c \cap NEG_{(\alpha,\beta)}(P)| + |P^c \cap POS_{(\alpha,\beta)}(P)|} \times 100 \tag{3.10}$$

$$Balanced\ accuracy = \frac{TPR_{(\alpha,\beta)}(P) + TNR_{(\alpha,\beta)}(P)}{2} \times 100 \tag{3.11}$$

In above equations, $|.|$ is the cardinality of a set. Also, terms $POS_{(\alpha,\beta)}(P)$ and $NEG_{(\alpha,\beta)}(P)$ are the positive and negative regions, respectively. Terms $|P \cap POS_{(\alpha,\beta)}(P)|$ and $|P^c \cap NEG_{(\alpha,\beta)}(P)|$ are the correct classification of pneumonia-positive and pneumonia-negative chest X-ray images into positive and negative regions, respectively, whereas terms $|P \cap NEG_{(\alpha,\beta)}(P)|$ and $|P^c \cap POS_{(\alpha,\beta)}(P)|$ are the incorrect classification of pneumonia-positive and pneumonia-negative X-ray images into negative and positive regions.

3.5 Trade-off Analysis between Evaluation Measures

An analysis of the accuracy and coverage criteria of a classification system can be visualized using GTRS in the rough sets. The sections 3.3 and 3.4 illustrate the

relationship between evaluation measures and probabilistic threshold pair with an example. A different trade-off between evaluation measures may result in a different value for test-treatment and testing thresholds. In this thesis, a competitive game between accuracy and coverage criteria is used to arrive at an effective threshold pair for implementing the three classes of pneumonia classification. This threshold pair is expected to achieve an optimal score for the coverage and accuracy criteria.

According to Algorithm 1, GTRS can be applied to obtain a practical threshold pair for the pneumonia classification system in three stages. Three steps are involved: game formulation, repetitive learning mechanism, and threshold pair selection. The stages are outlined below,

3.5.1 Game Formulation

This step requires to define the three basic elements of GTRS (represented in Figure 3.6), namely, a set of game players P, a strategy profile S, and the payoff function for each player ϑ, such as $G = \{P, S, \vartheta\}$.

Players

In this study, a two-player competitive game is formulated to analyze indecisive chest X-ray images in pneumonia classification system using game-theoretic rough sets. The selected two players are evaluation measures, accuracy (acc) and coverage (cov) criteria as described in Equations (3.6) and (3.7), respectively. Hence, a set of players

<table>
<tr><td colspan="2" rowspan="2">Game Formulation
(2 player game)</td><td colspan="2">Accuracy</td></tr>
<tr><td>Decrease α</td><td>Increase α</td></tr>
<tr><td rowspan="2">Coverage</td><td>Decrease β</td><td colspan="2" rowspan="2">payoffs (accuracy, coverage)</td></tr>
<tr><td>Increase β</td></tr>
</table>

Figure 3.6: GTRS game formulation

can be denoted by $P = \{acc,\ cov\}$.

Strategies

The set of strategies for players can be given by $S = S_{acc} \times S_{cov}$, where $S_{acc} = \{\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n\}$ is the set of possible actions assigned to player (acc), and $S_{cov} = \{\beta_1,\ \beta_2,\ \ldots,\ \beta_n\}$ is the set of possible actions assigned to player (cov).

Every strategy of acc and cov indicates updating the value of test-treatment (α) and testing (β) thresholds, respectively. This thesis uses a threshold update factor for each player, such as g_a and g_c are the threshold update factor for players acc and cov. Its values are determined by continuously experimenting with data. If the initial threshold pair is set to $(\alpha,\ \beta) = (1.0,\ 0.5)$, where $0 \leq \beta < \alpha \leq 1$, the possible strategy profile performed by player acc includes decreasing the value of α, such as $S_{acc} = \{\alpha$ no change, α decrease g_a, α decrease $2 \times g_a$, α decrease $3 \times g_a$, α decrease $4 \times g_a\}$, and player cov can decrease the value of β according to its strategy profile, such as $S_{cov} = \{\beta$ no change, β decrease g_c, β decrease $2 \times g_c$, β decrease $3 \times g_c$, β decrease $4 \times g_c\}$. Players in GTRS are considered to be rational such that opposing players can determine a player's strategy [30].

Payoff Functions

Payoff for the player represents the gain achieved by performing an action. It can be given by the same evaluation measures as players in the game. Hence, payoffs are defined by Equations (3.6) and (3.7). It can be denoted by $\vartheta = \{\vartheta_{acc},\ \vartheta_{cov}\}$, where ϑ_{acc} is the payoff for player acc, and ϑ_{cov} is the payoff for player cov.

For a given strategy profile $p = (\alpha_1,\ \beta_1)$, such as player acc performing action α_1 and player cov performing action β_1, the payoff for players acc and cov can be determined by $Acc_{(\alpha_1,\ \beta_1)}(P)$ and $Cov_{(\alpha_1,\ \beta_1)}(P)$, respectively. These payoffs are the functions of threshold pair $(\alpha,\ \beta)$ that is interpreted from performing strategies α_1

and β_1. Thus, for strategy profile $p = (\alpha_1,\ \beta_1)$, the payoff for player *acc* can be defined by $\vartheta_{acc}(\alpha_1,\ \beta_1)$ and the payoff for player *cov* can be defined by $\vartheta_{cov}(\alpha_1,\ \beta_1)$ as follows,

$$\begin{aligned} acc(\alpha_1,\ \beta_1) \Rightarrow \vartheta_{acc}(\alpha,\ \beta) = Acc_{(\alpha_1,\ \beta_1)}(P), \\ cov(\alpha_1,\ \beta_1) \Rightarrow \vartheta_{cov}(\alpha,\ \beta) = Cov_{(\alpha_1,\beta_1)}(P) \end{aligned} \tag{3.12}$$

In Equation (3.12), the sign $\Rightarrow$ represents the player performing a strategy profile $(\alpha_1,\ \beta_1)$ that leads to the respective payoff of the player. The Equation (3.12) shows that players *acc* and *cov* in the game receive their payoffs by accuracy and coverage criteria for any set of actions performed by them. Every strategy affects the payoff of both players.

Table 3.3 represents the payoff table corresponding to the game formulated in this study. Sign $\downarrow$ depicts a decrease in the threshold value. Threshold values in

Table 3.3: Payoff table for pneumonia classification system

		cov				
		β	$\beta \downarrow g_c$	$\beta \downarrow 2g_c$	$\beta \downarrow 3g_c$	$\beta \downarrow 4g_c$
acc	α	($\vartheta_{acc}(\alpha, \beta)$, $\vartheta_{cov}(\alpha, \beta)$)	($\vartheta_{acc}(\alpha, \beta - g_c)$, $\vartheta_{cov}(\alpha, \beta - g_c)$)	($\vartheta_{acc}(\alpha, \beta - 2g_c)$, $\vartheta_{cov}(\alpha, \beta - 2g_c)$)	($\vartheta_{acc}(\alpha, \beta - 3g_c)$, $\vartheta_{cov}(\alpha, \beta - 3g_c)$)	($\vartheta_{acc}(\alpha, \beta - 4g_c)$, $\vartheta_{cov}(\alpha, \beta - 4g_c)$)
	$\alpha \downarrow g_a$	($\vartheta_{acc}(\alpha - g_a, \beta)$, $\vartheta_{cov}(\alpha - g_a, \beta)$)	($\vartheta_{acc}(\alpha - g_a, \beta - g_c)$, $\vartheta_{cov}(\alpha - g_a, \beta - g_c)$)	($\vartheta_{acc}(\alpha - g_a, \beta - 2g_c)$, $\vartheta_{cov}(\alpha - g_a, \beta - 2g_c)$)	($\vartheta_{acc}(\alpha - g_a, \beta - 3g_c)$, $\vartheta_{cov}(\alpha - g_a, \beta - 3g_c)$)	($\vartheta_{acc}(\alpha - g_a, \beta - 4g_c)$, $\vartheta_{cov}(\alpha - g_a, \beta - 4g_c)$)
	$\alpha \downarrow 2g_a$	($\vartheta_{acc}(\alpha - 2g_a, \beta)$, $\vartheta_{cov}(\alpha - 2g_a, \beta)$)	($\vartheta_{acc}(\alpha - 2 g_a, \beta - g_c)$, $\vartheta_{cov}(\alpha - 2 g_a, \beta - g_c)$)	($\vartheta_{acc}(\alpha - 2 g_a, \beta - 2g_c)$, $\vartheta_{cov}(\alpha - 2 g_a, \beta - 2g_c)$)	($\vartheta_{acc}(\alpha - 2 g_a, \beta - 3g_c)$, $\vartheta_{cov}(\alpha - 2 g_a, \beta - 3g_c)$)	($\vartheta_{acc}(\alpha - 2 g_a, \beta - 4g_c)$, $\vartheta_{cov}(\alpha - 2 g_a, \beta - 4g_c)$)
	$\alpha \downarrow 3g_a$	($\vartheta_{acc}(\alpha - 3g_a, \beta)$, $\vartheta_{cov}(\alpha - 3g_a, \beta)$)	($\vartheta_{acc}(\alpha - 3 g_a, \beta - g_c)$, $\vartheta_{cov}(\alpha - 3 g_a, \beta - g_c)$)	($\vartheta_{acc}(\alpha - 3 g_a, \beta - 2g_c)$, $\vartheta_{cov}(\alpha - 3 g_a, \beta - 2g_c)$)	($\vartheta_{acc}(\alpha - 3 g_a, \beta - 3g_c)$, $\vartheta_{cov}(\alpha - 3 g_a, \beta - 3g_c)$)	($\vartheta_{acc}(\alpha - 3 g_a, \beta - 4g_c)$, $\vartheta_{cov}(\alpha - 3 g_a, \beta - 4g_c)$)
	$\alpha \downarrow 4g_a$	($\vartheta_{acc}(\alpha - 4g_a, \beta)$, $\vartheta_{cov}(\alpha - 4g_a, \beta)$)	($\vartheta_{acc}(\alpha - 4 g_a, \beta - g_c)$, $\vartheta_{cov}(\alpha - 4 g_a, \beta - g_c)$)	($\vartheta_{acc}(\alpha - 4 g_a, \beta - 2g_c)$, $\vartheta_{cov}(\alpha - 4 g_a, \beta - 2g_c)$)	($\vartheta_{acc}(\alpha - 4 g_a, \beta - 3g_c)$, $\vartheta_{cov}(\alpha - 4 g_a, \beta - 3g_c)$)	($\vartheta_{acc}(\alpha - 4 g_a, \beta - 4g_c)$, $\vartheta_{cov}(\alpha - 4 g_a, \beta - 4g_c)$)

the second column and second row correspond to the actions of players *acc* and *cov*, respectively. The action profile of player *acc*, such as $S_{acc} = \{\alpha$ no change, α decrease g_a, α decrease $2 \times g_a$, α decrease $3 \times g_a$, α decrease $4 \times g_a\}$ are given in every row of the second column. The action profile of *cov*, such as $S_{cov} = \{\beta$ no change, β decrease g_c, β decrease $2 \times g_c$, β decrease $3 \times g_c$, β decrease $4 \times g_c\}$ are given in every column of the second row in the payoff table. In aforementioned actions, g_a and g_c are the threshold update factor for *acc* and *cov*.

Payoffs of the players corresponding to their action profiles are provided in a respective cell starting from the third row and third column. For instance, payoff $(\vartheta_{acc}(\alpha,\ \beta),\ \vartheta_{cov}(\alpha,\ \beta))$ in third row-third column is the result of threshold pair $(\alpha,\ \beta)$, where $\vartheta_{acc}(\alpha,\ \beta)$ is the payoff for player *acc* and $\vartheta_{cov}(\alpha,\ \beta)$ is the payoff for player *cov*. This particular cell also represents the payoff of the initial threshold pair in every game iteration.

Obtaining Pareto Optimal Solution

An optimal solution or balanced trade-off that provides the best cooperation between the two conflicting game players, accuracy and coverage, is determined according to the game solution strategies [60]. We have used the Pareto optimality strategy in this thesis to determine the optimal game solution based on the cooperation of players, considering that both evaluation measures are important, and we need to coordinate between them.

Additionally, this thesis experiments with Nash equilibrium strategy, but observes that Nash equilibrium strategy decreases the payoff of opponents while increasing the payoff of a player. Compared to Nash equilibrium, the Pareto optimality strategy achieves a better result for accuracy and coverage criteria. Accordingly, this thesis empowers Pareto optimality to find the game solution. According to Pareto optimality, a strategy profile is found in which no further changes in strategy can result

in a better payoff for one player without worsening the payoff for another at the same time. The strategy profile $p = (\alpha^*,\ \beta^*) \in S_{acc} \times S_{cov}$ is Pareto optimality $if\ \forall\ p\ \in P,\ \vartheta_{i(\alpha^*,\ \beta^*)}\ \geq\ \vartheta_{i(\alpha,\ \beta)}$ [30], i.e.,

$$\forall \alpha_i \in S_{acc}, \forall \beta_i \in S_{cov},$$
$$(\vartheta_{acc_{(\alpha^*,\ \beta^*)}} \geq \vartheta_{acc_{(\alpha_i,\ \beta_i)}}) \wedge (\vartheta_{cov_{(\alpha^*,\ \beta^*)}} \geq \vartheta_{cov_{(\alpha_i,\ \beta_i)}});$$
$$where\ \alpha_i \neq \alpha^* \wedge \beta_i \neq \beta^*$$

3.5.2 Repetitive Learning

The repetitive learning method is used to analyze the payoff of each player repeatedly. Essentially, the game is reformulated by repetitively modifying threshold values until the desired stopping criteria are reached. This method includes the following two notions:

Game Reformulation

Let's consider that the game is being played with the initial threshold pair as $(\alpha,\ \beta)$ and the threshold pair corresponding to the Pareto optimality strategy can be given as $(\alpha^*,\ \beta^*)$. If currently selected threshold pair $(\alpha^*,\ \beta^*)$ does not provide a satisfactory outcome for accuracy and coverage, the game is reformulated with the currently selected threshold pair as the initial threshold pair in the next iteration to check if any threshold pair near the current one provides a better outcome [60]. Hence, if there is an opportunity to improve the outcome without violating the stop criteria, the game is re-played with the threshold pair $(\alpha^*,\ \beta^*)$ as initial thresholds in the next iteration.

Stop Criteria

We need to stop reformulating the game when certain criteria have been reached. We select stop criteria based on continuous experimentation and its observations. A default stop criterion is $0 \leq \beta < \alpha \leq 1$. The game reformulation can also be stopped if the threshold pair $(\alpha^*,\ \beta^*)$ satisfies other stop criteria such as a specific value for the payoff. It is imperative to stop game reformulation if further improvements in threshold pair do not outweigh previous improvements and may negatively affect payoffs. The stop criteria for this research are therefore defined as players reaching their payoffs over some specified value, such as player *cov* achieving coverage over 65%, $\vartheta_{cov} > 65\%$, and player *acc* achieving accuracy over 96%, $\vartheta_{acc} > 96\%$. When any of the stop criteria is violated, players stop performing their actions. As the whole process proceeds, an optimal value is gradually reached for the final threshold pair $(\alpha,\ \beta)$ [7, 60].

3.6 Cost-Sensitive Three-way Pneumonia Classification

In decision making problems, especially in the medical field, we are expected to contemplate the costs of making different decisions or costs that occur due to different classification errors [77]. Another application of game-theoretic rough sets, the email spam filtering system, is also a cost-sensitive three-way approach [74]. Due to the loss of vital information in this case, incorrect classification of a legitimate email into a spam folder costs more than incorrect classification of a spam email into a legitimate folder [77].

