Faculty of Cognitive Sciences and Human Development

CARDIAC ARRHYTHMIA CLASSIFICATION USING SELF ORGANIZING MAP (SOM) – BASED ENSEMBLE MODEL

Dayang Yasmin Binti Abang Abdul Wahab

Bachelor of Science with Honours
(Cognitive Science)
2015

✓

UNIVERSITI MALAYSIA SARAWAK

Grade: A-

Please tick one

Final Year Project Report ☒

Masters ☐

PhD ☐

DECLARATION OF ORIGINAL WORK

This declaration is made on the17 day of JUNE year 2015.

Student's Declaration:

I, DAYANG YASMIN BINTI ABANG ABDUL WAHAB, 34799, FACULTY OF COGNITIVE SCIENCES AND HUMAN DEVELOPMENT, hereby declare that the work entitled,CARDIAC ARRHYTHMIA CLASSIFICATION USING SELF ORGANIZING MAP (SOM) – BASED ENSEMBLE MODEL is my original work. I have not copied from any other students' work or from any other sources with the exception where due reference or acknowledgement is made explicitly in the text, nor has any part of the work been written for me by another person.

17 JUNE 2015

Dayang Yasmin Binti Abang Abdul Wahab
(34799)

Supervisor's Declaration:

I,TEH CHEE SIONG, hereby certify that the work entitled,CARDIAC ARRHYTHMIA CLASSIFICATION USING SELF ORGANIZING MAP (SOM) – BASED ENSEMBLE MODELwas prepared by the aforementioned or above mentioned student, and was submitted to the "FACULTY" as a *partial/full fulfillment for the conferment of BACHELOR OF SCIENCE WITH HONOURS (COGNITIVE SCIENCE), and the aforementioned work, to the best of my knowledge, is the said student's work

Received for examination by: _____ Date: 17 JUNE 2015
(ASSOC. PROF DR. TEH CHEE SIONG)

I declare this Project/Thesis is classified as (Please tick (√)):

☐ **CONFIDENTIAL** (Contains confidential information under the Official Secret Act 1972)*

☐ **RESTRICTED** (Contains restricted information as specified by the organisation where research was done)*

☒ **OPEN ACCESS**

I declare this Project/Thesis is to be submitted to the Centre for Academic Information Services (CAIS) and uploaded into UNIMAS Institutional Repository (UNIMAS IR) (Please tick (√)):

☐ **YES**

☐ **NO**

Validation of Project/Thesis

I herebyduly affirmed with free consent and willingnessdeclaredthat this said Project/Thesisshall be placed officially in the Centre for Academic Information Services with the abide interest and rights as follows:

- This Project/Thesis is the sole legal property of Universiti Malaysia Sarawak (UNIMAS).
- The Centre for Academic Information Services has the lawful right to make copies of theProject/Thesisfor academic and research purposes only and not for other purposes.
- The Centre for Academic Information Services has the lawful right to digitize the content to be uploaded intoLocal Content Database.
- The Centre for Academic Information Services has the lawful right to make copies of the Project/Thesis if required for use by other parties for academic purposes or by other Higher Learning Institutes.
- No dispute or any claim shall arise from the student himself / herself neither a third party on this Project/Thesis once it becomes the sole property of UNIMAS.
- This Project/Thesis or any material, data and information related to it shall not be distributed, published or disclosed to any party by the student himself/herself without first obtaining approval from UNIMAS.

Student's signature: _____ Supervisor's signature: _____

Date: 17 JUNE 2015 Date: 17 JUNE 2015

Current Address:

1351B Lorong Bayor Bukit No. 10, Tabuan Jaya, 93350 Kuching, Sarawak.

Notes: * If the Project/Thesis is **CONFIDENTIAL** or **RESTRICTED**, please attach together as annexure a letter from the organisation with the date of restriction indicated,and the reasons for the confidentiality and restriction.

