COMPUTATIONAL INTELLIGENCE APPROACH FOR CLASSIFICATION AND RISK QUANTIFICATION OF METABOLIC SYNDROME

HABEEBAH ADAMU KAKUDI

FACULTY OF COMPUTER SCIENCE AND INFORMATION TECHNOLOGY
UNIVERSITY OF MALAYA
KUALA LUMPUR

2019

COMPUTATIONAL INTELLIGENCE APPROACH FOR CLASSIFICATION AND RISK QUANTIFICATION OF METABOLIC SYNDROME

HABEEBAH ADAMU KAKUDI

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

FACULTY OF COMPUTER SCIENCE AND INFORMATION TECHNOLOGY
UNIVERSITY OF MALAYA
KUALA LUMPUR

2019

UNIVERSITI MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate:

Registration/Matric No.:

Name of Degree:

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):

Field of Study:

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date:

Subscribed and solemnly declared before,

Witness's Signature Date:

Name:
Designation:

COMPUTATIONAL INTELLIGENCE APPROACH FOR CLASSIFICATION AND RISK QUANTIFICATION OF METABOLIC SYNDROME

ABSTRACT

Metabolic Syndrome (MetS) is clinically defined as the presence of three out of the following five abnormalities - ihyperglycaemia, raised waist circumference, low High-Density Lipoprotein-Cholesterol, ihypertriglyceridaemia and hypertension. MetS places individuals at an unhealthy disadvantage and iis associated with an increased risk of non-communicable diseases such as cardiovascular disease and diabetes. Currently used non-clinical methods are not able to diagnose the risk of MetS in patients that fall very close to the clinically defined threshold. Therefore, the aim of this study is to propose and develop a novel non-clinical technique for the early risk quantification and classification of MetS refered to as genetically optimized Bayesian adaptive resonance theory mapping (GOBAM). Genetic Algorithm(GA) is used to optimize the order of sequence of the input sample and the parameters of the Bayesian ARTMAP (BAM). The "Cohort study on clustering of lifestyle risk factors and understanding its association with stress on health and well-being among school teachers in Malaysia" (CLUSTer) dataset was used to compare the performance of the proposed Genetically Optimised Bayesian ARTMAP (GOBAM) model and three other classic Adaptive Resonance Theory Mapping (ARTMAP) models -Genetic Algorithm Fuzzy ARTMAP (GAFAM), Fuzzy ARTMAP (FAM), and Bayesian ARTMAP (BAM). GOBAM achieved higher of area under the receiver operating curve, sensitivity, specificity, positive predictive value, negative predictive value, and Fscore performance metrics of 91.45%, 96.3% , 88.3% , 98.32% , 85.71% , and 96.41% respectively. The proposed GOBAM model was able to diagnose the risk of MetS efficiently with borderline MRF measurements, by utilising a novel risk prediction index that ranged between 0 and 1.

ABSTRAK

Pada definisi klinikal Sindrom metabolik (MetS), ia ditakrifkan tiga daripada lima keabnormalan seperti - glukosa darah terjejas (hiperglisemia), meninggikan lilitan pinggang, Ketumpatan Tinggi Lipoprotein-Kolesterol yang rendah, peningkatan trigliserida (hypertriglyceridaemia) dan tekanan darah tings (hypertension). MetS menempatkan individu pada kelemahan yang tidak sihat dan ia juga dikaitkan dengan peningkatan risiko penyakit tidak berjangkit seperti penyakit kardiovaskular and diabetes.

Tambahan pula, sukar untuk mendiagnosis individu yang hadir dengan sempadan di MetS risk factor (MRF) ukuran. Kaedah bukan klinikal semasa yang digunakan untuk mendiagnosis risiko MetS juga tidak dapat mempertimbangkan pengukuran MRF yang terlibat pada ambang klinikal yang ditetapkan. Oleh itu, matlamat kajian ini adalah untuk mencadangkan dan membangunkan ramalan risiko dan kuantifikasi model untuk kuantiti risiko awal dan ramalan MetS yang menggunakan semua lima keabnormalan factor MRF. Model yang dicadangkan adalah pemetaan teori resonans adaptif Bayesian genetik (GOBAM). Algoritma Genetik digunakan untuk mengoptimumkan susunan urutan sampel input dan parameter daripada Bayesian ARTMAP (BAM). Model ini dinilai dengan menggunakan kajian "Cohort pada clustering yang mengenai faktor risiko gaya hidup dan memahami hubungannya dengan tekanan kesihatan dan kesejahteraan di kalangan guru sekolah di Malaysia" (CLUSTer). Model GOBAM yang dicadangkan dapat mendiagnosis risiko sindrom metabolik dengan berkesan dengan pengukuran MRF. Model telah menggunakan kebarangkalian posterior yang berkisar antara 0 dan 1.

ACKNOWLEDGEMENTS

I would like express my profound gratitude to the Almighty Allah, The Greatest, The Most Beneficent, The Most Merciful and The Giver of all Bounties for giving me the wisdom, resolve and patience towards writing this research.

I would like to express my appreciation to my three supervisors, Prof. Dr. Loo Chu Kiong, Assoc. Prof. Foong Ming Moy, and Dr. Lim Chee Kau. I will always be grateful for your guidance and patience towards me. Thank you for taking your time and effort to get me here. Thank you for your constant positivity and belief in me. This is our beginning.

To my father, you started all of this. You are always in my heart and prayers. I wish you were here to see this.

To my mother, you continued where he left off and you are my pillar. May Allah increase you with health and grant you a very long life.

To loves of my life, my three boys, my dearest husband (I stand tall because you are by my side, always cheering me on...) and my two fine sons (Keep soaring, my loves), I could not have done it without your support and love. You mean the world to me.

To my siblings, your prayers and words of encouragement kept me going when things were really rough. You guys are the best and will always be. I love you all.

To all my lab mates, it was a real pleasure knowing and working with all of you. Let's keep in touch.

To all my colleagues and friends who were patient with me, prayed for me and wished me well, I pray that you are all granted your most favourable prayers by the Almighty.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

In this chapter, an introduction is presented to metabolic syndrome, its various prediction methods and the motivation behind this research. The problem statement, research questions, objectives and scope are then explained. A brief description of the research significance and contributions are further presented. Finally, a brief outline of the thesis is described.

1.1 Background

Hundreds of thousands die each year due to under diagnosed non-communicable diseases which could be prevented by the early diagnosis of metabolic syndrome. In the era of rising needs and challenges in the field of Non-Communicable Diseases (NCDs) care, a number of studies on the healthcare system, which can extract and synthesis patient data, have been carried out. The cluster of health risks that indicate metabolic and physiological abnormalities was first called "syndrome X" in 1988 (Reaven, 1988). Syndrome X was described as existence of several risk factors such as visceral obesity, glucose intolerance, increased Triglycerides (TG), hyperinsulinemia, decreased High-Density Lipoprotein Cholesterol (HDL-C), and hypertension. Several other terms have been used to describe this clustering of abnormal risk factors, such as the Insulin Resistance Syndrome (Haffner et al., 1992), the multiple MetS (Liese et al., 1997), and the MetS (S. Carroll, Cooke, & Butterly, 2000; Maison, Byrne, Hales, Day, & Wareham, 2001).

The first definition of MetS was first stated by the World Health Organisation (WHO) in the year 1999 (WHO, 1999). It was defined as the presence of Impaired Glucose Tolerance (IGT), diabetes mellitus (DM) or insulin resistance together with two or more of the components listed in column one of Table 1.1. Subsequently, the European Group for Study of Insulin Resistance European Group for the Study of Insulin Resistance (EGIR) suggested a variation to the WHO definition by defining MetS as insulin resistance syndrome (Balkau

& Charles, 1999). In this definition, the criteria for diagnosis include an elevated plasma insulin together with two other MRFs as stated in column two of Table 1.1. The EGIR gave more importance to the presence of abdominal obesity than WHO but removed Type II Diabetes Mellitus (T2DM) because they regarded insulin as mainly a risk factor for T2DM. The WHO definition was shortly followed by another definition released by the National Cholesterol Education Program Adult Treatment Panel III (NCEP ATP III) (Expert Panel on Detection, 2001).

NCEP ATP III changed the name 'insulin resistance syndrome' to 'metabolic syndrome' because central obesity was regarded in this definition as the major MRF. In this definition, MetS was defined as the presence of three or more of the MRFs itemised in column three of Table 1.1. This definition is less focused on insulin resistance compared to the WHO definition and replaced the term 'insulin resistance syndrome' with 'metabolic syndrome' However, this definition recognised central obesity as the major risk factor of MetS while removing the Body Mass Index (BMI), which is a parameter for generalised obesity. Central obesity is measured using Waist Circumference (WC) instead of the Waist – Hip Ratio (WHR) used in the WHO definition. The NCEP ATP III definition also separated body lipids as a low level of HDL-C and a high level of TG. The threshold for Systolic Blood Pressure (SBP), Diastolic Blood Pressure (DBP) and HDL-C were also lowered. The threshold for fasting blood glucose for the diagnosis of MetS was lowered to 5.6 mmol/L. The NCEP ATP III definition was more practically applicable in clinical settings because of the exclusion of the microalbuminuria. Other definitions of MetS were also proffered by the International Diabetes Federation (IDF) (G. Alberti, Zimmet, Shaw, Grundy, et al., 2006).

Table 1.1: Clinical definitions of metabolic syndrome

		WHO (1999)	EGIR (1999)	NCEP ATP III (2001)	American Heart Association — National Heart, Lung, and Blood Institute (AHA/NHLBI) (2003)	IDF (2006)	JIS (2009)
Diagnostic Criteria		Glucose intolerance, IGT or diabetes, and/or insulin resistance together with two or more of the following	Insulin resistance together with two of the following	Three or more of the following five MRFs	Three or more of the following MRFs	Central Obesity and two or more of the following four MRF	Three or more the following MRFs
MetS Risk Factors	Fasting Plasma Glucose (FPG)		≥ 110 mg/dl (6.01 mmol/L) but non-diabetic	≥ 100 mg/dl (5.6 mmol/L)	≥ 100 mg/dl (5.6 mmol/L) or treatment	≥ 100 mg/dl (5.6 mmol/L) or treatment	≥ 100 mg/dl (5.6 mmol/L) or treatment
	BP	≥ 140/90 mm Hg	≥ 140/90 mm Hg or treatment	≥ 130/ ≥ 85 mm Hg	>130/85 mm Hg	Systolic:≥ 130 mm Hg or Diastolic ≥ 85 mm Hg	Systolic:≥ 130 mm Hg and/or Diastolic ≥ 85 mm Hg
	TG	Raised plasma triglycerides: ≥ 150 mg/dL (1.7 mmol/L) *and/or*	≥ 178 mg/dL (2.0 mmol/L) or treatment	≥ 150 mg/dL (1.7 mmol/L)	≥ 150 mg/dL (1.7 mmol/L) or treatment	≥ 150mg/dL (1.7 mmol/L)	≥ 150mg/dL (1.7 mmol/L)
	HDL-C	< 35 mg/dL (0.9 mmol/L) in males: < 39 mg/dL (1.0 mmol/L) in females	< 39 mg/dL (1.0 mmol/L) or treatment	< 40 mg/dL (1.03 mmol/L) in males: < 50 mg/dL (1.29 mmol/L) in females	< 40 mg/dL (1.03mmol/L) in males: < 50 mg/dL (1.29 mmol/L) in females or treatment	< 40 mg/dL (1.03 mmol/L) in males: < 50 mg/dL (1.29 mmol/L) in females or treatment	< 40 mg/dL (1.0 mmol/L) in males: < 50 mg/dL (1.3 mmol/L) in females
	Obesity	WHR > 0.90 in Males; WHR > 0.8 in females and/or BMI > 30 kg/m^2	WC > 94 cm in Males; WC > 80cm in females	WC > 102 cm in Males; WC > 88 cm in females	WC > 102 cm in Males; WC > 88 cm in females	Population- and country-specific definitions	Ethnic- and country-specific definitions
	Micro albuminuria	Urinary albumin excretion rate ≥ 20 μg/min or albumin/creatinine atio ≥ 30 mg/g					

In view of all these disparities in MetS definitions and the lack of agreement in WC thresholds, a worldwide consensus definition of MetS which included race and gender specific WC thresholds was released by the IDF (K. G. M. M. Alberti, Zimmet, & Shaw, 2006). This definition considered central obesity as the major risk factor of MetS and diagnosed MetS as the presence of central obesity and any two of the MRFs as stated in column 5 of Table 1.1.

However, a major problem occurred because contrasts existed during data generation due to the adoption of the various MetS definitions (Hwu et al., 2008; Oda, Abe, Veeraveedu, & Watanabe, 2007; Hunt, Resendez, Williams, Haffner, & Stern, 2004). For instance, the sub-analysis of the nationwide survey of MetS in Malaysian adults identified MetS in 32.1% according to the WHO definition while 34.3% were identified as having MetS according to the NCEP ATP III definition and IDF definition identified the prevalence of MetS as 37.1% (Mohamud et al., 2012).

Finally, a harmonised JIS (K. Alberti et al., 2009) of the IDF task Force in Epidemiology and Prevention: the National Heart, Lung and Blood Institute: the American Heart Association (NHLBI); the World Heart Federation (WHF); the International Atherosclerosis Society; and the International Association for the Study of Obesity (IASO) suggested the use of the IDF global consensus definition, but excluding the central obesity as a compulsory parameter. They suggested the presence of any three or more out of the five parameters as identified in the last column of Table 1.1 (K. Alberti et al., 2009). This definition is also referred to as the harmonized definition of MetS and is at present the most up-to-date and most widely accepted definition. The harmonized dichotomized definition defines MetS as the presence of at least three out of the following five abnormalities:

1. Elevated WC: (ethnic- and country- specific);
2. Raised TG: TG $\geq$ 150 mg/dL(1.7 mmol/L);

3. Reduced HDL-C: HDL-C < 50 mg/dL in women and HDL-C (1.3 mmol/L) <40 mg/dL (1.0 mmol/L) in men;
4. Raised Blood Pressure (BP): - SBP ≥ 130 mmHg and/or DBP ≥ 85 mmHg; and
5. Raised FPG: FPG ≥ 100 mg/dL (5.6 mmol/L).

The MetS structure depicting the MRFs is presented in Fig. 1.1

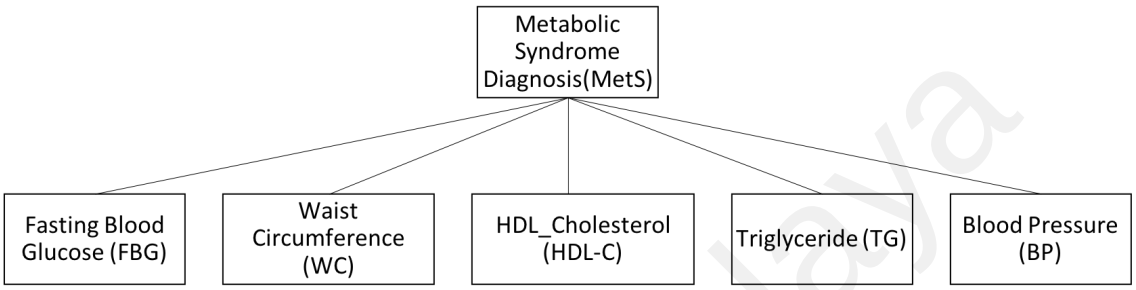

Figure 1.1: Metabolic syndrome hierarchy htructure.

Invariably, these MRFs are defined as continuous variables with differing measurement metrics. Let us assume that a middle-aged female presents with FPG of 5.4 mmol/L, WC of 78 cm, HDL-C of 1.1 mmol/L, TG of 1.8 mmol/L and SBP/DBP of 127/78 mm Hg, then according to the dichotomous definition of MetS, she will not be diagnosed as having MetS. The dichotomous definition will recognise only her TG and HDL-C as having exceeded the recommended threshold. In this regard, studies have shown that dichotomising the continuous variables of the MRFs based on cut-off points potentially leads to misclassification especially when the MRF values are at the borderline of the cut-off points (Ekelund et al., 2006). The dichotomous definition also assumes equal weighting for all the five MRFs during diagnosis despite the different indications that the MRFs represent (Simmons et al., 2010). Again, if we revisit the MRF values, it can be clearly seen that she was not diagnosed with MetS because only two out of five MRFs were considered in her diagnosis. The dichotomous method excludes any MRF that does not exceed or meet up to the threshold values e.g.HDL-C in this case. For these aforementioned

reasons, the dichotomous method incurs information loss in the its diagnosis. It has also been suggested that despite the growing number of research on finding the definition of MetS, an agreement of over a unified definition is yet to be apparent (Reaven, 1988). Thus, it is crucial to evaluate and select efficient MetS risk quantification methods to supplement the traditional binary/dichotomous method of MetS diagnosis.

1.2 Problem Statement

MetS is a continuing epidemic resulting from the concurrence of dyslipidemia, hyperglycaemia, visceral obesity, and hypertension. Despite the varying prevalence of MetS, studies have shown high global percentages of the existence of the abnormality. The adult prevalence of MetS is 7.3% in China (Lao et al., 2012), 29.6% in Brazil (de Carvalho Vidigal, Bressan, Babio, & Salas-Salvadó, 2013), 27.5 % in Malaysia (Rampal et al., 2012), 33.5% in India (Prasad, Kabir, Dash, & Das, 2012), and 34.7% in the USA (Aguilar, Bhuket, Torres, Liu, & Wong, 2015).

Its onset and existence significantly predisposes people to the risk of NCDs such Cardiovascular Disease (CVD) (Lakka et al., 2002; Meigs et al., 2003), T2DM (Laaksonen et al., 2002), premature mortality (Malik et al., 2004) and cancer. The prevalence of T2DM, CVD and obesity are increasing worldwide resulting in a cascade of health disturbances (Cani & Hul, 2015).

The clinical diagnosis of MetS is also known as the dichotomous definitions of MetS. Although it is accepted as the gold standard for the clinical diagnosis of MetS, it is faced with some drawbacks. The definitions of MetS shown in Table 1.1 are used to diagnose MetS by dichotomising the results of the risk factors. However, the MRFs in these definitions are more continuous than dichotomous. Therefore, loss of information is incurred by dichotomising the continuous MRFs (Kahn, 2007; Ferreira et al., 2007). Furthermore, dichotomising the MetS definition means that not all the MRFs contribute

to the clinical diagnosis of MetS. Some MRFs are isolated in the process, risking their important contributions to the diagnosis. For example, let's assume an Asian middle-aged female presents with MRF measurement in Table 1.2.

Table 1.2: MetS risk factor sample measurement for middle-aged Asian male

MRFs	Patient MRF measurements	JIS threshold	JIS MetS Assessment
FPG	5.0 mmol/L	≥ 5.6 mmol/L	No
WC	108 cm	> 102 cm	Yes
TG	1.9 mmol/L	≥ 1.7mmol/L	Yes
HDL-C	1.2 mmol/L	< 1.0 mmol/L	No
BP	129/80 mmHg	≥ 130/85 mmHg	No

Looking at Table 1.2, the patient has only two abnormalities - increased WC and raised TG according to the dichotomous definition. The other three MRFs are on the borderline of the measurement thresholds. Therefore, according the dichotomous definition, these MRFs are not abnormal. In this case, the patient is NOT diagnosed with MetS by any of the clinical definitions of MetS in Table 1.1. This means that the patient will not be well informed about her risk of both diabetes and hypertension due to the loss of information incurred by applying the dichotomous definition of diagnosing MetS. However, it is clear from the patient's measurement he is obviously at risk of both hypertension and diabetes given that her current measurements are close to the clinically defined cut-offs and she needs to consider a lifestyle change to prevent the onset of MetS and its associated diseases. Dichotomising the continuous MRFs therefore implies that not all them contribute equally to the diagnosis (Hillier et al., 2006). Furthermore, this dichotomisation reduces statistical power and correlative measurement of the MRFs (Wijndaele et al., 2006). However, every single MRF plays a major role in rise and contribution to MetS and its associated NCDs independent of the existence of the other MRFs (Sattar et al., 2003; McNeill et al., 2005). These shortcomings of the dichotomous definition will hinder the early diagnosis of MetS.

People with MetS are three folds more likely to have one form of cardiovascular disease

than those without the abnormality (Isomaa et al., 2001). MetS is thought to be one of the precursors to non-communicable diseases such as diabetes (Brozova, Cechurova, & Lacigova, 2016; Ford & Li, 2008; Demir et al., 2016), cardiovascular disease (Dekker et al., 2005; Dragsbæk et al., 2016), kidney disease (J. Chen et al., 2004; Domingos & Serra, 2014), obstructive sleep apnea (Drager, Togeiro, Polotsky, & Lorenzi-Filho, 2013). Therefore, it is mainly used to predict the risk of T2DM (Laaksonen et al., 2002) and CVD (Lakka et al., 2002; Galassi, Reynolds, & He, 2006). These plethora of health complications have made the early diagnosis of MetS a major concern by both researchers and medical practitioners alike.

Despite the significant of the early detection of the risk of MetS, very few researches have applied the use of non-clinical methods to predict and diagnose the risk of MetS. Principal Component Analysis (PCA) (Vikram, Pandey, Misra, Goel, & Gupta, 2009; Ayubi, Khalili, Delpisheh, Hadaegh, & Azizi, 2015; S. J. Carroll et al., 2014; Wiley & Carrington, 2016; Mochizuki, Miyauchi, Misaki, Ichikawa, & Goda, 2013), z-score (Batey et al., 1997; Eisenmann, 2008; Wijndaele et al., 2009; Okosun, Lyn, Davis-Smith, Eriksen, & Seale, 2010; Neto, de Campos, Dos Santos, & Junior, 2014; Heshmat et al., 2015), and Confirmatory Factor Analysis (CFA) (Vikram et al., 2009; Ayubi et al., 2015; S. J. Carroll et al., 2014; Wiley & Carrington, 2016; Mochizuki et al., 2013) are statistical techniques that were first used to diagnose the risk of MetS. These non-clinical techniques derived a value known as Continuous MetS (cMetS) score that represents the level of the risk of MetS based on the MetS risk factor measurement values. However, the statistical methods are sample specific and the same model cannot be applied to different populations. Also the impact of all the MRF assume to be equal leading to unaccountability in their different effects.

Machine learning techniques such as Multiple Logistic Regression (MLR), Support

Vector Machine (SVM) (Van Schependom et al., 2015), Artificial Neural Network (ANN) (Zhao et al., 2014; Hirose, Takayama, Hozawa, Hibi, & Saito, 2011), Bayesian Network (BN), MediBoost (Kakudi, Loo, & Moy, 2017) and Decision Tree (DT) (Romero-Saldana et al., 2016) have been used in the prediction of MetS. However, the DT, MLR, MediBoost and SVM are black-box techniques. This means that the models developed cannot be interpreted to predict the risk of MetS. Furthermore, these techniques tend to overfit data and may not recognize the risk of MetS in patients with borderline MRF measurements. Both SVM and ANN are black box techniques. This setback renders them inefficient for the early diagnosis and interpretation of the risk of MetS as required for prevention.

Thirdly, risk quantification techniques such as Areal Similarity Degree (ASD) (Jeong, Jo, Shim, Choi, & Youn, 2014), the "siMS score" (Soldatovic, Vukovic, Culafic, Gajic, & Dimitrijevic-Sreckovic, 2016), ASD with weights obtained by Quantum Particle Swarm Optimisation (QPSO) (Kakudi, Loo, & Pasupa, 2017) have been applied to derive the risk of MetS. However, these risk quantification techniques are data dependent and do not create models from learning.

Taking into context the limitations of the existing non-clinical methods previously used for the diagnosis of the risk of MetS and the capabilities of BAM and Genetic Algorithm (GA) as stated in Sections 3.2.4 and 3.3.11, we propose the GOBAM as the most appropriate and novel technique for the early diagnosis of the risk of MetS while taking into consideration patients with MRF measurements close to the clinically accepted thresholds of diagnosing MetS. GOBAM is able to predict the risk of MetS by generating a risk quantification index.

1.3 Aims of Research

The aim of this research is to apply GA in order to optimize the sequence order of input data and parameters of the BAM to build a GOBAM model for the early diagnosis of the risk of MetS while considering patients with MRF measurements that exceed the clinically recognized thresholds and those with borderline MRF measurements.

1.4 Research Objectives

The objectives required to overcome the limitations of the previously applied non-clinical approaches are as follows:

1. To identify non-clinical techniques and the challenges faced in diagnosing and quantifying the risk of MetS.
2. To propose and develop a novel risk prediction and quantification model for the early detection and diagnosis of MetS using the five clinically approved MRFs.
3. To evaluate the performance of the proposed model GOBAM with classical ARTMAP.

1.5 Research Questions

In order to achieve these research objectives, the following research questions will be answered:

i. What are the non-clinical techniques that have been used for the prediction, diagnosis and risk quantification of MetS?

ii. How can a risk prediction and quantification model for the diagnosis of MetS be developed?

iii. What is the predictive performance of the proposed GOBAM compared with other classic ARTMAP models?

iv. What is the significant difference between the predictive performance of the GOBAM model and the other comparative models?

v. Can GOBAM be applied for the risk quantification of MetS in patients who present with borderline MRF measurements?

1.6 Mapping the Objectives and the Research Questions

This section presents the connection between the objectives and the research questions of thesis in Table 1.3.

Table 1.3: Mapping the research objectives to the research questions

Research objectives	Research questions
1. To identify non-clinical techniques and the challenges faced in diagnosing and quantifying the risk of MetS.	i. What are the non-clinical techniques that have been used for the prediction, diagnosis and risk quantification of MetS?
2. To propose and develop a novel risk prediction and quantification model for the early detection and diagnosis of MetS using the five clinically approved MRFs.	ii. How can a risk prediction and quantification model for the diagnosis of MetS be developed?
3. To evaluate the performance of the proposed model GOBAM with classical ARTMAP.	iii. What is the predictive performance of the proposed GOBAM compared with other classic ARTMAP models? vi. What is the significant difference between the predictive performance of the GOBAM model and the other comparative models? v. Can GOBAM be applied for the risk quantification of MetS in patients who present with borderline MRF measurements?

1.7 Motivation

The prime motivation of this research is the global epidemic spread and economic burden of MetS and its associated diseases. MetS is a pre-cursive abnormality to various types of NCDs which include CVDs such as strokes, ischemic heart disease, atherosclerosis and T2DM. The abnormalities of MetS hypertension, insulin resistance, dyslipidaemia,

hypertriglyceridaemia and central obesity are leading causes of disability and mortality independently or collectively. For example, hypertension leads to 12.8% and 3.7% of the global premature death and total Disability Adjusted Life Years (DALY) respectively (WHO, 2018). The primary reason for the diagnosis of MetS is to identify patients who are at long term risk of developing CVDs and T2DM. The clinical definition of MetS provides a threshold for the existence of abnormality in any of the five MRFs. However, even individuals that present MRF measurements that are very close the specified thresholds will still be diagnosed as NOT being at risk of MetS and its associated diseases. However, disregarding these borderline MRF results will to to the progression and onset of Metabolic Syndrome (MetS). Therefore, the early detection of MetS for both individuals diagnosed with abnormality and those that have not been diagnosed according to the clinical dichotomous definition is relevant for clinical practitioners and individuals alike. Early knowledge of the nature and onset of MetS will enable the timely implementation of health management and intervention schemes that will either prevent or reduce progression to its associated diseases.

1.8 Methodology

This research begins with conducting a systematic literature review of all the existing non-clinical approaches that have been used to diagnose the risk of MetS and identifies their limitations from the literature. These limitations were used to formulate the problem statement of this study. Subsequently, the aims, research objectives and questions were defined form the problem statement. The data was collected as part of the CLUSTer (Moy et al., 2014) which consists of MRF measurements. The data was cleaned and preprocessed prior to input.

In this framework, we build a GOBAM that allows new data to be processed without discarding existing knowledge. The GA is used to optimise the sequence order of the input

sample of the MetS dataset and the parameters of the BAM (Vigdor & Lerner, 2007). The trained model can then be used to diagnose the risk of MetS for patients. FAM (Carpenter, Grossberg, & Reynolds, 1991), GAFAM (S. M. Liew, Doust, & Glasziou, 2011), BAM models were also built with the same dataset in order to evaluate the performance of the proposed GOBAM model. The FAM (Carpenter, Grossberg, & Reynolds, 1991), GAFAM (S. M. Liew et al., 2011), BAM algorithms were the ART and ARTMAP algorithms most recently used for model performance comprisons (W. S. Liew, Seera, & Loo, 2016; Masuyama, Loo, & Dawood, 2018). Their results were compared using the Friedman test to explore the significant difference in the performance of the other models against the GOBAM. The built model was tested to see if it was able diagnose MRF measurements of a patient with borderline measurement as being at risk of MetS.

1.9 Scope of Research

The focus of this research work is develop and implement a novel non-clinical model risk prediction using BAM and GA. This model will be used to support the clinical diagnosis of MetS in predicting the risk of MetS for patients with and without the abnormality according to the clinical dichotomous definition. The dataset used for the evaluation of the proposed model is collected specifically from the Malaysian population. The performance of the proposed GOBAM model is compared to FAM, GAFAM, and BAM.

1.10 Thesis Contribution

This research make its contribution to two different fields: the machine learning community and the medical practitioners.

The machine learning community

The primary contribution of this thesis is the development of a novel non-clinical model that can be used to support the clinical diagnosis of MetS. Furthermore, this research work

aims to identify significant advances and contributions of non-clinical methods used to diagnose metabolic syndrome. The proposed approach developed in this research work is an addition to other non-clinical techniques that have been proposed by previous researches. Our approach improves the early diagnosis of MetS for patients who are at risk and for those who present with MRF measurements close to the clinically defined thresholds.

The medical practitioners

Our proposed approach can be used to support the clinical diagnosis of MetS as an early and accurate detection tool. This information will enable medical practitioners to implement timely intervention systems for prevention and treatment of MetS and its associated diseases

1.11 Thesis Organisation

Chapter 1 presents the background of the thesis. It also includes the problem statement, aims, research objectives and questions, scope and contribution of the thesis. The research methodology of the proposed research model is briefly described.

In Chapter 2, we present a systematic literature review in order to identify the non-clinical approaches that have been previously used to diagnose MetS. Machine learning, statistical and risk quantification techniques are described and their strength and limitations are also presented.

The theoretical framework of the research is explained in Chapter 3. A detailed description of Adaptive Resonance Theory (ART), ARTMAP, FAM, GAFAM, BAM and GA is provided. The mathematical formulations, technique pseudo-code, flowchart and diagrams of the algorithms are presented.

The description of the dataset, its collection, cleaning and preprocessing methods are presented in Chapter 4. Subsequently, the development of the GOBAM for the

diagnosis of the risk of MetS is explained in the detail. The pseudo-code, flowchart and implementation of the proposed GOBAM model is presented. The experimental setup of the implementation is also described in detail.

Chapter 5 presents the results of the experiments conducted in Chapter 4.7. The performance comparison between ARTMAP, FAM, GAFAM, BAM, and the proposed GOBAM is made. Detailed discussion of the results are presented.

The conclusions made from the empirical findings in Chapter 5 is presented Chapter 6. Contributions of the study and future related works are highlighted.

CHAPTER 2: SYSTEMATIC LITERATURE REVIEW

2.1 Introduction

This chapter presents a systematic literature review on the diagnosis of MetS using non-clinical techniques. The diagnosis of MetS is a motivation for individuals towards implementing healthy life style choices. Therefore, existing literature reviews have answered several research questions relating to the characteristics and associated diseases of MetS. Both Timar and colleagues, (Timar, Sestier, & Levy, 2000) and Lopez-Candales (Lopez-Candales, 2001) present a summary of MetS by reviewing each MetS risk factor, cut-off thresholds, and people most susceptible to the abnormalities. The review by Palomo and colleagues (Palomo, Alarcon, Moore-Carrasco, & Argiles, 2006) showed that MetS is characterised by alterations in hemostatis and fibrinolysis which is attributed to metabolic abnormalities. In their survey, Gami and co workers (Gami et al., 2007) found thirty-seven studies that ascertained the association between MetS and CVDs . They concluded that individuals who present with the risk of MetS are highly at risk of developing CVDs which could increase mortality rate if lifestyle and preventive interventions are not applied. Xue and Michels, (Xue & Michels, 2007) reviewed available evidence that proved the existence of a clear relation between MetS, T2DM, and the onset of breast cancer. More recently, Wong and colleagues (Wong, Cook, Roderick, & Somani, 2016) noted a definite association between breast cancer, kidney stone and MetS caused by the presence of abnormalities such as obesity, hypertension, hyperinsulinemia and insulin resistance. Another study (Hert, Schreurs, Vancampfort, & Winkel, 2009) noticed the increased risk of developing MetS in patients treated with antipsychotic agents due sedentary lifestyle habits, unhealthy food choices and high rate of smoking. These factors subsequently led to weight gain, and increased prevalence of T2DM and CVDs. Motillo et al., (Mottillo et al., 2010) found

that MetS is related with a 2-times increase in CVD and a 1.5 times increase in overall mortality rate while Kaur et al., (Kaur, 2014) in their extensive review summarised existing literature related to the definition of MetS, its epidemiology and intervention approaches.

In summary, all these reviews have only analysed the current state-of-the-art of MetS and some of its associated diseases; however, none of them have examined the existing non-clinical methods that support the clinical diagnosis of MetS. We consider that the analysis of research activity in this domain is of utmost importance in order to investigate more research possibilities aimed at early diagnosis and prevention of MetS.

This systematic review aims to identify and assess non-clinical techniques which support the clinical diagnosis of MetS. The justification behind this systematic review is based on the requirement of knowledge acquisition that could assist in improving the quality of non-clinical methods for the diagnosis of MetS and subsequently to promote the management of MetS in clinical practice.

2.2 Pathophysiology of MetS

Metabolic syndrome is not a disease. It is a constellation of metabolic abnormalities which include central obesity, impaired glucose tolerance (IGT), hypertension and dyslipidemia (K. Alberti et al., 2009). Insulin resistance is theorized as the primary indicator of MetS (Eckel, Grundy, & Zimmet, 2005) which is the inability of the body to competently respond to endogenous and exogenous insulin.