Similarly, we can analyze the costs incurred when making different decisions in three-way pneumonia classification. We attempt to include costs associated with each decision made by the three-way pneumonia classification system in this section.

Table 3.4: Cost matrix for three-way pneumonia classification

	Pneumonia-positive	Pneumonia-negative
Treatment	$\lambda(treatment \mid pneumonia - positive)$	$\lambda(treatment \mid pneumonia - negative)$
Further test	$\lambda(further\ test \mid pneumonia - positive)$	$\lambda(further\ test \mid pneumonia - negative)$
No treatment	$\lambda(no\ treatment \mid pneumonia - positive)$	$\lambda(no\ treatment \mid pneumonia - negative)$

According to three-way pneumonia classification, medical experts have three options, treatment, no treatment, and further testing.

As shown in Table 3.4, a 3×2 cost matrix can be implemented for pneumonia classification. Six types of costs can be presented in this matrix [77]. In the event that a patient has pneumonia and medical experts determine that treatment is necessary, the cost may be calculated by using $\lambda(treatment \mid pneumonia - positive)$. Patient does not have pneumonia, and medical experts make the right decision to not provide treatment to the patient, the cost can be expressed by $\lambda(no\ treatment \mid pneumonia - negative)$. Both of these are correct decisions that cost actual associated risk.

The cost of incorrect decisions is usually higher. If the patient does not have pneumonia but medical experts decide to provide treatment of pneumonia, the cost can be given by $\lambda(treatment \mid pneumonia - negative)$. This type of cost occurs when pneumonia-negative X-ray images are classified as pneumonia-positive. The other type of misclassification when pneumonia-positive are classified as pneumonia-negative, usually costs more than the previous type of misclassification. So, if the patient has pneumonia but medical experts decide not to provide treatment, the cost for not having treatment or delayed treatment can be expressed by $\lambda(no\ treatment \mid pneumonia - positive)$.

There is also cost associated with deferred decisions (further tests) that are caused from X-ray images being classified as pneumonia-indecisive. The cost of performing diagnostic tests and postponing treatment for the patient with possible pneumonia is $\lambda(further\ tests \mid pneumonia - positive)$, and for the patient without pneumonia is $\lambda(further\ tests \mid pneumonia - negative)$.

In Bayesian decision theory [77], the cost incurred for each of these six decisions is not equal. For example, the cost associated with not providing treatment to a pneumonia patient is highest as it can result in death or serious injury. Then, providing treatment to non-pneumonia patients could be the next as it can lead to the loss of monetary value caused due to the treatment. In deferment scenarios, conducting tests rather than performing treatment on pneumonia patients may cost higher as it would delay the treatment procedure. Deferring further testing for a non-pneumonia patient would result in the loss of monetary value.

We can calculate the cost for making each decision using the above types of loss functions or cost incurred with the conditional probability of the equivalence classes [74],

$$\begin{aligned} C(treatment \mid E_i) &= \lambda_{(treatment \mid pneumonia-positive)} Pr(P|E_i) + \\ &\quad \lambda_{(treatment \mid pneumonia-negative)} Pr(P^C|E_i) \\ C(further\ test \mid E_i) &= \lambda_{(further\ test \mid pneumonia-positive)} Pr(P|E_i) + \\ &\quad \lambda_{(further\ test \mid pneumonia-negative)} Pr(P^C|E_i) \\ C(no\ treatment \mid E_i) &= \lambda_{(no\ treatment \mid pneumonia-positive)} Pr(P|E_i) + \\ &\quad \lambda_{(no\ treatment \mid pneumonia-negative)} Pr(P^C|E_i) \end{aligned} \tag{3.13}$$

Using binary classification to make a decision on every instance may lead to increased misclassification of X-ray images, thus, resulting in higher costs. By classifying chest X-rays according to three classes, we can minimize the expected cost associated with each decision. We attempt to minimize the highest cost by avoiding misclassifying pneumonia-positive images as negatives.

3.7 Summary

In this chapter, we present the implementation of the proposed model in two phases. The first phase extracts a set of features from chest X-ray images that the second phase uses as the decision table. The second phase deals with uncertainty in classification of indecisive chest X-ray images. GTRS is applied to derive three pneumonia classification categories based on an optimal threshold pair. Furthermore, at every step of the process, the chapter provides well-explained examples.

Chapter 4

Implementation and Empirical Evaluation

We describe the practical implementation of the proposed model in this chapter. It elaborates on determining the conditional attributes and constructing a decision table for rough sets analysis in Section 4.1. Section 4.2 applies GTRS to find an optimal threshold pair. In Section 4.3, we identify the three regions of pneumonia classification, analyze the test results of the proposed model, and compare the results with other related models.

4.1 Dataset

The proposed model is evaluated on a chest X-ray image dataset that was published institutionally by the University of California San Diego, and the Guangzhou Women and Children's Medical Center [26]. The dataset is publicly available on `https://data.mendeley.com/datasets/rscbjbr9sj/3` website and consists of 4,281 (72.83%) pneumonia-positive and 1,597 (27.17%) pneumonia-negative chest X-ray images.

There are more than twice as many pneumonia-positive instances in the dataset as

pneumonia-negative instances. Thus, image augmentation technique is used to produce a balanced version of raw data. It is possible to reproduce pneumonia-negative X-ray images by updating the brightness and contrast level of the images [47]. As part of image processing, we have gain α and bias β parameters that affect contrast and brightness respectively [37]. We increase the brightness level of the copied images by applying bias β, factors such as 10, 20, and 25, and adjust the contrast by multiplying the pixel by gain α factors 1.1, 1.2, 1.3. These techniques were chosen because medical experts also examine chest X-ray images under light and increasing contrast makes edges and certain regions more visible. A replica of approximately 50% of pneumonia-positive X-ray images is generated in order to approximate the number of images of both types.

After image augmentation, the experimental dataset contains 11,522 X-ray images. Of these, 5,514 are pneumonia-negative images while 6,008 are pneumonia-positive images. Furthermore, image processing techniques such as bilateral filtering and histogram equalization are used to reduce noise and improve contrast of X-ray images.

The whole set of chest X-ray images is divided into 10,548 images for training and remaining images for testing the model. In addition to testing the proposed model on two different test sets, this study also employs a standard 10-fold cross validation to measure the effectiveness of the model. As mentioned in Section 3, the training set is considered to be the universal set (U), with each chest X-ray image constituting an object of U.

4.1.1 Determining Attributes for the Decision Table

To perform further experiments, a set of statistical features is extracted from each chest X-ray image in U. As mentioned in Section 3.2.2, after feature selection step, we selected keypoint descriptors such as oriented fast and rotated binary robust independent elementary features (Orb), binary robust invariant scalable keypoints ($Brisk$),

and accelerated-kaze (*Akaze*) for rough sets analysis [29, 31, 39]. These features are treated as input (C) to train the proposed model. An output attribute (D) ("Label") is also considered to provide the output label for each X-ray image in the set. The domain of D is {0, 1}, where "0" represents pneumonia-negative and "1" illustrates pneumonia-positive X-ray images. Hence the set of attributes in the feature vector is {Img_Name, Orb, Kaze, Brisk, Label}, where Img_Name is the object identifier. Three features (Orb, Kaze, and Brisk) are conditional attributes and Label is the decisional attribute of the decision table (DT) in rough sets.

4.2 Empirical Implementation with GTRS

A threshold pair is obtained by employing GTRS after building a decision table for the experiment.

4.2.1 Decision Table Implementation

Conditional attributes are combined with the decisional attribute to form the decision table according to Equation (2.1). Thus, the decision table (DT) contains the following columns: Img_Name, Orb, Kaze, Brisk, and Label. All image instances in the table are transformed into a three-dimensional representation, every dimension corresponds to a conditional attribute.

Each conditional attribute is filled with continuous values, making interpretation of the image feature difficult. Any value from the conditional attribute is less likely to be repeated for any other instance in the table. Binning is a data processing technique that reduces the effect of trivial errors in observation by grouping a set of data into small bins [46]. Thus, the equal-frequency binning method is used to discretize the continuous value for each attribute.

To begin with, each conditional attribute is sorted ascendingly. The sorted values

Table 4.1: An example of the decision table from experimental data

Img_Name	Orb	Akaze	Brisk	Label
Img_1	1	2	1	1
Img_2	1	2	1	0
Img_3	2	2	3	0
Img_4	1	2	1	0
Img_5	5	6	4	1
Img_6	2	2	3	0
Img_7	5	6	4	1
Img_8	4	4	3	1
Img_9	2	5	3	1
Img_10	2	5	3	0

Table 4.2: Conditional probability of equivalence classes from experimental data

E_i	E_1	E_2	E_3	E_4	E_5	E_6	E_7	E_8	E_9	E_{10}	E_{11}	E_{12}
$Pr(P\|E_i)$	0.99	1.00	1.00	0.96	0.75	0.93	1.00	0.59	0.75	0.67	1.00	1.00
E_i	E_{13}	E_{14}	E_{15}	E_{16}	E_{17}	E_{18}	E_{19}	E_{20}	E_{21}	E_{22}	E_{23}	E_{24}
$Pr(P\|E_i)$	0.96	0.94	1.00	0.86	1.00	0.98	0.99	1.00	1.00	1.00	1.00	0.96
E_i	E_{25}	E_{26}	E_{27}	E_{28}	E_{29}	E_{30}	E_{31}	E_{32}	E_{33}	E_{34}	E_{35}	E_{36}
$Pr(P\|E_i)$	0.91	1.00	0.94	0.48	0.44	0.22	0.36	0.43	0.58	0.52	0.57	0.69
E_i	E_{37}	E_{38}	E_{39}	E_{40}	E_{41}	E_{42}	E_{43}	E_{44}	E_{45}	E_{46}	E_{47}	E_{48}
$Pr(P\|E_i)$	0.97	0.99	0.33	1.00	0.95	0.96	1.00	1.00	1.00	0.98	1.00	1.00
E_i	E_{49}	E_{50}	E_{51}	E_{52}	E_{53}	E_{54}	E_{55}	E_{56}	E_{57}	E_{58}	E_{59}	E_{60}
$Pr(P\|E_i)$	1.00	0.96	1.00	1.00	0.93	0.81	0.27	0.00	0.38	0.02	0.03	0.10
E_i	E_{61}	E_{62}	E_{63}	E_{64}	E_{65}	E_{66}	E_{67}	E_{68}	E_{69}	E_{70}	E_{71}	E_{72}
$Pr(P\|E_i)$	0.00	0.28	0.28	0.44	0.24	0.00	0.55	0.67	0.90	0.11	0.03	1.00
E_i	E_{73}	E_{74}	E_{75}	E_{76}	E_{77}	E_{78}	E_{79}	E_{80}	E_{81}	E_{82}	E_{83}	E_{84}
$Pr(P\|E_i)$	1.00	0.96	0.14	0.00	0.90	1.00	0.94	0.95	0.65	0.00	0.30	0.39
E_i	E_{85}	E_{86}	E_{87}	E_{88}	E_{89}	E_{90}	E_{91}	E_{92}	E_{93}	E_{94}	E_{95}	E_{96}
$Pr(P\|E_i)$	0.24	0.04	0.00	0.14	0.34	0.35	0.06	0.00	0.38	0.58	0.64	0.05
E_i	E_{97}	E_{98}	E_{99}	E_{100}	E_{101}	E_{102}	E_{103}	E_{104}	E_{105}	E_{106}	E_{107}	E_{108}
$Pr(P\|E_i)$	0.02	1.00	1.00	1.00	0.00	0.02	0.80	1.00	0.77	0.86	0.65	0.46
E_i	E_{109}	E_{110}	E_{111}	E_{112}	E_{113}	E_{114}	E_{115}	E_{116}	E_{117}	E_{118}	E_{119}	E_{120}
$Pr(P\|E_i)$	0.46	0.48	0.00	0.00	0.31	0.37	0.51	0.05	0.00	0.40	0.44	0.01
E_i	E_{121}	E_{122}	E_{123}	E_{124}	E_{125}	E_{126}	E_{127}	E_{128}	E_{129}	E_{130}	E_{131}	E_{132}
$Pr(P\|E_i)$	0.01	1.00	0.00	0.00	0.71	0.41	0.66	0.68	0.38	0.44	0.47	0.00
E_i	E_{133}	E_{134}	E_{135}	E_{136}	E_{137}	E_{138}	E_{139}	E_{140}	E_{141}	E_{142}	E_{143}	E_{144}/ E_{145}
$Pr(P\|E_i)$	0.27	0.45	0.00	0.00	0.36	0.55	0.51	0.04	0.00	0.00	0.92	0.26/ 0.00

are then divided into six bins with an equal number of instances, such that each bin contains exactly 1,758 X-ray image instances. Thereafter, each bin is assigned a label from 1, 2, 3, 4, 5, 6. Following the binning process, each conditional attribute has optimized values. An example of a decision table can be seen in Table 4.1, where each image instance has a value between 1 and 6 (inclusive) for each conditional attribute. The attribute *Label* indicates the decisional attribute and has a value either 0 (pneumonia-negative) or 1 (pneumonia-positive) per instance.

4.2.2 Equivalence Class and Conditional Probability

The Decision Table (DT) is partitioned into equivalence classes (i.e., instances with the same bin label across all conditional attributes are grouped together). A total of 145 equivalence classes are created during the experiment. For each equivalence class, the conditional probability $Pr(P|E_i)$ is calculated as per Equation (3.5). Table 4.2 represents the conditional probability of all 145 equivalence classes.

4.2.3 Game Development

As mentioned earlier in Subsection 3.5.1, the competitive game for rough sets analysis is formulated between accuracy and coverage criteria. As an initial threshold pair, we tried (1, 0) and (0.5, 0.5). We did not obtain satisfactory results from this approach and after a point, we stopped improving and started gaining constant payoffs. As a result, we selected (α, β) = (1, 0.5) as the initial threshold pair for the beginning of the game, which is continued until $0 \leq \beta < \alpha \leq 1$ constraint is violated or stop criteria is achieved.

The value of g_a and g_c are set to 0.01 and 0.02 respectively. These values were determined by continually experimenting with data. Player *acc* can change the value of α according to its strategy profile, such as $S_{acc} = \{\alpha$ no change, α decrease 0.01, α decrease 0.02, α decrease 0.03, α decrease 0.04$\}$. Player *cov* can change the value of

β as per its strategy profile, that is S_{cov} = {β no change, β decrease 0.02, β decrease 0.04, β decrease 0.06, β decrease 0.08}.

The corresponding set of actions (i.e., threshold pairs (α, β)) is {(1.0, 0.5), (0.48, 0.99), (0.46, 0.98), (0.44, 0.97), (0.42, 0.96)} with initial threshold pair (1, 0.5). The objective of the game is to obtain an effective threshold pair that can provide maximum payoff for players *acc* and *cov*.

4.2.4 Finding Threshold Pair with GTRS

We first try to obtain the Nash equilibrium solution in this research. We observe that the obtained threshold pair at every iteration of the game provides a better payoff for player *acc* than previous iteration, but it worsens off the payoff of player *cov* at the same time. This phenomena can be seen in Figure 4.1.