CARDIAC ARRHYTHMIA CLASSIFICATION USING SELF ORGANIZING MAP (SOM) – BASED ENSEMBLE MODEL

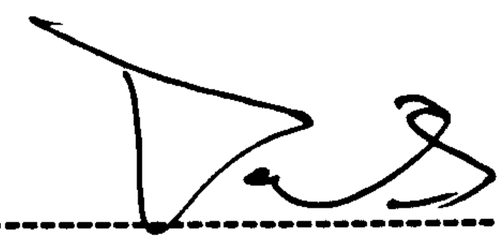
DAYANG YASMIN BT ABANG ABDUL WAHAB

This project is submitted
in partial fulfilment of the requirement for a
Bachelor of Science with Honours
(Cognitive Science)

Faculty of Cognitive Sciences and Human Development
UNIVERSITI MALAYSIA SARAWAK
(2015)

The project entitled 'Cardiac Arrhythmia Classification Using Self Organizing Map (Som) – Based Ensemble Model' was prepared by Dayang Yasmin Bt Abang Abdul Wahab and submitted to the Faculty of Cognitive Sciences and Human Development in partial fulfillment of the requirements for a Bachelor of Science with Honours (Cognitive Science)

Received for examination by:



(TEH CHEE SIONG)

Date:
8th June 2015

Grade
A-

ACKNOWLEDGMENTS

First and foremost I thank Allah SWT for giving me strength and the opportunity to see this project through.

I would like to express my utmost gratitude to my supervisor, Associate Professor Dr. Teh Chee Siong for his endless guidance, advice and patience throughout the project. Without his aid I would not have been able to complete this project. I am greatly indebted to him for everything he has done for me, for spending his precious time on me, for giving me endless advice and motivation and for having faith in me.

I would also like to thank my parents and family members for their love and support in these past few years. I am also thankful to my friends, colleagues and lecturers who have been generously supportive of me and for giving me words of encouragement when I was in a bind.

I am very grateful for everything that everyone has done for me, even the smallest gestures that has kept on pushing me forward. Without all of the support and encouragement, this project would not have seen completion. Thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Many clinical decision support systems have been using data mining techniques for prediction and diagnosis of various diseases with good accuracy. This is due to its ability to distinguish various patterns of data from its background, and make conclusions about the categories of the patterns. A large number of such systems have been widely used in the diagnosis of heart diseases. One of the heart diseases in concern is cardiac arrhythmia. Most systems used in diagnosing cardiac arrhythmia uses data mining techniques, like Artificial Neural Networks, particularly in the form of a single classifier. In this project, a Self Organizing Map (SOM) – Based Ensemble model is proposed for the classification of cardiac arrhythmia disease dataset. An ensemble is a model that applies multiple learning models and combining the outputs or predictions to solve a particular problem. An ensemble is stated to predict or classify datasets more accurately than some single classifier models. The ensemble consists of three SOM classifiers trained with different number of dimension. For the ensemble, a voting technique is used to average the prediction of each single SOM classifier to obtain the final prediction. The results displayed show that the SOM ensemble model has higher classification accuracy than that of single SOM classifiers. Ensemble learning eliminates errors of single classifiers by averaging the prediction of each classifier, thus resulting in a more accurate output.

Keywords: data mining, classification, self organizing map, ensemble model

ABSTRAK

Pada masa kini, banyak sistem sokongan keputusan klinikal yang menggunakan teknik perlombongan data untuk ramalan dan diagnosis pelbagai jenis penyakit dengan peratusan ketepatan yang tinggi. Ini adalah disebabkan oleh keupayaan teknik perlombongan data untuk membezakan pelbagai corak data dan membuat kesimpulan tentang kategori corak itu.

Sebilangan besar sistem tersebut telah digunakan secara meluas dalam diagnosis penyakit jantung. Salah satu penyakit hati dalam kebimbangan adalah aritmia jantung. Kebanyakan sistem yang digunakan dalam mendiagnosis aritmia jantung menggunakan teknik perlombongan data, seperti Rangkaian Neural Buatan, terutamanya dalam bentuk pengelas tunggal. Dalam projek ini, Self Organizing Map (SOM) – based Ensemble model dicadangkan untuk klasifikasi set data penyakit jantung aritmia. Ensemble adalah suatu model yang mengguna lebih daripada satu model pembelajaran dan menggabungkan hasil atau ramalan untuk menyelesaikan masalah tertentu. Satu ensemble dinyatakan dapat meramal atau mengelaskan set data dengan lebih tepat daripada beberapa model pengelas tunggal. Ensemble yang terdiri daripada tiga penjodoh SOM dilatih dengan dimensi yang berbeza. Untuk ensemble, teknik mengundi digunakan untuk purata ramalan setiap pengelas SOM tunggal untuk mendapatkan ramalan akhir. Keputusan yang diperolehi menunjukkan bahawa model SOM ensemble mempunyai ketepatan klasifikasi yang lebih tinggi daripada penjodoh SOM tunggal. Pembelajaran Ensemble menghapuskan kesilapan penjodoh tunggal dengan purata ramalan setiap pengelas, menghasilkan pengeluaran yang lebih tepat.