2.2.1 Central Obesity

Central obesity which is measured by the size of the waist circumference is a reliable substitute for the estimation of abdominal fat adiposity besides using computer tomography (CT) scans and dual-energy X-ray absorptiometry (DEXA) scans (Direk et al., 2013). Its simplicity and cheapness makes it a convenient tool for use in routine clinical practice

and health screen to assist in targeting early intervention of MetS. Central obesity and insulin resistance are claimed to be closely related (Zadeh-Vakili, Tehrani, & Hosseinpanah, 2011). The excess visceral tissue in abdominal obesity which is in close proximity to the liver releases free fatty acids (FFA) into the liver through the splanchnic circulation (Ebbert & Jensen, 2013). This leads to insulin resistance (Eckel et al., 2005; U. J. Jung & Choi, 2014) causing a failure to suppress gluconeogenesis in the liver, resulting to a hyperglycemic state (M. T. Sheehan & Jensen, 2000). Some studies have suggested central obesity as the earliest stage of MetS development and early weight decrease could prevent the development of MetS and its ailments (Esposito et al., 2003; L. Palaniappan et al., 2004). Central obesity is also a strong indicating factor of cardiovascular diseases (Seidell et al., 1992; Wildman et al., 2005) and type II diabetes mellitus (Balkau et al., 2007).

2.2.2 Impaired Glucose Tolerance

A well established relationship between impaired glucose tolerance and insulin resistance has been investigated because insulin assists the muscle, fat and liver cells to absorb glucose. Insulin resistance is identified as the core metabolic abnormality in MetS (Falkner & Cossrow, 2014). Insulin resistance is the deficiency of insulin-mediated glucose uptake which results in hyperglycemia (De Luca & Olefsky, 2008). There is also a failing to suppress gluconeogenesis in the liver (J. P. Sheehan, 2004). To make up for the imperfection in insulin activity, the insulin discharge is metrically increased to maintain normal glucose levels. In any case, if the level of insulin secretion still fails to meet the normal requirement, impaired glucose tolerance would still arise (Eckel et al., 2005).

2.2.3 Hypertension

Studies have associated essential hypertension with obesity and glucose intolerance (Ferrannini & Cushman, 2012; Turpin et al., 2014; Morris, 2018). Obesity is a major

risk factor in hypertension in MetS (Anderson et al., 2001; Morse, Zhang, Thakur, & Reisin, 2005; Falkner & Cossrow, 2014). The hyper-insulinemia existing in obese people excites the sympathetic nervous system (SNS) so as to preserve energy balance by elevating metabolic rate (Landsberg, 2001). Therefore, the SNS leads to hypertension by applying pro-hypertensive effects such as vasoconstriction, increased cardiac output and sodium re-absorption in the blood vessels, kidneys and heart.

2.2.4 Dyslipidemia

An increase in visceral adiposity increases the release of FFA into the liver through splanchnic circulation (Montague & O'rahilly, 2000), thereby, causing an elevated production of very low-density lipoprotein (VLDL) and triglycerides (M. T. Sheehan & Jensen, 2000). VLDL functions as a carrier of triglycerides from the liver to the circulatory system (Ginsberg & MacCallum, 2009). In a normal state, insulin inhibits the production of VLDL and triglycerides in the liver. However, in an abnormal insulin resistance state, insulin looses this ability which causes hyper-triglyceridemia. The VLDL composed of triglycerides will be converted to low density lipoprotein (LDL) and the high-density lipoprotein (HDL) in exchange for cholesterol. Thereafter, the triglyceride filled with LDL and HDL will be hydrolysed into small, dense LDL particles which atherogenic. This results in dyslipidemia in MetS and is characterised by an elevated triglycerides and low HDL-C factors.

2.3 Systematic Review Method

The systematic review method adopted in this chapter follows the procedures given by the PRISMA guidelines (Moher, Liberati, Tetzlaff, & Altman, 2009). Therefore, we developed the study protocol by designing the search strategy, enumerating inclusion/exclusion criteria, extracting, and synthesising extracted data. Initially, we carried out the search

strategy by identifying the search terms and the electronic database to carry out our search. Relevant primary studies were selected as depicted in the PRISMA flow chart in Fig 2.1. Next, we identified the quality of the selected studies based on a quality assessment questionnaire designed to assess the quality of all the selected studies.

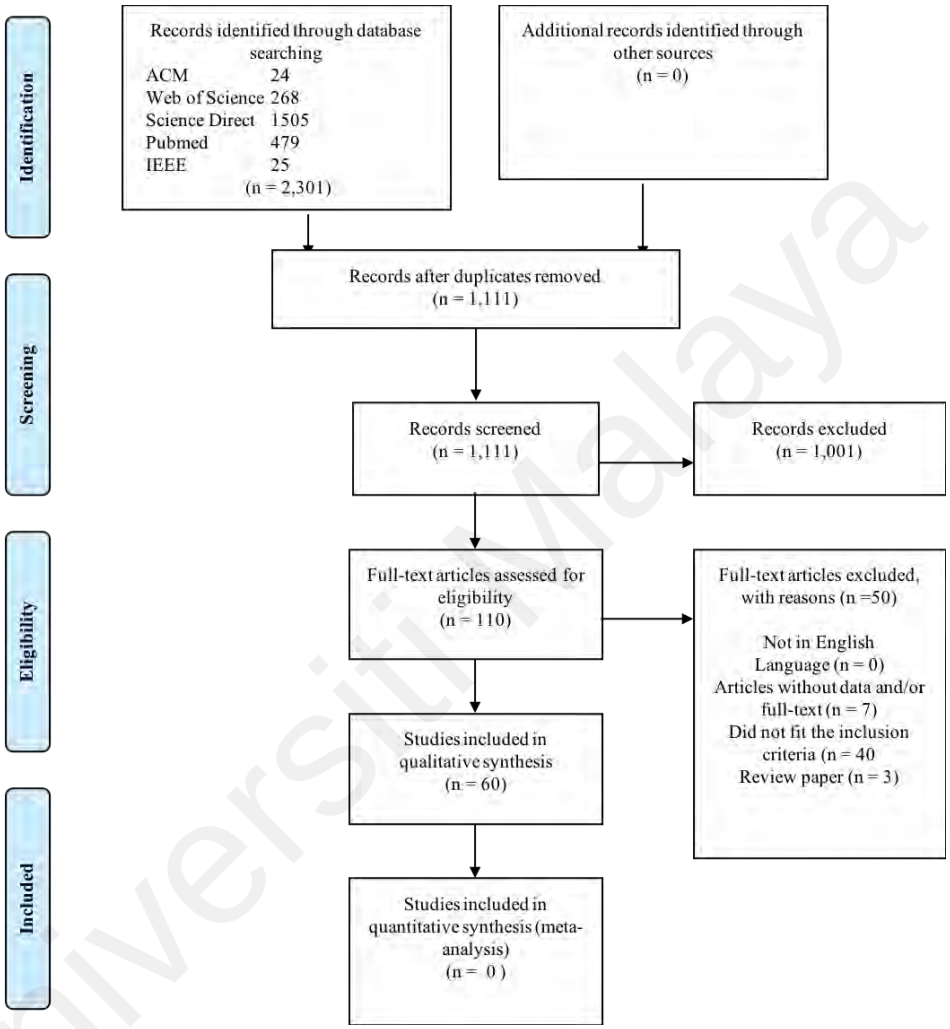

Figure 2.1: PRISMA flow chart for included and excluded studies in the systematic review on metabolic syndrome diagnosis methods

2.3.1 Search Strategy

All searches were carried out from the earliest date in the database to July 2019 in the following electronic data sources: Association of Computing Machinery (ACM) Digital Library, PubMed, Science Direct, Web of Science, and IEEE Xplore Digital Library. These electronic databases are the most frequently searched and widely accepted by various

research communities (Saeed, Ab Hamid, & Mustafa, 2016).

The primary search term of the systematic review was "Metabolic Syndrome". The secondary search terms consisted of relevant machine learning keywords with "Metabolic Syndrome" using the "AND" and "OR" boolean concatenators in the Web of Science, PubMed and IEEE Xplore digital libraries. However, to remove ambiguity of publications results after searching, the search secondary keywords were broken into 6 sections. Table 2.1 shows the digital databases and the search terms.

Table 2.1: List of search terms conducted in the various databases

Databases	**Search terms**
IEEE Xplore Digital Library PubMed Science Direct Web of Science Association of Computing Machinery (ACM) Digital Library	1. ("Metabolic Syndrome") AND ("Naive Bayes" OR "Bayesian Network" OR "Decision tree" OR "Classification and regression tree" OR "Chi-squared Automatic Interaction Detection" OR "Linear discriminant analysis" OR "K-nearest neighbors algorithm"); 2. ("Metabolic Syndrome") AND ("Learning vector quantization" OR "Self-organizing map" OR "K-means" OR "Expectation-maximization" OR "Hierarchical Clustering" OR "Fuzzy clustering" OR "Hierarchical clustering" OR "Hidden Markov models"); 3. "Metabolic Syndrome") AND ("Logistic regression" OR "Ordinary least squares regression " OR "Linear regression" OR "Stepwise regression" OR "AdaBoost" OR "Boosting" OR "Random Forest" OR "Gaussian mixture models" OR "k-Means clustering"); 4. ("Metabolic Syndrome") AND ("machine learning" or "deep learning" OR "data mining" OR "predictive models" OR "Artificial neural network" OR "Back-Propagation" OR "Multilayer Perceptron" OR "association rule mining"); 5. ("Metabolic Syndrome") AND (“Confirmatory factor analysis” OR "Factor analysis" OR "Principal component regression" OR “Principal Component Analysis”) 6. ("Metabolic Syndrome") AND (“Z score” OR “Risk Score” OR “Continuous risk score” OR “Continuous Metabolic syndrome risk score“ OR "Severity Score" OR "Risk quantification")

A total of 42 search terms were used to collect the articles in this systematic review. These search terms cut across the different types of statistical and machine learning techniques that are most frequently used for classification and prediction. The whole search was carried out through the University of Malaya library.

Only fully published, peer-reviewed papers reported in English were included.

2.3.2 Study Selection

The whole initial search from the electronic databases revealed 2301 studies. These studies were screened based on the objective of the review which sought to appraise the empirical quality of the studies as well as evaluate the effectiveness of applying non clinical methods to diagnose MetS. After removing duplicates, the first selection step included screening the titles, abstracts, and keywords of the studies for eligibility in relation to the following inclusion criterion:

i. The study participants were humans with or without metabolic syndrome and the participant characteristics were clearly defined.
ii. The study objective involved the diagnosis of metabolic syndrome using a non clinical approach in addition to or as opposed to the clinical dichotomous definition.
iii. At least an outcome for the non-clinical diagnosis of metabolic syndrome exists in the study.

110 primary studies were left after the first screening. We were able to retrieve the full text of 106 articles by searching online databases. We were not able to retrieve the full text of 7 articles. The inclusion/exclusion criteria were independently tested by two reviewers (Habeebah Adamu Kakudi (HAK) and Foong Ming Moy (FMM)) and a conclusion was agreed upon after comprehensive discussions. The entire full text of the 107 primary studies were screened by the two reviewers (HAK and FMM) based on the following exclusion criteria:

i. Non-peer reviewed studies such as tutorials, reports, conference, and editorial papers.
ii. Studies with adolescent and children participants.

iii. Studies based on conceptual frameworks or structures without empirical analysis and results.

The two researchers compared their results and where discrepancies existed between two results of the same paper, a joint assessment was carried out with all the researchers and an agreement was made to either include or exclude the study from the systematic review. Finally, 50 studies were excluded and a total of 60 primary studies were identified for inclusion into this systematic review.

2.3.3 Data Extraction

Reviewer HAK independently extracted data from the primary studies by filling a data extraction form and gathering general information from each study such as author name, publication year, title, study of dataset, country, study population, study design, MetS diagnosis techniques and evaluation metrics results. Researchers HAK and Chu Kiong Loo (CKL) verified the soundness of data by making sure that information extracted from each study justified the aim of the research.

2.4 Results

This sections presents a comprehensive analysis on the taxonomy and results obtained from the 61 primary studies included in this systematic review. First, we present a general analysis of the primary studies including their quality scores obtained from the quality assessment process. Subsequently, we present a detailed investigation of the results based on the research question of the systematic review.

2.4.1 Characteristics of Primary Studies

The non-clinical methods identified from the 61 included studies are categorised as follows:

- Statistical techniques: Confirmatory Factor Analysis (CFA), Principal Component Analysis (PCA), Principal Components Logistic Regression (PCLR), Generalized Multifactor Dimensionality Reduction (GMDR) and Z-score;
- Machine learning techniques: Decision Trees (DT), Decision Trees Chaid (DT Chaid), Random Forest (RF), Support Vector Machines (SVM), Gradient Boosted Trees (GBT), Logistic Regression (LR), Neural Networks (NN), and Bayesian Network (BN), Naive Bayes (NB), Association Rules; and
- Risk quantification techniques: Areal Similarity Degree (ASD), simScore, and Framingham Risk Score (FRS).

Out of the 60 included studies, 27(45%) studies investigated the use of machine learning techniques to predict MetS, while 30(48%) studies applied statistical techniques to calculate a continuous MetS srisk score. Lastly, 4(7%) studies derived risk quantification formulas for quantifying the risk of MetS. Demographic data extracted from the 60 included studies are presented in Table 2.2 while the type of non-clinical technique applied and the performance metrics of the included studies are presented in Table 2.3.

Table 2.2: Demography and MRFs of all the included studies

S/N	Study	Population (M, F)	Age (Mean ± SD) years	MetS Components				
				Obesity	Glucose	Lipids	Blood Pressure	Others
1	(Franks, Ekelund, Brage, Wong, & Wareham, 2004)	868 (378 M, 490 F)	53.3 ± 10.4	BMI + WHR/2	Insulin, 2HPG	HDL-C, TG	SBP +DBP/2	N/A
2	(Ekelund et al., 2006)	605 (246 M, 359 F)	56 ± 2.8	BMI+WHR/2	2HPG, Insulin	HDL-C, TG	SDP + DBP /2	N/A
3	(Hillier et al., 2006)	5,024 (2,467 M, 2,557 F)	30-65	WC, BMI	FPG, Insulin	HDL-C , TG	SBP, DBP	N/A
4	(Wijndaele et al., 2006)	1,021 (571 M, 449 F)	18-75	WC	FPG	HDL-C , TG	SBP, DBP	N/A
5	(Ferreira et al., 2007)	313 (160 M, 153 F)	20 - 25	WC	FPG	HDL-C, TG	MAP, SBP + DBP/2	N/A
6	(Johnson et al., 2007)	171 (91 M, 80 F)	53 ± 7	WC	FPG	HDL-C, TG	MAP	N/A
7	(de Edelenyi et al., 2008)	1,754	58.2± 0.18	N/A	N/A	N/A	N/A	Genetic risk factors: SNP, PFA, DFV
8	(C. C. Lin et al., 2010)	383 (254 M, 129 F)	47.5 ±13.5	WC, BMI	N/A	N/A	SBP, DBP	SGA medication use history, Weight, Height, Age, Gender
9	(Wijndaele et al., 2009)	992 (559 M, 433 F)	18 - 75	WC	FPG	HDL-C , TG	SBP, DBP	N/A
10	(Chang, Yen, & Chen, 2010)	8,908 (4,712 M, 4,196 F)	35.46 ± 7.72	WC, WtHR	FBG	HDL-C , TG	SBP, DBP	N/A
11	(NIHMS501834, 2011)	86 (45 M, 41 F)	18 - 70	WC	FBG	HDL-C, TG	MAP	N/A
12	(Hirose et al., 2011)	410 (254 M, 0 F)	44.2 (30-59)	WC, BMI, FFA	FPG, Insulin, HOMA-IR	TG, HDL-C, LDL-C, CHOL	SBP, DBP	AST, ALT, HMW-A, TA, GAI, SM
13	(Agarwal et al., 2012)	6,780 (3,210 M, 3,567 F)	44 - 84	WC	FPG	HDL-C , TG	SBP, DBP	N/A
14	(Kang et al., 2012)	3,598 (1,451 M, 2,147 F)	35 - 74	WC	FPG	HDL-C, TG	SBP, DBP	N/A
15	(Potteiger et al., 2012)	21 (21 M, 0 F)	27 ± 48	WC	FBG	HDL-C, TG	MAP	N/A
16	(Ushida et al., 2012)	229 Males (0 F, 229 M)	31.4± 7.5, 30.6 ± 4.6	N/A	N/A	N/A	N/A	SM, Hematocrit, Blood urea nitrogen, RBC, Creatinine, WBC, Uric Acid, Urine urobilinogen, γ-GTP, Urine protein, Hemoglobin, Urine sugar, GOT, Urinary occult blood, GPT, Alcohol habit
17	(Dusseault-Belanger, Cohen, Hivert, Courteau, & Vanasse, 2013)	7213 (3877 M, 3336 F)	65 ± 14.4	WC	FPG	HDL-C , TG	SBP, DBP	N/A
18	(Gomez-Marcos et al., 2013)	636(258 M, 378 F)	20 - 80	WC, BMI, WHtR, AI	FPG, Insulin, HOMA-IR	HDL-C, TG	SBP, DBP, MAP	N/A
19	(Huang, 2013)	1216 (N/A M, N/A F)	40-60	N/A	N/A	N/A	N/A	N/A
20	(Huo et al., 2013)	7472 (2666 M, 4806 F)	40-60	WC	FPG	HDL-C, TG	SBP, MAP	N/A
21	(Mochizuki et al., 2013)	308 (308 M, 0 F)	40 - 69	BMI	FPG	HDL-C , TG	SBP, DBP	N/A

Table 2.2 – **Continued**

Table 2.2 – **Continued from previous page**

S/N	Study Country	Population (M, F)	Age (Mean ± SD) years	MetS Components				
				Obesity	**Glucose**	**Lipids**	**Blood Pressure**	**Others**
22	(Worachartcheewan, Nantasenamat, Isarankura-Na-Ayudhya, & Prachayasittikul, 2013)	5,638 Males (2,661 M, 2,977 F)	20-99	BMI	FPG	HDL-C, TG	SBP, DBP	N/A
23	(Povel et al., 2012)	5118 (1761 M, 3357 F)	51.8	WC, BMI	FPG, HbA1C	HDL-C, TG	SBP, DBP	N/A
24	(S. J. Carroll et al., 2014)	3,993 (1,898 M, 2,095 F)	50.49± 16.39	WC	FPG	HDL-C , TG	MAP	N/A
25	(H. Chen, Xiong, & Ren, 2014)	2,074 (1,495 M, 579 F)	46.93± 47.06	WC, BMI, WHR	N/A	N/A	SBP, DBP	HC, Gender, Age
26	(Faria et al., 2014)	210 (24 M, 186 F)	39.6	WC	FPG	HDL-C, TG	SBP, DBP	N/A
27	(Gaio et al., 2014)	206 (87 M, 119 F)	56.43 ± 16.23	WC	FBG	HDL-C , TG	DBP, SBP	N/A
28	(Gurka, Lilly, Oliver, & DeBoer, 2014)	6084 (N/A M, N/A F)	20-64	WC	FPG,	HDL-C, TG	SBP	N/A
29	(Hosseini et al., 2014)	8,313 (4,148 M, 4,165 F)	38.54± 15.86	WC, BMI	FPG, 2hPG	TG, HDL-C, LDL-C, CHOL	MAP	Age, PA, CRP, GDI
30	(Jeong et al., 2014)	5355 (2276 M, 3079 F)	47.11 ± 14.30	WC	FPG	HDL-C, TG	SBP or DBP	BMI
31	(Z. Lin et al., 2014) (S. M. Liew et al., 2011)	163 (133 M, 30 F)	25-70	N/A	N/A	N/A	N/A	Serum Samples-2-Hydroxy-butanoic acid, D-Glucose, Hexadecanoic acid, Inositol, 9,12-Octadecadienioc acid
32	(Miller, Fridline, Liu, & Marino, 2014)	745 (335 M, 410 F)	56.7 (29.3 ± 5.8)	BMI, WC	FPG	HDL-C, TG	SBP, DBP	Age, Weight, Height, ethnicity
33	(Zhao et al., 2014)	2082 (908 M, 1,173 F)	55.35 ± 10.62	WC, BMI	FPG	HDL-C, TG	SBP, DBP	SNPs Age, gender, Adiponectin
34	(Tsou, Chang, Huang, & Hsu, 2014)	636 (432 M, 204 F)	45.1±8.8	WC, BMI	FPG,	HDL-C, TG	SBP, DBP	
35	(Steinberg, Church, McCall, Scott, & Kalis, 2014)	36,944 (N/A M, N/A F)	18-60	WC, BMI	FPG	HDL-C, TG	SBP, DBP	Age, gender ethnicity, SM
36	(Ayubi et al., 2015)	4567 (1861 M, 2706 F)	20–60	WC, BMI	FPG,	HDL-C , TG	SBP, DBP	N/A
37	(Dai et al., 2015)	100 (N/A M, N/A F)	N/A	BMI	FPG	HDL-C, TG, FFA	DBP	Free Fatty Acids
38	(NIHMS735680, 2015)	354 (142 M, 212 F)	20 - 64	WC	FBG	HDL-C, TG	SBP, DBP	N/A
39	(Graziano et al., 2015)	8102 (3485M, 4617F)	49.28 ± 17.59	WC, BMI	FPG	HDL-C, TG	SBP, DBP	N/A
40	(Obokata et al., 2015)	6817 (2605 M, 4212 F)	50.8 ± 8.8	WC, BMI	FPG	HDL-C, TG	SBP, DBP	Age, Gender
41	(Worachartcheewan et al., 2015)	5,646 (2,028 M, 3,618 F)	18 - 78	WC, BMI	FPG	CHOL, TG, LDL-C, HDL-C	SBP, DBP	Age, Gender WBC, Hb, Hct, PLT, SM, AC

Table 2.2 – **Continued**

Table 2.2 – **Continued from previous page**

S/N	Study Country	Population (M, F)	Age (Mean ± SD) years	MetS Components				
				Obesity	Glucose	Lipids	Blood Pressure	Others
42	(Yousefzadeh, Shokoohi, Najafipour, & Shadkamfarokhi, 2015)	4,192 (952 M, 3,238 F)	44.34 ± 16.32	WC	FPG	HDL-C, TG	SBP or DBP	N/A
43	(Van Schependom et al., 2015)	605 (404 M, 201 F)	34.7± 11.5, 39.8 ± 11.7	WC, BMI, WHR, hip	N/A	N/A	SBP, DBP	Height, Weight, AL, SM, HR, GAF
44	(Drehmer et al., 2016)	9,835 (4,445 M, 5,390 F)	50.7 ± 8.7	WC	FBG	HDL-C, TG	MAP	N/A
45	(Jiang et al., 2016)	1,053 (411 M, 642 F)	53.72 ± 12.46	WC	FPG	HDL-C, TG	MAP	N/A
46	(Karimi-Alavijeh, Jalili, & Sadeghi, 2016)	2,107(N/A M, N/A F)	48.07 (34.0-86.0)	WC, HC, WHR, BMI	FPG, 2hBG	HDL-C, LDL-C,TG, CHOL	SBP, DBP	Age, Gender, Weight, HYP, anti-HYP medication use, MCV, MCH
47	(Ivanovic, Kupusinac, Stokic, Doroslovacki, & Ivetic, 2016)	2,928 (1434 M, 1494 F)	43.4 (18-76)	BMI, WtHR	N/A	N/A	SBP, DBP	N/A
48	(Janghorbani & Amini, 2016)	1,869 (N/A M, N/A F)	30 - 70	WC	FPG	HDL-C, TG	MAP	N/A
49	(Romero-Saldana et al., 2016)	636 (432 M, 204 F)	45.1±8.8	WC, BMI, WtHR, BFP, WHR	N/A	N/A	SBP, DBP	N/A
50	(Wiley & Carrington, 2016)	2,125 participants (907 M, 1,218 F)	18 - 94	WC	FBG	HDL-C , TG	DBP, SBP	N/A
51	(Soldatovic et al., 2016)	528 (182 M, 346 F)	7 – 77	WC	FPG	HDL-C, TG	SBP or DBP	Height, Gender
52	(NIHMS865762, 2017)	13,094 (5644 M, 7450 F)	18 - 94	WC	FBG	HDL-C , TG	DBP, SBP	N/A
53	(Je, Kim, & Park, 2017)	5508 (2616 M, 2891 F)	40.35 ± 0.22	WC, BMI	FPG	HDL-C, TG	SBP, DBP	N/A
54	(K. J. Jung, Jee, & Jee, 2017)	441,411 (191095 M, 253847 F)	43.4 ± 10.9	WC, BMI	FPG	HDL-C, TG	SBP, DBP	N/A
55	(DeBoer, Filipp, & Gurka, 2018)	2476 (770 M, 1706 F)	N/A	WC	FPG	HDL-C, TG	SBP, DBP	N/A
56	(Jia et al., 2018)	22,584 (14,250 M, 8334 F)	18 - 88	WC	FPG	HDL-C, TG	SBP, DBP	N/A
57	(Miyauchi & Nishimura, 2018)	745 (489 M, 646 F)	20 - 30	WC, BMI	FPG	HDL-C, TG	DBP, SBP	N/A
58	(Shimoda, Ichikawa, & Oyama, 2018)	13,996 (10,573 M, 3423 F)	35 – 75	WC, BMI	FPG	HDL-C, TG	DBP, SBP	N/A
59	(Zou et al., 2018)	4395 (1365 M, 1565 F)	41.75 ± 14.70	BMI	FPG	HDL-C	DBP	N/A
60	(Perveen, Shahbaz, Keshavjee, & Guergachi, 2019)	667,907 (287,964 M, 379561 F)	48.35 ± 24.95	WC	FPG	HDL-C, TG	SBP, DBP	N/A

Table 2.3: Metabolic syndrome diagnosis using non-clinical approaches

S/N	Study	Classification Technique(*) best amongst multiple techniques	Performance								
			accuracy (ACC)(%)	SEN(%)	SPEC(%)	AUC	PPV(%)	NPV(%)	Others(%)	Strengths	Weaknesses
1	(Hillier et al., 2006)	PCA	N/A	N/A	N/A	N/A	N/A	N/A	50%(2.00)†		
2	(Wijndaele et al., 2006)	PCA	N/A	N/A	N/A	N/A	N/A	N/A	N/A		
3	Wijndaele et. al. (Wijndaele et al., 2009)	PCA	N/A	N/A	N/A	N/A	N/A	N/A	N/A	PCA has the ability to reduce high dimesional data to lower dimension while still retaining the relevant information imbedded in the data.	The covariance matrixin PCA is difficult to handle thereby making the principal component difficult to interprete as a risk prediction score. There is still the risk of informtion loss.
4	(Chang et al., 2010)	PCA	Female: 0.864 (0.844, 0.884)	Female: 78.829	Female: 80.19	N/A	N/A	N/A	N/A		
5	Agarwal et. al. (Agarwal et al., 2012)	PCA	N/A	54	65	N/A	N/A	N/A	cMetS cutoff: 0.475		
6	(Dusseault-Belanger et al., 2013)	PCA	N/A	N/A	N/A	N/A	N/A	N/A	PC1: 33.1, PC2: 20.2		
7	(Mochizuki et al., 2013)	PCA	N/A	N/A	N/A	N/A	N/A	N/A	PC1: 36.6		
8	(S. J. Carroll et al., 2014)	PCA	N/A	N/A	N/A	N/A	N/A	N/A	N/A		
9	(Gaio et al., 2014)	PCA N/A	N/A	N/A	N/A	N/A	N/A	N/A			
10	(Ayubi et al., 2015)	PCA	N/A	N/A	N/A	0.81	N/A	N/A	N/A		
11	Wiley et.al. (Wiley & Carrington, 2016)	PCA	69.5	51	83	N/A	N/A	N/A	N/A		
12	Gurka et.al. (NIHMS865762, 2017)	PCA	N/A	N/A	N/A	N/A	N/A	N/A	N/A		
13	(Franks et al., 2004)	Z-Score	N/A	N/A	N/A	N/A	N/A	N/A	N/A		
14	(Ekelund et al., 2006)	Z-Score	N/	N/A	N/A	N/A	N/A	N/A	N/A		
15	(Ferreira et al., 2007)	Z-Score	N/A	N/A	N/A	N/A	N/A	N/A	N/A		
16	(Johnson et al., 2007)	Z-Score	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Minimal changes in values of MetS components are measurable and accounted for in continuous MetS risk scores.	MetS risk scores derived using the Z score is sample specific and cannot be used across other populations.
17	(NIHMS501834, 2011)	Z-Score	N/A	N/A	N/A	N/A	N/A	N/A	N/A		
18	(Potteiger et al., 2012)	Z-Score	N/A	N/A	N/A	N/A	N/A	N/A			
19	(Drehmer et al., 2016)	Z-Score	N/A	N/A	N/A	N/A	N/A	N/A	N/A		
20	(Janghorbani & Amini, 2016)	Z-Score	0.613 (0.578 - 0.649)	74.1	56.4	N/A	N/A	N/A	N/A		
21	(Jiang et al., 2016)	Z-Score	N/A	N/A	N/A	N/A	N/A	N/A	N/A		
22	(K. J. Jung et al., 2017)	Z-Score	N/A	N/A	N/A	N/A	N/A	N/A	N/A		

Continued

Table 2.3 – **Continued from previous page**

S/N	Study	Classification Technique(*) best amongst multiple techniques	Performance								
			ACC(%)	**SEN(%)**	**SPEC(%)**	**AUC**	**PPV(%)**	**NPV(%)**	**Others(%)**	Strenghs	Weaknessess
23	Huo et. al. (Huo et al., 2013)	CFA	N/A	N/A	N/A	N/A	N/A	N/A	CFI > 0.96; SRMR < 0.80		
24	(Povel et al., 2012)	CFA	N/A	N/A	N/A	N/A	N/A	N/A	CFI: 0.98; SRMR: 0.039; RMSEA: 0.07	CFA uses the MetS risk factors to develop single factor models with the best goodness of fit for the diagnosis of MetS.	It requires a large sample size to describe a better generalisation of the MetS risk factors in the single model. It also requires a previous model hypothesis error.
25	(Gurka et al., 2014)	CFA	N/A	N/A	N/A	N/A	N/A	N/A	CFI:91.7	Mostly a quantitative MetS index is developed.	
26	DeBoer et.al. (NIHMS735680, 2015)	CFA	0.89	N/A	N/A	N/A	N/A	N/A	N/A		
27	(Graziano et al., 2015)	CFA, mathematical model		80	80	0.88	N/A	N/A	N/A		
28	(Musani et al., 2017)	CFA	N/A	N/A	N/A	N/A	N/A	N/A	N/A		
29	(DeBoer et al., 2018)	CFA	N/A	N/A	N/A	N/A	N/A	N/A	N/A		
30	(C. C. Lin et al., 2010)	ANN(*), MLR	81.2	85.2	78	0.908	71.9	89.2	N/A	ANNs have the ability to learn and model non-linear associations even with incomplete information.	ANN is a black box technique. Risk of MetS cannot be clinically interpreted.
31	(Hirose et al., 2011)	ANN(*), MLR	N/A	93	91	N/A	N/A	N/A	N/A		
32	(Ushida et al., 2012)	FNN	77.4	N/A	N/A	N/A	N/A	N/A	N/A		
33	(H. Chen et al., 2014)	ANN (*), PCLR	N/A	88.43	83.7	90.43	56.61	96.77	N/A		
34	(Zhao et al., 2014)	ANN (*), MLR	N/A	N/A	N/A	N/A		N/A	N/A		
35	(Ivanovic et al., 2016)	ANN	N/A	N/A	N/A	N/A	85.79	83.19	N/A		
36	(Faria et al., 2014)	MLR	N/A	94.7	94	0.84	N/A	N/A	N/A	MetS risk scores were developed using MLR as predictive scores of MetS.	MLR models assumes uniform relationships between all the MetR factors. Therefore prediction of MetS using MLR models not reliable is HDL-C has an opposite relationship with MetS.
37	(Hosseini et al., 2014)	MLinR	95.36(94.83-95.83)	89	87.93	N/A	N/A	N/A	cMetS cutoff:-1.151		
38	(Tsou et al., 2014)	MLR	N/A	N/A	N/A	N/A	N/A	N/A	N/A		
39	(Obokata et al., 2015)	MLR	N/A	78	54	80	N/A	N/A	N/A		
40	(Je et al., 2017)	MLR	N/A	81	61	0.78	14	98	N/A		
41	(Zou et al., 2018)	MLR	N/A	N/A	N/A	0.69	N/A	N/A	N/A		

Continued

Table 2.3 – **Continued from previous page**

S/N	Study	Classification Technique(*) best amongst multiple techniques	Performance								
			ACC(%)	**SEN(%)**	**SPEC(%)**	**AUC**	**PPV(%)**	**NPV(%)**	**Others(%)**	Strenghs	Weaknessess
42	(Huang, 2013)	Data Cutting and Inner Product Method (DCIP); Association Rules	N/A	N/A	N/A	N/A	N/A	N/A	N/A		
43	(de Edelenyi et al., 2008)	RF	N/A	71.4	N/A	N/A	N/A	N/A	N/A		
44	Lin et. al. (Z. Lin et al., 2014) (S. M. Liew et al., 2011)	RF	86.5	89.86	84.04	N/A	N/A	N/A	N/A	RF gives a more generisable model solution.	RF model is not easily interpretaable and cannot be used for the clinical diagnosis of MetS.
45	(Dai et al., 2015)	RF	86.32	86.22	86.44	N/A	N/A	N/A	N/A		
46	(Worachartcheewan et al., 2015)	RF	98.11	94.8	99.38	N/A	N/A	N/A	MCC: 95.0		
47	(Steinberg et al., 2014)	REFS	N/A	71.8	N/A	0.88	N/A	N/A	N/A		
48	(Jeong et al., 2014)	Areal Similarity Degree (ASD)	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Develops a risk score for the diagnosis of MetS.	The model is not stable and needs to be validated with other populations.
49	(Soldatovic et al., 2016)	simple Metabolic Syndrome score (siMS score)	N/A	N/A	N/A	N/A	N/A	N/A	N/A	siMS is simple to use.	The model is population dependent.
50	(Van Schependom et al., 2015)	SVM(*), ANN, DT	83.9	78.8	86.8	N/A	N/A	N/A	N/A	SVM is less prone to overfitting and is able to handle outliers.	SVM is a black-box technique which cannot be used to predict the risk of MetS in clinical diagnosis.
51	(Karimi-Alavijeh et al., 2016)	SVM (*), DT	75.7	77.4	74	N/A	N/A	N/A	N/A		
52	(Worachartcheewan et al., 2013)	DT (*), ANN (*), SVM, PCA, AA	99.98	99.87	99.85	N/A	99.87	99.85	MCC: 0.9972	DT models are interpretable and can handle colinearity.	The performance of DT models is higher with categorical dataset whereas the MetS risk factors are continous.
53	(Miller et al., 2014)	CHAID(*), MLR	92.3	71.8	N/A	N/A	69.8	N/A	PPV: 69.8		
54	(Romero-Saldana et al., 2016)	DT	94.2	91.6	95.7	N/A	79.2	98.5	N/A		
55	(Shimoda et al., 2018)	GBT(*), RF ,LR	N/A	87.7	78.3	89.3		N/A	N/A		

Continued

Table 2.3 – **Continued from previous page**

S/N	Study	Classification Technique(*) best amongst multiple techniques	Performance								
			ACC(%)	**SEN(%)**	**SPEC(%)**	**AUC**	**PPV(%)**	**NPV(%)**	**Others(%)**	Strenghs	Weaknessess
56	(Jia et al., 2018)	Markov Model	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Generates a model by using all the posssible states of the input data.	Markov model cannot generate a risk score for the clinical diagnosis of MetS.
57	(Miyauchi & Nishimura, 2018)	BN	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Perform very well on both small and large sample datasets.	BN have a limited capacity in developing models with continuous datasets.
58	(Perveen et al., 2019)	Naïve Bayes(*), J48		71.1	28.7	0.73	64.9	N/A	F-Measure: 69.3; MCC: 14.7		
59	(Kang et al., 2012)	Framingham risk score	0.80 (0.75 - 0.84)	87.8	58	N/A	N/A	N/A	cMetS cutoff: 13	FRS has been validated as tool for the diagnosis of CVD.	FRS is population dependent and is meant for only the prediction of CVD.
60	Yousefzadeh et. al. (Yousefzadeh et al., 2015)	Framingham risk score	N/A	N/A	N/A	N/A	N/A	N/A	N/A		

2.4.1.1 Metabolic Syndrome Risk Factor Distribution

The MetS risk factors identified by the clinical definitions (See Table 1.1) are generally classed into invasive and non-invasive measures (Romero-Saldana et al., 2016). Fasting blood glucose (FBG), HDL-C , and Triglycerides (TG) are considered as invasive variables because they require the drawing of blood sample and analysis. The non-invasive risk factors are waist circumference (WC) and blood pressure (BP). WC is used to identify central obesity in MetS . Other non-invasive measures used to identify central obesity in the primary studies include waist-to-hip ratio, body mass index, and waist-to-height ratio. Blood pressure is identified by measuring systolic and diastolic blood pressure readings. In addition to the clinically identified MetS risk factors, some of the studies included in this systematic review have identified and used other risk factors in the non-clinical methods such as such as age, sex, smoking habits, literacy rank, physical activity, and alcohol consumption. A single study, Lin et al., 2014 (Z. Lin et al., 2014) collected serum samples and extracted single nucleotide polymorphisms (SNPs) associated with MetS traits for use as input parameters in their studies. The remaining 33 studies used all the five clinically defined MetS risk factors in Table 1.1.