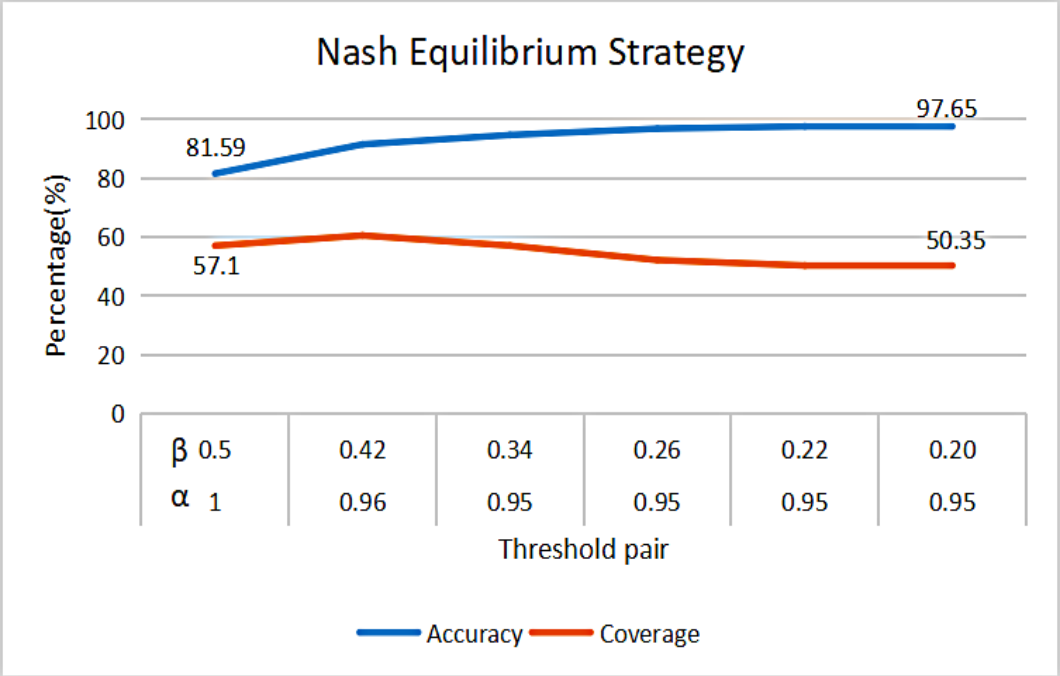

Figure 4.1: Nash equilibrium outcome of repetitive learning mechanism

After the fifth iteration (i.e., (α, β) = (0.95, 0.20)), it can be observed that Nash equilibrium does not improve payoffs of players further. The results of this experiment

indicate that Nash equilibrium obtains an accuracy score of 97.65% and a coverage score of 50.35% at the fifth iteration, and that coverage continues to be close to constant as iterations progress.

The threshold pair obtained using Nash equilibrium could not provide a satisfactory result for coverage. Focusing on accuracy reduces the coverage of chest X-ray images into positive and negative regions. As this strategy does not obtain a balanced trade-off between accuracy and coverage, we also applied Pareto optimality strategy in this study, expecting to obtain a better result than Nash equilibrium strategy. This section elaborates on achieving an optimal threshold pair using the Pareto optimality strategy.

Payoffs for the first iteration are given in Table 4.3. The right up corner is the game Pareto optimality where payoffs for players are (acc, cov) = (85.00%, 73.27%) and strategy profile is (α decrease 0.04, β no change). Hence, the outcome of the first iteration is the threshold pair (0.96, 0.5).

The stop criteria for the game is set to payoffs greater than a specified value, such as $\vartheta_{cov} > 65\%$, and $\vartheta_{acc} > 96\%$, and threshold pair $0 \leq \beta < \alpha \leq 1$. According to the stop criteria, there is still the possibility of improving payoffs by reformulating the game with updated thresholds after the first iteration. Therefore, the game is

Table 4.3: Game outcome with initial threshold

Accuracy						
Coverage	$S_c \downarrow S_a \rightarrow$	$\alpha = 1.0$	$\alpha \downarrow 0.01$	$\alpha \downarrow 0.02$	$\alpha \downarrow 0.03$	$\alpha \downarrow 0.04$
	$\beta = 0.5$	(81.59%, 57.10%)	(81.59%, 57.10%)	(83.01%, 62.28%)	(83.19%, 63.06%)	**(85.00%, 73.27%)**
	$\beta \downarrow 0.02$	(83.28%, 54.14%)	(83.28%, 54.14%)	(84.63%, 59.32%)	(84.79%, 60.11%)	(86.45%, 70.32%)
	$\beta \downarrow 0.04$	(85.36%, 50.63%)	(85.36%, 50.63%)	(86.60%, 55.80%)	(86.75%, 56.59%)	(88.19%, 66.80%)
	$\beta \downarrow 0.06$	(87.49%, 47.30%)	(87.49%, 47.30%)	(88.60%, 52.47%)	(88.73%, 53.26%)	(89.93%, 63.47%)
	$\beta \downarrow 0.08$	(89.52%, 44.40%)	(89.52%, 44.40%)	(90.48%, 49.57%)	(90.59%, 50.36%)	(91.53%, 60.57%)

Table 4.4: Game outcome of repetitive learning

	Initial (α, β)	Strategy	Solution (α, β)	Payoff (ϑ_{acc}, ϑ_{cov})
1	(1, 0.5)	($\alpha \downarrow 0.04$, β)	(0.96, 0.5)	(85.00%, 73.27%)
2	(0.96, 0.5)	($\alpha \downarrow 0.04$, β)	(0.92, 0.5)	(86.46%, 87.99%)
3	(0.92, 0.5)	($\alpha \downarrow 0.04$, β)	(0.88, 0.5)	(86.51%, 89.41%)
4	(0.88, 0.5)	(α, $\beta \downarrow 0.08$)	(0.88, 0.42)	(91.92%, 76.71%)
5	(0.88, 0.42)	(α , $\beta \downarrow 0.08$)	(0.88, 0.34)	(94.39%, 70.84%)
6	(0.88, 0.34)	(α , $\beta \downarrow 0.08$)	(0.88, 0.26)	(96.07%, 66.98%)
7	(0.88, 0.26)	(α , $\beta \downarrow 0.04$)	**(0.88, 0.22)**	**(96.66%, 65.45%)**

repeated by setting (0.96, 0.5) as the initial threshold pair for the next iteration. By updating thresholds from (1, 0.5) to (0.96, 0.5), accuracy and coverage increase to 86.46% and 87.99%, respectively. The outcome of the repetitive game is shown in Table 4.4.

We continue to reformulate the game until we reach stop criteria as mentioned earlier in Subsection 3.5.2. We repeat the game seven times. Table 4.4 shows the outcome of the repetitive learning mechanism. When we achieve a coverage score of over 65%, we focus on achieving a high accuracy until we reach the maximum accuracy of over 96% while constraining that the stop criteria of the game are not violated.

In the seventh iteration, it is noticed that players *acc* and *cov* have achieved their stop criteria (i.e., coverage greater than 65% and accuracy over 96%). In the experiment, it is observed that further repetition of the game violated the stop criterion for *cov*. At this iteration, repetitive learning from the game is stopped and the final threshold pair (α, β) = (0.88, 0.22) is selected.

4.2.5 Nash vs. Pareto

The result of the Pareto optimal strategy to find the threshold pair is shown in Figure 4.2. At the seventh iteration, it is observed that the game is stopped since it does not provide any major improvements in payoffs after reaching the threshold pair

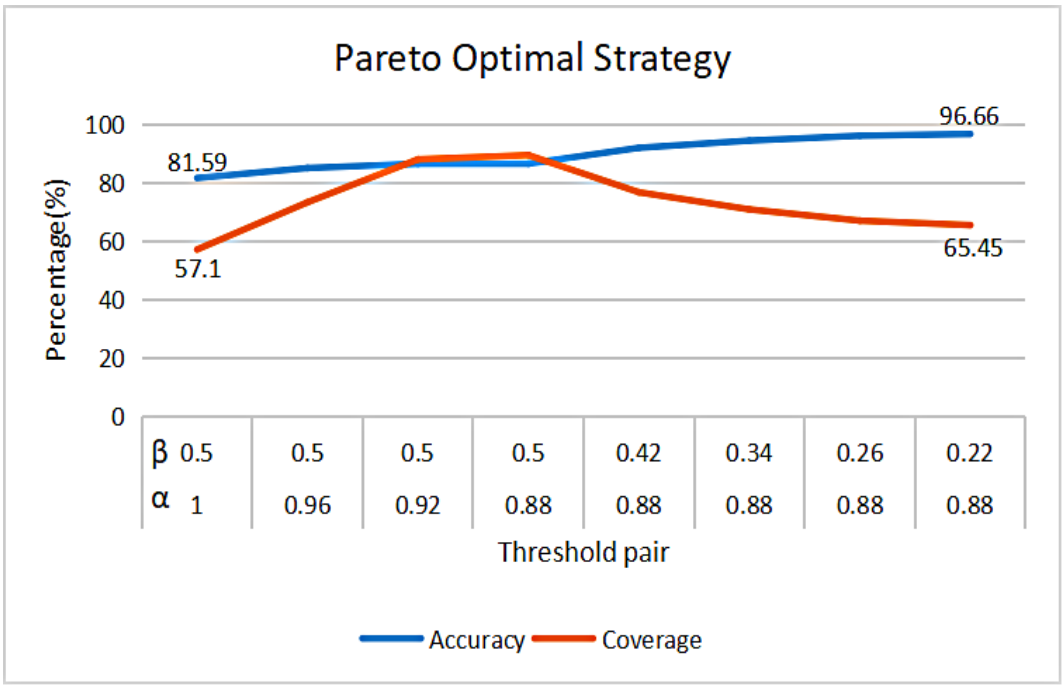

Figure 4.2: Pareto optimality outcome of repetitive learning mechanism

$(\alpha, \beta) = (0.88, 0.22)$, but negatively affects the coverage criterion.

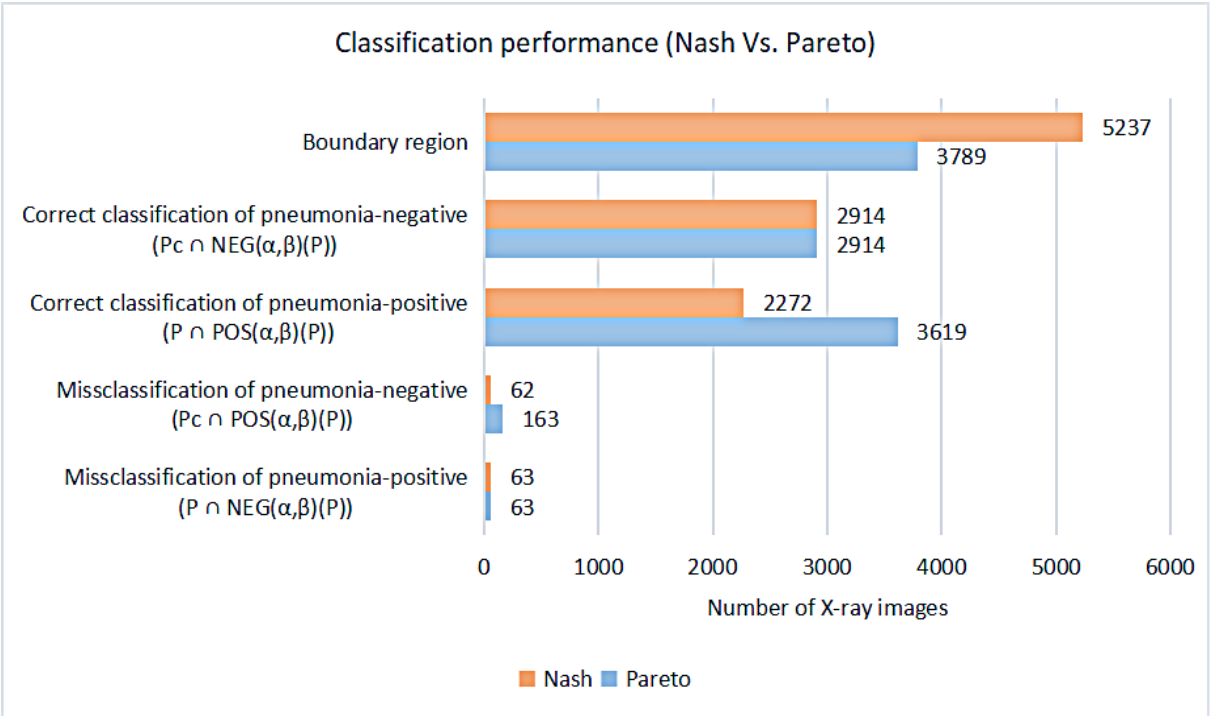

Figure 4.3: Classification performance of Nash and Pareto

Table 4.5: Comparison between Nash and Pareto strategies

Strategy	(α, β)	Accuracy	Coverage	Recall	Precision	TNR	Balanced Accuracy
Nash	(0.95, 0.20)	97.65%	50.35%	97.30%	97.34%	97.92%	97.61%
Pareto	(0.88, 0.22)	96.66%	65.45%	98.29%	95.69%	94.70%	96.50

The Pareto optimality strategy reaches 65.45% of coverage, which is 15.10% higher than the Nash equilibrium strategy. There is a slight decrease (0.99%) $\downarrow$ in the accuracy score as compared to the Nash equilibrium outcome. Further, Pareto optimality does not compromise the ability to detect the actual positive cases of pneumonia. This strategy provides a higher recall response (0.99% $\uparrow$) than Nash equilibrium strategy (Table 4.5).

In addition, the Pareto optimality strategy covers more instances of X-ray images into positive and negative regions (Figure 4.3). The Nash equilibrium strategy can classify 5,311 X-ray images, whereas the Pareto optimality strategy can classify 6,759 X-ray images into positive and negative regions, as shown in the figure. As a result, the system becomes more useful by providing definite decisions on a greater number of cases.

The table 4.5 compares the classification performance of two models whose threshold pair has been selected using Nash equilibrium and Pareto optimality. Using actual 2,335 pneumonia-positive X-ray images, Nash equilibrium can identify 2,272 of them correctly as pneumonia-positive, resulting in a true positive rate of 97.30% (Equation (3.8).

As compared to Nash equilibrium, Pareto optimality identifies 3,619 of 3,682 pneumonia-positive X-ray images as pneumonia-positive, yielding a true positive rate of 98.29% (Equation (3.8). Nash equilibrium provides a lower true positive rate than Pareto optimality, but it does not cover as many X-ray images into definite regions as Pareto optimality.

This comparison shows that although Pareto optimality decreases precision and

overall accuracy of the classification model, it outperforms Nash equilibrium strategy when it comes to classification efficiency (i.e., with regard to the number of chest X-ray images classified into definite classes and the true positive rate for pneumonia-positive X-ray image classification). Due to a higher misclassification rate of pneumonia-negative X-ray images in contrast to Nash equilibrium strategy, the Pareto optimal method loses its balanced accuracy.

As Pareto optimality provides greater coverage and better accuracy in classification of pneumonia-positive images than Nash equilibrium, it can be concluded that the proposed classification model based on the threshold pair obtained using Pareto optimality strategy rarely misleads the actual pneumonia-positive cases. Accordingly, we select the final threshold pair $(\alpha, \beta) = (0.88, 0.22)$ according to the Pareto optimality strategy.