Kata kunci: data perlombongan, klasifikasi, self organizing map, model ensemble

CHAPTER ONE

INTRODUCTION

Heart disease is one of the lethal diseases affecting people around the world. The heart beats through stimulation of electric signals flowing in from the heart's "natural pacemaker" called Sino Atrial node (SA), located at the top of the right chamber or Atrium (RA) of the heart (Kohli & Verma, 2011). Any disturbance to the rhythm of the heart's pacemaker will result in cardiac arrhythmia which is a disease defined as having irregular or abnormal heartbeats (Kastor, 2002). Arrhythmia affects the body blood circulation causing an ineffective heart beat (Kohli & Verma, 2011). While most arrhythmias are harmless, there are certain life-threatening arrhythmias that can lead to stroke, heart failure or sudden cardiac arrest (SCD) (*Heart arrhythmia,* 2014).

Cardiac arrhythmia can be diagnosed by an electrocardiogram procedure. It records the timing of both atrial and ventricular electrical signals. It measures the time taken for the electrical impulses to travel through the atria, the heart's upper chambers, the atrioventricular (AV) conduction system and the ventricles, the heart's two lower, pumping chambers (*Common test for arrhythmia,* 2014). The results of the procedure are ECG signals which are comprised of P wave, QRS complex, and T wave which are represented by capital letters P, Q, R, S, and T respectively. A typical normal ECG signal is shown in figure 1. The beats of the right and left atria or upper chambers make the first wave that is the P wave which subsides into a flat line when the electrical impulse moves to the bottom chambers. The beats of the right and left bottom chambers or ventricles make the next wave which is the QRS complex. The final wave or T wave represents the return of the electrical impulses to a resting state for the ventricles. Normal resting heart rate is between 60 and 100 bpm. Any changes in the waves signify an illness of the heart (Kohli & Verma, 2011).

Cardiac arrhythmias can cause sudden, unpredictable heart attack that usually results in death. Prevention is always better than cure and this can be done using various data mining techniques. Data mining techniques have been widely applied in clinical decision support systems for prediction and diagnosis of various diseases with good accuracy. These techniques have been very effective as they are able to discover hidden patterns and relationships in complex medical data (Amin, Argawal & Beg, 2013). The application of data mining such as Artificial Neural Networks (ANN) in medicine can help diagnose or predict if a person has arrhythmia by analyzing patient's heart's electrocardiogram (ECG) signals and classify them into respective classes.

Therefore in this project, Self- Organizing Map (SOM) Based Ensemble model is proposed to identify the presence or absence of cardiac arrhythmia by analyzing a given dataset of patients and classify them into the many types of arrhythmia. It will also greatly assist doctors in analyzing complex clinical data across an extensive range of medical applications.

Background of Study

Pattern Recognition. Pattern recognition is the study of how machines can learn to discern various patterns of data from its background, and make conclusions about the categories of the patterns (Sharma & Kaur, 2013). It involves the process of taking in raw data and classifying the data to a prescribed category based on its pattern (Duda, Hart & Stork, 2001). Pattern recognition involves three main stages of processes which are data acquisition and preprocessing, feature extraction and lastly decision making. In the first stage, data are collected from the environment and are presented as input to a pattern recognition model and undergoes preprocessing in order to become readable by the model. The second stage involves the extraction of related features of the input to form the entity of object for

the purpose of classification. Lastly in the third stage, recognition or classification of the input is done based on the extracted features.

To solve a pattern recognition problem, various computational intelligence techniques can be used such as Fuzzy Logic, Artificial Neural Network (ANN), Neuro-Fuzzy system, Hybrid Intelligent system and many more. One of the popular computational intelligence techniques used for pattern recognition is the Artificial Neural Network (ANN).