2.4.1.2 Performance Metrics Used in Included Studies

The most frequently used metrics for comparing and evaluating the performance of the non-clinical methods in the included studies are ACC, SEN, SPEC, PPV, AUC, and NPV. Table 2.4 below presents the definitions of each performance metric.

2.5 What is the current state of art in non-clinical methods for the diagnosis of metabolic syndrome?

The specification of a comprehensive taxonomy is a valid indication of the detailed assessment and quality of extracted data in a systematic literature review (Malhotra, 2015).

Table 2.4: Description of performance metric measures

Performance Metric	Description
ACC	The proportion of individuals correctly classified as having or not having the diseases over the total number of all the individuals examined.
SEN	The ability to correctly classify individuals with the disease as having the disease.
SPEC	The ability to correctly classify individuals without the disease as not having the disease.
AUC	It is the area under the ROC curve. ROC is plotted with SEN (true positive rate) on the Y-axis and 1-SPEC (false positive rate) on the x-axis.
PPV	The probability of individuals predicted as having the disease that actually have the disease.
NPV	The probability of individuals predicted as not having the disease that actually do not have the disease.
Comparative Fitness Index (CFI)	An incremental fit index which analyses the model fit by examining the variance between the data and the hypothesized model.

In this study, we discuss the state of art by iteratively examining and extracting relevant information from the primary studies.

Although the 36 studies included in this systematic review have developed diagnosis models using various types of techniques for the non-clinical diagnosis of MetS , the reason why these non-clinical methods are required to support the existing clinical dichotomous diagnosis approach should be visited. The rationale for developing these alternative non-clinical diagnosis methods need to be identified in order to fully understand and analyse the current state of art in the area of MetS diagnosis.

The clinical definitions of MetS shown in Table 1.1 are used to diagnose MetS by dichotomising the measurement values of the five clinically recognised MetS risk factors (FPG, WC, HDL-C , TG, BP-SBP, DBP). MetS is considered present when the requisite number of MetS risk factors that exceed certain threshold is met (WHO, 1999; Balkau &

Charles, 1999; Expert Panel on Detection, 2001; K. G. M. M. Alberti et al., 2006). However, the MetS risk factors' measurement values in these definitions are more continuous than categorical and this results in loss of information on the outcome of the MetS diagnosis (Kahn, 2007; Ferreira et al., 2007). Therefore, the clinical definition reduces statistical power (Ragland, 1992; Wijndaele et al., 2006) and patients that have MetS may be excluded. Furthermore dichotomising the continuous MetS risk scores implies that all the MetS risk factors contribute equally to the diagnosis (Hillier et al., 2006), however the predictive ability of some MetS risk factors towards CVD is higher than others (Simmons et al., 2010). Dichotomising the continuous MetS risk factors based on ad hoc thresholds could lead to mis-classification of the disorder, thereby reducing both statistical power and correlative measurement of the MetS risk factors (Ragland, 1992; Wijndaele et al., 2006). Also, summing up the MetS risk factors into a unitary value assumes that all the risk factors contribute equally, yet some MetS risk factors are known to have more importance than others (Simmons et al., 2010). Research has shown that there is a progressive relation between MetS and CVD which might be unidentified by dichotomising the MetS risk factors (Kahn, Buse, Ferrannini, & Stern, 2005; Wijndaele et al., 2006). Furthermore, both CVD and T2DM increase progressively as the number of MetS risk factors that exceed the threshold increase, thus eliminating to apply the dichotomous definition in the diagnosis of MetS (Klein, Klein, & Lee, 2002; Kahn et al., 2005; Wijndaele et al., 2006).

From the primary studies included in this systematic review, the earliest solution used to solve the limitation of the dichotomous clinical definition was the development of a unitary continuous score (Franks et al., 2004; Hillier et al., 2006) referred to as the continuous MetS risk (cMetS R) score. The cMetS R score was developed using three types of statistical techniques, the principal component analysis (PCA), standardised zscore and the confirmatory factor analysis (CFA). Most of the primary studies that derived a

cMetRS score using statistical techniques use the score to measure the association between interventions, and MetS because the use of the dichotomous definition limits association detection power (Franks et al., 2004; Johnson et al., 2007). Therefore, the cMetS R score is argued to represent and detect overall changes in the MetS dues its sensitivity to small changes in the values of the MetS risk factors (Johnson et al., 2007). However the cMetS R score was limited by its over dependence on the population sample from which it was built. Subsequently, machine learning techniques were then used in the primary studies to develop models for the non-clinical diagnosis of MetS . Intermittently, mathematical quantification techniques were also developed as quick measures of diagnosing MetS .

From this systematic review, we have identified a detailed taxonomy of the techniques that have been used to develop various non-clinical methods to support the clinical diagnosis of MetS . This taxonomy and classification is presented in Fig 2.2.

2.5.1 Statistical Techniques

In this section, we discuss the state of art of the 17 studies that used statistical techniques. These techniques include Confirmatory Factor Analysis (CFA), Principal Component Analysis (PCA), principal components logistic regression (PCLR), generalized multifactor dimensionality reduction (GMDR) and Z-score.

2.5.1.1 Principal Component Analysis

Dimension reduction models are approaches for data integration that best explains the structure of datasets, and the variance both within and between variables (Meng et al., 2016). Existing data is reduced into new variables known as components. These components best explain the difference in observations of a dataset. Two dimension reduction models were identified from the included studies: principal component analysis

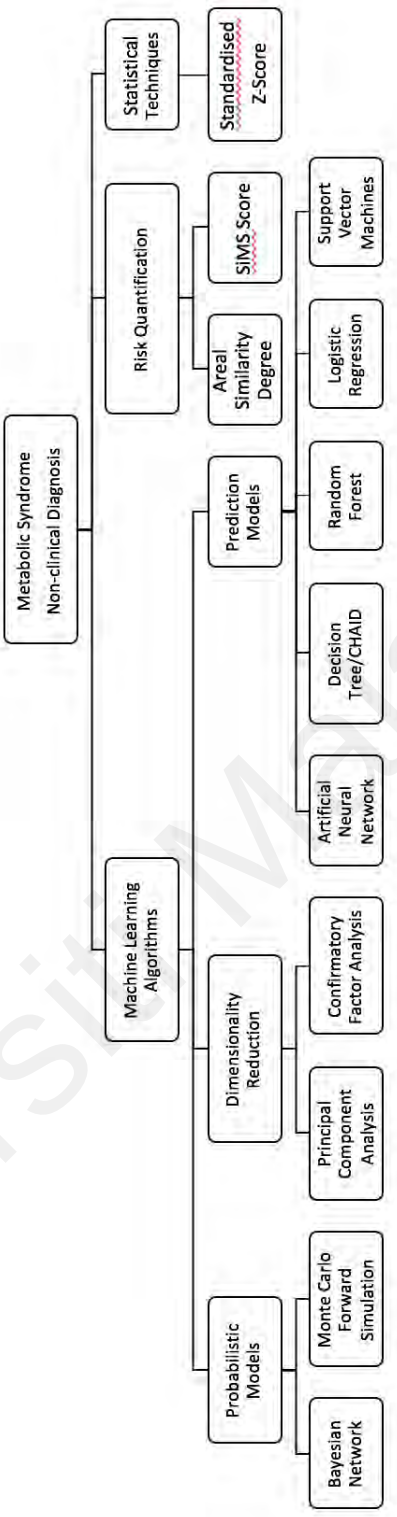

Figure 2.2: Classification and taxonomy of non-clinical methods

and confirmatory factor analysis. From the included studies, 6 studies used PCA, and 5 studies used CFA to calculate continuous MetS risk scores (cMetS R scores). One study used PCA and logistic reg as PCLR. The continuous MetS risk score is a statistical score

that was proposed to support the diagnosis of MetS (Eisenmann, 2008) by tackling the limitations of the dichotomous definition.

PCA is a multivariate statistical technique that reduces the dimensions of observations in a dataset (Jolliffe & Cadima, 2016). PCA is able to project the observations into a one-dimensional space that explains the variance-covariance structure of the variables in a data. The variance of the data is explained by extracting the most important information (attribute) from the observation. PCA is used to predict cumulative risk scores which enable the informative description of disease history and the development of appropriate prevention and management strategies (Vikram et al., 2009).

In the studies included in this review paper, the PCA is applied to analyse the structure of the MetS risk factors and the variability of their associations with MetS . Because the first principal component of the PCA is the linear sum of measures with the maximum possible variance, researchers have used it to identify the cMetS R score (Hillier et al., 2006) in MetS diagnosis. Using PCA, the continuous MetS risk score (cMetS R score) is derived as the linear sum of the principal component (PC) that explains the maximum total variance, which is mostly the first principal component (PC1).

Hillier et al., (Hillier et al., 2006) created a nomogram by combining the sum of six standardised MetS risk factor values weighted by the cMetSR score. The cMetSR score is the first principal component defined from the principal component analysis which explained 50% of the variance among the MetS risk factor values. They concluded that their PCA derived cMetSR score was able to predict the incidence of diabetes even in people whose FPG values were below the clinical threshold. Moreover, Wijndaele et. al. in (Wijndaele et al., 2006) and (Wijndaele et al., 2009) also concluded that their PCA derived cMetSR score is an effective measure of MetS analysis. Chang et al., (Chang et al., 2010) developed an easy-single parameter screening index called the first principal

component score (FPCS) by reducing only obesity (WC and BMI) and age into a single variable using PCA. This index precludes the need for using all the five MetS risk factors to test for the risk of MetS in the clinical definition. Even though the FPCS was found to yield an AUC of 0.864, the computation of this parameter is inefficient for application in clinical settings. Furthermore, Agarwal and colleagues, (Agarwal et al., 2012) investigated the use of PCA in deriving a cMetSR score as a summary of the MetS risk factors and its relation to the incident of T2DM and CVD. The PC1 from their PCA model explained 33% of the total variance of the MetS risk scores. They observed that the binary cMetSR score was a better predictor of CVD than the dichotomous definition (ATP-NCEP III). The cMetSR score in (Kang et al., 2012) was calculate by summing up the scores derived by assigning points to each MetS risk factor based on the size of its regression coefficient from a Cox proportional hazards model. The cMetSR was found to be useful in the prediction of CVD using MetS risk factors. However,the cutoff point used is specific to only their study's participants. Mochizuki et al., (Mochizuki et al., 2013) perform PCA on five metS risk factors. They extracted two principal components, PC1 and PC2, explained 336.6 and 21.9% of the variance in the population, respectively. Carroll and colleagues, (S. J. Carroll et al., 2014) constructed a continuous clinical index of cardiometabolic risk (cCICR) by first standardising each MetS risk factor using the Z-score to obtain the cCICR-Z. Then, they applied PCA with orthogonal rotation on the standardised MetS risk factors before finally taking the weighted sum of the two principal components (PC1 and PC2) to obtain a cCICR-PCA. This study identified that the total variances explained by their two PCs was higher in men (61.69%) than women 60.14%. They also demonstrated that both their cCICR-Z and cCICR-PCA had higher accuracy in predicting the risks of CVD , T2DM and MetS more adequately the clinical dichotomous method. The cMetSR score was also computed by Gaio and colleagues, (Gaio et al., 2014) using PCA to identify the genetic

factors associated with MetS. The first two principal component score, weighted by their relative contribution in the explained variance was summed up to obtain the cMetSR score. The score was able to explain the association of over 50% of the genetic phenotype with MetS and establishing a significant association.

The cMetS R score was computed by Gaio and colleagues, (Gaio et al., 2014) using PCA to identify the genetic factors relevant to the risk of MetS . The first two principal component score, weighted by their relative contribution in the explained variance was summed up to obtain the cMetS R score. The score was able to explain the association of over 50% of the genetic phenotype with MetS and establishing a significant association. Wiley and colleagues, (Wiley & Carrington, 2016) developed a MetS severity score (MetS SS) also using PCA stratified by age, gender, medication and work overtime. PCA was applied to MetS risk factor values standardised against clinical thresholds. The MetS SS was able differentiate between adult with and without MetS by correctly identifying 82% of adults with MetS . However, MetS SS requires further validation in different population groups. The cMetS R was found to be useful in the prediction of CVD using MetS risk factors. However,the cutoff point used is specific to only their study's participants.

In (Ayubi et al., 2015), the PCA on MetS components identified components with Eigenvalues ≥ 0.9, with 75% and 75% of the variance in males and females, respectively. Dusseault-Belanger et al., (Dusseault-Belanger et al., 2013) examined the correlation structure of MetS using principal component analysis. The first dimensionality explained 30% of the variance an was used for as the continuous MetS risk score. The score identified MetS with a higher predictive value over the clinical definition. Even though the principal components show associations of MetS with the clinically recognised risk factors, it is still not clear if it is a measure of progression of MetS and its related diseases (Mochizuki et al., 2013) Gurka et. al. (NIHMS865762, 2017) developed a sex and race-specific MetS

severity score (MetSSS) to assist in the prediction of T2DM over a time frame of 7.8 years. The MetSSS. The study found out that there an association between the increase of MetS measured by the MetSSS and risk of T2DM if observed progressively over time.

2.5.1.2 Confirmatory Factor Analysis

CFA is a statistical technique that enables the possibility of assessing existing associations between measured variables by measuring a model's goodness of fit (Babyak & Green, 2010). The CFA constructs a hypothesized model by linking various risk factors with hypothesized latent variables (Babyak & Green, 2010). In the included studies, the CFA was used to develop single factor models using MetS risk factors for the identification and diagnosis of MetS. CFA was used to establish the relationships between MetS risk factors and MetS in order to ascertain the validity of a composite MetS risk score construct.

Gómez-Marcos and co-workers, developed four different cardio metabolic risk index models to diagnose MetS using CFA. All the four models consisted of TG/HDL-C ratio, HOMA-IR index, MAP and each with a different measure of central obesity - WC, Waist-to-Height ratio, BMI, adiposity index. The model with waist circumfrerence had the best metabolic syndrome index with an average value of -0.022 ±1.29 (-3.36 - 4.57) in men and the model with the body mass index showed the best goodness of fit with a metabolic index of 0.0001 ± 1.53(-3.17 - 5.55) in women. The risk index was used to find associations between the MetS and physical activity.

In (Huo et al., 2013) the CFA was used to compare two models of MetS in a chinese population. Both models consisted of WC, TG/HDL-C ratio, FPG, but with different measures of blood presssure - mean arterial pressure (MAP) in Model 1 and systolic blood pressure in Model 2. WC had the highest loading both models. This reiterates the significance of central obesity in the diagnosis of MetS (K. G. M. M. Alberti et al., 2006). Model 1 showed the highest good fitness with a comparative fit index of

more than 0.96 and a standardised root mean square residual of less than 0.8. Smits and colleagues (Smits et al., 2013), CFA was aplied to link new variables - adipocytokines, CT-measured intra-abdominal fat (IAF), and insulin sensitivity (SI) as underlying factors of MetS . The model the one factor model consisting of 6 MetS risk factors - CT-measured intra-abdominal fat (IAF), insulin sensitivity (SI), systolic blood presure, diastolic blood pressure, triglycerides and HDL-C had the best CFI of 0.99. However, the use of IAF and SI makes their model difficult and slow to implement in clinical settings due to the high costs and increased accessibility time of the these risk factors. In Povel et. al. (Povel et al., 2012) CFA was used to develop a good MetS model-fit for the prediction of CVD and T2DM. The CFA model was able to predict T2DM and CVD better than the clinical definition with an integral discrimination index (IDI) of : 0.34 and 0.07) respectively. Gomez-Marcos et. al. (Gomez-Marcos et al., 2013) developed a single factor model using CFA to predict the risk of MetS. Gurka et al. (Gurka et al., 2014), DeBoer and coworkers, (NIHMS735680, 2015), Gurka and colleagues, (NIHMS865762, 2017), and Musani et al. (Musani et al., 2017) all used CFA to calculate a MetS severity score. This MetS severity score is a continuous risk score developed from a one factor CFA model consisting of the factor loadings of all the five clinically recognised MetS risk factors. In Gurka et al. (Gurka et al., 2014), the MetS severity score revealed multiple variations in how each MetS risk factor contributes to the overall MetS score based on racial/ethinic grouping. The CFA was performed on z-score standardised MetS risk factor values. With an AUC ranging from 0.77, the cMetSR score was able to predict diabetes in (NIHMS865762, 2017). The linear associations between the MetS severity score in childhood and adulthood identified the MetS severity score as a tool for the prediction of CVD (DeBoer, Gurka, Woo, & Morrison, 2015) and as predictor of T2DM (NIHMS735680, 2015). The cMetS R score also showed a high genetic correlation with MetS in (Musani et al., 2017). However, the

diagnosis of MetS using the cMetS R score is heavily dependent on the score cut-off to either emphasise specificity or sensitivity. Out of the five studies that applied the CFA, four studies ((Huo et al., 2013), (Povel et al., 2012), (Smits et al., 2013), and (Gurka et al., 2014)) reported the root mean square error of approximation (RMSEA), the standardised root mean square residual (SRMR), and the comparative fit index (CFI) as presented in Table 2.4. (Gomez-Marcos et al., 2013) reported an AUC of 89%. CFI, RMSEA and SRMR are indices that evaluate the goodness of fit a statistical model. A model is said to have a good fit if the CFI > 0.96, the RMSEA < 0.050 and the SRMR < 0.080 (Hu & Bentler, 1999). The CFI reported for the four studies ranged from 0.917 to 0.991. The SRMR ranged from 0.134 to 0.0212 and the RMSEA was from 0.125 to 0.045. These results show that all the four studies had good model fits.

2.5.1.3 Z-Score

The Z score is another statistical technique identified from the included studies in this systematic review. It is used to compute a continuous MetS risk (cMetS R) score which represents the presence of MetS in a population sample. The Z score is computed by subtracting the sample mean from each sample value and dividing by the sample standard deviation. The higher the cMetS R score, the less favourable the MetS profile. In some studies, each of the risk factors have been regressed on age, race, and gender to account for age, race and gender-related differences in the risk factors (Eisenmann, 2008). It was widely used to calculate a cMetSR score in previous studies (Franks et al., 2004; Ferreira et al., 2007; Eisenmann, Wickel, Welk, & Blair, 2005; Johnson et al., 2007; NIHMS501834, 2011; Potteiger et al., 2012). Franks et al., (Franks et al., 2004), derived a cMetSR score by summing up Z scores derived from each MetS risk factor in order to identify the relationship between physical activity and MetS. The cMetSR score confirmed that an increase in the physical activity of individuals significantly decreased their risk of MetS.

Similarly, Ferriera et al. (Ferreira et al., 2007) observed the association of arterial stiffness and MetS using a cMetSR score. The cMetSR score was derived by calculating the Z scores of each MetS risk factor and taking the average. They concluded that MetS is a significant predictor of arterial stiffness. Johnson et al. (Johnson et al., 2007), Bateman and colleagues, (NIHMS501834, 2011) Potteiger et al., (Potteiger et al., 2012) also derived a cMetSR z-score using the NCEP ATP III thresholds and the standard deviation (SD) of the data from the entire population sample. The resulting equation is $cMetSRz - score = [(40 - HDL)/SD] + [(TG - 150)/SD] + [(fastingbloodglucose - 100)/SD] + [(waistcircumference - 102)/SD] + [(meanarterialpressure - 100)/SD]$ for men and $cMetSRz - score = [(50 - HDL)/SD] + [(TG - 150)/SD] + [(fastingbloodglucose - 100)/SD] + [(waistcircumference - 88)/SD] + [(meanarterialpressure - 100)/SD]$ for women. The cMetSR z-score showed that MetS was significantly lower ($p \leq 0.05$)in participants with higher levels of exercise(Johnson et al., 2007). (Janghorbani & Amini, 2016) and (Jiang et al., 2016) investigated the utility of the cMetS R score computed as the standardised residuals of each MetS risk factors. The cMetS R score in (Janghorbani & Amini, 2016) ranged from -8.98 to 17.57 with the upper bound indicating a higher risk of MetS. (Jiang et al., 2016) showed that the cMetRS score was able to identify higher levels of MetS in the future based on preserved information from the history of the same participants. Thus, this suggests that cMetS R score can be applied for progressive monitoring of MetS over time. In (Drehmer et al., 2016), Drehmer and co-workers calculated a cMetS R score as the means of Z scores of the continuous MetS risk factors. The cMetS score was used to identify if any association exists between dairy consumption, and fat intake, and MetS.

2.5.2 Machine Learning Techniques

Fifteen studies used different types of machine learning techniques to diagnose MetS. In this section, the methods extracted from the primary studies which include DT, DT

Chi-Squared Automatic Interaction Detection (CHAID) Random Forest (RF), SVM, ANN and BN will be described.

2.5.2.1 Logistic Regression

Logistic regression is a statistical technique used to build prediction models for predicting values of a dependent variable from independent variables of a set of predictor values. It is a linear regression variant often applied for prediction problems where the dependent variable is a. The value for prediction is the probability of an event, ranging from 0 to 1. Logistic regression also estimates risk prevalence ratios of presence of a disease and differences variable contributions to the prediction model.

In (Hosseini et al., 2014), binary logistic was used to construct and validate a CMetSR score for the diagnosis of MetS in Iranian Adults. In a bottom - up, CMetSR score models were built from using two risk factors incrementally util all the fiver risk factors were included in the prediction model. The CMetSR score that includes all the 4 MetS risk factors including age and gender showed the highest performance with an AUC of 95.5%. (Obokata et al., 2015), used MLRto calculate a composite risk score for MetS prediction using data from Japanese employees. The MetS risk score was used to identify the incidence of MetS in a three year follow up of the population sample. A univariate logistic regression was carried out to identify variables that would be included derivation of the composite risk score. In citeTan2016, logistic regression was used to derive a cMetSR score using data collected from questionnaires. The resulting cMetSR score was able to predict MetS with an AUC of 94.2%, sensitivity of 90% and specificity of 74%. (Tsou et al., 2014) applied MLR to identify the relationships between the MetS risk factors and the risk of having MetS in elderly people. The results show a significant positive association between all the MetS risk factors and MetS. Thus, measurement of central obesity and high blood pressure may result in the early prevention of MetS.

2.5.2.2 Artificial Neural Networks

ANN is an abstract computational structure which models nonlinear problems based on the human brain. An ANN model consists of nodes called the artificial neuron which are interconnected in a network of layers. The neuron is a simple structure capable of receiving multiple input signals via its connections but it can only output one signal. The typical ANN network constitutes of an input layer, an output layer and sometimes with several intermediary layers (hidden layers) in-between. The neurons between layers are linked by weighted connections which pass signals from one neuron to another. Six studies out of the fourteen primary studies applied ANN in their MetS prediction.

In both studies, the ANN models were trained and evaluated by dividing the dataset into training and testing sets respectively (Hirose et al., 2011; H. Chen et al., 2014). The ANN model developed by Chen and colleagues was compared with a statistical technique, the principal component logistic regression classification model and the ANN model had a higher AUC value of 90.43% (H. Chen et al., 2014) . The sensitivity of the ANN model was also higher with 88.49% than the 52.89% of the PCLR. Hirose et al. compared their ANN model with a multiple logistic regression (MLR) model and the ANN model outperformed the MLR model with a sensitivity of 93% (Hirose et al., 2011). Overall, we see than the ANN model of Hirose et al. outperformed that of Chen and colleagues by 4.51%. This could be attributed to the isolation of three important MetS risk factors in Chen et al. -FPG, HDL-C and TG (H. Chen et al., 2014). This information loss will result in a model with low sensitivity.

Lin et al. (C. C. Lin et al., 2010), Murguía-Romero and colleagues, (Murguia-Romero, Jimenez-Flores, Mendez-Cruz, & Villalobos-Molina, 2013) and Ivanovic et al., (Ivanovic et al., 2016) proposed ANN models trained with the back propagation algorithm to predict MetS. They explored a 3 layer network consisting of between 1 to 100 neurons in the

middle layer. The network with the number of neurons with the highest PPV was chosen as the best model. Lin et al. (C. C. Lin et al., 2010) compared their ANN model with a multiple logistic regression model and the ANN model had a higher AUC of 93.4% and a higher PPV of 67.5%. The ANN model proposed by Murguía-Romero and colleagues had a PPV of 43.4% (Murguia-Romero et al., 2013). The highest PPV in Murguía-Romero and colleagues was 45.4% in the network with 6 hidden neurons while in Ivanovic et al., the highest PPV was 85.79% in the network with 96 hidden neurons. However, in both (Murguia-Romero et al., 2013) and (Ivanovic et al., 2016), comparison with other models was not evident in the literature. It suffices to say the model built by Lin et al. (C. C. Lin et al., 2010) has the highest PPV amongst these three studies because it included all the five MetS risk factors as input parameters.

.

In a slightly different approach, Ushida et al., (Ushida et al., 2012) applied a fuzzy neural network (FNN) to detect the significant combination of MetS risk factors that are most highly associated with MetS. The FNN is a multilayer feed-forward network with fuzzy logic as the activation function. The structure of a simple FNN consists of input nodes, membership nodes, rule nodes and output nodes in a four layer network structure. They were able to identify that the combination of γ-GTP level and the white blood count was significantly associated with MetS. The comparison of the FNN result with that of a multiple logistic regression model confirmed the significance of the MetS risk factor combination obtained from the FNN model. The FNN model outperformed the multiple linear regression model by 1%. However, the study data in Ushida et al. consists of male subjects only (Ushida et al., 2012).

(Zhao et al., 2014) and (Ivanovic et al., 2016) developed various ANN network models trained with the back propagation algorithm to predict MetS. BP enables small repetitive

consistent adjustments of the weights to reduce overall error in the network. The ANN model was trained and evaluated by dividing the dataset into training and testing sets respectively (H. Chen et al., 2014). The ANN model developed by Chen and colleagues was compared with a statistical technique, the principal component logistic regression classification model and the ANN model had a higherAUC value of 90.43% (H. Chen et al., 2014). The SEN of the ANN model was also higher with 88.49% than the 52.89% of the PCLR. However, three important MetS risk factors - FPG, HDL-C and TG were isolated from the ANN model in Chen et al. (H. Chen et al., 2014). In (Zhao et al., 2014), the back propagation ANN model was used to select single nucleotide polymorphisms (SNPs) that were associated with MetS. Their ANN model when compared with another model built with MLR showed a higher significance in the prediction of MetS. However, obtaining SNPs may not be cost and time efficient in clinical settings. (Ivanovic et al., 2016) explored a 3 layer network consisting of 1 to 100 neurons in the middle layer. The model with 96 hidden neurons had the highest PPV of 85.79%. However, comparison with other models was not evident in the literature and the model was built with only two MetS risk factors - waist-to-hip ratio and blood pressure (Ivanovic et al., 2016). The back propagation ANN analysis results in black box model which has limited ability in explicitly identifying possible causal relationship (Tu, 1996).

2.5.2.3 Decision Trees

Three studies investigated the use of decision trees to diagnose MetS and identify combinations of the MetS risk factors significantly associated with its prediction. DT are trees that classify data by recursive partitioning into hierarchical or sequential structures (Murthy, 1998). DTs reduce the volume of data into an accurate informative summary consisting of the most important characteristics of the data. The DT consist of nodes and each node represents a decision rule that may split into two or more partition. These rules

are automatically constructed and can be used for inferential decision making in clinical diagnosis. Splitting of nodes occurs iteratively until a stopping criterion has been reached for the terminal node which gives the final predicted response. Despite the ability of DT models graphically display the tree, this classification algorithm does not support online learning. This means every time a new input comes in, the model will have to be generated from scratch, making it computationally expensive.

In (Romero-Saldana et al., 2016), Romero-Saldańa and colleagues used only non-invasive MetS risk factors - WtHR and blood pressure. WtHR and blood pressure were identified as the MetS risk factors with the highest association to the risk of MetS in the decision tree model. They validated the method against the NCEP ATP III MetS definition and reported an accuracy of 94.2%, a sensitivity and specificity of 91.6% and 95.7% respectively. However, (Miller et al., 2014) utilised all the five clinically accepted MetS risk factors used to define MetS . They reported a higher classification accuracy of 92.3% than Romero-Saldańa and colleagues. This could be attributed to their adoption of more MetS risk factors (5) which should expectedly yield a better performance decision tree model.

A Naive tree model for predicting MetS was formed by (Van Schependom et al., 2015) in psychiatric patients. The naive tree model outperformed ANN, MLR and SVM) models with an accuracy, sensitivity and specificity of 83.9%, 78.8%, 86.8% respectively. The naive tree model was based on only simple non-invasive MetS risk factors - obesity and blood pressure. They conclude that MetS can be diagnosed using less complicated machine learning techniques with non-invasive risk factors. However, the performance of their model is highly dependent on schizophrenic patients with decreased central obesity.

2.5.2.4 Random Forest

Paul et al., (Paul, Ueno, Iwata, Hayashi, & Honda, 2008) proposed the probabilistic model building genetic algorithm (PMBGA) together with k-nearest neighbour (kNN) and DT(C4.5) to identify the health risk and two risk factors of MetS: blood pressure and triglyceride. The PMBGA plus C4.5 hybrid model outperformed the baseline method with an AUC of 70%. Three studies, Szabo de Edelenyi et al., (de Edelenyi et al., 2008), Lin et. al. (Z. Lin et al., 2014) and Worachartcheewan et. al. (Worachartcheewan et al., 2015) applied the RF tree algorithm to predict the presence of MetS, determine its prevalence and find significant risk factors related with the presence of MetS. RF is an ensemble method which combines several individual decision trees for classification and prediction (Breiman, 2001). The DT are trained by generating several bootstrap samples from the training dataset Then each bootstrap sample is fit with an un-pruned DT. RF generalises by using the bagging strategy to build each decision tree independently, thereby decreasing variance (Strobl, Boulesteix, Zeileis, & Hothorn, 2007). The ensemble of individual trees makes adjustment for the instability of individual trees thereby increasing the efficiency of the RF method. Nevertheless, model interpretation from the RF trees is more complex than from individual decision trees because the influence of the risk factors do not directly correspond to the risk factor's position in the tree. The Gini index identified TG as the most important risk factor in the model. The model with 20 trees resulted in an accuracy, sensitivity, and specificity of 98.12%, 94.80%, and 99.15% respectively on an external dataset.

In the case of (Z. Lin et al., 2014), they investigate the metabolic profiling changes using serum samples in MetS. First the MetS serum sample was analysed using Gas chromatography–mass spectrometry (GC-MS). Then RF models were created using metabolites from the metabolic profiling. Each MetS risk factor's level of contribution

to the RF model was calculated. A proximity matrix which identifies the structure in the data was used to construct multi-dimensional scaling (MDS) plots. Similar samples have high proximity. The accuracy for the RF model in (Z. Lin et al., 2014) was 86.5% while the sensitivity and specificity were 89.86% and 84.04% respectively showing significant discrimination between individuals with MetS and healthy controls. In Worachartcheewan et. al. (Worachartcheewan et al., 2015) data was divided into 2 subsets using PCA. The first subset is an internal dataset used for training the RF model by applying the 10-fold cross validation procedure. The second subset is an external dataset for evaluation of the RF model and FPG, WC, and BMI were the MetS risk factors with the highest association to MetS according to the RF model. The importance of each MRF was evaluated using the Gini index. The accuracy, sensitivity and specificity of their RF are 98.02%, 94.81% and 99.07% respectively.