The decision to accept a chest X-ray image as pneumonia-positive is concluded from the positive region. Furthermore, the decision to reject a chest X-ray image as pneumonia-positive is made from the negative region. The proposed model defers making a decision on the X-ray image if it cannot be decided whether it is in the positive or negative region. Such images are placed into the boundary region. The model can be expressed as,

- Accepting as Pneumonia-positive: $P(Label = 1|E_i) \geq \alpha$
- Rejecting as Pneumonia-positive: $P(Label = 1|E_i) \leq \beta$
- Deferral decision: $\beta < P(Label = 1|E_i) < \alpha$
  where "*Label*" is the decisional attribute

According to Pawlak's rough sets method and from table 4.2, the pneumonia-positive class contains $\{E_2, E_3, E_7, E_{11}, E_{12}, E_{15}, E_{17}, E_{20}, E_{21}, E_{22}, E_{23}, E_{26}, E_{40}, E_{43}, E_{44}, E_{45}, E_{47}, E_{48}, E_{49}, E_{51}, E_{52}, E_{72}, E_{73}, E_{78}, E_{98}, E_{99}, \ E_{100}, \ E_{104}, \ E_{122}\}$ equivalence classes, and pneumonia-negative class contains $\{E_{56}, E_{61}, E_{66}, E_{82}, E_{87}, E_{92}, E_{111}, E_{112},$

$E_{117}, E_{123}, E_{124}, E_{132}, E_{135}, E_{136}, E_{141}, E_{142}, E_{145}\}$ equivalence calsses. Further, remaining equivalence classes are classified into the boundary region (i.e., pneumonia-indecisive class).

After Applying GTRS, equivalence classes $\{E_1, E_4, E_6, E_{13}, E_{14}, E_{18}, E_{19}, E_{24}, E_{25}, E_{27}, E_{37}, E_{38}, E_{46}, E_{50}, E_{53}, E_{69}, E_{74}, E_{79}, E_{80}, E_{143}\}$ are also derived into pneumonia-positive class, and equivalence classes $\{E_{30}, E_{58}, E_{59}, E_{60}, E_{70}, E_{71}, E_{75}, E_{86}, E_{88}, E_{91}, E_{96}, E_{97}, E_{102}, E_{116}, E_{120}, E_{121}\}$ are derived into pneumonia-negative class apart from the equivalence classes already classified into positive and negative regions by Pawlak's rough sets.

As a result, classification method based on GTRS allows more numbers of chest X-ray images to be classified for definite decisions (i.e., pneumonia-positive and pneumonia-negative from positive and negative regions, respectively) while reducing the size of the boundary region by using Pareto optimal threshold pair (0.88, 0.22) for test-treatment and testing thresholds.

4.3 Analysis and Test Results

This section examines the classification performance of the proposed three-way chest X-ray image classification method based on GTRS. First, the pneumonia classification rules are explained in Subsection 4.3.1. Afterwards, results and observations are discussed after the model has been tested on two separate test sets in Subsection 4.3.2. In this thesis, we also use 10-fold cross validation to test the predictive performance of the proposed model in Subsection 4.3.3.

4.3.1 Interpretation of Classification Rules

Based on the final outcome of the game, such as the test-treatment and testing threshold pair (α, β) = (0.88, 0.22), three classes are constructed from the training set (U).

These classes lead to the classification rules of the pneumonia classification system. Following are the classification rules that can be derived based on Equation (2.9),

- $POS_{(0.88,0.22)}(P) = \{a \in U \mid Pr(P|[a]) \geq 0.88\}$,
- $NEG_{(0.88,0.22)}(P) = \{a \in U \mid Pr(P|[a]) \leq 0.22\}$,
- $BND_{(0.88,0.22)}(P) = \{a \in U \mid 0.22 < Pr(P|[a]) \wedge Pr(P|[a]) < 0.88\}$

Where P is the target concept for deriving the three-way pneumonia classification system. It represents the proportion of pneumonia-positive X-ray images in an equivalence class. Term $Pr(P|[a])$ is the conditional probability of equivalence class $[a]$.

The first rule, POS, represents the positive region of the pneumonia classification system. It accepts chest X-ray images of an equivalence class ($[a]$) as pneumonia-positive if its conditional probability ($Pr(P|[a])$) is greater than or equals to 0.88. Accordingly, the presence of pneumonia is confirmed and treatment can be initiated immediately.

The second rule, NEG, represents the negative region of the pneumonia classification system. It rejects chest X-ray images of $[a]$ to be the part of pneumonia-positive class if its $Pr(P|[a])$ is less than or equals to 0.22. Such X-ray images are categorized as pneumonia-negative. Accordingly, the absence of pneumonia is confirmed and further testing for pneumonia is not required.

Third rule, BND, represents the pneumonia-indecisive class. Making a decision on X-ray images of $[a]$ is deferred if its $Pr(P|[a])$ lies between 0.22 and 0.88 (exclusive). Further diagnostic tests are performed to determine whether pneumonia is likely. In most cases, medical experts perform additional tests to confirm the presence or absence of pneumonia if it has not been confirmed. Attempting to follow the manual procedure of medical experts would allow the healthcare system to gain trust.

With the help of the POS and NEG rules it is possible to quickly decide whether to treat or not to treat pneumonia. BND creates a room for accounting for suspicious cases of pneumonia. This is considered a "possibly" case of pneumonia. Decisions on such X-ray images are not immediately made but are held until certain information is gathered.

4.3.2 Test Results and Observations

After gaining three classes from the training set, the performance of the proposed model is initially observed on two separate test sets. From the same set of features extracted from test X-ray images or any any new image, we follow the same binning procedure using the training set, then, try to identify the equivalence class and predict its class from the region of the equivalence class. We tried cover almost all possibilities of the equivalence classes. If the conditional attributes do not match with any equivalence class, we don't get outcome for that image. We also apply the 10-fold cross validation technique for model evaluation in Subsection 4.3.3.

Test Set 1

Using 970 unseen chest X-ray images from the experimental dataset, the model achieves 94.88% accuracy and 62.37% coverage of the test data when making concrete decisions. Table 4.6 statistics regarding the prediction count for test set 1. The

Table 4.6: Prediction stats from test set 1

		Predicted		
Actual		pneumonia-positive ($POS_{(\alpha,\beta)}(P)$)	pneumonia-negative ($NEG_{(\alpha,\beta)}(P)$)	Actual classified
	pneumonia-positive (P)	314	8	322
	pneumonia-negative (P^c)	23	260	283
	Actual classified	337	268	**605**

model correctly identifies 314 out of 322 pneumonia-positive chest X-ray images as pneumonia-positive and obtains a true positive rate of 97.52% (Equation (3.8)). In addition, it correctly identifies 260 out of 283 pneumonia-negative images as pneumonia-negative, resulting in a true negative rate of 91.87% (Equation (3.10).

Enunciating the overall performance of the model, it achieves 94.70% balanced accuracy (Equation (3.11)) and 93.18% positive predictive rate (Equation (3.9)) which is comparable to other rough sets and related models. The true positive classification rate suggests that the proposed model rarely misdiagnoses the patient with pneumonia.

Test Set 2

We observe some existing experiments that distinguish between pneumonia types and COVID-19 [22]. So, as part of this study, a second test set is used to evaluate the classification performance of the model on a completely separate set of chest X-ray images depicting another disease (other than pneumonia). In order to generate the second test set, the chest X-ray images from the National Institutes of Health (NIH) that visualize pneumonia and other diseases, such as Emphysema and Effusion (available on `https://nihcc.app.box.com/v/ChestXray-NIHCC` website and used in [55]) are combined. In this test set, the model is evaluated on its ability to distinguish pneumonia-positive chest X-ray images from other kinds of disease-infected images.

Table 4.7: Prediction stats from test set 2

		Predicted		
Actual		pneumonia-positive ($POS_{(\alpha,\beta)}(P)$)	pneumonia-negative ($NEG_{(\alpha,\beta)}(P)$)	Actual classified
	pneumonia-positive (P)	127	3	130
	pneumonia-negative (P^c)	114	125	239
	Actual classified	241	128	**369**

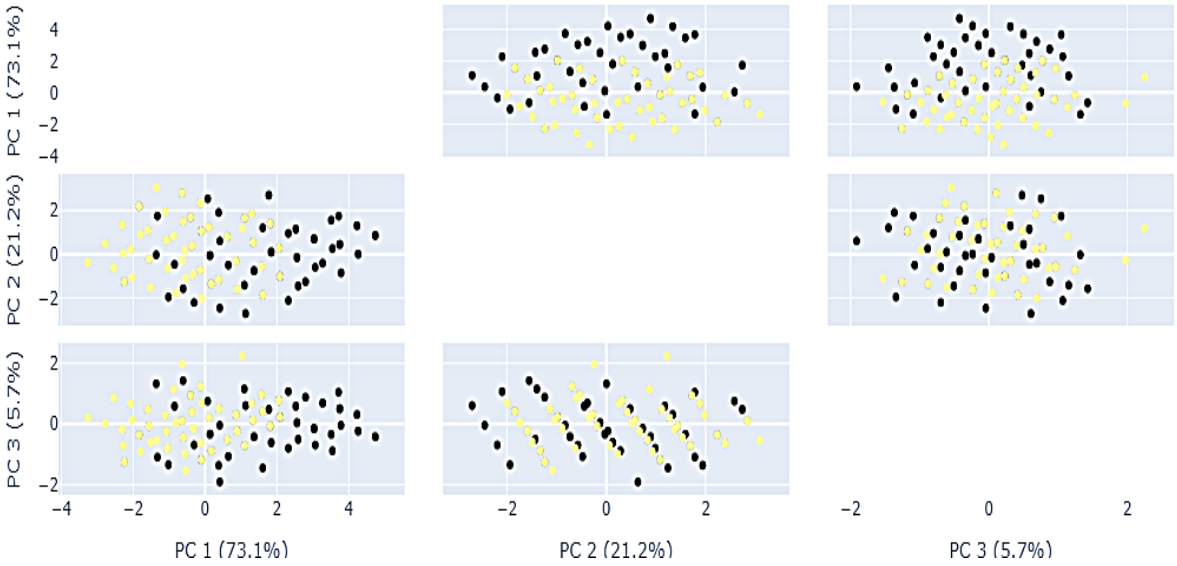

Figure 4.4: Principle component analysis on test set 2 (yellow: pneumonia-positive. black: other disease)

In the second set of images, there are no normal chest X-rays. Table 4.7 shows the results of the three-way pneumonia classification system on the second test set. The GTRS model correctly identifies 127 of the 130 actual pneumonia-positive images and obtains a true positive rate of 97.69% (Equation (3.8)).

On the basis of keypoint descriptors extracted from X-ray images, it is observed that other disease-infected chest X-ray images have similar features to pneumonia-positive chest X-ray images, such as a similar number of keypoint descriptors (Orb, Kaze, and Brisk). The Principle Component Analysis (PCA) can be observed in Figures 4.4 and 4.5.

PCA is a multivariate technique derived from probability theory and linear algebra. It generates a matrix to represent the variation in data [1, 33, 56]. Figure 4.4 shows the order of features that are the resulting principal components based on their ability to represent variance.

The plots between PC 2 and PC 3 are interconnected and do not completely represent the separation of two classes of chest X-ray images, whereas the plot of PC 1 does somewhat distinguish pneumonia-positive X-ray images from other disease-

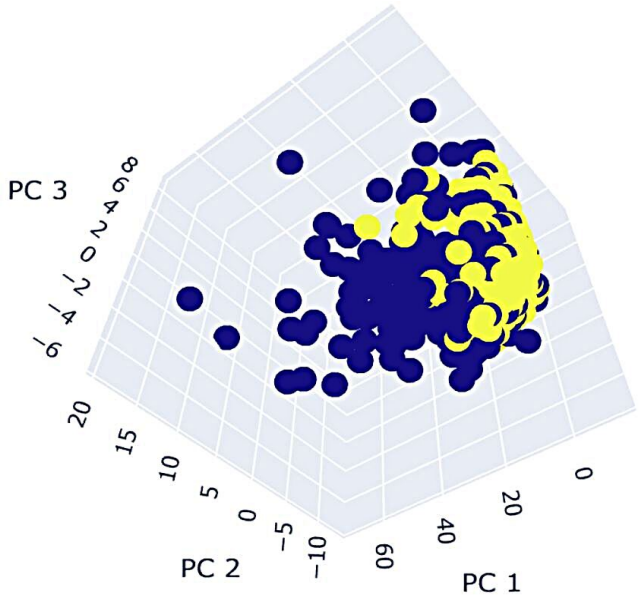

Figure 4.5: 3D representation of PCA on test set 2 (yellow: pneumonia-positive. dark blue: other disease)

infected images. A similar plot of PCA on the second test set can be observed in 3D format in Figure 4.5.

Due to this similarity, the proposed model identifies only 125 out of 239 actual other disease-infected images as pneumonia-negative (or other) and remaining as pneumonia-positive; causing the model to result into a low true negative rate of 52.30% (Equation (3.10)). The overall performance of the proposed model based on GTRS while differentiating pneumonia-positive from other types of diseases is 75% (Equation (3.11)).

In related work, we observe some existing experiments that differentiate between pneumonia types and COVID-19 [22]. So in order to check the ability of the proposed model to differentiate between two diseases, we perform testing on a completely different dataset and we observe that selected features are not sufficient for differentiation.

Test results of the proposed model from test set 1 and test set 2 are compared with 0.5-probabilistic and Pawlak's rough sets techniques in Table 4.8 and 4.9, respectively. The 0.5-probabilistic rough sets uses threshold pair (α,β) = (0.5, 0.5) such that it

Table 4.8: Comparison with other rough sets model on test set 1

Test Set 1 (pneumonia-positive and pneumonia-negative (normal) chest X-ray images)						
Model	(α, β)	Accuracy	Coverage	Recall	Precision	TNR
0.5-probabilistic rough sets	(0.5,0.5)	81.44%	100.00%	77.15%	85.37%	85.99%
Pawlak's rough sets	(1, 0)	95.76%	12.16%	95.83%	93.88%	95.71%
GTRS (Proposed model)	(0.88,0.22)	94.88%	62.37%	97.52%	93.18%	91.87%

Table 4.9: Comparison with other rough sets model on test set 2

Test Set 2 (pneumonia-positive and other disease-infected chest X-ray images)						
Model	(α, β)	Accuracy	Coverage	Recall	Precision	TNR
0.5-probabilistic rough sets	(0.5,0.5)	63.93%	100.00%	78.70%	50.84%	55.24%
Pawlak's rough sets	(1, 0)	71.05%	6.12%	100.00%	26.67%	67.65%
GTRS (Proposed model)	(0.88,0.22)	68.29%	59.42%	97.69%	52.70%	52.30%

nearly provides 100% coverage of the data in classification. In contrast, the latter technique uses threshold (α,β) = (1, 0) that tends to provide error-free pneumonia-positive and pneumonia-negative classes [41]. Both techniques are evaluated in the same manner as the proposed model (i.e., by dividing the training set (U) into three regions as per Equations (2.10) in Pawlak's rough sets and (2.11) in 0.5-probabilistic rough sets).

From test results, we can say that the classification model based on GTRS increases classification accuracy by 13.44% (test set 1) and 4.36% (test set 2) as compared to 0.5-probabilistic rough sets. In contrast to Pawlak's rough sets model, the proposed GTRS model surges coverage of the data by 50.21% (test set 1) and 53.3% (test set 2). This shows a greater usability of the system than Pawlak's rough sets,but at the cost of a slight decrease in overall classification accuracy whilst providing a similar true positive classification rate (1.69% $\uparrow$ on test set 1 and 2.31% $\downarrow$ on test set 2).