Artificial Neural Network. Artificial Neural Network is a computing paradigm for information processing that consists of a large number of highly interconnected processing elements also known as neurons that work together to solve data mining problems. The neural network paradigm is inspired by the operation mechanism of the human biological nervous system, the brain (Wadhonkar, Tijare & Sawalkar, 2013).

The structure of the neural network mainly consists of three layers; input layer, one or more hidden layer and output layer, and a large number of highly interconnected processing elements or neurons at each layer. Inputs are first inserted into the input layer where they are multiplied by the connection weight values which are the strength of each signal. Then, the overall product is computed by an activation function which will then yield the output at the output layer (Wadhonkar, Tijare & Sawalkar, 2013). Artificial Neural Network learns from examples based on two learning methods which are supervised learning and unsupervised learning. The neural network learns iteratively where input data are inserted to the network one at a time and the weights are adjusted in accordance to the respective input value. The learning process which is when the weights are adjusted, are repeated over and over again until the network is able to predict the correct class output (Rani, 2011).

Self Organizing Map. Self Organizing Map is an unsupervised learning neural network that aims to transform multidimensional data in a much lower dimensional space for a more proficient visualization and representation of complex or large quantities of data

(Guthikonda, 2005). The network is made up of two spaces which are the continuous high dimensional input space containing the input nodes and the discrete low dimensional output space containing the map nodes. The map nodes are connected to the input nodes where the SOM network will attempt to map the input nodes onto the output space based on their weights which are associated with a winning neuron from the output space. This will finally result in a coordinated or ordered representation on the output space (Kutlu & Kuntalp, n.d.) and concludes the learning process of SOM.

The detailed learning process of SOM is summarized as follows:

Data is first inserted into the input space. Each input node's weights are initialized. Then, a vector is chosen at random from the set of data from the input space and presented to the output space. The winning node, commonly known as the Best Matching Unit (BMU) is identified by calculating the Euclidean distance. Then, the radius of the neighborhood of the BMU is calculated as the radius reduces after every epoch. Finally, once the BMU is found, the BMU's neighboring nodes will update their respective weights.

Ensemble Learning Model. Ensemble learning is the concept of applying multiple learning models and combining the outputs or predictions to solve a particular problem (Polikar, 2009). An ensemble model is designed to improve the performance in analysis and classification in order to provide an accurate output. There are many existing ensemble learning algorithms such as bagging, boosting, stacking, error-correcting output codes, AdaBoost and many more (Sewell, 2008). Different subsets of training data and different parameters of the classifiers can be used to train an ensemble of classifiers (Polikar, 2009).

Statement of Problem

Heart disease is one of the lethal diseases affecting people around the world. One typeof heart disease is called cardiac arrhythmia a disease defined as having irregular or abnormal heartbeats (Kastor, 2002). Arrhythmia affects the body blood circulation causing an

ineffective heart beat (Kohli & Verma, 2011). While most arrhythmias are harmless, there are certain life-threatening arrhythmias that can lead to stroke, heart failure or sudden cardiac arrest (SCD) (*Heart arrhythmia,* 2014). A variety of computational intelligence based system has been applied in cardiac arrhythmia analysis and classification. A majority as listed below uses data mining techniques, like Artificial Neural Networks, particularly in the form of a single classifier:

- Classification of Heart Disease Dataset using Multilayer Feed forward backpropogation Algorithm (Wadhonkar, Tijare & Sawalkar, 2013)
- A Fuzzy Rule Base System for the Diagnosis of Heart Disease (Barman & Choudhury, 2012)
- Genetic Neural Network Based Data Mining in Prediction of Heart Disease Using Risk Factors (Amin, Agarwal & Beg, 2013)

There has not been any proposed ensemble based system that is used to analyze and classify heart disease datasets despite the fact that an ensemble model can produce a more accurate output than some single classifiers. Ensemble learning has the advantage to reduce small sample size problem by averaging and integrating multiple classification models to reduce the potential for over fitting the training data, enabling the data to be used more efficiently (Yang, Yang, Zhou & Zomaya, 2010). Every classifier model has limitations and tends to make errors. Ensemble learning is able to handle each models strengths and weaknesses, resulting in the best possible result (Brown, 2010).