2.5.2.5 Decision Trees with CHAID

Two studies Miller et al., (Miller et al., 2014), (Romero-Saldana et al., 2016) proposed the use of decision trees with CHAID methodology for the early detection of MetS . The chi-squared automatic interaction detection method is an algorithm used for finding patterns in datasets by merging, splitting and finally applying a user-specified stopping criteria(Kass, 1980). It is based on a stepwise regression for split selection procedure based on a chi-square test statistic like a sequential cross-tabulation (Loh, 2014). In clinical analysis, it classifies what risk factors are associated with the clinical outcome. The algorithm specifies the combination of risk factors that best predict the binary outcome of a disease. Search takes place sub-optimally to reduce computational time.

2.5.2.6 Support Vector Machines

(Karimi-Alavijeh et al., 2016) explored the use of both SVM and DT to predict the risk of MetS. SVM is a linear classifier that creates an optimal hyperplane which separates samples of two classes using least square regression (Vapnik, 2013; Cortes & Vapnik, 1995). It can work with mixture of both numerical and categorical data. Even though it is has accurate predictability, it is a black box technique that disallows interpretation of the classification model. The SVM model proves to outperform the DT model with accuracy, sensitivity and specificity of 75.7%, 77.4% and 74.0% respectively.

2.5.2.7 Bayesian Network, Markov Model and Naive Bayes

BN is a probabilistic modelling algorithm based on Bayes' Theorem which defines the probability of an event given the occurrence of another related event. It categorises data by monitoring the probabilities that specific features are related to specific classifications. The BN leverages on its ability to analyse results into meaningful information given an existing knowledge domain. It has the ability to handle uncertainty in complex problems. (Miyauchi & Nishimura, 2018) applied BN modelling to connect information from specific health check-up data from Japan. Information from the BN was used to provide lifestyle advice to patients identified as being at risk of MetS. The BN model was confirmed as being a useful support tool in specific checkup and guidance system. It empowers individuals to find problems in their lifestyle and appropriate medical and health solutions easily. Jia et. al. (Jia et al., 2018) developed a 12-state Markov model to predict the risk of MetS in a Chinese population. Their findings identified that the risk of MetS starts with the states of overweight/obesity, hypertension or hyperglycemia, and is followed by the state of dyslipidemia as MetRFs . J48 decision tree and Naïve Bayes methods were used to developed MetS models for the long term prediction of T2DM in Perveen et. al. (Perveen

et al., 2019). The models were developed using various sampling methods used to balance the training datasets. Random under-sampling, random over-sampling, and K-Medoids under-sampling were applied in then study. Both machine learning models were able to predict future T2DM in patients. However, the Naive Bayes model outperformed the J48 model with higher AUC of 86% in the prediction of diabetes using the MetSR factors.

2.5.3 Risk Quantification Models

The ASD was a mathematical technique proposed by (Jeong et al., 2014) for MetS risk quantification. The ASD is a similarity analysis between the MetS risk factor thresholds and MetS risk factor sample measurements in a weighted radar chart. The outcome of the similarity analysis is value which determines the presence or absence of MetS based on a defined cut-off value. Although the proposed model was able to diagnose MetS in individuals with borderline measurement presentations, it is sensitive to the frequency of the population sample and the positioning of the MetS risk factors on the weighted radar chart. (Soldatovic et al., 2016) developed and evaluated the siMS score, continuous MetS score: siMSscore $= 2 * Waist/Height + Gly/5.6 + Tg/1.7 + TA_{systolic}/130 - HDL/1.02$ or 1.28 (for male or female subjects respectively). There was a high correlation between the siMS score and the cMetSR scores derive from both Z-score and PCA. The siMS also outperformed the cMetSR scores with an AUC of 92.6%. (Yousefzadeh et al., 2015) and (Kang et al., 2012) investigated the use Framingham risk score (FRS) for the prediction of MetS. TheFRS is clinical tool used to asses the risk level of coronary artery disease and identifying the chance of developing any CVD in long-term (P. W. Wilson et al., 1998; S. M. Liew et al., 2011; Sullivan, Massaro, & D'Agostino Sr, 2004). There was a significant association between the FRS and the presence of MetS in predicting the risk of CVD in both men and women, 39.5% and 18% respectively. The odd ratio of risk of MetS was 6.7 in the high-risk FRS group ($P < 0.001$).

In 2018,

2.6 Discussion and Future Guidelines

The cMetSR score, developed using Z score, PCA and CFA, is a unitary score that has been determined to have a higher MetS risk diagnosis result than the clinical dichotomous definition (Olza et al., 2015). A high score indicates a high MetS risk while a lower score is an indication of a less alarming risk of MetS. It is also capable of predicting the incidence of T2DM (Janghorbani & Amini, 2016; Magnussen et al., 2016; S. J. Carroll et al., 2014) and CVD (Hillier et al., 2006) compared with the clinical dichotomous definition. Currently, the cMetSR score was frequently used to determine the association between MetS and other emerging risk factors. However, despite its ability to maximise statistical power (Ragland, 1992) on the cut-off point of the MetS risk factors by reducing loss of information, the cMetSR score is limited in its diagnostic capacity because all MetS risk factor measurements are assumed to have an equal contribution into the diagnosis of MetS (Thivel et al., 2009). Furthermore, the cMetSR score is constrained in its application because it is a sample - specific statistical measure. This indicates that the individual score of a single patient cannot be the same in two different studies. There is also no provision to compare mean scores derived from two different studies due to differences in demographics distributions, measures of central tendency and variabilities related to the sample data. Out of the 17 studies that used statistical techniques, only (Wiley & Carrington, 2016) reported the accuracy of their model. Five other studies, (S. J. Carroll et al., 2014), (Janghorbani & Amini, 2016), (NIHMS865762, 2017), (NIHMS735680, 2015), and (Gomez-Marcos et al., 2013) reported the AUC's of their models. The four studies, Huo et al. (Huo et al., 2013), (Povel et al., 2012), (Smits et al., 2013), (Gurka et al., 2014) that used the CFA reported the goodness of fit of their models. Only three studies (Wiley & Carrington, 2016), (Janghorbani & Amini, 2016), and (NIHMS865762, 2017)

reported both the sensitivity and specificity of their statistical models. However, there was no evidence of performance evaluation comparison with other statistical methods in these studies. The predictive performance of the cMetSR score against other non-clinical methods is required in order to ascertain its efficiency. More research is therefore required to evaluate the cMetSR score against other types of MetS indexes.

The fourteen studies that used machine learning techniques evaluated their methods using the performance metrics defined in Table 2.4. Only 10 studies ((Worachartcheewan et al., 2013), (Zhao et al., 2014), (Van Schependom et al., 2015), (Karimi-Alavijeh et al., 2016), (Tan et al., 2016), (Z. Lin et al., 2014), (Worachartcheewan et al., 2015), (Miller et al., 2014), (Steinberg et al., 2014), (Romero-Saldana et al., 2016)) reported the accuracy of their models. The average accuracy performance of each technique that was used in the included studies is presented in Fig 2.3. The other four studies (H. Chen et al., 2014), (Ivanovic et al., 2016), (Hosseini et al., 2014) and (Obokata et al., 2015) reported the AUC of their models. It is clear from the graph that RF and DT were reported as the most frequently accurate techniques and the least accurate is the GDMR. This is closely followed by the accuracies of the ANN, the MLR and the REF techniques.

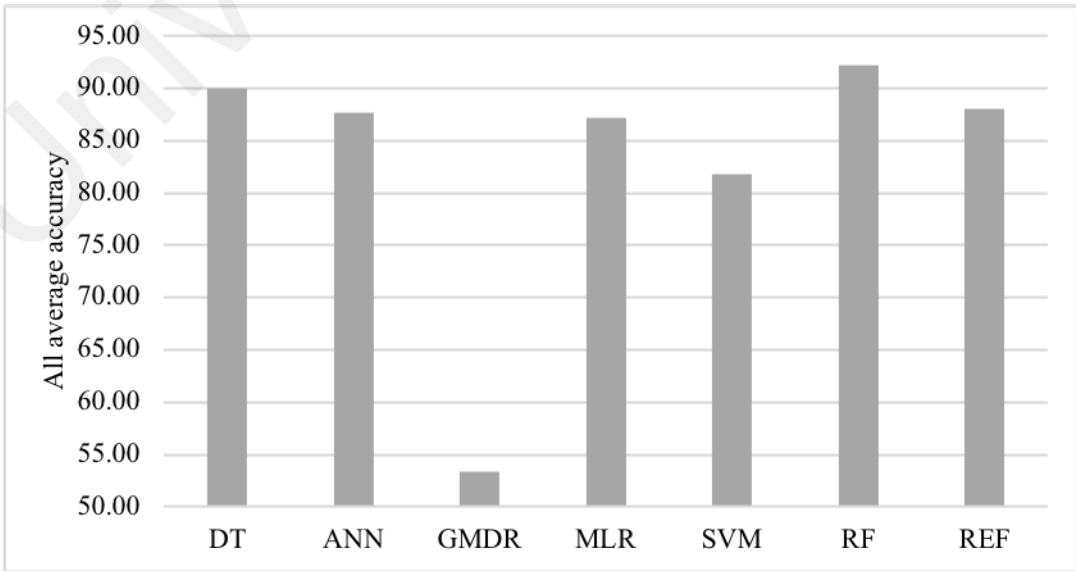

Figure 2.3: Overall average accuracy graph for each technique.

Most of the studies that applied statistical techniques used the PCA to develop the cMetS score. The PCA computes cMetS score using the principal component that explains the largest variance while maintaining the structure of the data. However, the Zcore is only a normalisation technique and data preservation less evident than in the PCA The most frequently used machine learning techniques are ANN, BN, and variants of DTs. The use of ANN could be ascribed to its efficiency in solving non-linear complex problems by being able to model any functional relationships and data structure. However, the back propagation algorithm which is used to train the neural network, is a gradient descent technique that is characterised by getting stuck in local minima and slow convergence, limiting its application to real life domains such as the prediction of MetS (Zweiri, Whidborne, & Seneviratne, 2003). Also the BPNN depends heavily on its learning parameter settings. The use of BNs for prediction and classification problems has been successfully applied in the medical domain (Park & Cho, 2012). This could be due to its ability to handle uncertainty and integrate previous knowledge to support causal relationships. Nevertheless, performing inference creates an expensive computational burden due to the inversion of finite elements. Additionally, DTs are favoured owing to their ability to generate interpretable results.

Significant conclusions can therefore be drawn from the performance measures applied to evaluate the prediction models. Consequently, the goal of machine learning models for predicting MetS should be to have a high predictive performance and generalisation capability that enables the diagnosis of the maximum number of individuals that have or are at risk of MetS. Out of the 4 techniques mentioned, ANN, DT, SVM, and MLR the machine models developed with ANN, DT, SVM can be said to consistently perform better than the MLR. However, a clear winner is difficult to ascertain as the techniques show varying performances with regards to their accuracy, sensitivity and specificity in different

studies. This variance could be attributed to the difference in size and dimension of the datasets. With regards to the RF, conclusions cannot be made until further studies have been conducted which compares it with other relevant machine learning techniques.

Accordingly, more number of studies which perform comparative evaluation between various machine learning techniques are required in order to ascertain more generalisable models that can quickly and accurately diagnose MetS . Oversampling methods such as SMOTE (Chawla, Bowyer, Hall, & Kegelmeyer, 2002) to balance the training dataset and also to remove noise from the whole dataset (He & Garcia, 2009) should be applied in preprocessing stages of MetS diagnosis. In addition, datasets should be cleaned using data cleaning methods such as Tomeks links (Tomek, 1976) and Wilson's Edited Nearest Neighbour Rule (D. L. Wilson, 1972), to remove any overlapping that may occur with the application of oversampling methods(He & Garcia, 2009). It is therefore essential to have a combination of at least the five clinical MetS risk factors (FPG, WC, TG, HDL-C , BP) as input features into a MetS prediction model (Murguia-Romero et al., 2013; Hosseini et al., 2014). Development of algorithms that consider the different weights and contributions of the MetS risk factors will contribute towards more accurate diagnosis of MetS in adults, adolescents and the paediatric population (Thivel et al., 2009). Furthermore, more research using different population samples is required for better generalizability of non-clinical methods.

The techniques applied to develop the machine learning models in this systematic review assume an outcome of either being at risk or not at risk of Metabolic syndrome. This binary prediction only agrees with a correct or incorrect outcome. However, for the prediction of the risk of MetS, it is recommended that machine learning algorithms that also predict the probabilities of the binary outcome should be used in determining the impact of the risk of MetS in the diagnosis. This probability will aid clinicians and individuals on the best

management guidelines to follow for prevention and treatment procedures.

Furthermore, the generalisation of a machine learning model can also be determined based on the number of algorithms used for its comparative evaluation. For the studies in this systematic review that have conducted comparative model evaluation, it can be argued that the number of machine learning algorithms used for evaluation are inadequate for determining the efficiency of the proposed models. From the primary studies in this category, only five studies ((Van Schependom et al., 2015), (Karimi-Alavijeh et al., 2016), (Miller et al., 2014)) compared their proposed models with other existing machine learning techniques. Even in these studies the number of comparable models for evaluation is less than sufficient to determine the efficiency of a proposed model. However in majority of the studies, appropriate procedures for evaluation, such as splitting the dataset into training, testing, and validation sets or the use of cross-validation was applied. Therefore, for proposed machine learning models which predict MetS, more number of existing machine learning models should be used for comparative evaluation and predictive generalisation.

In the case of the risk quantification techniques, no performance metrics were reported in any of the studies. Therefore, risk quantification techniques will need to be evaluated on other population samples due to their high dependency on the examination results of the population sample. Furthermore, the performance of these methods is required in order to ascertain how well they can generalised in quantifying the risk of metabolic syndrome. Majority of the primary studies are cross-sectional, making it difficult to identify the impact of the different MetS risk factors on the disease outcome. More longitudinal studies could be carried out to investigate how these factors interact with each other (Vikram et al., 2009). This information may be useful in ascertaining the individual effects of each risk factor in developing future algorithms. Finally, the growth in the non-clinical approaches is encouraging with studies showing promising results. However, there is a need for further

studies as follows:

1. Research that perform comparative performance evaluation using statistical techniques with various population samples should be carried out, so as to obtain generalisable statistical models.
2. Due to imbalance in MetS datasets, preprocessing techniques such as SMOTE should be applied before deriving prediction models.
3. Studies using machine learning techniques should perform comparative evaluation between new models derived and other relevant machine learning techniques in order to ascertain the efficiency of the non-clinical models.
4. Inference based machine learning techniques should be applied to derive non-clinical models due to their ability to present probabilistic values for MetS prediction.
5. Studies that evaluate mathematical quantification techniques using performance metrics should be conducted.

2.7 Limitations

Our analysis was limited to only studies published in the searched databases and written in English language. Secondly, due to the diverse algorithmic structures of the identified methods (Fig 2.2), a direct comparison between all the studies could not be carried out. For example, while machine learning algorithms have the ability to learn patterns from data, mathematical quantification techniques cannot learn. Thirdly, data extraction bias might have occurred because data was extracted by only one reviewer. Potentially, systematic reviews are also prone to selection bias. However, two of the reviewers independently selected the studies thereby minimising the risk of this bias. Other limitations beyond our control such as publication bias could also be present. Often, most studies published in peer reviewed journal tend to have positive results. .

2.8 Conclusions

Our review shows that the three main types of non-clinical methods used for the diagnosis of MetS are statistical, machine learning, and mathematical quantification techniques.

The statistical techniques include principal component analysis, confirmatory factor analysis, and standardised Z score. As the cMetSR score was mostly derived for the purpose of finding associations between MetS and its related diseases (CVDs and T2DM) and other risk factors, its effectiveness for use as a tool for the diagnosis of MetS in clinical setting has not been evaluated. Therefore more studies are required to evaluate the use of cMetSR score for clinical use.

The machine learning techniques used include artificial neural networks, decision trees, random forests, support vector machines, multiple logistic regression, reverse engineering and forward simulation, and Bayesian networks. The artificial neural network was the most frequently used machine learning technique, nonetheless, highlighted proof based on performance measures shows that the random forest technique is more applicable in the development of non-clinical methods for the diagnosis of MetS. However, the random forest model tends to create large trees which makes it inefficient for quick and easy clinical application. Therefore, more alternative non-clinical methods using machine learning techniques should be explored to develop applications that are readily and easily available to support the clinical diagnosis of MetS in practical clinical settings.

Three mathematical quantification techniques, areal similarity degree, simScore and Framingham risk score, were also used to develop the non-clinical methods. All the models are risk quantification models which are heavily reliant on the examination results of sample participants.

This study has several implications in personalised and public health management. It provides an opportunity for researchers and health care practitioners to gain an insight

into the current trends and development of existing machine learning, statistical and risk quantification methods used for the diagnosis of metabolic syndrome. Indexes derived from these non-clinical methods could be used as tools which serve as quick and early preventive indicators to guide the treatment and monitoring of ongoing management of MetS and its associated diseases such as T2DM and CVD . This study sought to equip researchers and clinicians with a comprehensive analysis on the different existing classifications and application of efficient algorithms for the early diagnosis, management, and prevention of MetS and its associated diseases in health care management systems. The future guidelines of this study will guide researchers in the process of developing and advanced.

CHAPTER 3: THEORETICAL FRAMEWORK OF THE RESEARCH

3.1 Introduction

One of the major goals of Artificial Intelligence (AI) is to design intelligent machines that simulate expert reasoning. These intelligent machines are driven by AI and Machine Learning (ML) algorithms to predict in various extents, events such as the trajectories of the planets and solar system, the weather patterns, the causes and trends of diseases, the fast commute route to and from a destination, the rise and fall of financial stock markets and economic growth, the spam filtering from personal emails, the grades and performances in academic institutions and a myriad of other cultural, natural, and social phenomena. AI is the broader term used to describe the process of building autonomous intelligent machines that are capable of skilled performances in a complex environment while machine learning refers to intelligent AI techniques developed for building computing systems that can learn and adapt from past experiences (Dietterich, 1999). Some of these machine learning techniques have been developed based on the idea of mimicking the capabilities of the human brain. These biological and computational ML techniques such as the ANN and genetic optimisations like GA were investigated aggressively by the research community in the 1980's (Mitchell, 1998).

The basic theories, structures, and processes of the ART and ARTMAP architectures, and the genetic algorithm is briefly described in this chapter.

3.2 Adaptive Resonance Theory (ART)

The stability-plasticity dilemma is a prominent problem for both biological and ANN. The principal concept is that intelligent systems once trained on a specific set of inputs is incapable of learning anything new from the environment. However, these systems require plasticity for the learning of new knowledge and also stability so as to restrain the

loss of previous knowledge. Excess plasticity will lead to past learnt knowledge being consistently disregarded, while excess stability will also inhibit the effective learning of the knowledge in ANN. Plasticity is a common problem that is inherent in neural networks where the learned network completely forgets any previous knowledge acquired when presented with new information to learn from (McCloskey & Cohen, 1989; Ratcliff, 1990). It is a common problem that occurs in various types of ANNs from standard feed forward back-propagation ANN to unsupervised neural networks such as self-organising maps (Richardson & Thomas, 2008). The feedforward back-propagation neural networks are highly sensitive to catastrophic forgetting due to their extremely distributed internal structure (French, 1992). However, a theory designed to enable an intelligent system to stay plastic by learning important events, and yet also remaining stable in reaction to insignificant occurrences is the ART (Grossberg, 1976b, 1976a). The concept is designed to adapt intelligent systems to new and dynamic environment that may be unpredictable without consequently forgetting existing knowledge. ART is based on how the brain learns, recognises, categorises, and predicts objects in the changing environment. In ART , top-down learned knowledge direct attention on bottom-up information, such that past learned experiences are protected from being removed by new learned experiences (Carpenter & Grossberg, 1988). Thus, enabling newly acquired information to be automatically integrated into the overall knowledge base of the ART system in a universally consistent method.

3.2.1 ART Architecture

The ART architecture has been developed as a physical theory which predicts data about cognitive information (Grossberg, 1982, 1987c, 1987a, 1987b). The adaptive pattern recognition model in ART is based on the competitive learning model (Carpenter & Grossberg, 1988). The progression of the ART characteristic neural architectures capable

of learning, pattern recognition and hypothesis testing were developed from ART1, ART2, and ART3 (Grossberg, 1982, 1987c, 1987a; Carpenter & Grossberg, 1988; Carpenter, Grossberg, & Reynolds, 1991). The ART architecture is a neural network that has self-organised stable perception capability in real-time when required to train random input data sequences. In such ART architecture, the system of adaptive pattern recognition enables a cognitive process of knowledge discovery, testing, searching, learning and classification. The basic structure of the ART1 architecture is presented in Fig 3.1.

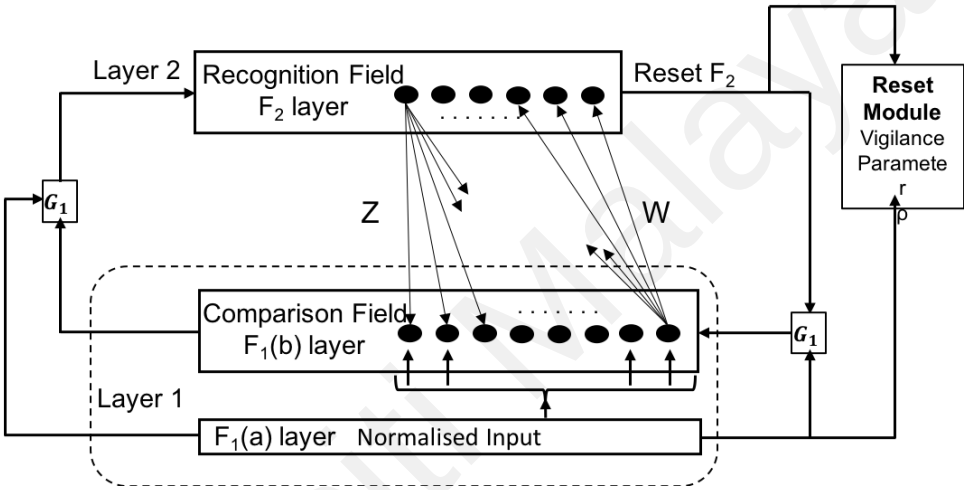

Figure 3.1: Adaptive resonance theory (ART) architecture

The ART1 network recieves a vectorised input and categorises it into one of the categories of the existing patterns it resembles the most. The ART network is made up of two units - computational and recognition units.

1. The computational unit: This unit further consists of two more sub-units, the input and cluster layers.

    a) Input layer: This layer is referred to as the F_1 layer. It consists of the input field, $F_1(a)$ sub layer and the comparison field, $F_1(b)$ sub layer.

        i. $F_1(a)$ sub layer: The $F_1(a)$ sub layer receives an input vector only. No computation process is carried out here. It sends the input it receives to the $F_1(b)$ sub layer which compares the input pattern and expectations.

ii. $F_1(b)$sub layer: The $F_1(b)$ sub layer is referred to as the comparison field which compares the input pattern and expectations. It connects to the F_2 sub layer through bottom-up weights and the F_2 sub layer is connected to the $F_1(b)$ sub layer via top-down weights.

b) Cluster Unit (F_2 layer): Competition relating which category a new input should be mapped into is carried out in this layer. Category mapping is carried out in a winner-take-all competition strategy. The cluster with the largest number of input wins in this competition and is triggered to learn the new input pattern. All the other non-winning activated clusters are then set back to 0.

c) Reset Mechanism: The reset mechanism is triggered when there is a mismatch between an expectation and an input pattern in layer 1. The level similarity between the top-down weight and the input vector determines whether or not a cluster is triggered to learn. If this similarity is less than the vigilanc parmeter, ρ, then the cluster is disallowed from learning the new pattern and a rest occurs.

2. The Supplement unit: Depending on the situation, the F_2 layer can either be inhibited from categorisation or be made available for categorisation when learning occurs. Therefore two supplemental units, G_1 and G_2 are have been added together with the reset unit R. These units are referred to as the gain control units. They recieve and send signals to the other layers in the network.

ART 1 is the basic ART network which only accepts binary inputs. ART2 differs with ART1 in that it accepts continuous inputs, while ART3 is a variant of both ART1 and ART2 architectures.

3.2.2 Predictive ART (ARTMAP)

The competency of ART to solve the stability-plasticity dilemma, i.e. the ability to learn new input patterns while remaining stable to insignificant patterns, prompted for its supervised learning adaptation called the ARTMAP (Carpenter, Grossberg, & Reynolds, 1991). ARTMAP architecture is a predictive ART architecture that integrates ART modules into its system to enable predictive learning. It consists of two unsupervised ART1 networks, linked by an associative learning network and an internal controller that can carry out fast and stable incremental learning as a supervised learning system. Given an n - dimensional input vector a, ARTMAP predicts an m - dimensional output vector b. This transformation from vectors in $\mathfrak{R}^n$ into vectors in $\mathfrak{R}^m$ defines a map which is learned from the correlated pairs a^n, b^n of sequential input vectors, $n = 1, 2, \ldots$ (Carpenter, 1989). Fig. 3.2 illustrates the ARTMAP architecture.

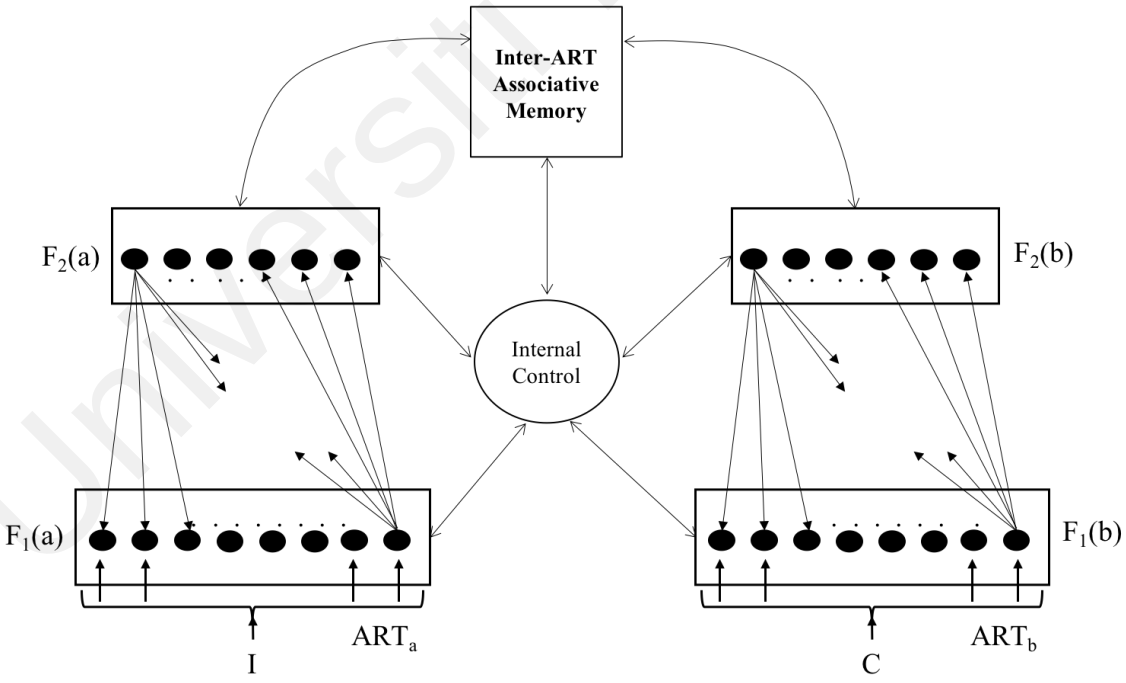

Figure 3.2: ARTMAP architecture adopted from (Carpenter, Grossberg & Reynolds, 1991).

Each ARTMAP network consists of a pair of unsupervised ART modules, ART_a and ART_b that respond to arbitrary sequences of input patterns by creating stable recognition

categories. In a supervised learning process, ART_a receives a sequence a^n of input patterns and ART_b receives a sequence of b^n target patterns, where b^n is the target prediction given the input pattern a^n. ART_a and ART_b are linked by an associative learning network and an internal controller in a map field F^{ab}. The controller is designed to ensure ARTMAP's autonomous operation in real time by creating the least number of ART_a recognition clusters needed to match the criteria in ART_b. Complement coding is a preprocessing method which transforms the N_a-vector a into the $2M_a$-vector, $A = a, a^c$ in the ART_a network, where A is the input vector ART_a field F_0^a. Likewise, the $2M_b$-vector, $B = b, b^c$ is the input sequence into the ART_b network.

When ART_b fails to make a prediction of a training input sequence from ART_a, the network, in the map field, creates connections between categories through learning and a match field inhibition is triggered which causes the ARTMAP match tracking rule. Match tracking raises the ART_a vigilance parameter ρ_a when predictive error at ART_b occurs (Carpenter, Grossberg, Markuzon, Reynolds, & Rosen, 1992)· ARTMAP minimises predictive error and maximises predictive generalisation by making a match between a winning category and a new input pattern using repetitive trials. The systems operates by raising the vigilance parameter ρ_a of ART_a with the least value required make the F_1^a to F_0^a match ratio $|x^a|/|A|$. This causes an ART_a search which generates an ART_a category that makes a correct prediction b or leads to a previously uncommitted category node in ART_a (Carpenter, Grossberg, & Reynolds, 1991).

The FAM (Carpenter et al., 1992) which is a combination of Fuzzy logic (Zadeh, 1965) and ART networks replaces the Fuzzy ART (FA) (Carpenter, Grossberg, & Rosen, 1991) for ART1. FAM is a generalised version of ARTMAP for classifying analogue as well as binary patterns. Fuzzy logic computations are incorporated into the ARTMAP systems. However, FAM is sensitive to statistical overlapping between clusters (Koufakou, Georgiopoulos,

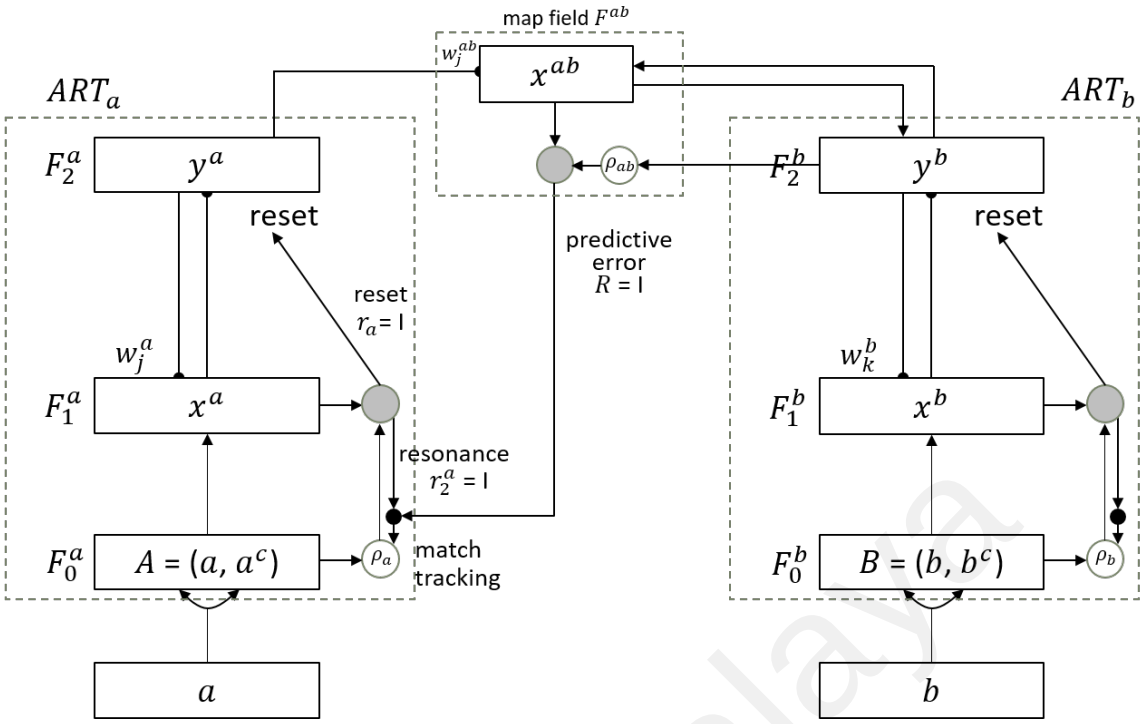

Figure 3.3: Fuzzy-ARTMAP architecture adapted from Carpenter et al., (Carpenter et al., 1992)

Anagnostopoulos, & Kasparis, 2001; Marriott & Harrison, 1995). This problem results in the creation of large number of pattern recognition clusters with overlapping allocation of nodes referred to as category proliferation. Category proliferation creates an output neural network architecture consisting of a redundant number of clusters which incurs a high computational and memory allocation cost with reduction in classifier performance. The ART-EMAP network (Carpenter & Ross, 1995) is a representation to improve the performance of the FAM. It is a combination of ART and spatial and temporal evidence for dynamic predictive mapping (EMAP). ARTMAP-IC (Carpenter & Grossberg, 1988) extends the distribution sequence with a sample counting procedure and a different match tracking algorithm that invariably enhances both predictive accuracy and code reduction than the classic ARTMAP and ART-EMAP networks. Gaussian ARTMAP (Williamson, 1996; Carpenter & Markuzon, 1998), PROBART (Marriott & Harrison, 1995), and ARTMAP-IC (Carpenter & Markuzon, 1998) are examples of FAM variants that have been proposed as a solution to category proliferation problem.

3.2.3 Bayesian ART

Progressively, a more successful approach to the problem associated with FA was proposed as the Bayesian ART (BA) and BAM (Vigdor & Lerner, 2007). BA and BAM architecture is an integration of the Bayes' theorem into the ART and ARTMAP neural network architecture. The BA makes changes in several characteristics of the FAM algorithm such as:

1. changing the hyper-rectangular categorisation with to a Gaussian categorisation;
2. restricting the mass of a chosen category, thereby enabling the growing and shrinking of categories;
3. relating patterns with clusters and clusters with probability classes to ensure ART and ARTMAP learning respectively; and
4. permitting class probability inference using all the related categories.