In comparison with 0.5-probabilistic rough sets, the GTRS model covers 37.63% less test data for definite decision. As illustrated in Section 3.4, 0.5-probabilistic rough sets provides almost binary classification, forcing the system to make decisions on uncertain images. This causes the system to misclassify 180 chest X-ray images to positive and negative regions, obtaining the classification error rate of 18.56%. Through the GTRS based three-way approach, uncertain images can be classified into a third category. This approach reduces the classification error rate to 5.12%. The chest X-ray images of uncertain quality should be re-examined in order to avoid making a mistake that could have grave consequences, both financially and in terms of life. Through the use of GTRS based three-way classification approach, this research provides doctors with the opportunity to re-examine cases of patients falling into the indecisive category rather than making decisions based on uncertain information.

Both test sets indicate that the GTRS technique simulates a decision support system that can quickly assist experts in identifying true positive pneumonia cases. Positive pneumonia cases are rarely misclassified by the system and may cause delayed treatment. Future work in this literature can concentrate on improving the true negative response and overall classification accuracy of the model.

4.3.3 The 10-fold Cross Validation

We also employ the 10-fold cross validation technique to evaluate the prediction performance of the model in order to ensure validity of this approach. The entire dataset of 11,522 chest X-ray images is used in this standard testing. According to this technique, we first apply $(\alpha, \beta) = (0.88, 0.22)$ on the training set to obtain a model by forming three classes. On the remaining test set, we predict the output label. Pawlak's rough sets and 0.5-probabilistic rough sets are both evaluated in the same way as GTRS.

The classification statistics for Pwalak's rough sets, 0.5-probabilistic rough sets,

and GTRS can be observed in Tables 4.10, 4.11, and 4.12, respectively. In each of the three tables, the first column is the fold number, the next four columns are the classification stats in positive and negative regions, and the remaining columns show different criteria for evaluation. Each criterion is calculated using equations defined by Section 3.4.

In tables 4.10, 4.11, and 4.12, the second column (TP) is $P \cap POS_{(\alpha,\beta)}(P)$, representing the correct classification of pneumonia-positive X-ray images to pneumonia-positive class; the third column (FN) is $P \cap NEG_{(\alpha,\beta)}(P)$, representing the incorrect classification of pneumonia-positive X-ray images to pneumonia-negative class; the fourth column (FP) is $P^c \cap POS_{(\alpha,\beta)}(P)$, representing the incorrect classification of pneumonia-negative X-ray images to pneumonia-positive class; the fifth column (TN) is $P^c \cap NEG_{(\alpha,\beta)}(P)$, representing the correct classification of pneumonia-negative X-ray images to pneumonia-negative class.

In Pawlak's rough sets (Table 4.10), ten runs achieve accuracy scores ranging from 96.30% to 100.00% and coverage scores ranging from 9.11% to 14.58%. On average, Pawlak's rough set, based on testing set of nearly 1,151 X-ray images per fold, identifies 28 pneumonia-positive and 97 pneumonia-negative X-ray images for

Table 4.10: 10-fold result with Pawlak's rough sets (TP: $P \cap POS_{(\alpha,\beta)}(P)$. FN: $P \cap NEG_{(\alpha,\beta)}(P)$. FP: $P^c \cap POS_{(\alpha,\beta)}(P)$. TN: $P^c \cap NEG_{(\alpha,\beta)}(P)$)

	Positive		Negative		Evaluation measures				
Fold	TP	FN	FP	TN	Accuracy	Coverage	Recall	Precision	TNR
1	23	0	0	100	100.00%	10.68%	100.00%	100.00%	100.00%
2	38	1	4	97	96.43%	12.14%	97.44%	90.48%	96.04%
3	22	1	0	82	99.05%	9.11%	95.65%	100.00%	100.00%
4	28	1	1	79	98.17%	9.48%	96.55%	96.55%	98.75%
5	31	0	0	78	100.00%	9.47%	100.00%	100.00%	100.00%
6	24	2	0	115	98.58%	12.26%	92.31%	100.00%	100.00%
7	20	1	0	147	99.40%	14.58%	95.24%	100.00%	100.00%
8	24	0	1	80	99.05%	9.12%	100.00%	96.00%	98.77%
9	32	0	1	78	99.10%	9.64%	100.00%	96.97%	98.73%
10	29	2	3	101	96.30%	11.73%	93.55%	90.63%	97.12%

definite diagnosis.

On average, based on 28 pneumonia-positive X-ray images, 27 of them were correctly identified as pneumonia-positive, obtaining a recall (true positive rate) score of 96.42% (Equation (3.8)). Also, it identifies 96 of 97 actual pneumonia-negative images correctly as pneumonia-negative and secures 98.96% true negative rate (Equation (3.10)). These two measures lead to a balanced accuracy (Equation (3.11)) score of 97.69% on overall performance of the model while making definite decisions.

In 0.5-probabilistic rough sets (Table 4.11), ten runs obtain accuracy scores ranging from 83.16% to 85.33% and 100.00% coverage scores. On average, 0.5-probabilistic rough sets, from testing set of nearly 1,151 X-ray images per fold, classifies 600 pneumonia-positive and 551 pneumonia-negative X-ray images for definite decisions.

On average, based on 600 actual pneumonia-positive X-ray images, 492 of them are correctly identified as pneumonia-positive, obtaining a true positive rate score of 82.00% (Equation (3.8)). Also, it identifies 475 of 551 actual pneumonia-negative images correctly as pneumonia-negative and secures 86.20% true negative rate (Equation (3.10)). These two measures lead to a balanced accuracy (Equation (3.11)) score of 84.10% on overall performance of the model while making definite decisions.

Table 4.11: 10-fold result with 0.5-probabilistic rough sets (TP: $P \cap POS_{(\alpha,\beta)}(P)$. FN: $P \cap NEG_{(\alpha,\beta)}(P)$. FP: $P^c \cap POS_{(\alpha,\beta)}(P)$. TN: $P^c \cap NEG_{(\alpha,\beta)}(P)$)

	Positive		Negative		Evaluation measures				
Fold	TP	FN	FP	TN	Accuracy	Coverage	Recall	Precision	TNR
1	491	114	68	479	84.20%	100.00%	81.16%	87.84%	87.57%
2	482	103	88	480	83.43%	100.00%	82.39%	84.56%	84.51%
3	465	124	70	493	83.16%	100.00%	78.95%	86.92%	87.57%
4	490	109	81	470	83.48%	100.00%	81.80%	85.81%	85.30%
5	506	103	82	460	83.93%	100.00%	83.09%	86.05%	84.87%
6	497	123	68	462	83.39%	100.00%	80.16%	87.96%	87.17%
7	515	105	64	468	85.33%	100.00%	83.06%	88.95%	87.97%
8	521	95	85	450	84.36%	100.00%	84.58%	85.97%	84.11%
9	468	114	68	501	84.19%	100.00%	80.41%	87.31%	88.05%
10	481	95	91	484	83.84%	100.00%	83.51%	84.09%	84.17%

Table 4.12: 10-fold result with GTRS (TP: $P \cap POS_{(\alpha,\beta)}(P)$.

FN: $P \cap NEG_{(\alpha,\beta)}(P)$. FP: $P^c \cap POS_{(\alpha,\beta)}(P)$. TN: $P^c \cap NEG_{(\alpha,\beta)}(P)$)

	Positive		Negative		Evaluation measures				
Fold	TP	FN	FP	TN	Accuracy	Coverage	Recall	Precision	TNR
1	382	9	23	325	95.67%	64.15%	97.70%	94.32%	93.39%
2	389	6	19	340	96.68%	65.39%	98.48%	95.34%	94.71%
3	377	5	19	327	96.70%	63.19%	98.69%	95.20%	94.51%
4	381	6	22	311	96.11%	62.61%	98.45%	94.54%	93.39%
5	379	8	18	299	96.31%	61.16%	97.93%	95.47%	94.32%
6	410	13	16	311	96.13%	65.22%	96.93%	96.24%	95.11%
7	415	10	6	317	97.86%	64.93%	97.65%	98.57%	98.14%
8	412	9	20	298	96.08%	64.21%	97.86%	95.37%	93.71%
9	392	14	17	331	95.89%	65.51%	96.55%	95.84%	95.11%
10	374	9	27	324	95.10%	63.77%	97.65%	93.27%	92.31%

In probabilistic rough sets using GTRS (Table 4.12), ten runs obtain accuracy scores ranging from 95.10% to 97.86% and coverage scores ranging from 61.16% to 65.51%. On average, from testing set of nearly 1,151 X-ray images per fold, the proposed model classifies 400 pneumonia-positive and 337 pneumonia-negative X-ray images for definite decisions.

On average, from 400 actual pneumonia-positive X-ray images, 391 of them are correctly identified as pneumonia-positive with a recall (true positive rate) score of 97.75% (Equation (3.8)). Also, it identifies 318 of 337 actual pneumonia-negative images correctly as pneumonia-negative and secures 94.36% true negative rate (Equation (3.10)). These two measures lead to a balanced accuracy (Equation (3.11)) score of 96.06% on overall performance of the model while making definite decisions.

A comparison of coverage and accuracy among all three models can be seen in Figure 4.6. Vertical bars represent coverage criteria, while dashed lines indicate accuracy criteria. In Figure 4.7 and Figure 4.8, we present the trend of true positive rate and true negative rate for all three models.

Average accuracy of ten folds is used to determine the model's average classification accuracy. Similarly, usability of the model is determined from the average

coverage of the ten folds. In addition, other criteria are averaged. Table 4.13 compares the results of 10-fold cross validation (on average) of GTRS with other rough sets techniques.

Observations:

- The recall score (True positive rate) of pneumonia-positive X-ray images in pneumonia classification system entails the robustness of the system against misleading actual pneumonia-positive cases. GTRS model obtains a true positive rate of 97.75% which is 1.33% ↑ higher than Pawlak's rough sets, and

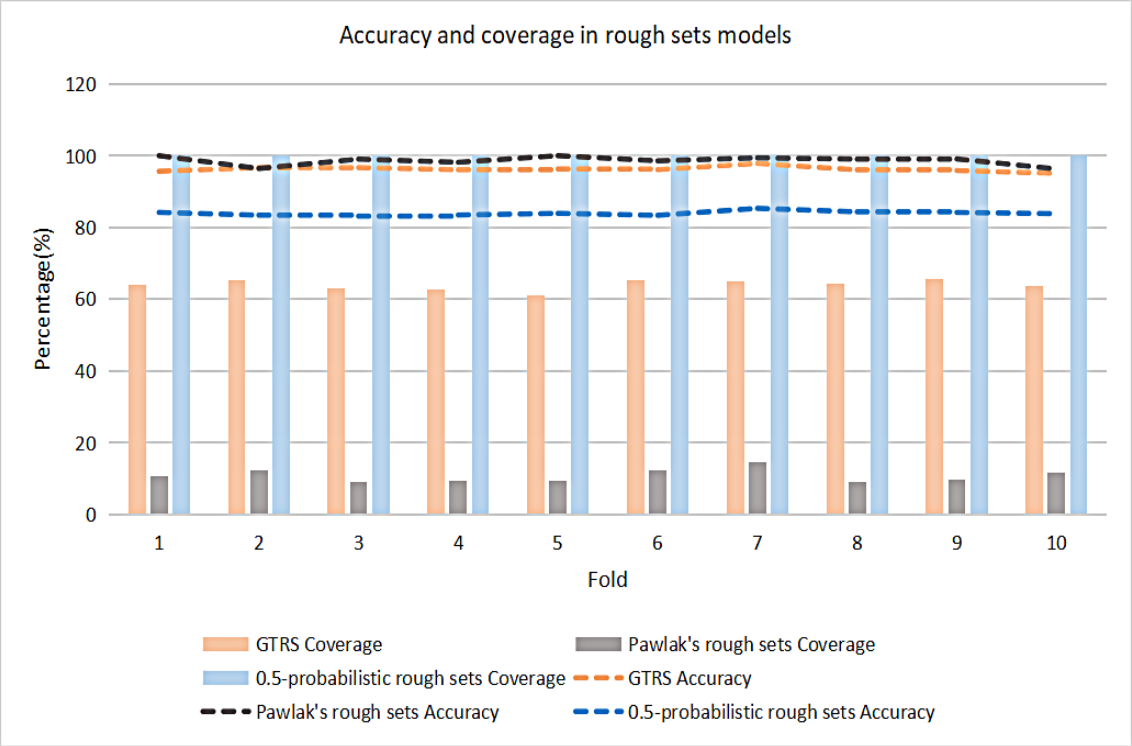

Figure 4.6: Comparison of accuracy and coverage in rough sets models

Table 4.13: Comparison of 10-fold result among rough sets model

Model	(α, β)	Accuracy	Coverage	Recall	Precision	TNR	Balanced Accuracy
Pawlak's rough sets	(1, 0)	98.61%	10.82%	96.42%	97.06%	98.96%	97.69%
0.5-probabilistic rough sets	(0.5, 0.5)	83.93%	100.00%	82.00%	86.55%	86.20%	84.10%
GTRS (Proposed model)	(0.88, 0.22)	96.25%	64.01%	97.75%	95.42%	94.36%	96.06%