The generalization ability of an ensemble model is usually much stronger than single classifiers where by combining the predictions of multiple classifiers, a more accurate classification can be obtained (Brown, 2010). Several theoretical and empirical results have shown that the accuracy of an ensemble can significantly exceed that of a single classifier

model. Hence in this paper, an ensemble based model is proposed to analyze and classify a cardiac arrhythmia dataset.

Objective of Study

The objectives of this study are as follows:

1. To study in depth about machine learning models used in data analysis and classification.

2. To design and develop an ensemble model based on SOM classifier.

3. To investigate the accuracy of the ensemble model in analyzing and classifying cardiac arrhythmia dataset.

4. To compare the performance of the ensemble model with the performance of single classifier SOM

Scope of Research

- This research is done only within the field of heart disease; cardiac arrhythmia.
- The data used is a benchmark dataset obtained from the Cleaveland database from UCI Machine Learning Repository: Heart Disease Dataset.
- The research focuses on the use of Self Organizing Map (SOM) model and ensemble learning.

Contribution/significance of research

This study is expected to contribute towards the knowledge of the investigated proposed ensemble model, in addition to improving the performance in the analysis and classification process of a dataset.

Definition of terms

Artificial Neural Networks :

A computing paradigm for information processing that consists of a large number of highly interconnected processing elements also known as neurons that work together to solve data mining problems (Wadhonkar, Tijare & Sawalkar, 2013).

Pattern Recognition :

Pattern recognition is the study of how machines can learn to discern various patterns of data from its background, and make conclusions about the categories of the patterns (Sharma & Kaur, 2013).

Classification :

Classification one of the important methods in data mining which involves the process of predicting group membership for data instances or distinguishing data classes or concepts (Wadhonkar, Tijare & Sawalkar, 2013).

Ensemble learning :

The process by which multiple models, such as classifiers or experts, are employed and their outputs are strategically generated and combined to solve one particular problem (Sewell, 2008).

Self Organizing Map (SOM) :

An unsupervised learning neural network that transforms multidimensional data into a much lower dimensional space (Guthikonda, 2005). It is used to visualize and interpret high-dimensional data sets (Kutlu & Kuntalp, n.d.).

CHAPTER TWO

LITERATURE REVIEW

The analysis of a heart disease dataset is crucial in the medical field in order for doctors to decide the appropriate treatment and care for their patients. In the case of cardiac arrhythmia, different types of this disease have different symptoms, different effects on the patient and different treatment. Therefore, it is very beneficial for doctors and patients alike to be able to classify the possible class of cardiac arrhythmia that the patient may have.

A variety of computational intelligence techniques have been applied in heart disease dataset analysis. Many researchers have proposed works based on Neural Networks, usually single classifiers. The many related works on heart disease analysis using computational intelligence techniques are discussed in this paper.

Related Works

Analysis of Heart Diseases Dataset Using Neural Network Approach. Rani (2011) proposed a feed forward backpropagation learning algorithm neural network model to analyze heart disease dataset. Backpropogation algorithm with momentum and variable learning rate is used to train the networks. The network consists of 13 input neurons to represent 13 attributes based on the database used. The number of classes is four: 0 – normal person, 1- first stroke, 2- second stroke and 3- end of life. The output layer consists of two neurons to represent the four classes.

The backpropagation algorithm is used to train the neural network. Parallelism is implemented at each neuron in all hidden and output layers to speed up the learning process. The experimental result obtained through the neural network technique was satisfactory for the classification task.

Classification of Heart Disease Dataset using Multilayer Feed forward backpropogation Algorithm. Wadhonkar, Tijare and Sawalkar (2013) also proposed a

multilayer feed forward backpropagation algorithm to classify heart disease dataset. In this paper, the Cleaveland dataset involving 13 attributes with the addition of two other attributes, smoke and obesity, was used to carry out the classification process. The structure of the network consists mainly of three layers which are the input layer, hidden layer and the output layer. The input layer of the network consists of 15 neurons to represent each attribute as the database consists of 15 attributes. The number of classes is four: 0 – normal person, 1- first stroke, 2- second stroke and 3- end of life. The output layer consists of two neurons to represent these four classes.

In this multilayer feed forward network, neurons are arranged into the layers where neurons from each layer receive input from the previous layer and transmit it to the next layer. The outputs of these nodes are then input into another layer of nodes and so on until the output of the final layer of nodes is the output of the network. Then, the network is trained using backpropogation algorithm. The back propagation calculates the error value and passes error signals backwards through the network to update the weights of the network. The modification of the weights is to minimize the mean squared error made in the network desired and the actual target value. The classification results of the dataset showed 94% accuracy for 13 attributes and 100% accuracy for 15 attributes.