The BAM also computes class posterior probabilities and category, thereby ensuring the insertion of various deficits into the classification task. In BA network, clusters are represented by multidimensional Gaussian distributions with each cluster parametrised by its mean, prior probability, and co-variance matrix. These three parameters give a richer description about the Gaussian hyper-volume category as opposed to the FA hyper-rectangular category's weight vector. Therefore, instead of the FA's category shallow representation using a weight vector that is characterised by values of the hyper-rectangle's two corner ends, the Gaussian hypervolume category is distinctively characterised by its column, structure of distribution and its superiority in comparison with other equivalent categories in the same field. For instance, the FA category is most likely to have the same weights independent of the number of samples (3 patterns or 3000 patterns) it clusters, while in the case of the BA, the prior probabilities of a category will depend on the number of samples it clusters. Additionally, the data dispersion in FA is not specified while the

Table 3.1: Symbol notation of BAM

Symbol	Description
x	input vector for training sample
j	index of cluster
D	number of feature in each sample
ω_j	number of clusters
$P(\omega_j)$	assumed prior probability of the j-th cluster
$P\left(\omega_j \mid x\right)$	posterior probability of category ω_j given x
N_C	number of clusters
$p(x\mid\omega_j)$	conditional probability density of x given cluster ω_j
μ_j	D-dimensional mean vector
Σ_j	$D \times D$ co-variance matrix of j-th cluster
G	winning cluster
V_G	volume of the winning Gaussian hypersphere G
$V_{\max}$	maximum volume of a hypersphere
Σ_G	product of each dimensions' variances
$P_{\min}$	posterior probability threshold
$P(y_i\mid G)$	winning class category
N_G	number of samples that are categorised by the G-th cluster

data dispersion for the BA is specified using a Gaussian function in the form of the data co-variance matrix. Furthermore, complement coding is which needed in the FA input is not required in the BA. Generation of the BAM structure consists of the following three procedures (i) cluster choice, (ii) cluster match, and (iii) cluster learning. The symbol notations of BAM are presented in Table 3.1.

3.2.3.1 Cluster Selection

All current clusters are eligible to be selected during training phase. The j-th cluster of each D-dimensional sample x is represented by a posterior probability of category ω_j given x defined as follows:

$$P\left(\omega_j \mid x\right) = \frac{p\left(x\mid\omega_j\right) P(\omega_j)}{\sum_{n=1}^{N_C} p(x\mid\omega_n) P(\omega_n)}, \tag{3.1}$$

where x is the sample input, ω_j represents the number of cluster count, $P(\omega_j)$ is the assumed prior probability of the j-th cluster n_j, defined as

$$P(\omega_j) = \frac{n_j}{\sum_{j=1}^{N_C} n_j} \tag{3.2}$$

where N_C is the number of clusters. $p(x|\omega_j)$ represents the conditional probability density of x given cluster ω_j defined as

$$p(x|\omega_j) = \frac{1}{(2\pi)^{\frac{D}{2}} |\Sigma_j|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}\left(x - \mu_j\right)^T \Sigma_j^{-1} \left(x - \hat{\mu}_j\right)\right], \tag{3.3}$$

where μ_j and Σ_j are the D-dimensional mean vector and $D \times D$ co-variance matrix of j-th cluster, respectively.

The cluster with the highest posterior probability is selected as the winning cluster G, computed as

$$G = \arg\max_{j \in N} \left(P\left(\omega_j|x\right)\right). \tag{3.4}$$

The efficiency of the BA is enforced by selection of a winning, G cluster– ω_j that either has the largest value of the prior probability $p(\omega_j)$ or the cluster with the closest distance to the current sample input or both. The addition of the Bayes' theorem as another condition for selection also enables the accurate selection of a winning cluster.

3.2.3.2 Cluster Match (Vigilance Test)

The vigilance test is carried out to constrain the size of the winning cluster, G. The test ensures that the volume, V_G of the winning Gaussian hypersphere G does not exceed the the maximum volume of a hypersphere, $V_{\max}$ set for a winning cluster as follows:

$$V_G \leq V_{\max}. \tag{3.5}$$

where V_G is the determinant of the Gaussian covariance matrix, Σ_G computed as the product of each dimensions' variances, σ as follows:

$$V_G \triangleq \det(\Sigma_G) = \prod_{i=1}^{d} \sigma_{G_d}^2, \tag{3.6}$$

If the vigilance test (3.5) is passed by the winning cluster, then match tracking criterion will be checked. However, if G fails the vigilance test, then it is removed from the current competition by setting its posterior probability with (3.1) to zero, and the search continues until a winning cluster that passes test is selected. If no cluster from the existing clusters passes the vigilance test then a new cluster is created with a hypervolume that meets (3.5).

3.2.3.3 Match Tracking

In the map field of the BAM network, the class posterior probability is used to update the winning class category $P(y_i|G)$. A class posterior probability threshold, $P_{\min}$, is set as follows

$$P(y_i|G) \leq P_{\min} \tag{3.7}$$

such that if it is less than the class posterior probability, then the i-th class will be associated with the winning G category; $P(y_i|G)$ is derived from the BAM matrix $K = [N_{ij}]_{Y \times N_C}$, where Y is the number of classes, N_C is the number of clusters, and the ijth input into the BAM matrix K, N_{ij}, is the number of training input samples that are associated to he jth cluster and belonging to the ith class of the Y classes. $P(y_i|G)$ is calculated as

$$P(y_i|G) = \frac{N_{ij}}{\sum_{i=1}^{Y} \sum_{j=1}^{N_C} N_{ij}}, \tag{3.8}$$

where $i = 1, \ldots, Y$, and $G = 1, \ldots, N_C$. The full derivation of (3.8) using Bayes' theorem can be found in (Vigdor & Lerner, 2007).

Match tracking will occur if $P(y_i|G)$ satisfies (3.7) and the winning cluster G will be associated with the class Y_i. However, if the match tracking criterion is not met, then matching tracking is triggered by decreasing the maximum hypervolume, $V_{\max}$, by a small value δ as follows

$$V_{\max,\text{new}} = V_G - \delta, \quad 0 < \delta \ll V_G. \tag{3.9}$$

$V_{\max,\text{new}}$ should be small enough to remove the current winning category V_G from the competition and initiate a search for a new winning cluster. The new winning cluster must have a new hypervolume that is less than $V_{\max}$ as in the vigilance test (3.5). This search iterates until the new winning category is found.

3.2.3.4 Cluster Learning

If the selected category G meets the vigilance test (3.5) and match tracking criterion (3.7) requirements, then the category parameters, mean vector μ_G and covariance matrix Σ_G, will be updated as follows:

$$\hat{\mu}_{G_{\text{new}}} = \frac{N_G}{N_G + 1}\hat{\mu}_{G_{\text{old}}} + \frac{1}{N_G + 1}x, \tag{3.10}$$

$$\hat{\Sigma}_{G_{\text{new}}} = \frac{N_G}{N_G + 1}\hat{\Sigma}_{G_{\text{old}}} + \frac{I}{N_G + 1}\left(x - \hat{\mu}_{G_{\text{new}}}\right)\left(x - \mu_{G_{\text{new}}}\right)^T, \tag{3.11}$$

where $N_G^{\text{new}} = N_G^{\text{old}} + 1$. N_G is the number of samples that are categorised by the G-th cluster and I is an identity matrix. Then the process will go back to cluster selection step 3.2.3.1 to learn the next input sample.

3.2.3.5 Cluster Creation

A new cluster is created using

$$N_{C_{\text{new}}} = N_{C_{\text{old}}} + 1, \quad n_{N_C} = 1, \quad \hat{\mu}_{N_C} = x,$$
$$\hat{\Sigma}_{N_C} = \eta(V_{\max})^{1/D} I \tag{3.12}$$

such that the parameters of the cluster (3.10), the mean of the cluster $\hat{\mu}_{N_C}$, and the variance of the cluster $\hat{\Sigma}_{N_C}$, are initialised with the training input sample x, and $\eta(V_{\max})^{1/D} I$, where η is a small positive value and I is a $D \times D$ identity matrix, respectively. The count of the new cluster n_{N_C} which is an entry into (3.8) is initialised as 1.

3.2.3.6 Inference in BAM

Inference in the BAM network involves the use of all clusters associated to a class in order to define the class label during testing. Therefore, the class selected for a test sample x is defined as

$$y_i = \arg\max_i P(y_i|x) \tag{3.13}$$

where y_i is the class label for the test sample x and $P(y_i|x)$ is defined as

$$p(y_i|x) = \frac{\sum_{j=1}^{N_C} P(y_i|G)p(x|\omega_j)P(\omega_j)}{\sum_{i=1}^{Y}\sum_{i=1}^{N_C} P(y_i|G)P(x|\omega_j)P(\omega_j)} \tag{3.14}$$

where $P(y_i|G)$, $p(x|\omega_j)$, and $P(\omega_j)$ are defined in (3.8), (3.3), and (3.2), respectively. The pseudo-code for the BAM algorithm is represented in algorithm 1. BAM has the advantage of solving the category proliferation problem by using Bayesian theory and probability to create and update the clusters (Vigdor & Lerner, 2007; Masuyama et al., 2018). The ategory proliferation is solved by allowing the formed clusters the ability to shrink and

Algorithm 1 Algorithm for BAM

Require:
the samples : $A = (x_1, x_2, \ldots, x_N)$ $(x_n \in \Re^d)$,
the class label for each sample: y
a maximal hypervolume : $V_{\max}$.
a vigilance parameter bias: δ_V, and
a match tracking class probability threshold: $P_{\min}$,

Ensure: the class posterior probability $P(y_i|G)$ as in (3.1)

Input a vector x_n
if No clusters exist in ART network **then**
  Create a new cluster as $N_{C_{\text{new}}}$ as in (3.12).
else
  Compute a cluster posterior probability $P\left(\omega_j|x\right)$ as in (3.1) .
  Compute the index of winning cluster G as in (3.4).
  Compute V_G as (3.5).
  if $V_{G_k} \geq V_{\max}$ **then**
    if All the categories do not pass the vigilance test **then**
      Create a new category $N_{C_{\text{new}}}$ as in (3.12).
    else
      Remove y_G from selection and continue from step 6 with the next chosen cluster.
    end if
  end if
end if
Update the category frequency count N_{ij}.
Compute the class posterior probability $P(y_i|G)$ as in (3.8).
if $P(y_i|G) \geq P_{\min}$ **then**
  The map field learning takes place.
  Update $\hat{\mu}_{G,\text{new}}$ as in (3.10).
  Update $\hat{\Sigma}_{G,\text{new}}$ as in (3.11).
else
  if Match tracking fails with all categories **then**
    Create a new category as $N_{C_{\text{new}}}$.
  else
    Remove $\hat{\mu}_{G,\text{new}}$ from selection.
    Remove $\hat{\Sigma}_{G,\text{new}}$ from selection.
    Continue from step 17 with a next candidate category.
    Update a maximal hypervolume $V_{\max}$ as in (3.9).
  end if
end if
if $j < G$ **then**
  Continue from step 1 with $j \leftarrow j + 1$
end if

expand while learning. Categories or clusters are created in Gaussian spheres which greatly reduces thier number and thereby increasing the performance of BAM. BAM achieves inference through the posterior probability. Clusters are illustrated as Gaussian hyper

volumes. However, the maximum hyper volume parameter, V_{max}, of the BAM is sensitive to reiterative search that leads to high computational time and the creation of more clusters. These resulting clusters reduces the predictive performance of the BAM. Moreover, BAM is also sensitive to the sequence order of its input patterns. Therefore, these factors should be considered to get the optimal solution.

3.2.4 Advantages of Bayesian ARTMAP

The advantages of BAM over the ANN include:

- The formation of clusters in the form of Gaussian hyperspheres shapes has the ability to better represent multidimensional real distributed data.
- The Gaussian hypershere clusters have the ability to grow and shrink during training. This reduces the unnecessary creation of small categories which increases predictive accuracy (Williamson, 1996).
- The cluster prior probability resulting from the adoption of the Bayes' theory ensures that each class cluster is distinctively represented in relation to all the existing clusters. Repetition of class clusters during inference is mininmised.
- It has a small number of parameters, requires no problem-specific system crafting or choice of initial weight values

3.3 Genetic Algorithm

Evolutionary computation is a group of computational methods that simulate natural evolution by generating a population consisting of individuals. GA was founded by John Holland in the early 1970s as one of the methods of evolutionary computation (Holland, 1992). GA are abstract search algorithms that represent of the natural process of selection where only the fittest individuals who survive in a population are selected. These fit individuals are used to generate a new population using genetic operations and this cycle is

repeated until only the fittest individuals survive (Negnevitsky, 2005). Genetic algorithms have been applied to solve different types of optimisation problems such as numerical , discrete, multi-objective, and global optimisation problems (Sakawa, 2012). A concise definition of GA is evident in the academic literature (Mitchell, 1998). However, GA is described as a process that evaluates the fitness of a candidate solutions in a population through a repetitive cycle of evaluation, selection, mutation, and crossover (Grefenstette, 1986).

3.3.1 Gene, Chromosome, Allele, Phenotype, Genotype and Breeding

The essential foundation of the GA is composed of Genes. A gene is a binary encoding of model variables. Each gene comprises of different forms of allele that encode the several characteristics of a gene. In abstract terms, a gene is encoded as a string of binary digits- 0 or 1 in the form of an allele. A string of genes form the chromosome. The chromosome represents a candidate solution of a problem. A typical chromosome is represented as an array of elements. In this work, the terms chromosomes and candidate solutions will be interchange where necessary. The encoded structure of genes, which is the chromosome, is also known as the genotype while phenotype is the physical description or decoded structure of the genotype as a model (Sivanandam & Deepa, 2008; Franz, 2006). Representations of these structures are presented in Fig. 3.5 and Fig. 3.6.

3.3.2 Encoding and Decoding

Encoding is the representation of the chromosomes as a string of characters in an array. Different forms of encoding techniques have been used to represent candidateeee solutions in the search space of GA. Different types of encoding representations such as binary, octal, hexadecimal, permutation, value, and tree encoding are applied to genes in GA. The binary encoding scheme is the most commonly used representation in GA. In Binary encoding,

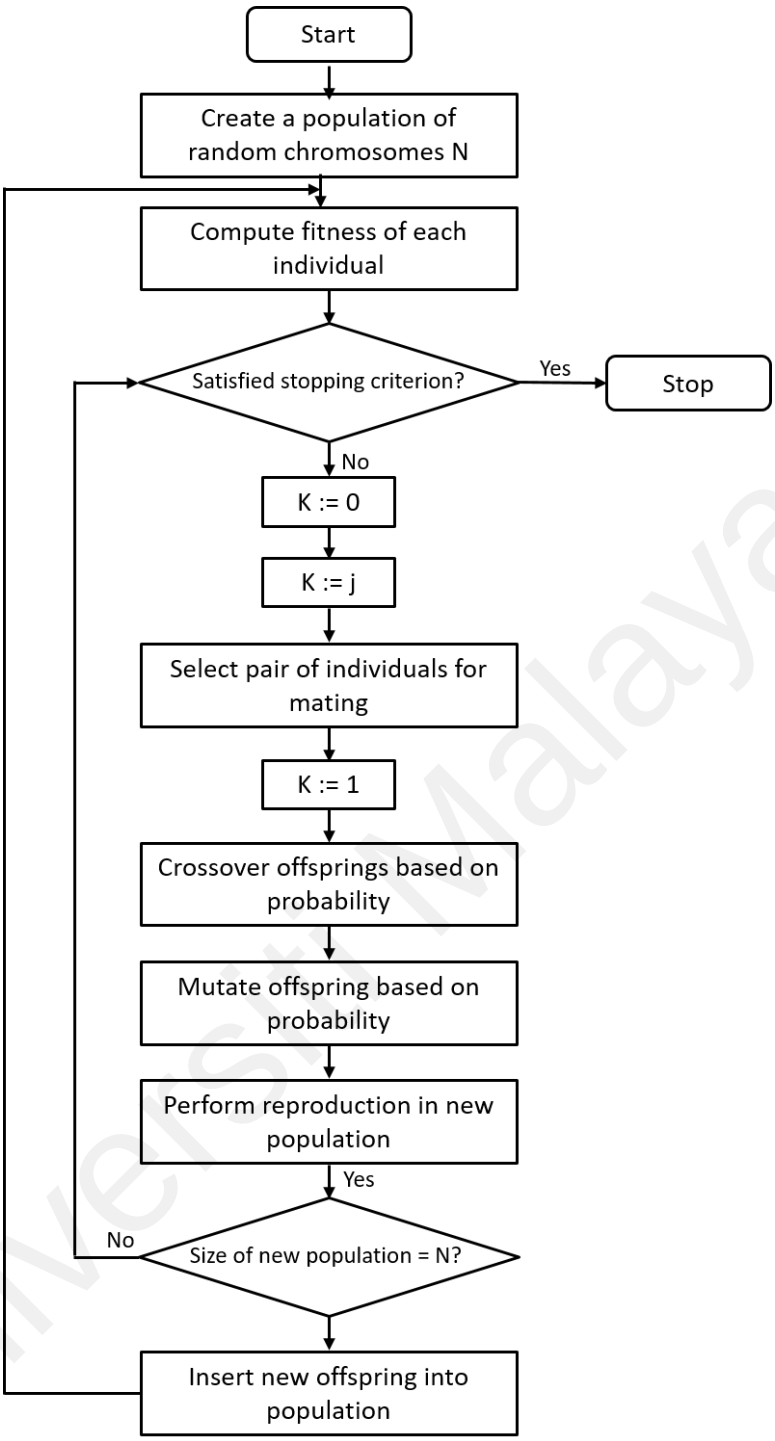

Figure 3.4: The Genetic Algorithm process

each gene is represented using a binary (bit) string. Binary encoding with 0s and 1s enables a large number of chromosomes to be comfortably represented using a small amount of alleles. Moreover, binary encoding makes it easy to represent integers and variations of real numbers. These flexibility of the binary encoding scheme enabled its application in this proposed algorithm. Other forms of encoding schemes include permutation encoding,

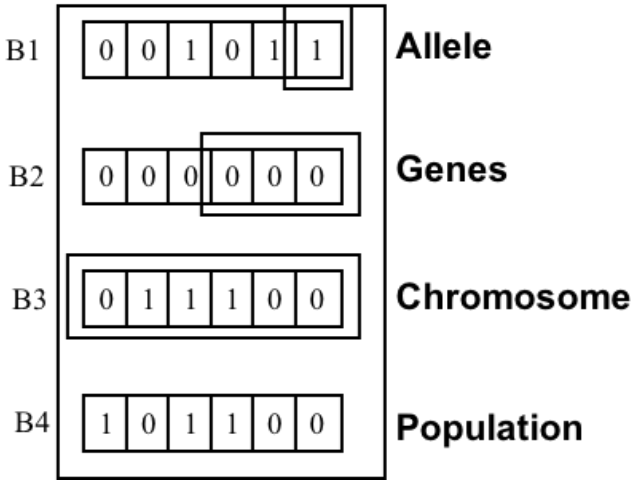

Figure 3.5: Representation of allele, gene, chromosome and population

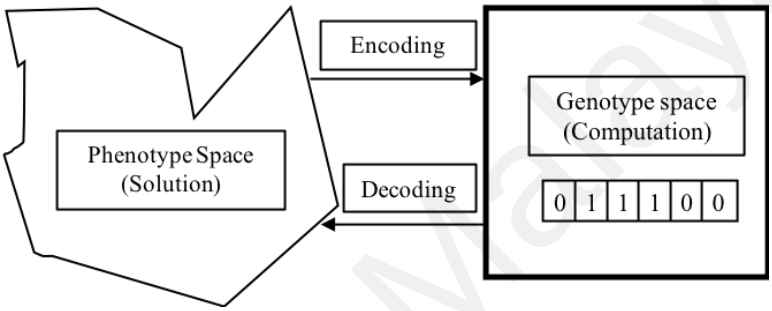

Figure 3.6: Representation of genotype and phenotype

value encoding, tree encoding and value encoding (Sakawa, 2012).

3.3.3 Fitness Function

In natural evolution, only the fittest in the population survives. GA follows this law such that only the fittest chromosomes are chosen for the evolution processes. The fitness landscape represents the space of all chromosomes defined by fitness function (Mitchell, 1998). The fitness function is used to assign a fitness score to each chromosome in the population. This fitness value determines the ability of chromosome to form a candidate solution for the problem to be solved. The fitness of a chromosome determines its chances of selection from the population. The higher the fitness value, the more superior the chromosome. Chromosomes with low fitness values are discarded from the evolution process (Sakawa, 2012).

3.3.4 Population

A population in GA consists of a set of chromosomes that represent the candidate solutions to the problem in the search space. The population consists N_{pop} is represented as a matrix of $N_{pop} \times N_{bits}$ filled with chromosome vectors. In the GA process, an initial set of population is randomly chosen for the first generation. The formation of a new population requires the reproduction operators to be chosen depending on the choice of the reproduction probabilities, P_c (crossover probability), P_m (mutation probability), and P_r (direct reproduction probability), where $P_c + P_m + P_r = 1$. These are the parameters that control the genetic algorithm and are determined by the user.

3.3.5 Selection

Selection occurs when chromosomes are selected to serve a parents in reproduction operations. Various selection methods such as rank selection, roulette wheel selection, tournament selection, and Boltzman selection, are used to select the fittest chromosomes (Eiben, Smith, et al., 2003). Typically, a selection probability P_j is assigned to each candidate solution j in the population according to the its fitness value. These sequence of numbers is compared with the the population's cummulative probability $C_i = \sum_{j=1}^{i} P_i \; N$ (Pencheva, Atanassov, & Shannon, 2009). The fittest individual i is chosen to participate in the reproduction process is selected. The fitness of a chromosome determines its chances of being selected as a parent chromosome for reproduction. The selection mechanism enables the choosing of genetic components inside the population with an aim to improve future generations of offspring using the reproduction operators. The selection cycle illustrated in Fig. 3.7.

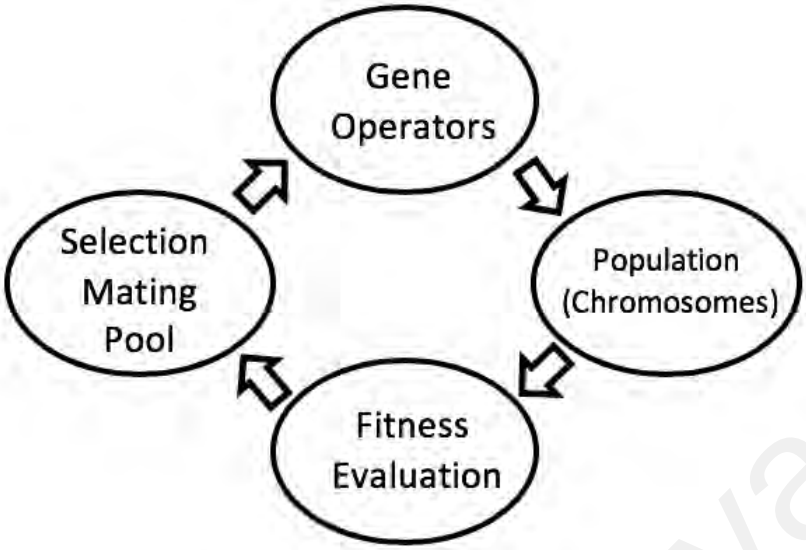

Figure 3.7: The selection cycle

3.3.6 Crossover (bit-string)

The locus is a particular location of a gene in a chromosome. In a locus, an allele is represented as either a 0 or a 1. The crossover and mutation are reproduction operators which are tasked with the movement of genotype in the fitness search space determined by the fitness function (Mitchell, 1998). The crossover operator allows genes to be exchanged at a particular locus between a pair of parent chromosomes to form two new chromosomes known as the offspring. These offspring or new individuals are entered into the population for possible mutation and fitness evaluation. In bit-string or single point crossover, a sub-sequence from a randomly selected locus in two randomly selected chromosomes are swapped to form two new offspring. This process is presented in Figure 3.8.

The number of times a crossover operation is carried out in the GA is dependent up the crossover probability. If the crossover probability is 100%, then all offspring of the new generation are formed by crossover, otherwise if the probability is 0%, then all the offspring of the new generation will be formed as the exact replicas of the parent generation. The crossover operation is used to produce the best chromosomes. This offspring chromosomes

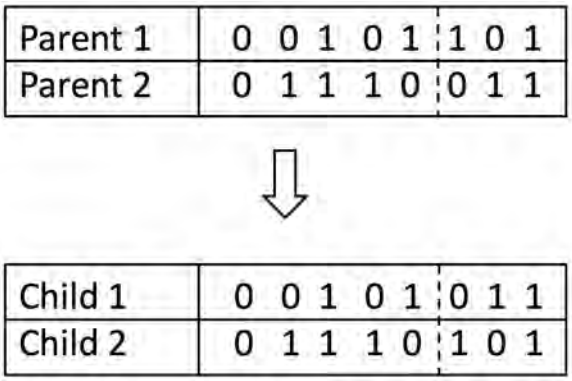

Figure 3.8: Single point crossover

are subjected to mutation by the mutation operator. The crossover operation explain in this section is the single-point crossover. Other crossover operations include the two-point, multi-point, three parents and reduced surrogate crossover amongst others. However, the simplicity and efficiency of the single-point crossover makes it the mostly commonly adopted choice.

3.3.7 Mutation

The mutation operator enables the flipping of bits at a random point in a new offspring chromosome. Genes in an offspring chromosome are subjected to mutation depending on a probabilistic value. Mutation occurs to enforce uniqueness and hinders early convergence. Mutation, however, does not introduce new information into the search space. After running many generations of the GA, there is a possibility that the same bits may be selected over again. This may lead to premature convergence which generate a satisfactory but not the best solution to a problem (Whitley, 1994). Therefore, mutation is introduced to preserve particular bits and alleles in fit chromosomes. Mutation works in the background by sometimes making the flips of bits at random points.

3.3.8 Search Space

The space which consists of all candidate solutions of a problem is known a the search space of the GA. The candidate solution represents a point in the search space. Each

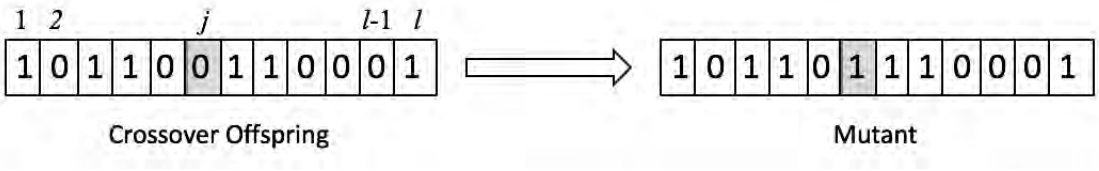

Figure 3.9: Chromosome mutation

candidate solution in the search space is defined by fitness function. During searching, the value of the fitness function is used to evaluate each candidate solution for selection as a solution. Due to complexity of searching, different types of searching techniques such as the genetic algorithm are employed to find an appropriate solution. GA searching evolves the populations of candidate solutions until an optimal solution is obtained. The genetic process optimises the chance of reaching the optimal solution by providing a evaluating the fitness of each candidate solution.

3.3.9 Reproduction

Depending on their fitness value, offspring bred by selected individuals are chosen such that the population can either grow or shrink. The fittest individuals are passed over to form the population of the next generation while those offspring with low fitness are discarded from the process (Sivanandam & Deepa, 2008). These fit individuals are immune to further recombination by the reproduction operators. They are, however, selected for the next reproduction cycle. The reproduction cycle of selection, crossover and mutation will be carried out on the low fit individuals to generate higher fit offspring.

3.3.10 Convergence and Search Termination

Convergence occurs when a suitable solution to the problem has been found. The conditions for convergence are set by some stopping criteria such as when a certain given percentage of the offspring are identical with the same fitness value and the best fitness value stops changing for some specified number of generations (Coley, 1999). Other

convergence occurrence stopping criteria includes the reaching a of maximum number of generations by a GA process, the elapsing of a specified amount of computation time, and the lack of improvement of the fitness values within a specified time limit (Sivanandam & Deepa, 2008). Furthermore, a best fitness value could be set such that convergence occurs if the least fitness in the population goes down. GA has the ability to jump out of any local minimum because it proffers several solutions to a problem and chooses the best out of the solutions. This feature of GA that enables to find an optimal minima is albeit one of its major advantages.

3.3.11 Advantages of Genetic Algorithm

GA has the advantage of making a thorough global search in the search space using the value of a fitness function. The use of the reproduction operators gives GA the ability to evolve initial individuals in a population of the best candidate solutions. By choosing best or fittest candidate solution, GA provides the best solution to problem solving. The GA search always finds the best solution to the problem because it performs a parallel search distributed evenly across the search space (Haupt & Haupt, 2004). It is able to perform optimisation with both discrete and continuous variables without needing any derivative knowledge. Thus, making it suitable for our proposed algorithm because all the variables are continuous.

3.3.12 Summary

The theoretical frameworks and background of the ARTMAP, FAM BAM and GA were discussed in this chapter. Furthermore, their mathematical equations and pictorial representations were also presented. The theory of GA and its descriptive operators such as population, reproduction, crossover and mutation were discussed and illustrated with descriptive diagrams. The advantages of both the BAM and GA for diagnosis of the risk

of MetS were explained.

CHAPTER 4: DATASET AND EXPERIMENTAL SETUP

4.1 Introduction

The dataset used in this research is a real world data collected on site from teachers in a prospective cohort study. Thus, the data requires cleaning and pre-processing in order to provide an improved and optimised model performance. Data preprocessing activities such as data collection, data cleaning and integration, and normalisation are presented in this chapter. Finally, the chapter discussed the proposed approach and experimental setup based on the pre-processed dataset.

4.2 Experimental Data Collection

The dataset used for the evaluation of the proposed model in this research was collected in a prospective cohort study, the "Cohort study on clustering of lifestyle risk factors and understanding its association with stress on health and well-being among school teachers in Malaysia (CLUSTer)" (Moy et al., 2014) from March 2013 to December 2014. This study was aimed at extensively studying the interaction of work related stress and the clustering of lifestyle risk factors on teachers' health and well-being in Malaysia. The study population consisted of primary and secondary school teachers in Peninsular Malaysia. In a multi-stage sampling method, five states out of the 12 states in Peninsular Malaysia are selected as depicted in Figure 4.1.

Seventy percent (70%) of primary and secondary schools from the selected states were invited for the study. Teachers from the schools that accept the invitation are then invited to a voluntary participation in the study. All permanently employed teachers that do not have any form of psychiatric illness were eligible to participate in the data collection. Schools qualified for the study were ranked based on the statistics of primary and secondary schools from the Ministry of Education, Malaysia for the year 2013. Participants were asked

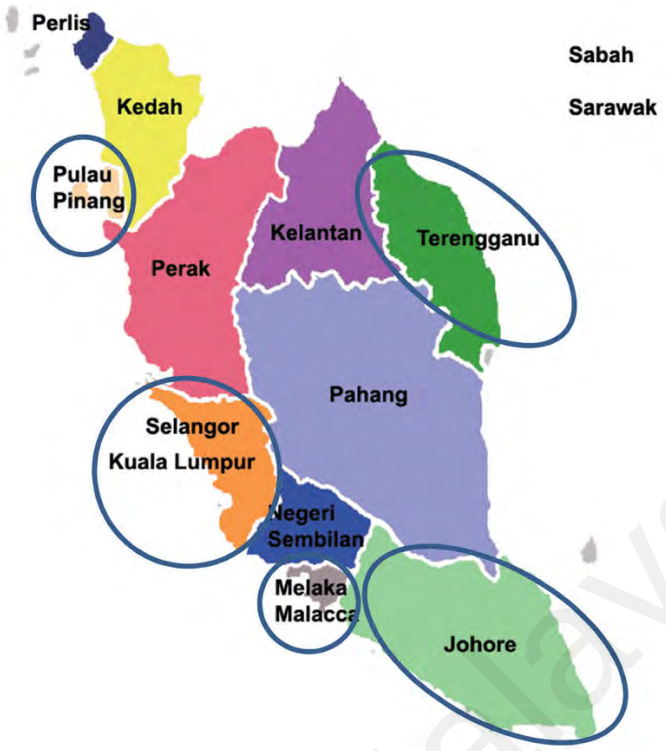

Figure 4.1: Selected states of Peninsular Malaysia for data collection. Adapted from (Moy et al., 2014)

to answer a questionnaire and also engage in the required health screening procedures. Measurements of five MRFs were selected for the experiments in this study. These MRFs - FPG, WC, HDL-C, TG, SBP, DBP - were selected based on their prominence in diagnosing the risk of MetS in all the clinical definitions as presented in Table 1.1 . In the anthropometric assessment, weight was measured without shoess to the nearest 0.1kg using a digital calibrated floor scale (SECA 813, Hamburg, Germany). Height was measured using a portable stadiometer (SECA 217, Hamburg, Germany) while waist circumference was measured using a flexible tape measure (SECA 203, Hamburg, Germany) at the umbilicus to the nearest 0.1cm. Hip circumference, fat mass and muscle mass were also measured. The threshold for identifying visceral obesity among Malaysians was set according to the Asian standards (Zimmet, Shaw, & Alberti, 2005) where the waist circumference for males and females is 90cm and 80cm respectively.

A questionnaire was handed out to all participants. The required information included

family and medical history of chronic diseases, and socio-demographic characteristics of participants. Other questionnaires include a structured lifestyle questionnaire on alcohol consumption, smoking and physical activities (IPAQ) (Craig et al., 2003); questions on fruits, vegetables, and fats and oil consumption; the job content questionnaire (JCQ) (Hadi, Naing, Daud, & Nordin, 2006); the depression, anxiety and stress scale (DASS21) (Gandek et al., 1998); the health related quality of life (SF12-V2)(Ware Jr, Kosinski, & Keller, 1996); and the Voice Handicap Index 10 (VHI 10) (Rosen, Lee, Osborne, Zullo, & Murry, 2004) on voice disorder. Questions on chronic pain, sleep duration and obstetric history for female participants were also asked.

The clinical health assessments that were conducted included systolic and diastolic blood pressure measurement and biochemical analysis (fasting blood glucose, renal function test and blood lipid profile). Blood pressure was measured in a sitting position, once on the left arm using a validated oscillometric blood pressure monitor (Omron HEM 907, Japan) (DuBose & McKune, 2014). For the biochemical analyses, four tubes of blood samples were collected from each participant. The baseline measurements (health assessment) of fasting blood glucose, blood lipids and renal function tests were analysed using the Dimension clinical chemistry system known as in-vitro diagnostic test. A total of 12,429 samples was collected.

The Medical Ethics Committee of the University Malaya Medical Centre (UMMC) granted the ethical approval (Reference Number: 950.1) for this study using human subjects.

4.3 Data Description

The experimental dataset consists of 11,237 samples at baseline, (N = 2133) and female (N =9104). The characteristics of the dataset is shown in Table 4.1. The dataset consists of six attributes FPG, WC, HDL-C, TG, SBP, and DBP. The prevalence of MetS in Malaysian

women is higher than that of the men at 30.1% and 24.8%, respectively. Therefore, we found it necessary to divide the dataset was into two groups based on to gender. The prevalence increase in MetS is directly proportional to age (Walker, Gurka, Oliver, Johns, & DeBoer, 2012). As such the severity of MetS is said to increase over time in individuals (Vishnu, Gurka, & DeBoer, 2015).