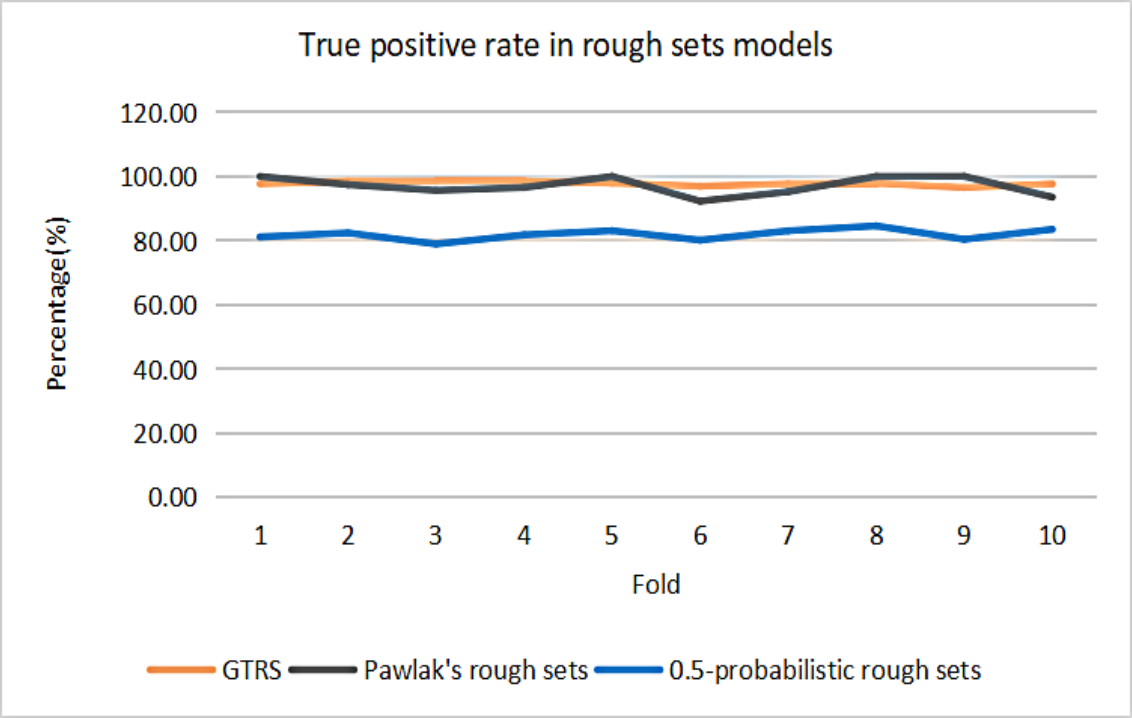

Figure 4.7: Comparison of true positive rate in rough sets models

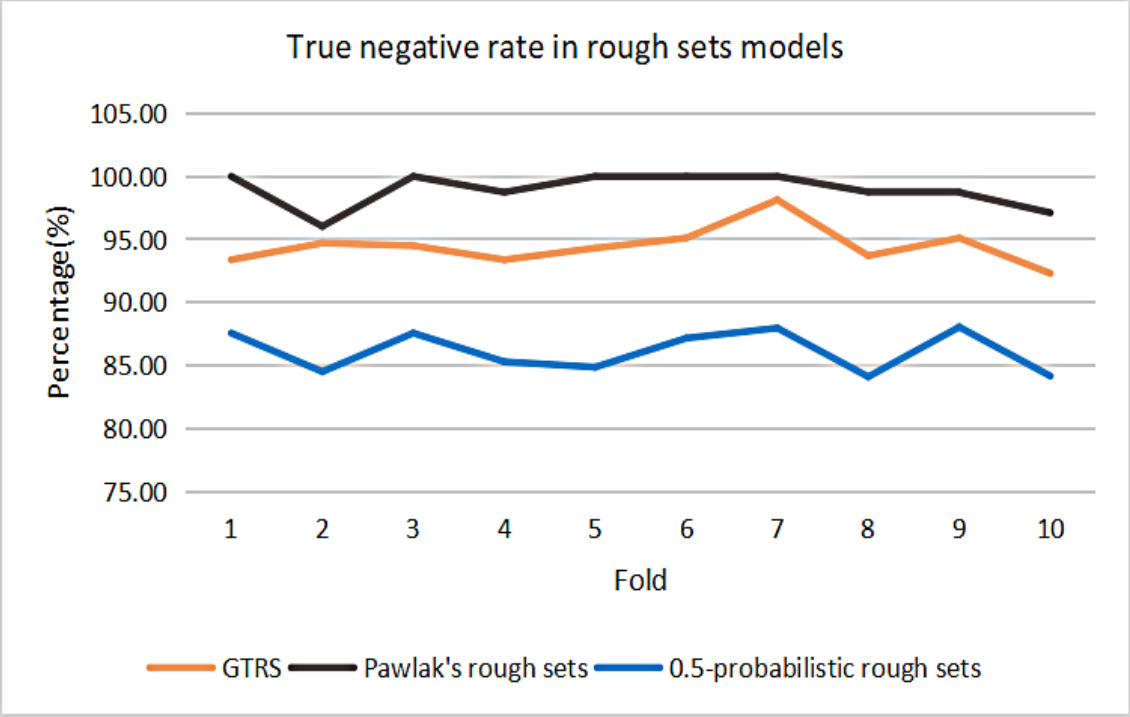

Figure 4.8: Comparison of true negative rate in rough sets models

15.75% ↑ higher than 0.5-probabilistic rough sets (Table 4.13 and Figure 4.7). This result implies that the pneumonia classification model based on GTRS is unlikely to mislead the positive cases of pneumonia.

- Figure 4.6 represents that GTRS compromises on accuracy (2.36% ↓) as compared to Pawlak's rough sets model. The reduced accuracy of the proposed model is due to the incorrect classification of pneumonia-negative X-ray images to positive class as shown in Figure 4.8. Conversely, GTRS obtains a higher true positive rate (1.33% ↑) than Pawlak's rough sets. The true positive rate reflects the accuracy of the classification of pneumonia-positive X-ray images into pneumonia-positive class.

- Figure 4.6 also illustrates the coverage of the data for definite decisions. The GTRS model can cover more X-ray images in both positive and negative regions. It covers 64.01% of the data, which is (53.19% ↑) higher than Pawlak's rough sets. The higher the coverage of data for definite decisions, the more usable the model. Pawlak's rough sets also give the impression that a low proportion of X-ray images is classified would rarely lead to any misclassification of X-ray images. This might lead to a higher accuracy than GTRS.

- Compared with 0.5-probabilistic rough sets, GTRS achieves a lower level of data coverage for definite decisions. In the former rough sets technique, threshold pair (0.5, 0.5) is used to classify X-ray images into boundary regions if features exactly match with an equivalence class that has a conditional probability of 0.5; otherwise, X-ray images are classified into positive and negative regions. Probability of an equivalence class with an equal number of pneumonia-positive and pneumonia-negative X-ray images is less. We observe the same phenomenon in Table 4.2, where we do not find an equivalence class with 0.5 conditional probability.

0.5-probabilistic rough sets nearly classifies all X-ray instances to positive and negative regions and tends to achieve 100.00% coverage of the data in most cases [77]. Although it increases usability of the model, it fails to contemplate the uncertainty of the data in the classification process. As a result, the system cannot re-examine uncertain chest X-ray images, leading to an increased misclassification rate of 16.07%. The GTRS-based three-way approach, on the other hand, allows for handling uncertain chest X-ray images by classifying them to the third category and reduces classification error rate to 3.75%. Uncertain chest X-ray images must be re-examined in order to avoid an associated risk of making the wrong decision (Section 3.6). In other words, by using GTRS based three-way classification approach, this research offers medical experts an opportunity to re-examine the case of patients falling into indecisive category instead of making concrete decisions based on incomplete information and increasing the expected cost.

- From test results in Table 4.13, it can be inferred that GTRS achieves a better recall, precision, and true negative scores than 0.5-probabilistic rough sets (i.e., accuracy by 12.32% ↑, recall (true positive rate) by 15.75% ↑, precision by 8.87% ↑, true negative rate by 8.16% ↑, and balanced accuracy by 11.96% ↑). The 0.5-probabilistic model essentially inclines towards the binary classification approach and may have the shortcoming of increased misclassification rate and nearly no ability to deal with uncertain information of the X-ray images. This increases the expected incurred cost. This research aims to decrease the expected cost (Section 3.6)

According to the above observations, a classification approach based on GTRS can overcome the shortcoming of Pawlak's rough sets and 0.5-probabilistic rough sets.

Additionally, the GTRS based pneumonia classification model is compared with other classification models discussed in related work. The comparison is shown in

Table 4.14: Comparison with existing experiments

Model	Dataset	Coverage	Accuracy	TPR	TNR
Proposed 5-layer CNN architecture [44]	5,856	100%	88.90%	87.66%	91.21%
Proposed 4-layer CNN architecture [50]	5,856	100%	93.01%	-	-
Proposed 6-layer CNN architecture [12]	5,856	100%	92.89%	-	-
Pre-trained AlexNet model Viral pneumonia vs. normal [22]	4,983	100%	94.05%	-	-
Pre-trained AlexNet model Bacterial pneumonia vs. normal [22]	4,870	100%	91.43%	-	-
GTRS based three-way (Proposed model in this thesis)	5,878	64.01%	96.25%	97.75%	94.36%

Table 4.14. The compared experiments [44, 50, 12] have used the same dataset and the rest have used a different dataset. Although the latter class of approaches achieves 100% coverage and represents a higher usability of the model than the GTRS based three-way method, it follows binary classification. However, binary classification does not provide the ability to deal with uncertain, inconsistent, and incomplete information of indecisive chest X-ray images. Similar response can be observed in 0.5-probabilistic rough sets, as shown in Tables 4.8 and 4.11. Binary classification approach leads to the increased misclassification rate which is very critical in the medical field. This increases the expected cost of each decision (Section 3.6).

In contrast, GTRS based three-way approach provides a provision to handle uncertain chest X-ray images by classifying them to a third category and reduces classification error rate to 3.75%. Chest X-ray images with uncertainty should be re-examined in order to avoid making wrong decisions that could result in death, poor health, and financial losses. Hence, with the aim of dealing with uncertainty in classification, avoiding making incorrect decisions over suspicious cases of pneumonia, and reducing the expected cost of making a wrong decision, this thesis employs GTRS to supplement the classification model with a third class that accounts for indecisive cases of pneumonia disease. Medical experts can re-examine the cases of patients falling under indecisive category rather than making concrete decisions based on incomplete information.

Test results from both testing (Subsections 4.3.2 and 4.3.3) represent a simulation of the decision support system that is based on GTRS. It tends to provide a balanced trade-off between accuracy and criteria. Using GTRS, one can obtain a balanced test-treatment and testing threshold pair that allows effective control of the three pneumonia classification classes - pneumonia-positive, pneumonia-indecisive, and pneumonia-negative.

The research focuses on the use of game-theoretic rough sets in developing a pneumonia classification system. Three classes are derived based on the concept of acceptance, rejection, and deferment that are equivalent to positive, negative, and boundary regions in three-way classification. This thesis also highlights how three regions can be effectively controlled by a threshold pair. The proposed model is implemented based on the principles of probabilistic rough sets technique that overcomes the stringent criteria of Pawlak's rough sets. It benefits from considering the degree of overlap between the target concept and equivalence classes.

This thesis shows that probabilistic rough sets can classify more chest X-ray instances into positive and negative regions than Pawlak's rough sets. It shows the impact of different threshold pairs (α, β) on accuracy and coverage criteria of the proposed model and concludes that obtaining a suitable threshold pair is vital to derive three classes of the proposed model.

The classification method leverages GTRS for identifying the test-treatment and testing threshold pair by formulating a two-player competitive game between accuracy and coverage and establishing a fair trade-off between them. Accuracy and coverage are evaluation criteria in the three-way model that conflict, where increasing one criterion lowers the score for the other.

Furthermore, this thesis examines different artifacts of GTRS technique such as game players, game strategies, repetitive learning mechanism, game stop criteria, and payoff functions within the context of pneumonia classification. The use of the

repetitive learning mechanism to reach a balanced threshold pair gradually is also examined.

The thesis also highlights the main advantages of GTRS. Firstly, it provides a trade-off based approach to finding an effective threshold pair for the three-way pneumonia classification model. Secondly, it simulates a meaningful interpretation of the threshold pair and derives three classes, provided that the pneumonia-indecisive class deals with incomplete information of X-ray images. This research brings an extension to the application of GTRS and helps to deliver a high quality three-way pneumonia classification model.

4.4 Summary

The chapter describes the step-by-step implementation of the proposed model. It implements the decision table for rough sets analysis from a chest X-ray image dataset, then discusses the application of GTRS to find an optimal threshold pair in Section 4.2. The obtained threshold pair is then used to determine three regions of the pneumonia classification system in Section 4.3. The analysis also highlights the results of testing the model on two different test sets. It also employs the 10-fold cross validation technique to evaluate classification performance of the proposed model and compares the test results with other related models.

Chapter 5

Conclusions

This chapter concludes the research work of this thesis to develop pneumonia classification system, summarizes its contribution and highlights the future research work. In this study, a method for classifying chest X-ray images is presented to support early diagnosis of pneumonia. In order to overcome the shortcoming of binary classification in existing work, the thesis proposes a pneumonia classification model using GTRS and three-way decisions technique.

The three-way pneumonia classification system reduces misclassification rates by dividing X-ray images into three categories (instead of two classes as in binary classification). The three classes are; pneumonia-positive containing pneumonia-infected X-ray images; pneumonia-negative containing normal or without pneumonia X-ray images; pneumonia-indecisive consists of chest X-ray images, representing incomplete and inconsistent information.

The experiment uses a chest X-ray image dataset that is publicly available on Mendeley for research purposes. The employed technique is tested on two different test sets. Test statistics from both sets demonstrate that the GTRS model helps deliver a decision support system for pneumonia that can 97.61% accurately bolster the true positive decisions of experts. It also shows that GTRS improves usability of

the model (51.76% ↑) while achieving a similar accuracy rate as Pawlak's rough sets.

This thesis also employs standard 10-fold cross validation technique to evaluate classification performance of the proposed model and compares its results with Pawlak's rough sets and 0.5-probabilistic rough sets. It shows that the true negative rate of the GTRS model is compromised, lowering the overall accuracy by (2.36% ↓). Yet, it achieves a higher true positive rate (1.33% ↑) than Pawlak's rough sets and accuracy (12.32% ↑ than 0.5-probabilistic rough sets. The higher true positive rate of the GTRS model suggests that this model rarely leads to misdiagnosis of the true cases of pneumonia.

GTRS simulates a meaningful interpretation of the threshold pair that determines the three classes of the classification model. The third class deals with uncertain and incomplete information of X-ray images. This research contributes to the application of GTRS to deliver a pneumonia classification system that overcomes the shortcoming of binary classification.

5.1 Summary of Contributions

Major contributions of this thesis are described below,

- Implementing a three-way pneumonia classification system:

  The three-way classification technique contributes to a framework for developing a pneumonia classification system. The implementation of three disjoint regions; positive, negative, and boundary regions results in three classes: pneumonia-positive class showing chest X-ray images with pneumonia infection; pneumonia-negative class showing normal X-ray images; pneumonia-indecisive class showing chest X-ray images that are suspicious of pneumonia, but can be classified either pneumonia-positive or pneumonia-negative until certain information is collected.

- A Provision to Re-examine Chest X-ray Images with Incomplete Information:
  In the three-way pneumonia classification system, the three classes represent distinct levels of information uncertainty in X-ray images. This research contributes to the definition of pneumonia-indecisive class, which includes X-ray images with incomplete and uncertain information. By doing so, we have the opportunity to gather and analyze additional information about the patient before making a concrete decision on indecisive chest X-ray images. This would allow the system to align with manual approach of the medical expert to deal with doubtful cases of pneumonia and reduce the expected cost by avoiding to make incorrect decision on them.

- Defining GTRS related artifacts in the context of three-way pneumonia classification system:
  We determine the final threshold pair based on game theory in rough sets for three-way classification. To do so, we need to define a set of elements for the game. A two-player game is developed between accuracy and coverage, in which their strategies are to increase or decrease the value of test-treatment and testing thresholds. Payoffs of the player are defined by accuracy and coverage itself. Stop criteria are defined based on players, accuracy and coverage, reaching some specified payoffs.

- Applying GTRS for three-way pneumonia classification system:
  Accuracy and coverage are two contradictory evaluation measures. An increasing accuracy score reduces the coverage score, and vice versa. Additionally, this thesis demonstrates how different thresholds affect accuracy and coverage. Thus, GTRS is used to formulate trade-offs between accuracy and coverage. By repeatedly modifying thresholds in the game, it gradually reaches an optimal threshold pair. Test results suggest that GTRS improves the classification

performance of the three-way pneumonia classification system.

- Approaching Pareto optimality strategy for obtaining game solution:
  The payoff of accuracy is affected when coverage plays a strategy to maximize its payoff, and vice versa. To establish a balanced trade-off, we approach both Nash equilibrium and Pareto optimal strategies. Nash equilibrium achieves low coverage while focusing on high accuracy. In order to have a balanced trade-off between evaluation measures, this study applies the Pareto optimality strategy to identify an optimal solution (i.e., a pair of test-treatment and testing thresholds). Thus, this research uses a cooperative game structure rather than a non-cooperative game structure. When compared to the Nash Equilibrium strategy, Pareto optimality tends to establish a more favorable compromise between accuracy and coverage. This contributes to the application of Pareto optimality for obtaining the optimal threshold pair.

5.2 Future Research

This section indicates the possible research work as follows,

- As mentioned earlier in Section 5.1, in order to achieve a balanced trade-off between accuracy and coverage, this research employs cooperative game structure and obtains the game solution by Pareto optimality strategy. Other types of game, such as sequential game [23], combinatorial game [8], and differential game [9] can also be applied to obtain an optimal threshold pair. We can also approach decision-theoretic rough sets that aims to find the threshold pair by minimizing the expected cost.