Genetic Neural Network Based Data Mining in Prediction of Heart Disease Using Risk Factors. Amin, Agarwal and Beg (2013) proposed a hybrid system involving genetic algorithm and multilayered feed forward neural networks. In this hybrid system, the genetic algorithm is first used to optimize the neural network weights. Then, the neural networks use the Levenberg-Marquardt backpropagation algorithm to train the networks using the weights optimized by the genetic algorithm. The genetic algorithm is implemented as it solves the problem of finding the globally initialized neural network weights and the slow process of network convergence faced by the backpropagation algorithm.

12 nodes representing risk factors of heart disease are used as the input data in the input layer and 2 nodes in the output layer represent the presence and absence of heart disease. The learning is fast, stable and accurate. The system was implemented in Matlab and results in 89% for accuracy on prediction and 96.3% for the accuracy on validation of data. The hybrid system proposed which involves genetic algorithm and neural network approach resulted in better average prediction accuracy than the traditional Artificial Neural Networks.

Comparative Study of Data Mining Classification Methods in Cardiovascular Disease Prediction. Kumari and Godara (2011) proposed a comparative study on data mining classification techniques RIPPER classifier, Decision Tree, Artificial neural networks (ANN), and Support Vector Machine (SVM) on the analysis of cardiovascular disease dataset. The data used was obtained from Cleveland cardiovascular disease dataset which consists of 14 attributes and 303 records. All four data mining classification techniques were applied on the dataset and their performance were assessed based on the comparison in terms of their sensitivity, specificity, accuracy, error rate, True Positive Rate (TPR) and False Positive Rate (FRP). 10-fold cross validation method was used to measure the unbiased estimate of these prediction models. Results obtained from the study are as shown in Table 1.0.

Table 1.0

Sensitivity, Specificity, Accuracy, Error rate, TPR and FPR value of different classification techniques

	Sensitivity	**Specificity**	**Accuracy**	**Error Rate**	**TPR**	**FPR**
Ripper	86.25%	75.82%	81.08%	0.2756	0.8625	0.2410
Decision Tree C4.5	83.12%	74.26%	79.05%	0.2755	0.8312	0.2573
ANN	83.75%	75.73%	80.06%	0.2248	0.8375	0.2426
SVM	90.0%	77.20%	84.12%	0.1588	0.9000	0.2279

The analysis of the results shows that the SVM predicts cardiovascular disease best with the least error rate and highest accuracy as compared to other three classification techniques.

Enhanced Prediction of Heart Disease with Feature Subset Selection using Genetic Algorithm. Anbarasi, Anupriya and Iyengar (2010) proposed an enhanced prediction of heart disease with feature subset selection using genetic algorithm with a comparative study on Naïve Bayes, Clustering, and Decision Tree methods. 13 attributes were involved but the number was reduced to 6 attributes as a result of the application of the genetic algorithm. The genetic algorithm is applied in the Feature Extraction process to detect and eliminating attributes that does not really relate or contribute more to heart disease. The Naïve Bayes, Clustering, and Decision Tree methods are used to predict the presence and absence of heart disease based on the 13 attributes and also the reduced, 6 attributes. The prediction of the presence of heart disease was found to be more accurate with reduced number of attributes. The results obtained are as shown below.

Table 2.0

Comparative table of the three classifiers

	Accuracy	**Mean Absolute Error**
Naïve Bayes	96.5%	0.044
Decision Tree	99.2%	0.00016
Clustering	88.3%	0.117

The decision tree method was concluded to have the best performance as compared to the other two methods where the Naïve Bayes method performed consistently with both 13 and 6 attributes while the clustering method performed poorly on both occasions.

A few researches done in the implementation of computational intelligence techniques in heart disease analysis were discussed. Each research proposed different application of computational intelligence techniques which results in different performances in terms of prediction and classification accuracy. Such difference in performance shows which systems can perform better in prediction and classification of heart disease dataset.

CHAPTER THREE

METHOD

Research Design

This research is a system design and development research for data mining classifications.