Table 4.1: Characteristics of CLUSTer dataset after cleaning (Moy et. al, 2014)

	Male		**Female**	
Number of Subjects, *n* (%)	2133 (19.00)		9104 (81.00)	
Age, years	44.33±9.43		42.95±8.4	
20-39 (young)	1397(19.71)		3410(82.25)	
40-64 (middle-aged)	736(17.75)		5691(80.29)	
65 and above (old)	0		3	
Body Mass Index (kg/m^2)	44.43±7.45	Mn 23.9 Mx 81.7	40.38±7.78	Mn 19.3 Mx 100
FPG (mmol/L)	5.00±0.58	Mn 3.1 Mx 7.6	4.82±0.50	Mn 3.1 Mx 7.1
WC (cm)	89.55±10.99	Mn 57 Mx 137	79.68±10.69	Mn 48 Mx 121
HDL-C (mmol/L)	1.24±0.26	Mn 0.52 Mx 2.3	1.49±0.34	Mn 0.52 Mx 3
TG (mmol/L)	1.56±0.77	Mn 0.6 Mx 5	1.12±0.48	Mn 0.6 Mx 3.3
SBP (mm Hg)	132.37±15.13	Mn 76 Mx 200	123.43±16.51	Mn 65 Mx 198
DBP (mm Hg)	81.64±10.66	Mn 48 Mx 124	75.15±10.97	Mn 41 Mx 121
T2DM *n* (%)	0		0	
Hypertension [1] *n* (%)	182(8.50)		454(5.00)	

Values are means ± SD or n(%); Mn, Minimum; Mx, Maximum.
Fasting blood glucose ≥ 5.5 (mmol/L) and/or physician diagnosed diabetes mellitus.
Systolic BP ≥ 130 mm Hg and/or diastolic BP ≥ 85 mm Hg and/ or physician diagnosed hypertension

Therefore, the dataset was further divided into six subgroups based on the age of

the participants. The classification into six subject categories by gender and by age is categorised as: young (between 20 and 39 years old), middle-aged (between 40 and 64 years old), and old (exceeding 65 years old). However, there were no male participants over the age of 65 and only three female participants were identified. Hence, only four categories of datasets will be used for the training and evaluation of our proposed GOBAM model–Young Male, Middle-age Male, Young Female, and Middle-age Female.

The prevalence of MetS is 22% and 34% in the young and middle-aged male groups respectively. The young and middle-aged female groups have a MetS prevalence of 11% and 21% respectively. The number of individuals with MetS are less than those without MetS. This situation where the number of samples of one class is more dominant than the other class in a dataset is referred to as the class imbalance (Japkowicz & Stephen, 2002). Class imbalance has been known to occur in real world datasets that have been collected in different areas of scientific research (Kubat, Holte, & Matwin, 1998; Chawla, Japkowicz, & Kotcz, 2004; Rao, Krishnan, & Niculescu, 2006). It is obvious from Fig. 4.2 that the CLUSTer dataset is skewed largely towards non-MetS in all the four subgroups of the Cohort Study on Clustering of Lifestyle Risk Factors and Understanding its Association wth Stress on Health and Wellbeing Among School Teachers in Malaysia (CLUSTer) dataset. We shall see how this data characteristics affects the predictive performance of the GOBAM classifier.

4.4 Analysis and Cleaning of Experimental Data

Collection of data from the real world is prone to human errors especially during recording and input. Therefore, data is required to be cleaned and preprocessed to ensure optimal quality for an efficient model performance. Although the data collection was carried out meticulously with care and precision, thorough data exploration for outliers, missing values and other inconsistencies was conducted. After conducting data exploration,

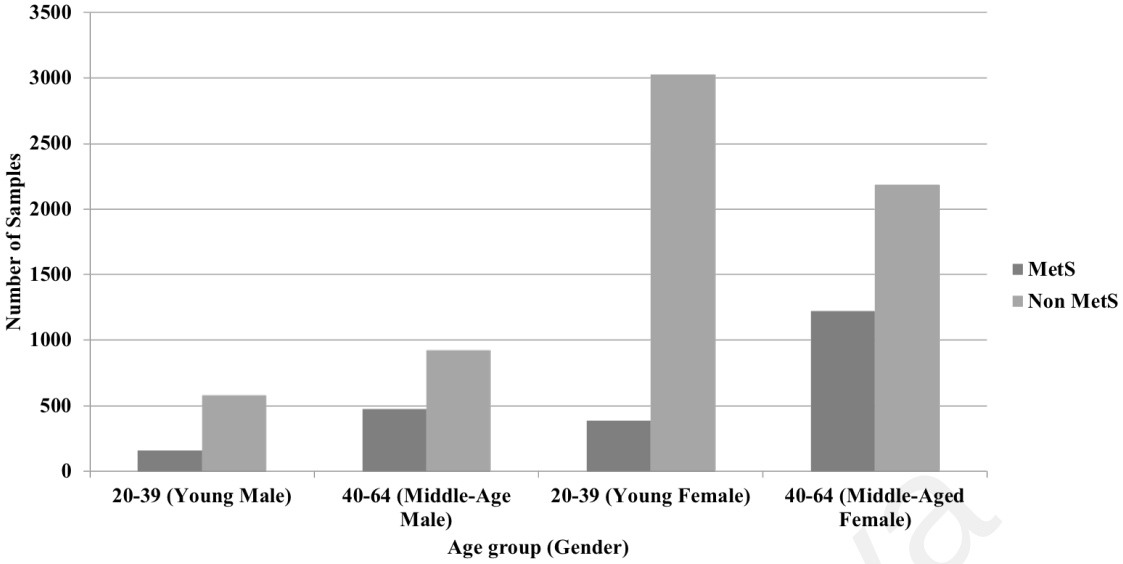

Figure 4.2: Pictorial representation of class imbalance in the CLUSTer dataset for all the four subgroups

the dataset was found to contain outliers in the attributes. These outliers may be as a result of human errors while collecting the data. For example, the WC attribute had a minimum value of 1cm measurement. The samples with these extremely low measurements do not match the characteristics of real world measurement of MRFs. They fall outside the range of MRFs for adults. Therefore, the outliers samples were identified and removed using the Mahalanobis distance (De Maesschalck, Jouan-Rimbaud, & Massart, 2000). The Mahalanobis distance is a measure of distance between the attributes and their means. The Mahalanobis distance assumes a Gaussian distribution for each attribute x, $N(\mu, M)$ and is calculated as

$$d(x) = \sqrt{(x-\mu)M^{-1}(x-\mu)^T} \tag{4.1}$$

where μ is the mean of the vector x and M is the covariance matrix.

It was necessary to remove the outliers before normalising due to the disparity between the samples. The initial data collected consisted of 12,429 participants. This data set was categorised based on gender resulting in a total of 11, 237 samples of males (N = 2432)

and females (N =9997). Then the Mahalanobis distance was run on each group. For the female group, the Mahalanobis distance detected 893 outliers in 8 runs while 299 outliers were detected and removed from the male group in the same number of runs.

4.5 Normalisation of Experimental Data

Preliminary analysis of the dataset was carried out to determine the most suitable data scale for the highest predictive performance. WC is measured in Centimeter (cm), TG, HDL-C and FPG are measured in Millimoles Per Liter (mmol/L) while SBP and DBPis measured in Millimeter Of Mercury (mmHg). For example, in Table 4.1, the range for WC MRF is 57 - 137 cm while the range of HDL-C is 0.52 - 2.3 mmol/L. The order of these two MRFs is different.

Due to these different scales of the MRFs measurements, it was necessary to normalise the attributes to fit into a new range from 0 to 1 as follows:

$$x_{i_{new}} = \frac{x_i - x_{i_{\min}}}{x_{i_{\max}} - x_{i_{\min}}}. \tag{4.2}$$

where x_i is the i-th metabolic syndrome risk factor, $x_{i_{new}}$ is the new normalised input value, and $x_{i_{\max}}$ and $x_{i_{\min}}$ are the maximum and minimum values, respectively. By normalisation, the disparity between the attributes will be nullified and this will lead to better prediction performance of the proposed classifier. Furthermore, the vigilance parameter V_G of the BAM is unable to handle a large numeric disparity between sample features.

4.6 Genetically Optimised Bayesian ARTMAP (GOBAM)

GA can enable rapid convergence and a reduction of generalisation errors in classifiers. Parameter optimisation has been proposed to tackle problems associated with FAM using various evolutionary computation techniques (Cervantes, Lee, & Lee, 2007; Granger,

Henniges, Oliveira, & Sabourin, 2006; R. Palaniappan & Eswaran, 2009; Loo, Liew, Seera, & Lim, 2015). However, to the best of our knowledge, no one has attempted to optimise the parameter settings and order of input sequence in the BAM neural network for medical diagnosis. Therefore, we propose GOBAM to diagnose MetS in this thesis. As previously mentioned, a biased training sequence (Masuyama et al., 2018) and under-tuned parameters affects the BAM's classification performance and its stability. Here, we utilise GA to search for an optimal combination of parameter values and training sample sequence in order to increase the predictive performance of BAM. The GOBAM flowchart is illustrated in Fig. 4.5. The parameter values that affect the performance of BAM are as follows:

(i) maximal hyper-volume: $V_{\max}$, and

(ii) vigilance parameter bias: δ_V.

4.6.1 Chromosome design

The sequence of training samples and parameter settings of the BAM network will be optimised using GA. Each chromosome g contains the following:

1. The training samples sequence, $g_T^1 \sim g_T^N$, where N is the total number of training sample. $g_T^1 \sim g_T^N$ is encoded using permutation encoding. Each element of this chromosome subset must be a unique element representing the index of a training sample input.
2. The maximum hyper-volume $g_{V_{\max}}$ which is encoded using real values ranging between [0] and [100].
3. The vigilance parameter bias: $g_{\delta V}$ represents the values of parameters δ_V.

Ten population of chromosomes $g(pop)$, $pop = 1, ..., 10$, are randomly initialised as possible candidate solutions. The sequence numbers of the training samples are randomly ordered and also the values of the BAM parameters. The initialised chromosomes are

passed into the BAM network for training and testing. The illustration in Fig 4.3 is a representation of each chromosome in the search space.

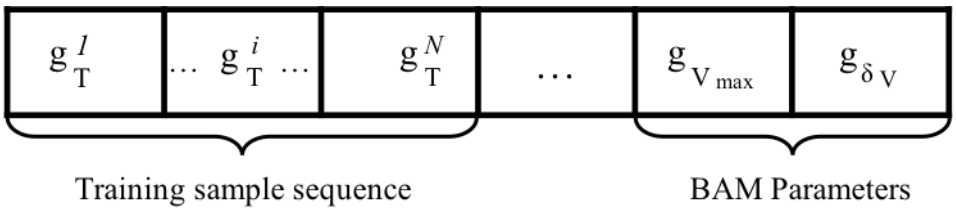

Figure 4.3: GOBAM chromosome representation

4.6.2 Fitness Evaluation

Next, GOBAM evaluates the fitness of each candidate solution in the BAM network and determines its fitness values using the fitness function. Our proposed algorithm is driven by the fitness function which enables the GA's search for optimal parameters and sequence of input patterns for the BAM. The fitness evaluation of the proposed GOBAM algorithm uses the AUC instead of accuracy because accuracy measures performance in relation to the total number of only the correct predictions while AUC is a summary measure of accuracy derived from the ROC curve. A detailed explanation of the AUC is presented in Section 4.8.7.

4.6.3 Selection

The selection approach adopted in our proposed GOBAM is the roulette wheel developed by Holland (Holland, 1992). The probability of selecting an individual using the roulette wheel is given as

$$P_i = \frac{f_i}{\sum_{j=1}^{N} f_j} \tag{4.3}$$

where P_i is the selection probability value of selecting a chromosome i, f_j is the fitness value of chromosome i, and N is the number of chromosomes in the population. The

chromosomes with the best fitness relative to the fitness values of the other chromosomes in the population have a higher chance of being selected. the roulette wheel operator The selected parent chromosomes will be used to create new chromosomes using the GA operations - crossover and mutation. The roulette wheel selection process is illustrated in Fig 4.4.

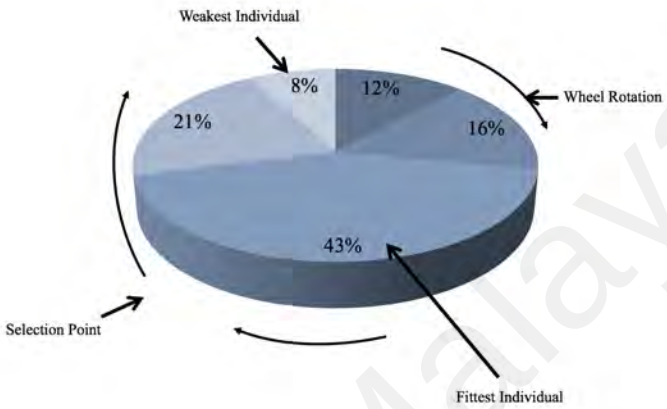

Figure 4.4: The roulette wheel selection process

4.6.4 Crossover and Mutation

Crossover and mutation for the order of sample sequence and the BAM parameters is carried out differently because of the difference in the encoding process.

As depicted in Fig. 4.6, Partially Mapped Crossover (PMX) (Goldberg & Lingle, 1985) technique is applied for generating the offspring of the sequence order chromosome $g_T^1 \sim g_T^N$ in order to get the sequence order with the best fitness value. This technique is chosen because it does not allow for tie ranks in the offspring chromosome. After crossover, each element in the offspring chromosomes has to be a unique entry as is required in the sequence order of the training sample of the BAM network. Given two selected parents, Parent *A* and Parent *B*, PMX generates two offspring chromosomes, Offspring *A* and Offspring *B*, by uniformly selecting two random points in each of the parent chromosomes and swapping the elements within the bounds of those points. Each element in Parent *A*

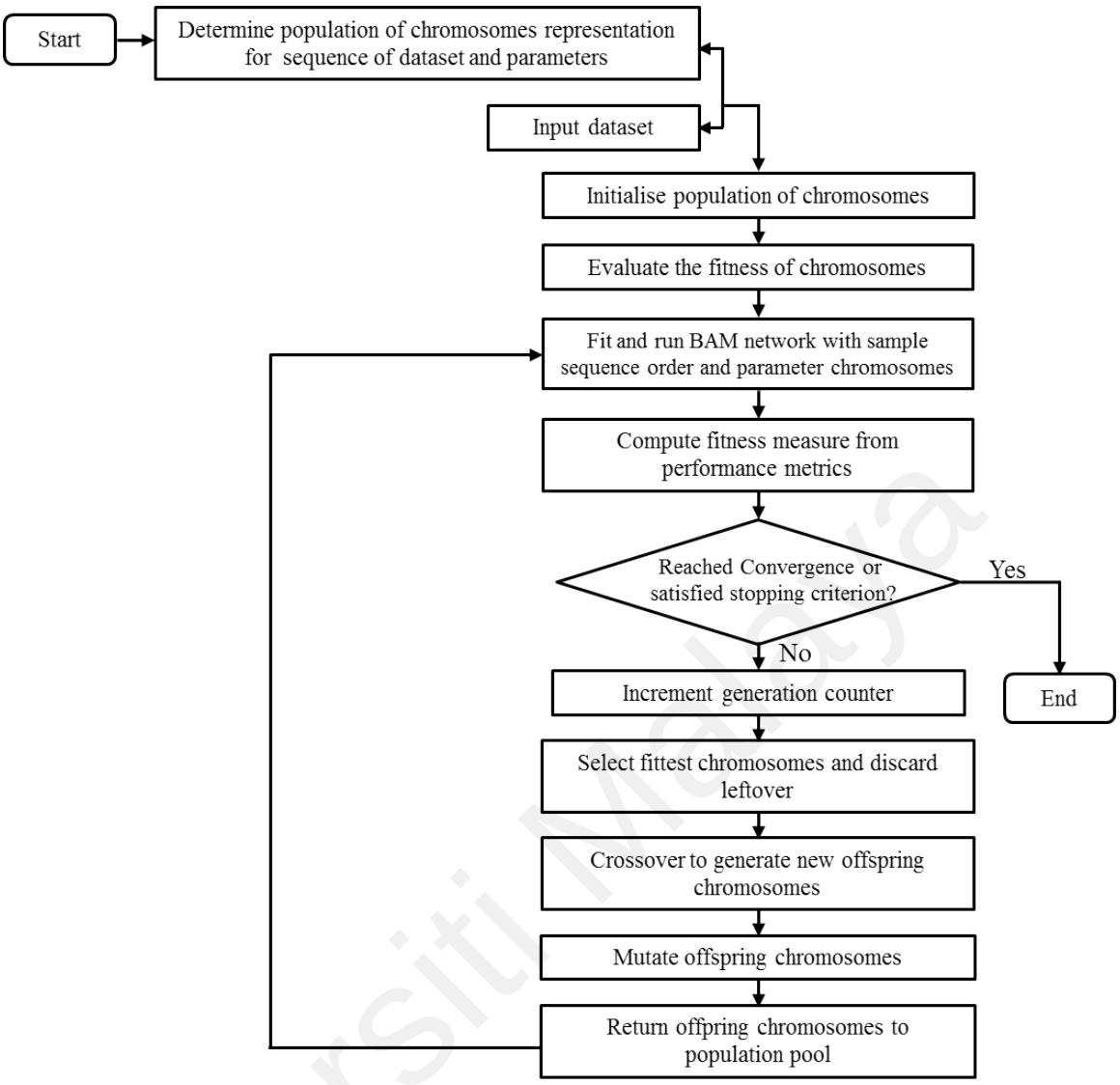

Figure 4.5: Flow chart of genetically optimised Bayesian ARTMAP

is mapped to the element in the same position in Parent *B*. Subsequently, the remaining blank fields in Offspring *A* are filled with elements from Parent *A*. However, if the element to be copied from Parent *A* to Offspring *A* is already present in Offspring *A*, then that element from Parent *A* will be exchanged with the element from Parent *B* which is mapped to the element in Parent *A*. This pattern continues until all the blank chromosome fields in Offspring *A* have been filled. Then the blank fields in Offspring *B* will be filled in by elements in Parent *B* not already present in Offspring *B* in the same order. Fig 4.6 illustrates the crossover and mutation process of generating an offspring chromosome. Mutation is then carried out by randomly swapping two elements of each offspring chromosome.

The single point crossover is applied to the BAM parameter which is encoded as a real valued number. As shown in Fig. 4.6, this crossover technique generates offspring chromosomes by selecting one crossover point and swapping all the elements from that point between the parent chromosomes. The mutation of the BAM parameter offspring chromosome involves the addition and subtraction of some random float numbers.

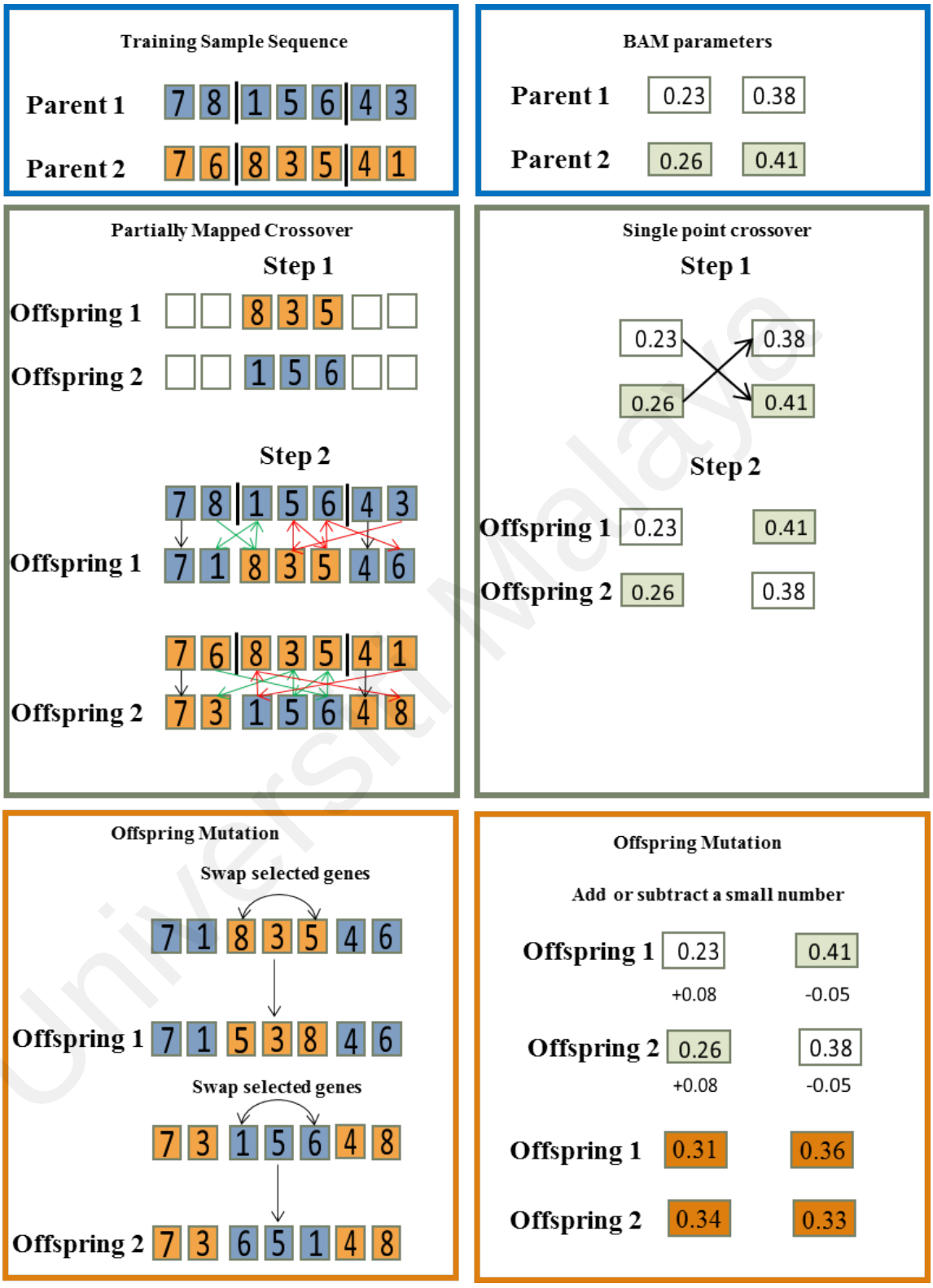

Figure 4.6: Chromosome crossover and mutation of the training sequence and BAM parameters.

The GOBAM algorithm is described in algorithm 2.

Algorithm 2 Algorithm for GOBAM

Require:
the samples: $(< x_1, y_1 >, < x_2, y_2 >, \ldots, < x_N, y_N >)$ $(x_N \in \Re^D)$,
the training sample sequence chromosome: $g_T^1 \sim g_T^N$,
the maximum hypervolume chromosome: $g_{V_{\max}}$,
the vigilance parameter bias chromosome: g_{δ_V},
the size of the population: pop,
the maximum number of generations $maxGen$.

Ensure: the best chromosome g_{best} from the population $g(pop)$

Initialise individual chromosomes $g_T^1 \sim g_T^n$, $g_{V_{\max}}$, g_{δ_V} in population $g(pop)$
Evaluate the fitness f_{AUC} of each chromosome in population $g(pop)$
while maximum iteration $iter$ is not reached **or** convergence is not reached **do**
  for x = 1 to $maxGen$ **do**
    Perform k-fold cross-validation of BAM
    Calculate fitness f_{AUC} for each chromosome g
  end for
  if Convergence occurs **then**
    Return best chromosomes g_{best}
  else
    Select the best pair of chromosomes based on the value of fitness function f_{AUC}
    Perform crossover and mutation as described in section 4.6.1 to obtain new chromosomes g_{new}
  else
    Increase selection value to include optimum solutions
    Decrease mutation value to converge onto solutions
  end if
  Update population $g(pop)$ with new chromosomes g_{new}
  Update count $iter$
  repeat
    From step 2
  until Convergence or stopping criteria reached
end while

4.7 Implementation

We implemented the proposed GOBAM algorithm in Matlab R2014a (8.3.0.532) on an Intel(R) Core(TM) i5-4200U CPU @ 1.6GHz 64-bit computer with 8.00 GB RAM. The algorithm was compared to standard ARTMAP algorithms such as FA, BAM, and GAFAM on the CLUSTer (Moy et al., 2014) dataset shown in Table 4.1.

4.7.1 Experimental Design

Cross-validation is widely acceptable for evaluating the performance of machine learning algorithms and building prediction models. The cross-validation process is used to certify the efficiency and generalisation of our proposed model. We used the K-fold cross validation technique (Hastie, Tibshirani, Friedman, & Franklin, 2005) to calculate the classification performance measures of our model. The K-fold cross-validation randomly divides the training dataset into K equal-sized parts as shown in Fig. 4.7. Cross-validation works by fitting the model into $K - 1$ parts of the training dataset while holding out k th part for testing. The overall performance of the model is averaged on the results of the K iterations as presented in Eq. (4.4).

$$R = \frac{1}{K} \sum_{k-1}^{K} R_k \tag{4.4}$$

where R_k is a performance metric for the kth partition. The K-fold cross validation is the most reliable method for evaluating the performance of the proposed algorithm because it provides an unbiased estimation of the model's performance. It is used to evaluate the efficiency of a model using any kind of performance metric. The experiments carried out were averaged over 5 trials of 10-fold cross validation on each data set, while recording the balanced crossed validation performance measure on the hold-out test fold.

The number of generations was set to 200. This is to reduce the bias of random sampling from the training dataset. Two different criteria were applied to terminate the training: (i) training was stopped if the mutation operator value is less than 0.01 and (ii) the training was stopped if the crossover operator value was greater than 0.9. The mutation rate should not be low in the GA search because a low mutation rate prevents exploration in search space (Loo et al., 2015). The crossover rate should be large enough to allow the GA to reach global optimum. This combination of mutation and crossover parameters allows the

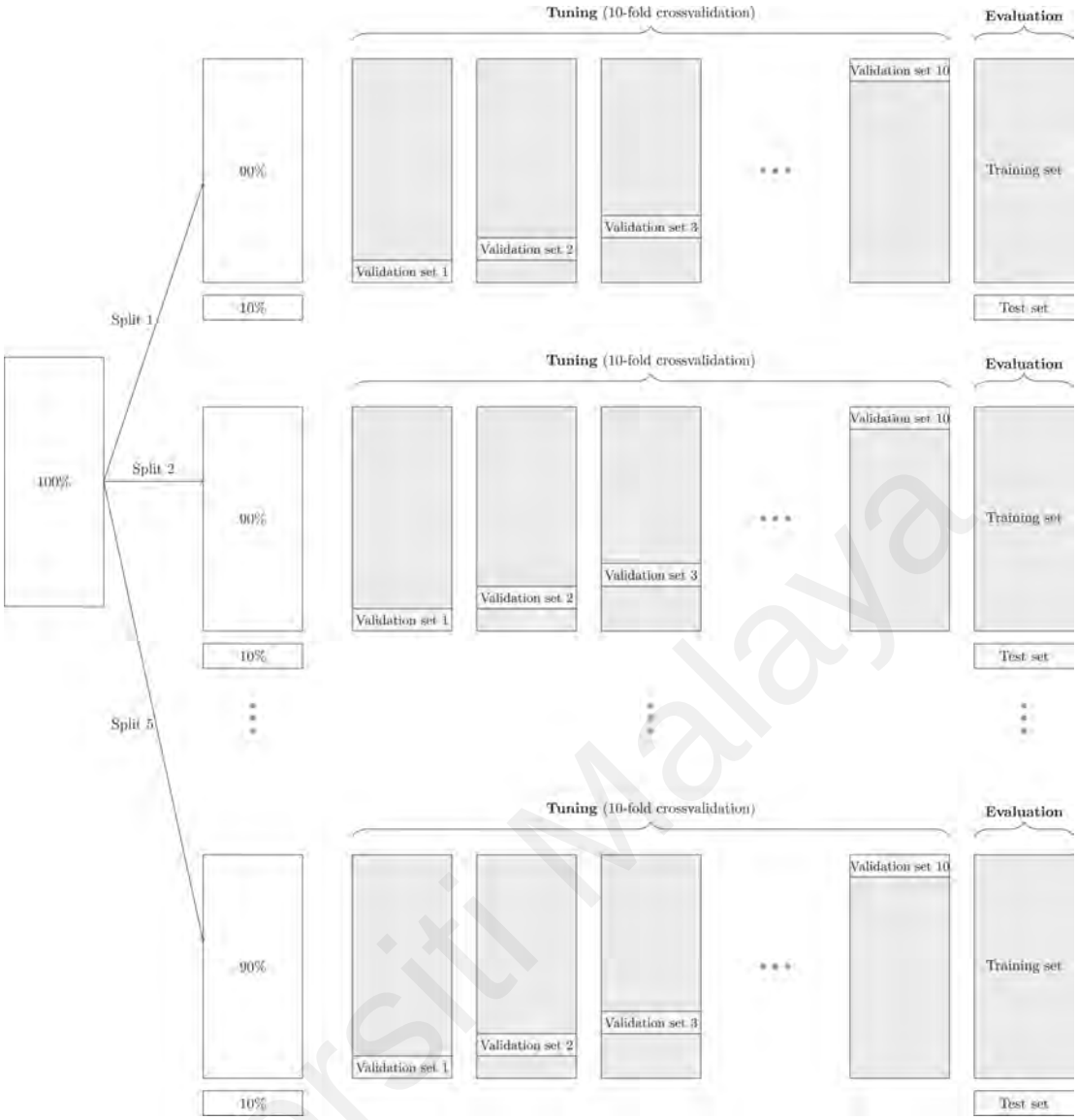

Figure 4.7: 10-fold cross-validation partitioning of datasets. Adopted from Baunmann et al. (Baumann et al., 2018)

GA to avoid local optima convergence.

4.8 Performance evaluation methods

For machine learning classifiers where the problem requires a binary decision solution, as is the case for the MetS diagnosis in this thesis, the performance of classifiers is evaluated using performance metrics known as predictive bio-markers (Pepe, 2003): AUC, SEN, SPEC, PPV, NPV and FSCORE. These quantitative performance metrics are derived using the metric values from the confusion matrix in Table 4.2. The confusion matrix shows the comparison between each predicted class with its actual class in four types of metrics

values:

Table 4.2: Calculation of sensitivity and specificity for a specific cut-off point of the predicted4 probability P.

		Actual		Total cases
		Metabolic syndrome	Non-metabolic syndrome	
Predicted	Metabolic syndrome	TP (true positive) A	FP (false positive) B	$TP + FP$
	Non-metabolic syndrome	FN (false negative) C	TN (true negative) D	$FN + TN$
	Total cases	$TP + FN$	$FP + TN$	

1. Cell A contains the number of individuals from the test dataset who have been correctly classified as having MetS by the proposed diagnostic method GOBAM and also by the JIS harmonised criteria adopted for MetS diagnosis in this study. This means that all individuals in cell A have the same diagnosis as the traditional MetS. This value is referred to as the True Positives (TPs). TP is the count of actual people with MetS and correctly classified as having MetS.
2. Cell B contains the number of individuals that GOBAM classifies as having MetS but do not have MetS according to the clinical diagnosis of MetS. These are referred to as the False Positives (FPs). FP is the count of actual people without MetS but incorrectly classified as having MetS.
3. In cell C, the number of individuals who have been diagnosed as having MetS by the proposed method but do not have MetS according to the clinical diagnosis. According to the clinical diagnosis definition, these individuals have been wrongly classified as having the abnormality. These are known as the True Negatives (TNs). TN is the count of actual people with MetS but incorrectly classified as not having MetS.
4. Cell D contains the number of individuals are classified as not having MetS by GOBAM and also have the same diagnosis according to the clinical diagnosis

definition. These individuals are not at risk of MetS according to both methods. These are referred to as the TNs. False Negative (FN) is the count of actual people without MetS incorrectly classified as having MetS.

Note that the actual classification of MetS is based on any of the clinical definition of MetS such as the NCEP ATP III (Expert Panel on Detection, 2001). The metric values from the confusion matrix in Table 4.2 are used to compute the other performance measures that will be used to evaluate the proposed GOBAM. Below are the formulas for the various performance evaluation methods .

4.8.1 Sensitivity

Sensitivity is the ability of a classifier to correctly diagnose MetS amongst when the condition in present. It is the probability that the diagnosis result is positive given that the abnormality exists in the individual.

$$Sensitivity = \frac{TP}{(TP + FN)} \times 100 \tag{4.5}$$

4.8.2 Specificity

Specificity is the ability of a classifier to correctly diagnose non-MetS amongst all actual non-diseased classes. It the probability that the diagnosis result is negative given that the abnormality does not exists in the individual.

$$Specificity = \frac{TN}{(FP + TN)} \times 100 \tag{4.6}$$

4.8.3 Positive Predictive Value

PPV is the probability that people classified by the proposed model as having MetS, actually have the abnormality.

$$PPV = \frac{TP}{(FP + TP)} \times 100 \tag{4.7}$$

4.8.4 Negative Predictive Value

NPV measure is the probability that the people classified as non-MetS actually do not have the abnormality.

$$NPV = \frac{TN}{(FN + TN)} \times 100 \tag{4.8}$$

4.8.5 FScore

FSCORE is a weighted average of the SEN and the PPV.

$$NPV = \frac{2TP}{(2TP + FP + FN)} \times 100 \tag{4.9}$$

4.8.6 Receiver Operating Characteristic Curve

The ROC curve is a standard summary of accuracy used for the analysis of diagnostic test performance (Heagerty, Lumley, & Pepe, 2000). The ROC curve is a graphic plot of SEN on the y-axis against 1-SPEC on the x-axis. The ROC curve compares SEN against SPEC across all the range of values in order to predict a dichotomous outcome (Fawcett, 2006). This outcome is ensured because the SEN and SPEC constitute the basic measures of performance evaluation for diagnostic models as described equations (4.5) and (4.6). An attractive feature of the ROC curve for medical diagnosis include its ability to internally discriminate the capacity of a diagnostic test by attribution across all threshold ranges

(Heagerty et al., 2000). It is a valid comparison approach even when the attributes of a model differ in measurement scales.