- The pneumonia classification system based on GTRS obtains a higher true positive rate but a lower true negative rate. Future work could be to examine a

competitive game between true positive rate and true negative rate to improve precision and accuracy of the model.

- The proposed classification model is trained on a dataset of pneumonia-positive and normal chest X-ray images, but it is also evaluated on a completely different test set that combines pneumonia-positive with other disease-infected chest X-ray images to check the usefulness of the proposed model in identifying different diseases. Although the model provides a satisfying true positive rate, it compromises the balanced accuracy score on this test set due to similarity in extracted features of all types of X-ray images. Future exploration could be to experiment with a different set of features, such as histogram of oriented gradient descriptor.

- The approach to implement the decision table in this study can also be used to avoid minor observation errors in other data mining and machine learning applications.

References

[1] H. Abdi and L. J. Williams. 2010. Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(4), 433–459.

[2] I. Abraham, L. Alvisi, and J. Y. Halpern. 2011. Distributed computing meets game theory: combining insights from two fields. *ACM SIGACT News*, 42(2), 69–76.

[3] N. Azam. 2014. *Investigating Decision Making with Game-theoretic Rough Sets*. PhD thesis. University of Regina, SK, CAN.

[4] N. Azam and J. T. Yao. 2014. Analyzing uncertainties of probabilistic rough set regions with game-theoretic rough sets. *International Journal of Approximate Reasoning*, 55(1), 142–155.

[5] N. Azam and J. T. Yao. 2014. Game-theoretic rough sets for recommender systems. *Knowledge-Based Systems*, 72, 96–107.

[6] M. Bell. 2003. The use of game theory to measure the vulnerability of stochastic networks. *IEEE Transactions on Reliability*, 52(1), 63–68.

[7] M. J. Fischer and R. N. Wright. 1993. An application of game theoretic techniques to cryptography. *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, 13, 99–118.

[8] A. S. Fraenkel. 2004. Complexity, appeal and challenges of combinatorial games. *Theoretical Computer Science*, 313(3), 393–415.

[9] A. Friedman. 1971. *Differential Games*. John Wiley & Sons, New York, USA.

[10] D. Fudenberg and J. Tirole. 1989. Noncooperative game theory for industrial organization: an introduction and overview. *Handbook of Industrial Organization*, 1, 259–327.

[11] D. Fudenberg and J. Tirole. 1991. *Game Theory*. MIT Press, Cambridge, MA, USA.

[12] V. Garima and P. Shiva. 2020. Pneumonia classification using deep learning in healthcare. *International Journal of Innovative Technology and Exploring Engineering*, 9(4), 1715–1723.

[13] S. Greco, B. Matarazzo, and R. Słowiński. 2008. Parameterized rough set model using rough membership and bayesian confirmation measures. *International Journal of Approximate Reasoning*, 49(2), 285–300.

[14] S. Gurmail and K.-C. Yow. 2021. Object or background: an interpretable deep learning model for covid-19 detection from ct-scan images. *Diagnostics ((Basel, Switzerland)*, 11(9), 1732.

[15] B. Halalli and A. Makandar. 2018. Computer aided diagnosis-medical image analysis techniques. *Breast Imaging*, 85, 85–109.

[16] P. Hammerstein and R. Selten. 1994. Game theory and evolutionary biology. *Handbook of Game Theory with Economic Applications*, 2, 929–993.

[17] M. Hassaballah, H. A. Alshazly, and A. A. Ali. 2019. Analysis and evaluation of keypoint descriptors for image matching. In *Recent Advances in Computer Vision: Theories and Applications*. Cham, 113–140.

[18] J. Herbert and J. T. Yao. 2005. A game-theoretic approach to competitive learning in self-organizing maps. In *Proceedings of the 1st International Conference on Natural Computation*, LNCS 3610, 129–138.

[19] J. P. Herbert and J. T. Yao. 2008. Game-theoretic risk analysis in decision-theoretic rough sets. In *Proceedings of the 3rd International Conference on Rough Sets and Knowledge Technology*, LNCS 5009, 132–139.

[20] J. P. Herbert and J. T. Yao. 2011. Analysis of data-driven parameters in game-theoretic rough sets. In *Proceedings of the 6th International Conference on Rough Sets and Knowledge Technology*, LNCS 6954, 447–456.

[21] J. P. Herbert and J. T. Yao. 2011. Game-theoretic rough sets. *Fundamenta Informaticae*, 108(4), 267–286.

[22] A. U. Ibrahim, M. Ozsoz, S. Serte, F. Al-Turjman, and P. S. Yakoi. 2021. Pneumonia classification using deep learning from chest x-ray images during covid-19. *Cognitive Computation*, 1–13.

[23] J. R. Kale and T. H. Noe. 1990. Risky debt maturity choice in a sequential game equilibrium. *Journal of Financial Research*, 13(2), 155–166.

[24] S. Kannan and M. Sathik. 2011. Detection of pneumonia in chest x-ray images. *Journal of X-ray science and technology*, 19, (January 2011), 423–8. DOI: 10.3233/XST-2011-0304.

[25] S. Ke-Chen, Y. Yun-Hui, C. Wen-Hui, and X. Zhang. 2013. Research and perspective on local binary pattern. *Acta Automatica Sinica*, 39, 6, 730–744.

[26] D. S. Kermany, M. Goldbaum, W. Cai, C. C. Valentim, H. Liang, S. L. Baxter, A. McKeown, G. Yang, X. Wu, F. Yan, et al. 2018. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*, 172(5), 1122–1131.

[27] S. Kim. 2014. *Game theory applications in network design*. IGI Global, Hershey, PA, USA.

[28] S. A. Lashari and R. Ibrahim. 2013. A framework for medical images classification using soft set. *Procedia Technology*, 11, 548–556.

[29] S. Leutenegger, M. Chli, and R. Y. Siegwart. 2011. Brisk: binary robust invariant scalable keypoints. In *Proceedings of the 2011 International conference on computer vision*, 2548–2555.

[30] K. Leyton-Brown and Y. Shoham. 2008. Essentials of game theory: a concise multidisciplinary introduction. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 2(1), 1–88.

[31] D. Li, Q. Xu, W. Yu, and B. Wang. 2020. Srp-akaze: an improved accelerated kaze algorithm based on the sparse random projection. *IET Computer Vision*, 14.

[32] J. Li and N. M. Allinson. 2008. A comprehensive review of current local features for computer vision. *Neurocomputing*, 71(10), 1771–1787.

[33] S. P. Mishra, U. Sarkar, S. Taraphder, S. Datta, D Swain, R. Saikhom, S. Panda, and M. Laishram. 2017. Multivariate statistical data analysis-principal component analysis. *International Journal of Livestock Research*, 7(5), 60–78.

[34] H. Mohsen, E.-S. A. El-Dahshan, E.-S. M. El-Horbaty, and A.-B. M. Salem. 2018. Classification using deep learning neural networks for brain tumors. *Future Computing and Informatics Journal*, 3(1), 68–71.

[35] T. Navamani. [n. d.] Chapter 7 - efficient deep learning approaches for health informatics. In *Deep Learning and Parallel Computing Environment for Bioengineering Systems*. Academic Press, 123–137.

[36] R. Olli, D. R. Murdoch, E. Lahti, et al. 2011. Viral pneumonia. *Lancet*, 377, 1264–1275.

[37] OpenCV. 2015. Changing the contrast and brightness of an image. Retrieved 01/02/2022 from `https://docs.opencv.org/4.x/d3/dc1/tutorial_basic_linear_transform.html`.

[38] N. O'Mahony, S. Campbell, A. Carvalho, S. Harapanahalli, G. V. Hernandez, L. Krpalkova, D. Riordan, and J. Walsh. 2019. Deep learning vs. traditional computer vision. In *Proceedings of 2019 Computer Vision Conference*, 128–144.

[39] M. Partio, B. Cramariuc, M. Gabbouj, and A. Visa. 2002. Rock texture retrieval using gray level co-occurrence matrix. In *Proceedings of the 5th Nordic Signal Processing Symposium*. Volume 75.

[40] S. G. Pauker and J. P. Kassirer. 1980. The threshold approach to clinical decision making. *New England Journal of Medicine*, 302, 20, 1109–1117.

[41] Z. Pawlak. 1991. *Rough sets: Theoretical aspects of reasoning about data*. Volume 9. Springer Science & Business Media.

[42] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. 2011. Scikit-learn: machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.

[43] N. Petrick, B. Sahiner, S. G. Armato III, A. Bert, L. Correale, S. Delsanto, M. T. Freedman, D. Fryd, D. Gur, L. Hadjiiski, et al. 2013. Evaluation of computer-aided detection and diagnosis systems. *Medical Physics*, 40(8), 087001.

[44] L. Račić, T. Popović, S. Šandi, et al. 2021. Pneumonia detection using deep learning based on convolutional neural network. In *Proceedings of the 25th International Conference on Information Technology*, 1–4.

[45] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. Langlotz, K. Shpanskaya, et al. 2017. Chexnet: radiologist-level pneumonia detection on chest x-rays with deep learning.

[46] S. Ribeiro and J. T. Yao. 2019. Toward a three-way image classification model: a case study on corn grain images. In *Proceedings of the 21st IEEE International Symposium on Multimedia*, 177–1776.

[47] C. J. Saul, D. Y. Urey, and C. D. Taktakoglu. 2019. Early diagnosis of pneumonia with deep learning.

[48] R. E. Schapire. 2003. The boosting approach to machine learning: an overview. *Nonlinear Estimation and Classification*, 149–171.

[49] D. Slezak and W. Ziarko. 2005. The investigation of the bayesian rough set model. *International Journal of Approximate Reasoning*, 40(2), 81–91.

[50] O. Stephen, M. Sain, U. J. Maduh, and D.-U. Jeong. 2019. An efficient deep learning approach to pneumonia classification in healthcare. *Journal of Healthcare Engineering*, 2019, 1–7.

[51] L. Suganthi, K. Nirmala, S. Deepa, K. Nagalakshmi, and M. Santhya. 2021. Detection of pneumonia clouds from chest x-ray images. In *Proceedings of 2021 Data Intelligence and Cognitive Informatics*, 819–829.

[52] C. Tannert, H.-D. Elvers, and B. Jandrig. 2007. The ethics of uncertainty: in the light of possible dangers, research becomes a moral duty. *EMBO Reports*, 8(10), 892–896.

[53] V. Thambusamy and S. Perumal. 2018. Preprocessing by contrast enhancement techniques for medical images. *International Journal of Pure and Applied Mathematics*, 118(18), 3681–3688.

[54] D. van Ravenzwaaij and J. Ioannidis. 2019. True and false positive rates for different criteria of evaluating statistical evidence from clinical trials. *BMC Medical Research Methodology*, 19(1), 218–225.

[55] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers. 2017. Chestx-ray8: hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. *2017 IEEE Conference on Computer Vision and Pattern Recognition*, 3462–3471.

[56] S. Wold, K. Esbensen, and P. Geladi. 1987. Principal component analysis. *Chemometrics and intelligent laboratory systems*, 2(3), 37–52.

[57] X. P. Yang and J. T. Yao. 2010. A multi-agent decision-theoretic rough set model. In *Proceedings of the 5th International Conference on Rough Sets and Knowledge Technology*, LNCS 6401, 711–718.

[58] A. C.-C. Yao. 1977. Probabilistic computations: toward a unified measure of complexity. In *Proceedings of the 18th Annual Symposium on Foundations of Computer Science*, 222–227.

[59] J. T. Yao and N. Azam. 2015. Web-based medical decision support systems for three-way medical decision making with game-theoretic rough sets. *IEEE Transactions on Fuzzy Systems*, 23(1), 3–15.

[60] J. T. Yao and J. P. Herbert. 2008. A game-theoretic perspective on rough set analysis. *Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition)*, 20(3), 291–298.

[61] L. Yao, E. Poblenz, D. Dagunts, B. Covington, D. Bernard, and K. Lyman. 2017. Learning to diagnose from scratch by exploiting dependencies among labels.

[62] Y. Y. Yao. 2007. Decision-theoretic rough set models. In *Proceedings of the 2nd International Conference on Rough Sets and Knowledge Technology*, LNCS 4481, 1–12.

[63] Y. Y. Yao. 2008. Probabilistic rough set approximations. *International Journal of Approximate Reasoning*, 49(2), 255–271.

[64] Y. Y. Yao. 2010. Three-way decisions with probabilistic rough sets. *Information Sciences*, 180(3), 341–353.

[65] Y. Y. Yao. 2011. The superiority of three-way decisions in probabilistic rough set models. *Information Sciences*, 181(6), 1080–1096.

[66] Y. Y. Yao. 2011. Two semantic issues in a probabilistic rough set model. *Fundamenta Informaticae*, 108(4), 249–265.

[67] Y. Y. Yao. 2012. An outline of a theory of three-way decisions. In *Proceedings of the 8th International Conference on Rough Sets and Current Trends in Computing*, LNCS 5009, 1–17.

[68] Y. Y. Yao. 2015. Rough sets and three-way decisions. In *Proceedings of the 10th International Conference on Rough Sets and Knowledge Technology*, LNCS 9436, 62–73.

[69] Y. Y. Yao. 2016. Three-way decisions and cognitive computing. *Cognitive Computation*, 8(4), 543–554.

[70] Y. Y. Yao. 2018. Three-way decision and granular computing. *International Journal of Approximate Reasoning*, 103, 107–123.

[71] Y. Y. Yao and X. F. Deng. 2014. Quantitative rough sets based on subsethood measures. *Information Sciences*, 267, 306–322.

[72] Y. Y. Yao, S. Greco, and R. Słowiński. 2015. Probabilistic rough sets. In *Springer Handbook of Computational Intelligence*, 387–411.

[73] Y. Y. Yao and S. K. M. Wong. 1992. A decision theoretic framework for approximating concepts. *International Journal of Man-Machine Studies*, 37(6), 793–809.

[74] Y. Zhang, P. F. Liu, and J. T. Yao. 2019. Three-way email spam filtering with game-theoretic rough sets. In *Proceedings of the 8th International Conference on Computing, Networking and Communications*, 552–556.

[75] Y. Zhang and J. T. Yao. 2012. Rule measures tradeoff using game-theoretic rough sets. In *Proceedings of the 5th International Conference on Brain Informatics*, LNCS 7670, 348–359.

[76] Y. Zhang and J. T. Yao. 2017. Multi-criteria based three-way classifications with game-theoretic rough sets. In *Proceedings of the 23rd International Symposium on Methodologies for Intelligent Systems*, 550–559.

[77] B. Zhou. 2012. *A Cost-Sensitive Approach to Ternary Classification*. PhD thesis. University of Regina, SK, CAN.

[78] B. Zhou, Y. Y. Yao, and J. Luo. 2010. A three-way decision approach to email spam filtering. In *Proceedings of the 23rd Canadian Conference on Artificial Intelligence*, LNCS 6085, 28–39.

[79] W. Zhu, N. Zeng, and N. Wang. 2010. Sensitivity, specificity, accuracy, associated confidence interval and roc analysis with practical sas implementations. In *Proceedings of the 23rd NorthEast SAS Users Group, Health Care and Life Sciences*, 1–9.

[80] W. Ziarko. 2008. Probabilistic approach to rough sets. *International Journal of Approximate Reasoning*, 49(2), 272–284.