The data to be used is a benchmark data obtained from the Cleaveland database consisting of 13 feature attributes as input data and 5 class attribute as output data. The systems to be used to analyze the data set is the Self Organizing Map (SOM) where the outputs from the many SOM classifiers will be combined using one of the many ensemble combiner algorithm to produce the final output value.

The general overview of the procedure for this research is shown is Figure 1 below.

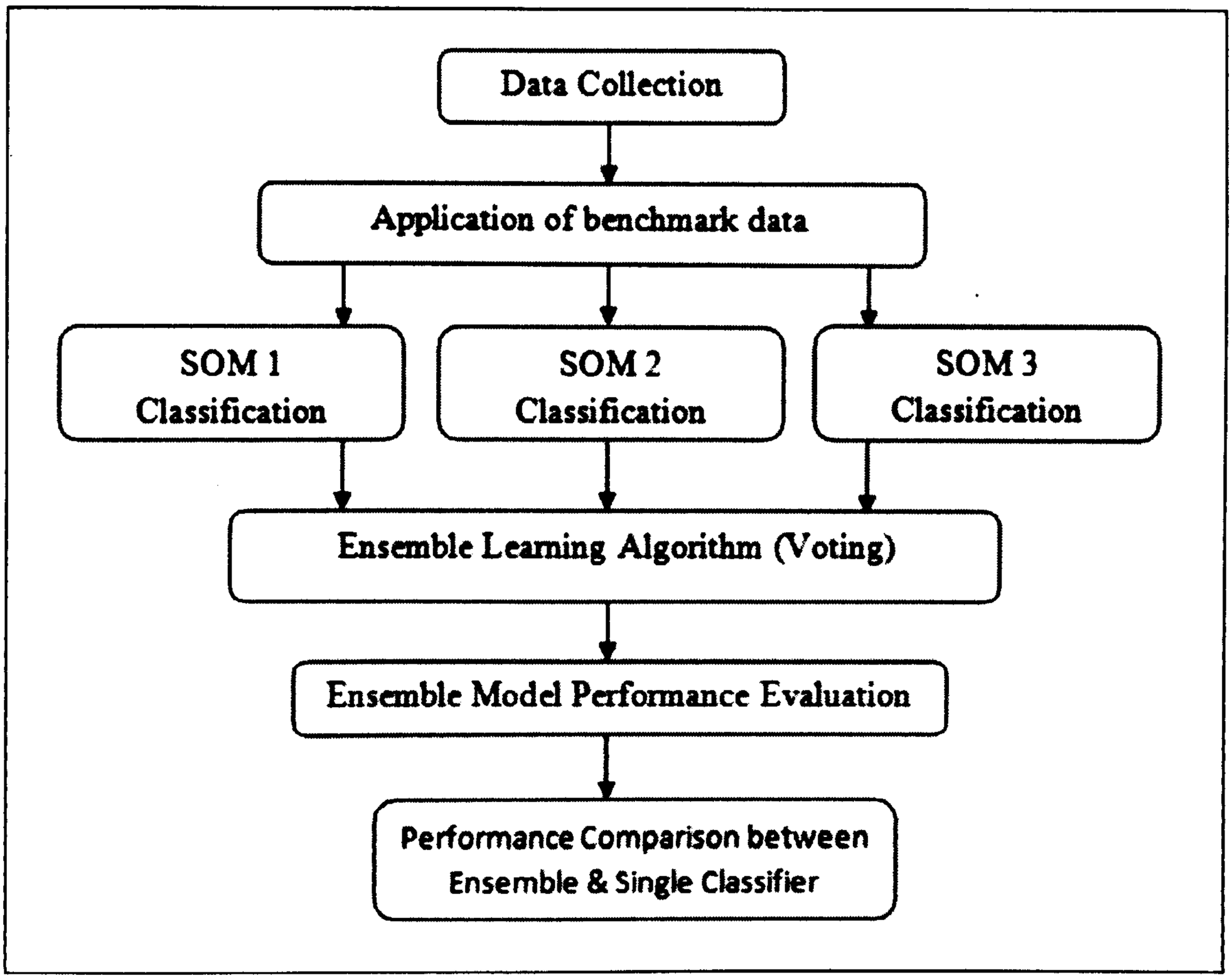


Figure 1 *Overall procedure of the research*