4.8.7 Area Under the ROC curve

One summary index associated with the ROC curve is the AUC. The AUC is a probabilistic interpretation that the diagnosis of a diseased individual chosen randomly exceeds that of a non-diseased individual which is a summary of the ROC curve (Swets, 1988). As mentioned in section 4.8.6, the ROC curve is independent of disease prevalence in the dataset. It is a performance measure of imbalanced datasets which makes is more effective in evaluating classifier performance than the ACC metric (Saito & Rehmsmeier, 2015).

As mentioned in section 4.6.2, the AUC is used as the fitness function of the proposed GOBAM model. The AUC is the estimated trapezoidal integration calculated in (4.10) as follow (Bradley, 1997):

$$f_{AUC} = \sum_{\varphi} \left\{ [s_{\varphi} \cdot \Delta(1-t)] + \frac{1}{2}[\Delta s \cdot \Delta(1-t)] \right\} \tag{4.10}$$

where r is the SEN, t is the SPEC, $\Delta(1-t) = (1-t)_{\varphi} - (1-t)_{\varphi-1}$, $\Delta s = s_{\varphi} - s_{\varphi-1}$, and φ is an index. r in equation (4.10) is referred to as the True Positive Rate (TPR) while $1-t$ is known as the False Positive Rate (FPR).

The AUC provides more information about the predictive performance of a classification model than the single ACC metric. It takes a value between 0 and 1 since both SEN and SPEC are also values between 0 and 1. The overall diagnostic performance of a model indicates reliable diagnosis when the AUC is closer to 1 and a model with an AUC value reading exactly 1 is considered perfect at its prediction. Practically, the lower bound of the AUC is set to 0.5. However, an AUC value of 0.5 interprets the diagnosis of the classifier

to be based on pure chance for balanced datasets (Zhou, McClish, & Obuchowski, 2009). Both the SEN and SPEC are independent of the prevalence of the disease in the population dataset (Zhou et al., 2009). Therefore, the AUC also has this same feature which makes it efficient as an evaluation metric when a classifier is implemented using a dataset that has class imbalance (Veganzones & Séverin, 2018). This justifies the use of the AUC as a fitness function for the proposed model.

4.9 Summary

This chapter presents a comprehensive description of experimental data collection process. A detailed description of the characteristics of the dataset is also given. This is followed by an analysis of the preprocessing methods applied to the dataset. A detailed description of the proposed GOBAM algorithm is presented together with the flowcharts and pseudo-codes of the algorithm. Subsequently, the experimental setup was described and the definitions of the performance evaluation methods were also presented.

CHAPTER 5: RESULTS AND DISCUSSIONS

This chapter presents the results of the experiments carried out for the proposed model GOBAM and a discussion on the comparison of the performance evaluation between the proposed model and three other classical Adaptive Resonance Theory Mapping (ARTMAP) models - GAFAM (S. M. Liew et al., 2011), BAM (Vigdor & Lerner, 2007) and FAM (Carpenter et al., 1992). The performance of the proposed model was evaluated from three aspects: efficiency, reliability and improvement over existing models. First, preliminary results using UCI benchmark datasets are presented in order to evaluate the efficiency of the proposed GOBAM. Then, the results of experiments using the 10-fold cross validation on the CLUSTer dataset are represented using ROC curves, AUC, ACC, SEN, SPEC, PPV, NPV and FSCORE performance metrics. The fitness graphs generated from the experiments of the two optimisation models GOBAM and GAFAM for the four subgroups of the CLUSTer dataset are presented and discussed. In all result tables, the highest performance metric values are highlighted in bold.

5.1 Analysis of Proposed Model's Performance

As mentioned in Section 4.6.2, the ACC as a model performance metric was not used in evaluating the performance of GOBAM using the CLUSTer dataset. The fact that the CLUSTer dataset is characterised by a class imbalance where the non-MetS class is more dominant makes the use of ACC ineffective as performance evaluation metric (Fawcett, 2006). However, an initial performance evaluation of the GOBAM was carried out in order to ascertain its better performance, reliability and generalisation using some 2-dimensional UCI benchmark datasets (Dheeru & Karra Taniskidou, 2017). The UCI datasets used for the preliminary experiments are the Glass, Wisconsin Breast Cancer, Pima Indian Diabetes, Iris, Thyroid, Crab, and Wine datasets. These UCI datasets are the popular datasets used

for benchmarking (W. S. Liew et al., 2016).

The characteristics of the UCI datasets used is presented in Table 5.1.

Table 5.1: Characteristic of UCI datasets

Dataset	Class	Attribute	Sample Size
Glass	7	9	214
Breast Cancer	2	9	569
Pima Indian	2	8	768
Iris	3	4	150
Thyroid	2	21	7 200
Crab	2	6	200
Wine	3	13	178

The AUC, ACC, SEN, SPEC, PPV, NPV and FSCORE results for the preliminary experiments using the 10-fold cross validation on GOBAM, GAFAM, FAM and BAM are presented in Table 5.2. In the Glass, Iris, Crab, and Wine UCI benchmark datasets, the proposed GOBAM model has an ACCof 100% while its slightly less in the remaining three datasets. The AUC results are the same as that of the ACC of the models.For the Iris, Wine and Crab datasets the results show 100% performance for GOBAM and some of the other three classic ARTMAP models.

GOBAM has also shown higher SEN values than the other ARTMAP models. Overall, the results in Table 5.2 indicate the better performance of the GOBAM model across all the seven performance metrics for the UCI benchmark datasets. These results confirm the generalisation of GOBAM model as a efficient and reliable classifier for binary classification.

Table 5.2: Performance metrics of some benchmark datasets from the UCI machine learning repository for preliminary evaluation of the proposed GOBAM algorithm using 10-fold cross validation

ARTMAP models	Dataset	AUC	ACC	SEN	SPEC	PPV	NPV	FSCORE
GAFAM	Glass	100	99.09 ±0.02	100	100	100	100	100
GOBAM		**100**	**100**	**100**	**100**	**100**	**100**	**100**
FAM (Carpenter, Grossberg, & Reynolds, 1991)		96.67 ±0.07	98.09 ±0.03	94 ±0.13	99.33 ±0.02	98.33 ±0.05	98.33 ±0.04	98.33 ±0.05
BAM (Vigdor & Lerner, 2007)		**100**	64.7 ± 9.6	**100**	**100**	**100**	**100**	**100**
GAFAM	Breast Cancer Wisconsin	92.41 ±0.05	92.81 ±0.04	90.77 ±0.05	94.06 0.04	90.52 ±0.06	94.5 ±0.05	90.42 ±0.06
GOBAM		**96.05 ±0.02**	**95.79 ±0.02**	**96.52 ±0.02**	**95.59 ± 0.03**	**92.72 ±0.05**	**97.66 ±0.03**	**94.48 ±0.03**
FAM (Carpenter, Grossberg, & Reynolds, 1991)		91.23 ±0.08	89.63 ±0.1	96.34 ±0.06	86.11 0.16	82.54 ±0.17	97.97 ±0.03	82.54 ±0.17

Continued

Table 5.2 – *Continued from previous page*

ARTMAP models	**Dataset**	**AUC**	**ACC**	**SEN**	**SPEC**	**PPV**	**NPV**	**FSCORE**
BAM (Vigdor & Lerner, 2007)		94.31 ±0.04	94.19 ±0.04	94.83 ±0.05	93.8 0.05	90.61 ±0.07	96.86 ±0.03	90.61 ±0.07
GAFAM	Pima Indian Diabetes 1991	65.84 ±0.06	67.27 ±0.06	71.49 ±0.06	60.19 0.1	77.08 ±0.07	53.13 ±0.1	73.89 ±0.05
GOBAM		**73.52 ±0.05**	**76.62 ±0.04**	**83.65 ±0.05**	**63.39 ± 0.09**	**81.13 ±0.04**	**68.02 ±0.09**	**82.23 ±0.04**
FAM citeCarpenter		62.43 ±0.06	65.02 ±0.06	71.08 ±0.06	53.77 ± 0.1	74.16 ±0.05	50.06 ±0.07	74.16 ±0.05
BAM (Vigdor & Lerner, 2007)		67.82 ±0.05	70.44 ±0.05	76.6 ±0.08	59.05 ± 0.11	77.92 ±0.04	58.05 ±0.07	77.92 ±0.04
GAFAM	Iris	100	98.67 ±0.04	100	100	100	100	100
GOBAM		**100**	**100**	**100**	**100**	**100**	**100**	**100**
FAM (Carpenter, Grossberg, & Reynolds, 1991)		100	100	100	100	100	100	100

Continued

Table 5.2 – *Continued from previous page*

ARTMAP models	Dataset	AUC	ACC	SEN	SPEC	PPV	NPV	FSCORE
BAM (Vigdor & Lerner, 2007)		100	100	100	100	100	100	100
GAFAM	Thyroid	100	95.49 ±0.01	100	100	100	100	100
GOBAM		**99.88 ±0.01**	**99.78**	100	**99.76 ± 0.05**	**97.08 ±0.02**	100	**98.5 ±0.01**
FAM (Carpenter, Grossberg, & Reynolds, 1991)		99.53 ±0.01	99.76	99.25 ±0.01	99.8 ± 0.01	97.61 ±0.01	99.94 ±	97.61 ±0.01
BAM (Vigdor & Lerner, 2007)		99.76 ±0.03	99.56	100	99.52 ±	94.41 ±0.02	100	94.41 ±0.02
GAFAM	Crab	100	98.5 ±0.02	100	100	100	100	100
GOBAM		**100**	**100**	**100**	**100**	**100**	**100**	**100**
FAM (Carpenter, Grossberg, & Reynolds, 1991)		98.5 ±0.03	98.5 ±0.03	99 ±0.03	98 ±0.06	98.33 ±0.05	99.09 ±0.03	98.33 ±0.05

Continued

Table 5.2 – *Continued from previous page*

ARTMAP models	**Dataset**	**AUC**	**ACC**	**SEN**	**SPEC**	**PPV**	**NPV**	**FSCORE**
BAM (Vigdor & Lerner, 2007)		100	100	100	100	100	100	100
GAFAM	Wine	100	100	100	100	100	100	100
GOBAM		**100**	**100**	**100**	**100**	**100**	**100**	**100**
FAM (Carpenter, Grossberg, & Reynolds, 1991)		98.2197 ±0.0314	97.5368 ±0.0434	96.4394 ±0.0629	100	100	93.8095 ±0.1048	98.089 ±0.0342
BAM (Vigdor & Lerner, 2007)		100	100	100	100	100	100	100

5.2 Training Performance Results

The performance of the learning curve for GOBAM and GAFAM is displayed as fitness curves for the four subgroups in Figure 5.1 (young male), Figure 5.2 (middle-aged male), Figure 5.3 (young female), Figure 5.4 (middle-aged female). The graphs depict learning curves of using GA in the optimisation parameters and order of sample sequence in GOBAM and GAFAM. The lines in all the graphs show a higher fitness function value for our proposed GOBAM model than the GAFAM model. Furthermore, it can be seen that GOBAM model converged earlier than the GAFAM model. These learning curves can therefore be used to signify GOBAM as a stable model for the prediction of MetS.

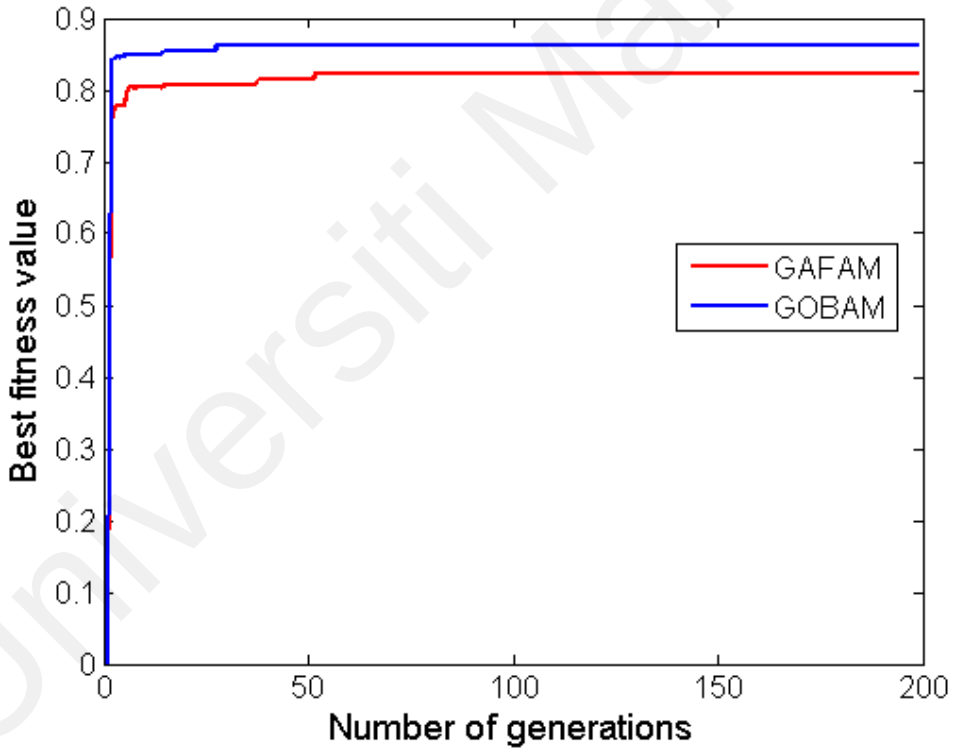

Figure 5.1: GOBAM vs GAFAM fitness learning curves for the diagnosis of MetS in the young male from the CLUSTer dataset

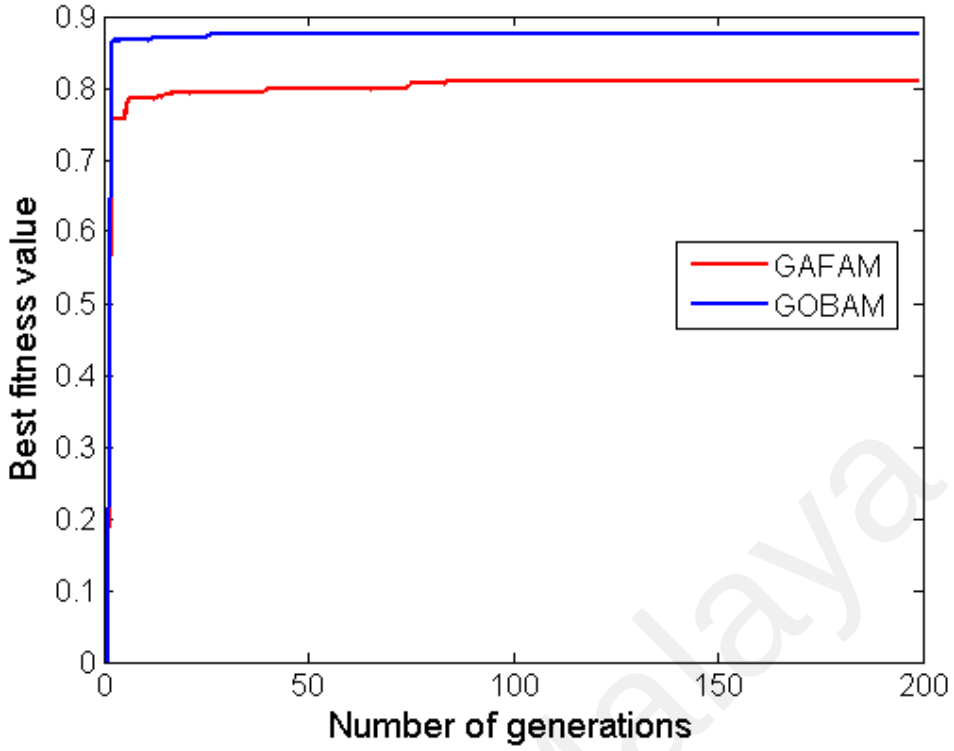

Figure 5.2: GOBAM vs GAFAM fitness learning curves for the diagnosis of MetS in the middle-aged male from the CLUSTer dataset

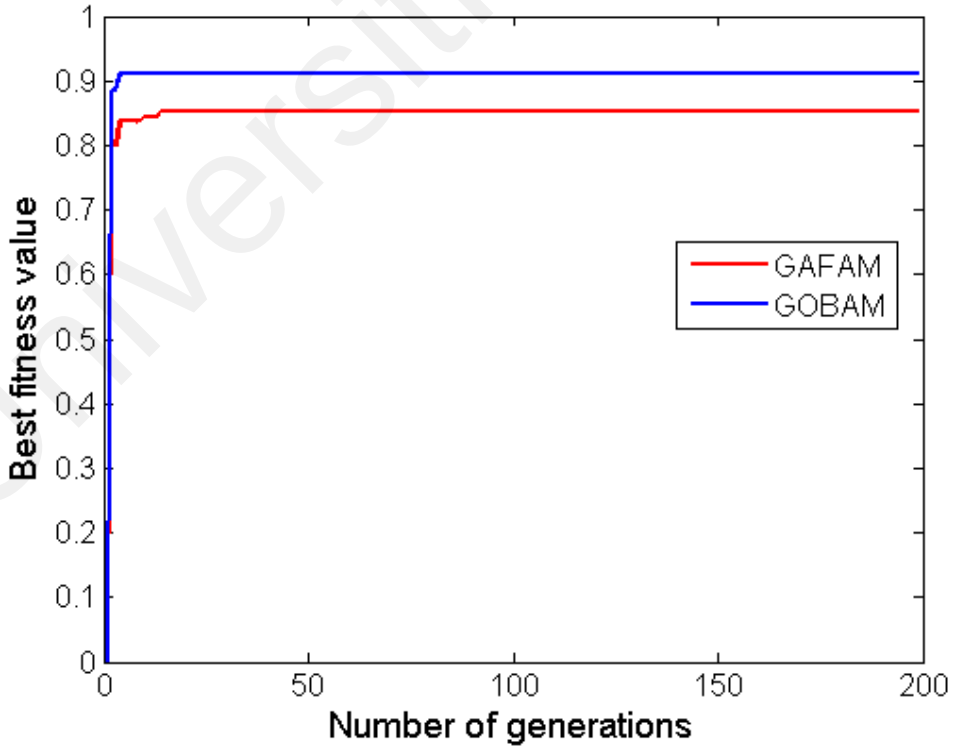

Figure 5.3: GOBAM vs GAFAM fitness learning curves for the diagnosis of MetS in the young female from the CLUSTer dataset

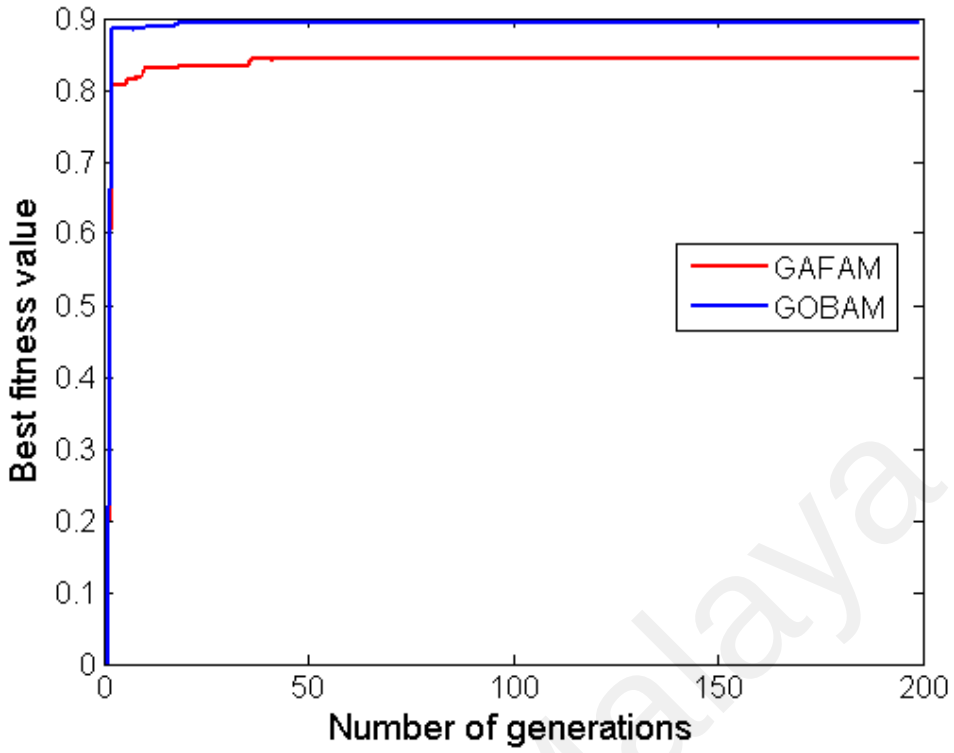

Figure 5.4: GOBAM vs GAFAM fitness learning curves for the diagnosis of MetS in the middle-aged female from the CLUSTer dataset

5.3 Model Evaluation using ROC Curve and the AUC

The ability of our proposed model GOBAM to discriminate between individuals with MetS and those who do not have the abnormality is displayed using the ROC curve. The ROC curves of GOBAM vs GAFAM,FAM, and BAM of CLUSTer dataset is displayed in Figure 5.5 (young male), Figure 5.6 (middle-aged male), Figure 5.7 (young female), and Figure 5.8 (middle-aged female). The ROC curve was adopted as one of the model evaluation tool due the class imbalance found the in the CLUSTer dataset as described in 4.3. The model with an ROC curve that moves towards the upper right corner of the graph is the model with the better predictive performance. An inspection of the ROC curves of the GOBAM compared with other ARTMAP models for all the four subgroups reveals that GOBAM has the highest AUC. These graphs show that the proposed GOBAM model outperforms all the other classical ARTMAP models in the prediction of Metabolic Syndrome (MetS).

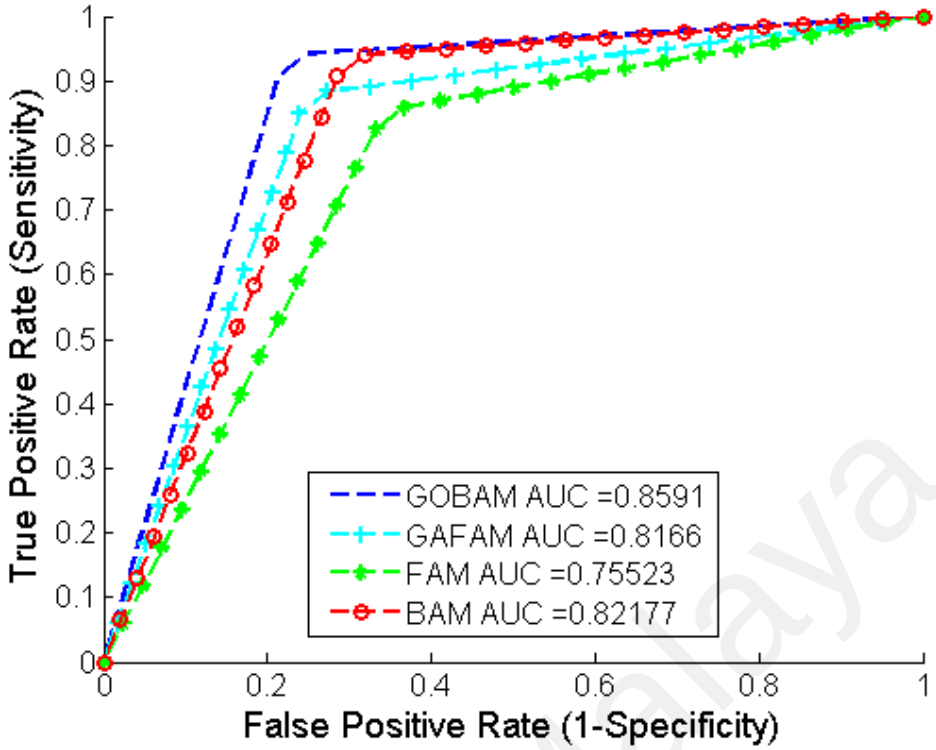

Figure 5.5: ROC curve comparing GOBAM with GAFAM, BAM, and FAM for young male

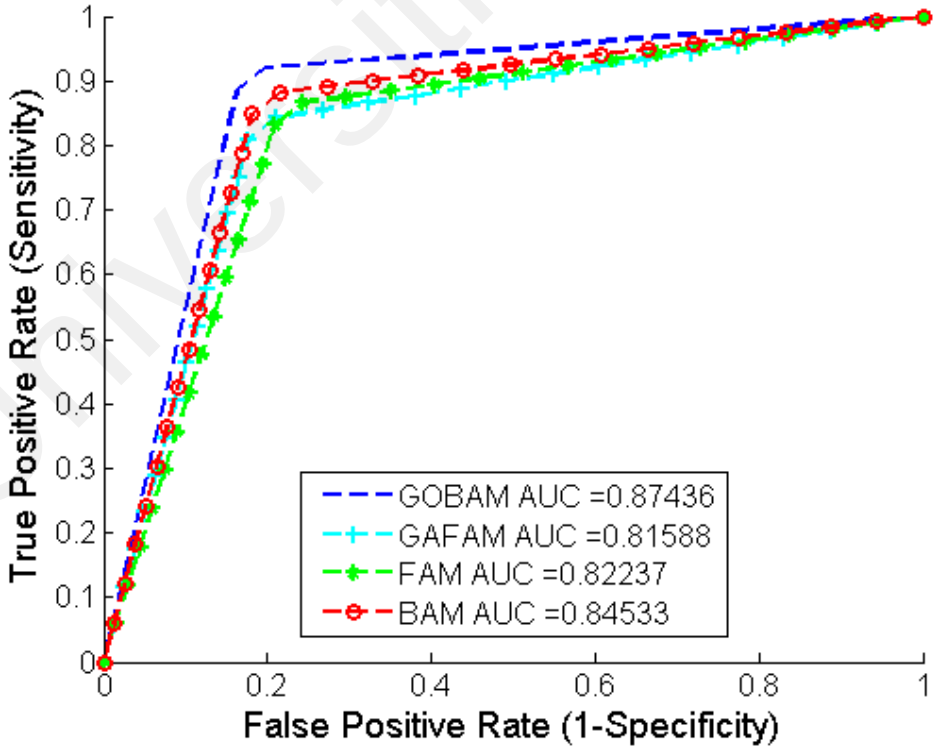

Figure 5.6: ROC curve comparing GOBAM with GAFAM, BAM, and FAM for middle-aged male

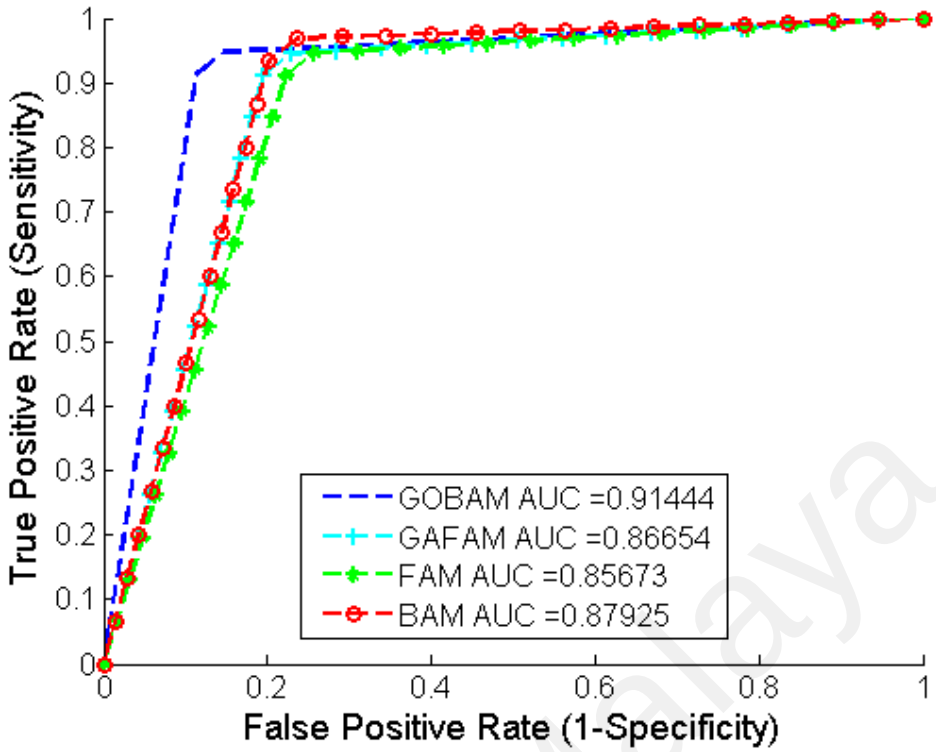

Figure 5.7: ROC curve comparing GOBAM with GAFAM, BAM, and FAM for young female

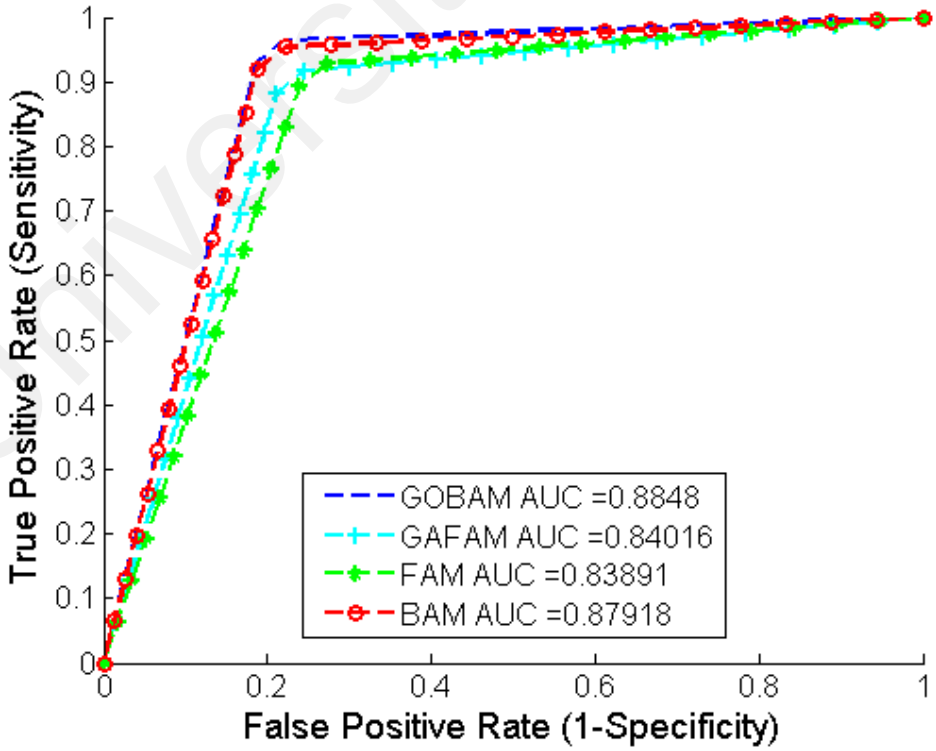

Figure 5.8: ROC curve comparing GOBAM with GAFAM, BAM, and FAM for middle-aged female

An AUC value of 0.5 indicates that a model cannot classify individuals with the abnormality and a value of 1.0 indicates the ability of the model carry out a perfect diagnosis. The AUC is a summary of ROC curve with respect to SEN and SPEC. An inspection of the AUC values in Figure 5.9 for all the prediction of MetS using GOBAM vs GAFAM, FAM, and BAM indicates that GOBAM outperforms the other ARTMAP models for all the young male, middle-aged male, young female and middle-aged female subgroups.

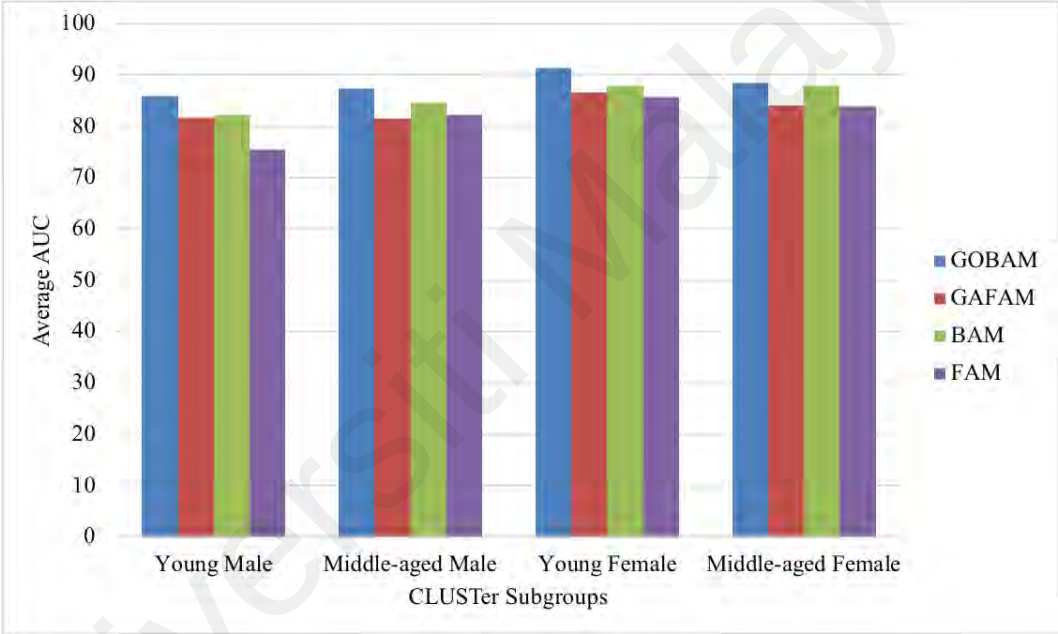

Figure 5.9: Comparative diagram of GOBAM with GAFAM, BAM, and FAM with respect to AUC for all the CLUSTer dataset subgroups

5.4 Confusion Matrix

The average results of the confusion matrix generated from the 10-fold cross validation experiment of GOBAM, GAFAM, BAM, and FAM are presented in Table 5.3 (young male), Table 5.4 (middle-aged male), Table 5.5 (young female), and Table 5.6(middle-aged female). The confusion matrices consists of the TP, FP, FN, and TN values which have already been explained in Section 4.8. For the prediction of MetS, the TPs is the number of

individuals who actually have MetS that GOBAM predicted as having MetS. The TNs are the number of individuals who actually do not have MetS according to the dichotomous definition that GOBAM also predicted as non-MetSs. However, FPs is the number of individuals who do not have MetS but have been predicted by GOBAM as having MetS. This value is also referred to as the Type I error. FNs are the number of individuals who actually have MetS but were predicted by GOBAM as non-MetS. This value is known as the type II error. For the purpose of predicting MetS in our study, one of our main objectives is to predict MetS in individuals who actually do not have MetS according to the clinical definition but present with RF measurements that are very close to the clinically defined thresholds as described in Table 1.1.