Appendix A

Implementation

A.1 Python Code Snippet

A.1.1 Feature Extraction

```
from skimage import exposure
from skimage import util
classes = ['NORMAL', 'PNEUMONIA']
skip = int(0)
def get_data(data_dir):
    dataset = []
    skip = int(0)
    for label in classes:
        path = os.path.join(data_dir, label)
        class_label = classes.index(label)
        for file in os.listdir(path):
            try:
                filename = os.path.basename(file)
                filenames = re.split('\.',filename)
                img = cv2.imread(os.path.join(path, file),0)
                img = cv2.resize(img, (150, 150), interpolation = cv2.INTER_AREA)
                img = cv2.bilateralFilter(img,10,75,75)
                img = exposure.equalize_hist(img)
                img = util.img_as_ubyte(img)
                brisk = cv2.BRISK_create()
                orb = cv2.ORB_create()
                akaze = cv2.AKAZE_create()
                kaze = cv2.KAZE_create()
                kp, des = orb.detectAndCompute(img,None)
                kp1, des1 = brisk.detectAndCompute(img, None)
                kp2, des2 = akaze.detectAndCompute(img, None)
                kp3, des3 = kaze.detectAndCompute(img, None)
                dataset.append([img, len(kp),len(kp1), len(kp2), filenames[0], class_label, len(kp3)])
```

Figure A.1: Keypoint descriptor feature extraction

```
from skimage import util
import cv2
csv_filename = "DatasetAugmented1.csv"
csv_file = os.path.normpath(os.path.join(os.getcwd(), csv_filename))
#main_folder = r'C:\Users\subys\CTS2R\Data\Raw\Frames'
print(path)
#os.chdir(path)
with open(csv_file, mode='a', newline="") as csv_f:
    fields = ['IMG_NAME', 'LBP_ENERGY', 'LBP_ENTROPY', 'CONTRAST', 'DISSIMILARITY', 'HOMOGENEITY', 'ENERGY', 'CORRELATION', 'C
    writer = csv.writer(csv_f, delimiter=',')
    writer.writerow(fields)
    for img in prepro_Data:
        features = []
        try:
            img_arr = img[0]
            img_arr = util.img_as_ubyte(img_arr)
            #img = cv2.imread(img_arr,0)
            #img_arr = crop_img(np.array(img_arr), 0.75)
            feat_lbp = local_binary_pattern(img_arr,8,1,'uniform') #Radius = 1, No. of neighbours = 8
            feat_lbp = np.uint8((feat_lbp/feat_lbp.max())*255) #Converting to unit8
            lbp_hist,_ = np.histogram(feat_lbp,8)
            lbp_hist = np.array(lbp_hist,dtype=float)
            lbp_prob = np.divide(lbp_hist,np.sum(lbp_hist))
            lbp_energy = np.sum(lbp_prob**2)
            lbp_entropy = -np.sum(np.multiply(lbp_prob,np.log2(lbp_prob)))
            # Finding GLCM features from co-occurance matrix)
            distance = [1, 2, 3, 4]
            angles = [0, np.pi/4, np.pi/2, 3*np.pi/4]
            gCoMat = greycomatrix(np.array(img_arr), distances = distance, angles = angles, symmetric = True, normed = True)
            #gCoMat = greycomatrix(np.array(img_arr), [2], [0],256,symmetric=True, normed=True) # Co-occurance matrix
            contrast = greycoprops(gCoMat, prop='contrast')
            dissimilarity = greycoprops(gCoMat, prop='dissimilarity')
            homogeneity = greycoprops(gCoMat, prop='homogeneity')
            energy = greycoprops(gCoMat, prop='energy')
            correlation = greycoprops(gCoMat, prop='correlation')
            moment = cv2.moments(np.array(img_arr))
            #print(moment['m00'])
            hu = cv2.HuMoments(moment)
```

Figure A.2: GLCM feature extraction

A.1.2 Feature Selection

```
import numpy as np
from sklearn.feature_selection import chi2
np.set_printoptions(precision=2)
X = tt
y = df['DETECT']
chi_scores = chi2(X,y)
print(type(chi_scores))
print(chi_scores) #here first array represents chi square
for i in chi_scores:
    print(i)
    for j in i:
        print("{:.5f}".format(j))
```

Figure A.3: Univariate feature selection

A.1.3 Decision Table Implementation

```
for z in range(10):
    z=z+1
    print(z)
    train_file = "TRAINFILE_"+str(z)+".csv"
    test_file = "TESTFILE_"+str(z)+".csv"
    df_train = pd.read_csv(train_file)
    print(len(df_train))
    df_test = pd.read_csv(test_file)
    print(len(df_test))
    subset1 = df_train.filter(['IMG_NAME','ORB','BRISK','KAZE','AKAZE','CONTRAST', 'DISSIMILARITY'
    print(len(subset1))
    subset1 = subset1[:10368]
    print(len(subset1))
    ln = 6
    subset_train = subset1.sort_values(by = ['ORB'])
    subset_train['ORB_L1'] = 0
    subset1_list = np.vsplit(subset_train, ln)
    print(len(subset1_list))
    subset_train['ORB_L1'] = 0
    k = 1
    for row in range(len(subset_train)):
        i = 1
        for df1 in subset1_list:
            df1['ORB_L1'] = i
            min1 = df1['ORB'].min()
            max1 = df1['ORB'].max()
            if k == 1:
                print("ORB", df1['ORB'].min(), df1['ORB'].max())
            #print("ORB", df1['ORB'].min(), df1['ORB'].max())
            ##print(df1.head())
            if (subset_train.loc[row, 'ORB'] >= min1) & (subset_train.loc[row, 'ORB'] < max1) :
                subset_train.loc[row, 'ORB_L1'] = i
            if (i==len(subset1_list)) &  (subset_train.loc[row, 'ORB'] >= min1):
                subset_train.loc[row, 'ORB_L1'] = i
            i += 1
        k = 2
```

Figure A.4: Binning process for ORB

```
subset_train = subset_train.sort_values(by = ['AKAZE'])
subset_train['AKAZE_L1'] = 0
subset1_list = np.vsplit(subset_train, ln)
k = 1
subset_train['AKAZE_L1'] = 0
for row in range(len(subset_train)):
    i = 1
    for df1 in subset1_list:
        df1['AKAZE_L1'] = i
        min1 = df1['AKAZE'].min()
        max1 = df1['AKAZE'].max()
        if k == 1:
            print("AKAZE", df1['AKAZE'].min(), df1['AKAZE'].max())
        #print("AKAZE", df1['AKAZE'].min(), df1['AKAZE'].max() )
        #print(df1.head())
        if (subset_train.loc[row, 'AKAZE'] >= min1) & (subset_train.loc[row, 'AKAZE'] < max1) :
            subset_train.loc[row, 'AKAZE_L1'] = i
        if (i==len(subset1_list)) &  (subset_train.loc[row, 'AKAZE'] >= min1):
            subset_train.loc[row, 'AKAZE_L1'] = i
        i += 1
    k=2


subset_train = subset_train.sort_values(by = ['BRISK'])
subset_train['BRISK_L1'] = 0
subset1_list = np.vsplit(subset_train, ln)

subset_train['BRISK_L1'] = 0
k = 1
for row in range(len(subset_train)):
    i = 1
    for df1 in subset1_list:
        df1['BRISK_L1'] = i
        min1 = df1['BRISK'].min()
        max1 = df1['BRISK'].max()
        #print("BRISK", df1['BRISK'].min(), df1['BRISK'].max() )
        #print(df1.head())
        if (subset_train.loc[row, 'BRISK'] >= min1) & (subset_train.loc[row, 'BRISK'] < max1) :
            subset_train.loc[row, 'BRISK_L1'] = i
```

Figure A.5: Binning process for AKAZE and BRISK

A.1.4 Equivalence Class and Conditional Probability

```
grouped_df = subset_new.groupby(['AKAZE_L1', 'ORB_L1','BRISK_L1'])
print ("Total groups = ", len(grouped_df))
cnt = 0
i = 0
probdict ={}
#print(type(grouped_df))
#print(type(subset_new), "\n")
yes_cnt = 0
no_cnt = 0
for key,item in grouped_df:
    a_group = grouped_df.get_group(key)
    cnt += len(a_group)
    probdict[i] = {}
    probdict[i]["df"] = a_group
    yes_cnt = 0
    no_cnt = 0
    for index, row in a_group.iterrows():
        #print(row['DETECT'])
        if row['DETECT'] == 1:
            yes_cnt += 1
        else :
            no_cnt += 1

    probdict[i]["yes_cnt"] = yes_cnt
    probdict[i]["no_cnt"] = no_cnt
    probdict[i]["total"] = yes_cnt + no_cnt
    probdict[i]["prob"] = yes_cnt/len(a_group)
```

Figure A.6: Equivalence class and conditional probability

A.1.5 Game-theoretic Rough Sets

```
a = 1
b = 0.5
fields = ['a_thrshld', 'b_thrshld', 'accuracy', 'coverage', 'prof', 'boundary']
a_strategy = [a, a-0.01, a-0.02, a-0.03, a-0.04]
b_startegy = [b, b-0.02, b-0.04, b-0.06, b-0.08]
gt_df = pd.DataFrame(columns=fields)
y_cnt = 0  ## accept count
n_cnt = 0  ## reject count
b_cnt = 0  ## boundary count
c_cnt = 0  ## correct decision
w_cnt = 0  ## wrong decision
i_cnt = 0  ## immediate decision
for j in b_startegy:
    b = j
    #print(b)

    if b > 1 or b < 0:
        continue
    y_cnt = 0  ## accept count
    n_cnt = 0  ## reject count
    b_cnt = 0  ## boundary count
    c_cnt = 0  ## correct decision
    w_cnt = 0  ## wrong decision
    i_cnt = 0  ## immediate decision
    for k in a_strategy:
        y_cnt = 0  ## accept count
        n_cnt = 0  ## reject count
        b_cnt = 0  ## boundary count
        c_cnt = 0  ## correct decision
        w_cnt = 0  ## wrong decision
        i_cnt = 0  ## immediate decision
        a = k
```

Figure A.7: Building game-theoretic rough sets (1/2)

```
if a > 1 or a < 0:
    continue
for d_id, d_info in probdict.items():
    if d_info['prob'] >= a:
        #print("accept")
        y_cnt += d_info['yes_cnt']
        c_cnt += d_info['yes_cnt']
        w_cnt += d_info['no_cnt']
        i_cnt += d_info['total']
        d_info['pred'] = 1
    elif d_info['prob'] <= b:
        #print("reject")
        n_cnt += d_info['no_cnt']
        c_cnt += d_info['no_cnt']
        w_cnt += d_info['yes_cnt']
        i_cnt += d_info['total']
        d_info['pred'] = 0
    elif d_info['prob'] > b and d_info['prob'] < a:
        #print("Can not decide")
        #print( d_info['total'])
        b_cnt += d_info['total']
        d_info['pred'] = 2

coverage = (trainlen - b_cnt)/trainlen
accuracy = (y_cnt + n_cnt)/i_cnt
gt_df = gt_df.append(pd.Series(["{:.4f}".format(a), "{:.4f}".format(b), "{:.4f}".format(accuracy), "{:.4f}".format(coverage
```

Figure A.8: Building game-theoretic rough sets (2/2)

A.1.6 Pareto Optimal Strategy

```
#gt_df_desc = gt_df.sort_values(by=['prof'], ascending=False)
pre_acc = gt_df.loc[0, "accuracy"]
pre_cov = gt_df.loc[0, "coverage"]
par_index = 0
for i in range(len(gt_df)):
    acc = gt_df.loc[i, "accuracy"]
    cov = gt_df.loc[i, "coverage"]
    if acc >= pre_acc and cov >= pre_cov:
        pre_cov = cov
        pre_acc = acc
        par_index = i

print("Coverage", ("{:.2f}".format(100 * float(pre_cov))))
print("Accuracy", ("{:.2f}".format(100 * float(pre_acc))))
print("boundary", gt_df.loc[par_index,"boundary"])
a = float(gt_df.loc[par_index,"a_thrshld"])
b = float(gt_df.loc[par_index,"b_thrshld"])
print("a = ", a, "b =", b)
```

Figure A.9: Pareto Optimality Strategy

A.1.7 Three Regions

```
a=0.88
b=0.22
print("a = ", a, "b =", b)
pd.set_option('expand_frame_repr', False)
for d_id, d_info in probdict.items():
    #print("\n\nID:", d_id)
    before = d_info['pred']
    if d_info['prob'] >= a:
        d_info['pred'] = 1
        name = "accept_"+str(d_id)+"_"+str(d_info['yes_cnt'])+"_"+str(d_info['no_cnt'])+"_"+str(d_info['total'])+".csv"
        #if d_id == 42 or d_id == 21:
        #d_info['df'].to_csv (name, index = False, header=True)
    elif d_info['prob'] <= b:
        d_info['pred'] = 0
        name = "reject_"+str(d_id)+"_"+str(d_info['yes_cnt'])+"_"+str(d_info['no_cnt'])+"_"+str(d_info['total'])+".csv"
        #if d_id == 109 or d_id == 111:
        #d_info['df'].to_csv (name, index = False, header=True)
    else:
        d_info['pred'] = 2
        name = "group_"+str(d_id)+"_"+str(d_info['yes_cnt'])+"_"+str(d_info['no_cnt'])+"_"+str(d_info['total'])+"_"+str(d_info["pro
        #d_info['df'].to_csv (name, index = False, header=True)
    after = d_info['pred']
```

Figure A.10: Building three regions

A.1.8 Testing

```
for row in range(len(sub_test)):
    if sub_test.loc[row, 'PRED'] != sub_test.loc[row, 'DETECT']:
        w_cnt += 1
        if sub_test.loc[row, 'PRED'] == 1:
            w_n_cnt += 1
            #print ("w_n_cnt", sub_test.loc[row, 'IMG_NAME'], "d_id", s
            no_wrong.append(sub_test.loc[row, 'IMG_NAME'])
        elif sub_test.loc[row, 'PRED'] == 0:
            w_r_cnt += 1
            #print ("w_r_cnt", sub_test.loc[row, 'IMG_NAME'], "d_id", s
            yes_wrong.append(sub_test.loc[row, 'IMG_NAME'])
        if sub_test.loc[row, 'PRED'] == 2:
            w_b_cnt += 1
            bon_re.append(sub_test.loc[row, 'IMG_NAME'])
            #print ("w_b_cnt", sub_test.loc[row, 'IMG_NAME'], "d_id", s
    else:
        r_cnt += 1
        if sub_test.loc[row, 'PRED'] == 1:
            y_cnt += 1
        elif sub_test.loc[row, 'PRED'] == 0:
            n_cnt += 1
    if sub_test.loc[row, 'PRED'] == 2:
        b_cnt += 1
```

Figure A.11: Testing

A.1.9 PCA Analysis on Second Test

```
import plotly.express as px
from sklearn.decomposition import PCA

#df = px.data.iris()
X = df_test[['ORB', 'BRISK', 'AKAZE']]

pca = PCA(n_components=3)
components = pca.fit_transform(X)

total_var = pca.explained_variance_ratio_.sum() * 100

fig = px.scatter_3d(
    components, x=0, y=1, z=2, color=df_test['DETECT'],
    title=f'Total Explained Variance: {total_var:.2f}%',
    labels={'0': 'PC 1', '1': 'PC 2', '2': 'PC 3'}
)
fig.update_layout(
    margin=dict(l=0, r=0, t=0, b=0),
    paper_bgcolor="white",
)
fig.show()
#fig.write_image("pc.png")
```

Figure A.12: PCA analysis on second test