Table 5.3: 10-fold cross validation average confusion matrix of GOBAM and the three classic ARTMAP algorithms for the young male CLUSTer subgroup

Young Male				
			Actual	
			MetS	**Non-MetS**
GOBAM	Predicted	Non-MetS	12.3	3
		MetS	3.9	54.8
GAFAM	Predicted	Non-MetS	12.6	8.1
		MetS	3.2	49.7
BAM	Predicted	Non-MetS	11.2	3.5
		MetS	4.7	54.2
FAM	Predicted	Non-MetS	10.3	8.3
		MetS	5.4	48.7

Table 5.4: 10-fold cross validation average confusion matrix of GOBAM and the three classic ARTMAP algorithms for the middle-aged male CLUSTer subgroup

Middle-aged Male				
			Actual	
			MetS	**Non-MetS**
GOBAM	Predicted	Non-MetS	37.4	7.9
		MetS	10.2	84.5
GAFAM	Predicted	Non-MetS	36.1	15.8
		MetS	9.7	73
BAM	Predicted	Non-MetS	36.6	12.8
		MetS	10.1	79
FAM	Predicted	Non-MetS	38.5	11.2
		MetS	8.9	81.1

Table 5.5: 10-fold cross validation average confusion matrix of GOBAM and the three classic ARTMAP algorithms for the young female CLUSTer subgroup

Young Female				
			Actual	
			MetS	**Non-MetS**
GOBAM	Predicted	Non-MetS	29.9	9
		MetS	8.5	293.6
GAFAM	Predicted	Non-MetS	27.6	15.1
		MetS	10.4	285.5
BAM	Predicted	Non-MetS	29.3	16.8
		MetS	8.8	283.6
FAM	Predicted	Non-MetS	30.4	9.8
		MetS	8	292.8

From Tables 5.3, 5.4, 5.5 and 5.6, the FN values are accurately higher than the TP values. This is because the number of non-MetS samples is higher than the number of MetS samples. However, the AUC which is the objective function used in the GOBAM model is impartial to the class imbalance problem.

Table 5.6: 10-fold cross validation average confusion matrix of GOBAM and the three classic ARTMAP algorithms for the middle-aged female CLUSTer subgroup

Middle-aged Female				
			Actual	
			MetS	**Non-MetS**
GOBAM	Predicted	Non-MetS	95.4	18.9
		MetS	26.9	428.8
GAFAM	Predicted	Non-MetS	27.6	42.5
		MetS	27.2	399.7
BAM	Predicted	Non-MetS	90.5	33.2
		MetS	29.8	409.7
FAM	Predicted	Non-MetS	98.6	21.4
		MetS	23.7	425.4

5.5 Validity and Stability

In this section, the predictive performance of GOBAM was compared with that of the three classical ARTMAP models - GAFAM, BAM and FAM. The predictive performance metrics are displayed in Figure 5.10, Figure 5.11, Figure 5.12, Figure 5.13, and Figure 5.14 for SEN, SPEC, PPV, NPV and FSCORE respectively. An inspection of Figure 5.10 shows that the SEN of GOBAM is higher than the SEN of all the other three ARTMAP models in all but one (young female) of the CLUSTer subgroups. This means 94%, 92%, 96% of individuals predicted with MetS by GOBAM are young males, middle-aged males and middle-aged females that actually have MetS according to the the clinical definition. The SEN of BAM is slightly higher than that of GOBAM for the young female subgroup.

Furthermore, the SPEC of GOBAM is higher for all the four subgroups where 78%, 83%, 88%, and 81% of individuals predicted by GOBAM as non-MetS are actually young males, middle-aged males, young females and middle-aged females that do not have MetS according to the clinical definition. In clinical settings, a valid test should have both high SEN and SPEC (Boyce, 2017). Interestingly, an inspection of the GOBAM results in Figure 5.10 and Figure 5.11 shows that a good balance between SEN and SPEC exists for

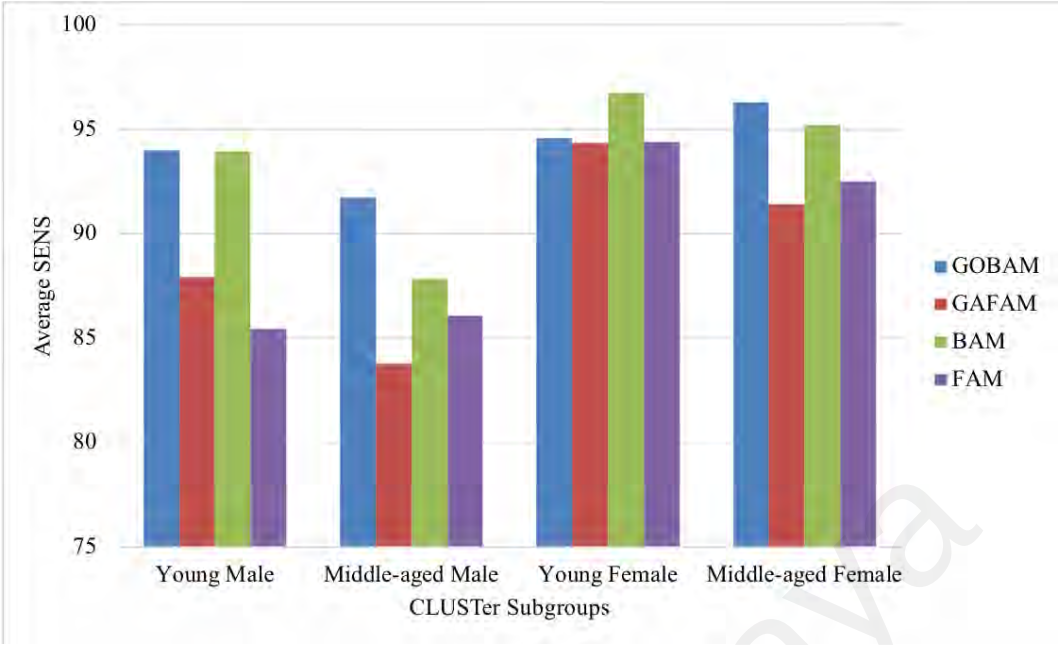

Figure 5.10: Comparative diagram of GOBAM with GAFAM, BAM, and FAM with respect to SEN for all the CLUSTer dataset subgroups

the prediction of MetS. The SEN and SPEC values of GOBAM is in line with previous studies that applied machine learning to predict MetS. (Worachartcheewan et al., 2013), (Hosseini et al., 2014), (J. Chen et al., 2004), (Z. Lin et al., 2014), (Obokata et al., 2015), (Van Schependom et al., 2015), (Worachartcheewan et al., 2015), and (Karimi-Alavijeh et al., 2016) where they report various SEN and SPEC values as 99.85% & 99.86%, 89% & 87.93%, 88.43% & 83.70%, 89.86% & 84.04%, 78% & 54%, 94.89% & 99.15%, and 77.4% & 74.0% respectively. Thus the GOBAM model with both high SEN and SPEC reflect its ability for the early prediction of MetS. However, It can be observed that the overall SEN is higher than the SPEC for all the ARTMAP models in all the subgroups. This is because SPEC is influenced by the prevalence of a condition in the population sample (Brenner & Gefeller, 1997). Therefore, in our experiments, the number of non-MetS samples in the CLUSTer dataset are more than the number of MetS samples as depicted in Figure 4.2.

The percentage of individuals that have been diagnosed with MetS by GOBAM and that actually have MetS according to the clinical definition is the PPV of the model.

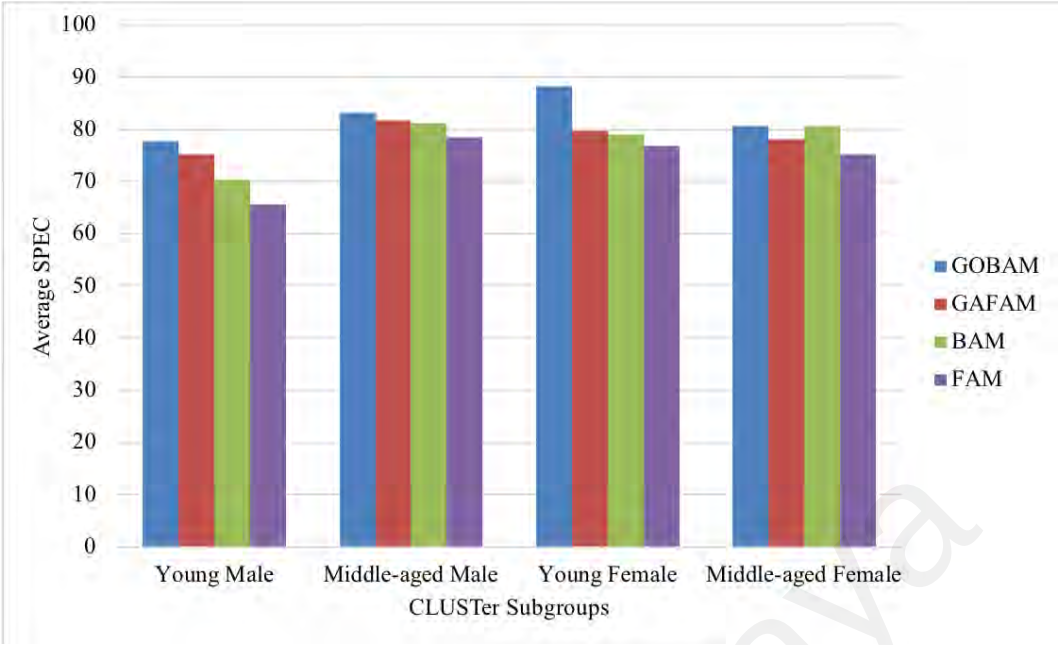

Figure 5.11: Comparative diagram of GOBAM with GAFAM, BAM, and FAM with respect to SPEC for all the CLUSTer dataset subgroups

GOBAM predicts correctly in Figure 5.12 that 93%, 91%, 98%, 95% of the young males, middle-aged males, young females and middle-aged females that have MetS according to the clinical definition are at risk of MetS. The NPV is the percentage of individuals that have been predicted as non-MetS which actually do not have MetS according to the clinical definition. So, GOBAM predicts correctly that 78%, 84%, 76% and 86% of the young males, middle-aged males, young females and middle-aged females do not have MetS actually do not have MetS according to the clinical definition. The results of AUC, ACC, SEN, SPEC, PPV, NPV, and FSCORE for GOBAM, ARTMAP models - GAFAM, BAM, and FAM are presented in Table B.1, Table B.2, Table B.3, Table B.4, Table B.5, Table B.6, Table B.1 are presented in Appendix B.

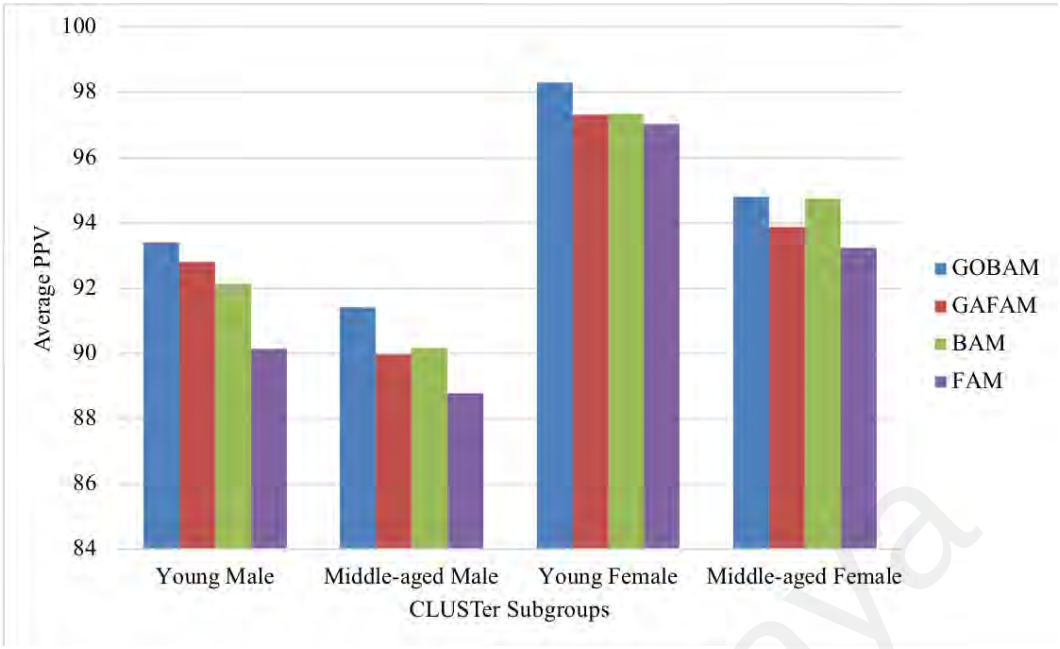

Figure 5.12: Comparative diagram of GOBAM with GAFAM, BAM, and FAM with respect to PPV for all the CLUSTer dataset subgroups

In a two class problem such as the prediction of the risk of MetS, the predictive model's ability to maximise the percentage of cases with MetS is directly proportional to maximising the PPV and/or sensitivity (true positive rate) percentage. MetS is not directly a disease but a constellation of abnormalities which may result to diseases that are fatal (K. Alberti et al., 2009). Therefore, the early diagnosis or prediction of MetS risk can be used for the early prevention, identification or treatment of the associated diseases of MetS. Consequently, a good predictive or diagnostic model is one with a high PPV or sensitivity that maximises the number of people correctly diagnosed as being at risk of MetS. The percentage of individuals diagnosed positively by GOBAM as having MetS who actually have MetS according to the clinical definition is the PPV (Boyce, 2017). The percentage of individuals diagnosed negatively by GOBAM as non-MetS who actually do not have MetS according to the clinical definition is the NPV (Boyce, 2017).

A high positive diagnosis of the number of people is more required than the accuracy of the prediction model. Hence, the PPV and/or sensitivity are seen as more efficient performance metrics in both statistical and machine learning models that predict or

diagnose the risk of MetS. Figure 5.15 presents the time complexity of the GOBAM, GAFAM, BAM and FAM for the all the four subgroups. The time complexity the optimised BAM and FAM is higher than that of the classic BAM and FAM. This is due to the GA's high processing time requirement. The dataset with largest number of samples has the highest computational cost. Generally, GOBAM has the highest computational cost and is consistent in all the datasets. This high time complexity is one of the limitations of this work.

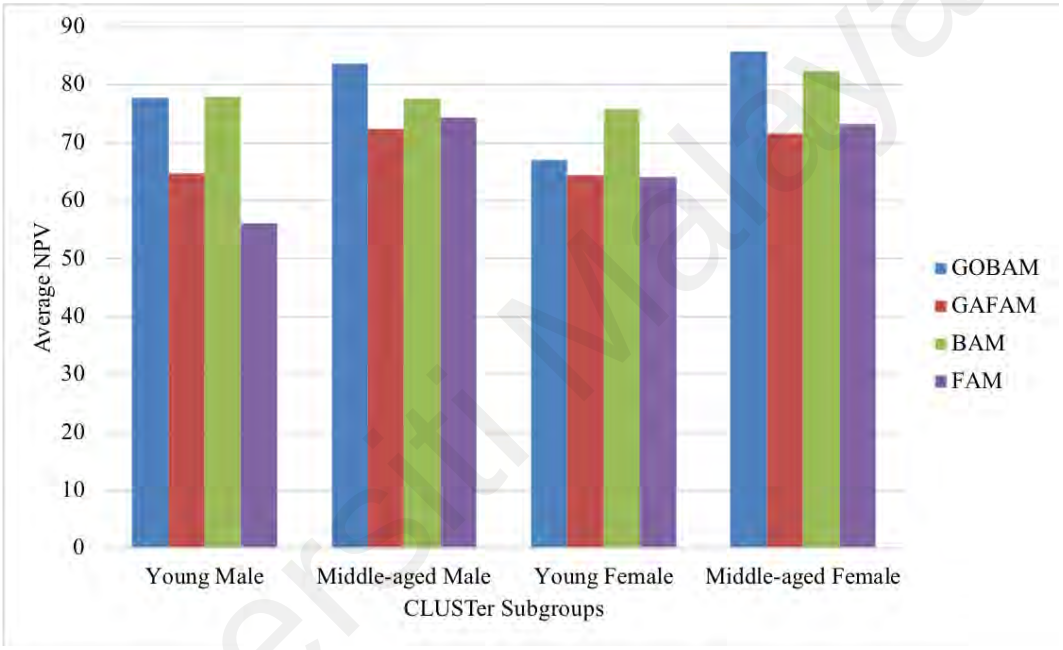

Figure 5.13: Comparative diagram of GOBAM with GAFAM, BAM, and FAM with respect to NPV for all the CLUSTer dataset subgroups

5.6 Statistical Test

In order to analyse the statistical significance of the performance the proposed GOBAM model, the Friedman test is applied (Friedman, 2001). The Friedman test is a non-parametric test that ranks a performance measure of multiple algorithms for each data set separately. The Friedman test can measure the statistical difference of machine learning algorithms based on their predictive performance measure rankings. The algorithm with the best performance measure is ranked highest, followed by the second best in rank

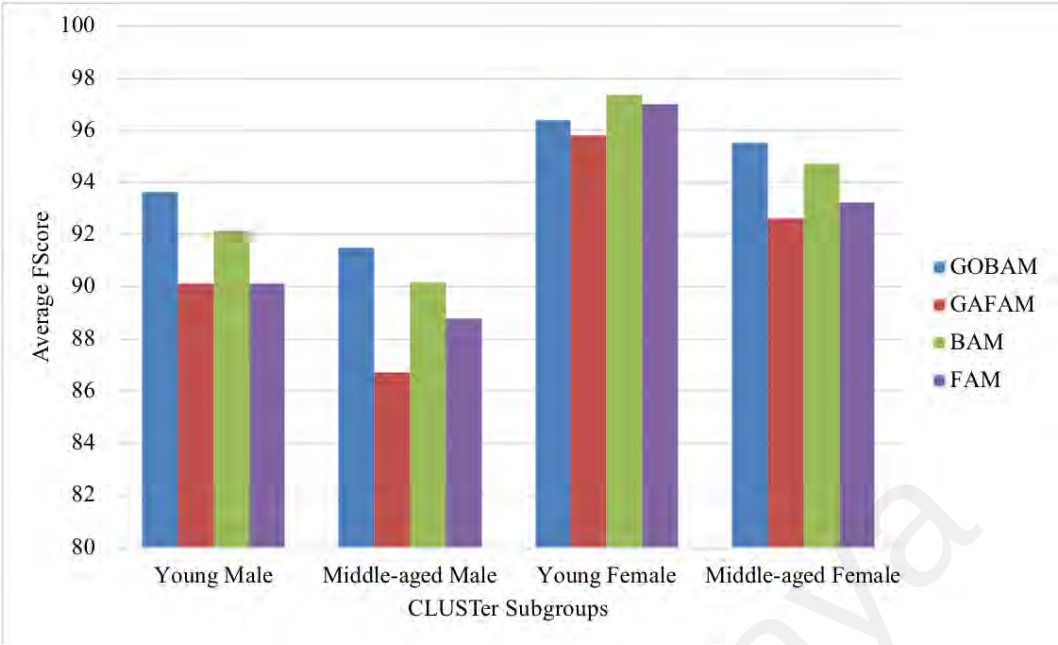

Figure 5.14: Comparative diagram of GOBAM with GAFAM, BAM, and FAM with respect to FSCORE for all the CLUSTer dataset subgroups

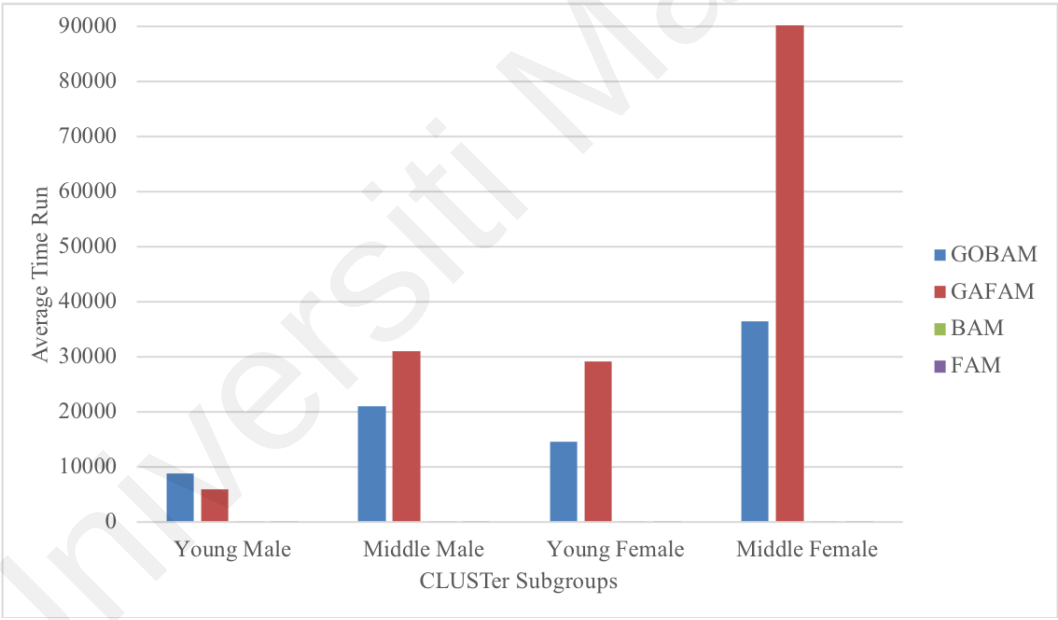

Figure 5.15: Comparative diagram of GOBAM with GAFAM, BAM, and FAM with respect to time complexity for all the CLUSTer dataset subgroups

and so on. In the case where algorithms are tied, an average ranks are assigned. This statistical test was conducted for each of the six predictive performance measures displayed in Figure 5.10 (SEN), Figure 5.11 (SPEC), Figure 5.12 (PPV), Figure 5.13 (NPV), and Figure 5.14 (FSCORE).

Table 5.7 shows the Friedman results for the average mean performance metrics of the

Table 5.7: Friedman Test Results for comparison of GOBAM, BAM, GAFAM, and FAM predictive performance metrics

Mean Ranks						
	AUC	**SEN**	**SPEC**	**PPV**	**NPV**	**FSCORE**
GOBAM	4	4	4	3.5	3.5	2.5
BAM	3	2.25	2.75	3.5	3.25	4
GAFAM	1.75	2.75	2.25	1.5	1.25	1
FAM	1.25	1	1	1.5	2	2.5

four CLUSTer subgroups, young male, middle-aged male, young female and middle-aged female for AUC, SEN, SPEC, PPV, NPV and FSCORE. Since there are four models, GOBAM, BAM, GAFAM, and FAM, to be compared, the rank "1" indicates the model with lowest predictive performance while the rank value "4" indicates the model with the highest predictive performance. From Table 5.7, it can be seen that the proposed GOBAM model ranks highest in AUC, SEN, SPEC, and NPV predictive performance metrics. However, it ties with BAM in the PPV performance metric and ties with FAM for the third place in th FSCORE performance metric. These results show the superiority of the predictive performance and generalisability of the proposed MetS model over the other ARTMAP models. Furthermore, it confirms validity of the proposed GOBAM as a tool for risk quantification and prediction of MetS.

5.7 Clinical Utility of Metabolic Syndrome Risk Quantification Using GOBAM

Having ascertained the high predictive performance of GOBAM, it is now necessary to discuss the clinical utility GOBAM. This research work also focuses on the prediction of MetS in individuals who present with MRF measurements that are close to the clinically defined thresholds as stated in Table 1.1. The proposed GOBAM model outputs a posterior probability that is associated with the class category of an input. This probability value ranges between 0 and 1. It is a probabilistic value that can be used as a risk quantification measure related to the prediction of MetS by GOBAM. The clinical utility of GOBAM

as risk quantification and prediction method of MetS can be explained in terms of this value. The posterior probability value derived from the GOBAM model has the same range as AUC, between 0 and 1. Therefore, this research work adopts the same value range to defined the risk quantification of MetS using GOBAM. The risk quantification of MetS is defined in Table 5.8.

Table 5.8: Risk quantification for MetS for borderline non-MetS diagnosis from the GOBAM model

MetS class	Risk Quantification Value	Description
0 (non-MetS)	0.8 - 1	Very low risk of MetS
	0.7 - 0.8	Low risk of MetS
	0.6 - 0.7	Questionable risk of MetS
	0.5 - 0.6	High risk of MetS
	< 0.5	Very High risk of MetS

The MRF measurements in Table 5.9 will be used to describe the clinical utility of risk quantification value derived from the GOBAM model. For example, let us assume that a middle-aged male presents with a BMI of 28 kg/m^2 and MRF measurements as shown in Table 5.9. According to the JIS (K. Alberti et al., 2009) definition, he will not be diagnosed as being at risk of MetS (See section 1.1) because only the WC and TG measurements have exceeded the clinical thresholds. However, the FPG, HDL-C and BP MRF measurement values are very close to the clinical thresholds and if unchecked, this individual will be at risk of MetS in the nearest future. To identify these risks, GOBAM generates probability estimations for each class prediction which can aid in ranking an individual with borderline values. After running this sample through our proposed GOBAM model, this middle aged male is classified as having non-MetS with 0.2567 posterior probability. This means that although he has been classified correctly as not having MetS, the risk quantification of having MetS is "very High" as stated in Table 5.9 and thus he is at risk of having MetS. This results can be used for the early diagnosis of MetS in order to administer early intervention

methods to prevent the progression of the MRF measurement values from exceeding the clinical threshold.

Table 5.9: MetS risk factor sample measurement for middle-aged Asian male and diagnosis according to JIS definition

MRFs	Patient MRF measurements	JIS threshold	JIS MetS Assessment
FPG	5.5 mmol/L	$\geq$ 5.6 mmol/L	No
WC	108 cm	> 102 cm	Yes
TG	1.9 mmol/L	$\geq$ 1.7mmol/L	Yes
HDL-C	1.2 mmol/L	< 1.0 mmol/L	No
BP	129/80 mmHg	$\geq$ 130/85 mmHg	No

5.8 Summary

The results of the initial preliminary experiments and the final experiments conducted in this research work were presented. The preliminary experiments were conducted to evaluate the overall efficiency and generalisability of the proposed GOBAM. A detailed discussion on the evaluation of the predictive performance of GOBAM against GAFAM, BAM and FAM was presented. The results of the performance metrics of the four CLUSTer subgroups, young male, middle-aged male, young female and middle-aged female for AUC, SEN, SPEC, PPV, NPV and FSCORE were presented and discussed in detail. The Friedman statistical test was also conducted to evaluate the generalisability and efficiency of the proposed GOBAM model. And finally an example of the risk quantification utility of the proposed GOBAM was presented.

CHAPTER 6: CONCLUSION

6.1 Introduction

This chapter presents a summary of the research findings on the computational intelligence approach risk quantification and prediction of MetS. The aim of this research is to investigate the development and application of machine learning techniques for the prediction and risk quantification of MetS. This research also considers the risk quantification and diagnosis of MetS in individuals who present with MRF measurements that are close to the clinically defined threshold in Table 1.1. The MRF measurements was collected from the CLUSTer dataset.

6.2 Summary of Research Findings

The challenges and demands in the area of diagnosis, monitoring and management of MetS are increasing. These problems have lead a number of studies to be carried out for the application of non-clinical techniques in order to support the clinical diagnosis of MetS. The early diagnosis of MetS is important because it usually preceeds the onset and development of major challenging NCDs such as T2DM. However, both the clinical definitions and the existing non-clinical methods used to diagnose MetS have some limitations as mentioned in Section 1.2. The clinical definition of MetS is mainly limited by loss of information because it diagnosis MetS by dichotomising the five clinical MRF measurements. Furthermore, it is difficult to diagnosis the risk of MetS in individuals who present with borderline MRF measurements that are close to the clinically defined MRF measurements. The non-clinical techniques used to diagnose the risk MetS are sample specific, inefficient for early diagnosis of the risk of Metabolic Syndrome (MetS) and cannot be generalised for use in clinical utility. In order to overcome these problems, this research work addressed the following research objectives as presented in Section 1.4. We

achieved the first objective by conducting an exhaustive literature review of existing studies that applied non-clinical techniques to predict the risk of MetS in various populations. Research Question 1 was answered in chapter 2 to achieve the Research Objective 1 in Section 1.4.

The findings revealed that three categories of non-clinical techniques have been adopted by various studies to predict the risk of MetS using various population samples. These methods include statistical techniques, risk quantification and machine learning techniques. Furthermore, it was realised that non of the studies identified used both prediction and risk quantification at the same time to diagnose the risk of MetS. Furthermore, the risk of MetS has not been predicted especially in individuals who present with MRF measurements that are close to the clinically defined MRF thresholds as defined in Table 1.1. Research Objective 2 was realised by proposing a non-clinical clinical approach for the risk quantification and prediction of MetS based on the BAM and GA known as GOBAM.

This is a novel machine learning model that genetically optimises the parameters and input sequence order of the BAM. The GOBAM model was able to quantify and predict the risk of MetS while considering individuals who present with borderline MRF measurements that are close to the clinically defined threshold.

In order to achieve the Research Objective Research Questions 3, Research Questions 4 and Research Questions 5 were answered. The CLUSTer dataset was used to conduct preliminary and main experiments on the proposed GOBAM model and three other classic ARTMAP models, GAFAM, FAM, and BAM. The six predictive performance metrics, namely, ROC curves, AUC, SEN, SPEC, PPV, NPV and FSCORE were generated. A comparative analysis was carried out to ascertain, the efficiency, validity and generalisability of the GOBAM model against the other classical ARTMAP models.

6.3 Research Contribution

This research work has applied GA and BAM in a computational intelligence approach to develop a novel machine learning model GOBAM which can be used for the risk quantification and prediction of MetS. The contributions of this research work include the following:

- The use of GA to optimise the parameters and the sample sequence order of the BAM has improved the predictive performance of the BAM as an variant of ARTMAP.
- The proposed GOBAM model was able to predict the risk of MetS with a maximum of 91.45% correctness probability.
- The computational intelligence of the GOBAM model was exhibited predicting the risk quantification of MetS for individuals who present with MRF measurements that are close to the clinically defined MRF thresholds.
- Comparison with classical ARTMAP models show that the GOBAM outperforms GAFAM, FAM, and BAM in terms of ROC curve, AUC, SEN, SPEC, and NPV predictive performance metrics.
- The GOBAM model exhibits clinical validity because its SEN and SPEC metric performance values are balanced. Therefore it is sensitive to individuals with MetS and specific to isolating non-MetS.
- Medical practitioners can use the proposed model to determine baseline risk of future disease for their patients.

6.4 Future Work

The practical utility of the risk quantification and prediction of MetS for the temporal progression and management of its two major associated diseases, T2DM and CVD will

be carried out in future works. The proposed machine learning model will be applied in a long term epidemiological study. The risk quantification and prediction of MetS will be used to monitor and identify the both collective and isolated effects of the MRFs to CVD and T2DM. The GA model used in this research work has many parameters that require fine tuning. Therefore, we intend to use less demanding optimisation techniques to improve the performance of the ARTMAP models in future research.

This research work proposed a novel MetS risk quantification and prediction model, GOBAM. GOBAM is able to support the clinical diagnosis of MetS by resolving its weaknesses. The proposed model also showed clinical validity by extensive performance evaluation using a large population sample MRF examination results for young males, middle-aged male, young females and middle-aged females. The evaluation results showed that the proposed model can quantify and predict the risk of MetS and effectively. The imbalance nature of the dataset will be investigated. Various imbalance techniques will be applied to the dataset and the results of the experiments will be analysed in order to further ascertain the efficiency of GOBAM.

Furthermore, GOBAM will be explored as an online and incremental learning model. This features of the algorithm will be applied in tracking the progression of the risk of MetS over time.

6.5 Limitations of the Study

This study has the following limitations. First, the cohort study for the dataset used was restricted to Malaysian school teachers. Secondly, the incidence of MetS was low in the dataset, hence the low rate of true positive predictions. This bias was decreased by performing the 10-fold cross validation and averaging the results. Moreover, results from the other classic ARTMAP algorithms are similar to our results from the proposed GOBAM algorithm. This further ensures the high performance of the results obtained.

Furthermore, the dataset is imbalanced with more samples not having MetS as the clinical definition indicates. The problem of imbalanced datasets in predominant in healthcare (Perveen et al., 2019). This is as a result of the real life nature of the dataset, in this case and its population characteristic being teachers. The imbalance problem can be tackled by preprocessing the dataset using various sampling methods (Melillo, De Luca, Bracale, & Pecchia, 2013) in order to improve the performance of the risk prediction model. The procedure may include testing the efficacy sampling methods like random undersampling, ramdom oversampling, and the K-Medoids undersampling (Perveen et al., 2019).

REFERENCES

Agarwal, S., Jacobs, D. R., Vaidya, D., Sibley, C. T., Jorgensen, N. W., Rotter, J. I., ... Kritchevsky, S. (2012). Metabolic syndrome derived from principal component analysis and incident cardiovascular events: the multi ethnic study of atherosclerosis (mesa) and health, aging, and body composition (health abc). *Cardiology Research and Practice*, *2012*.

Aguilar, M., Bhuket, T., Torres, S., Liu, B., & Wong, R. J. (2015). Prevalence of the metabolic syndrome in the united states, 2003-2012. *Jama*, *313*(19), 1973–1974.

Alberti, G., Zimmet, P., Shaw, J., Grundy, S. M., et al. (2006). The idf consensus worldwide definition of the metabolic syndrome. *Brussels: International Diabetes Federation*, *23*(5), 469–80.

Alberti, K., Eckel, R. H., Grundy, S. M., Zimmet, P. Z., Cleeman, J. I., Donato, K. A., ... Smith, S. C. (2009). Harmonizing the metabolic syndrome a joint interim statement of the international diabetes federation task force on epidemiology and prevention; national heart, lung, and blood institute; american heart association; world heart federation; international atherosclerosis society; and international association for the study of obesity. *Circulation*, *120*(16), 1640-1645.

Alberti, K. G. M. M., Zimmet, P., & Shaw, J. (2006). Metabolic syndrome - a new world-wide definition. a consensus statement from the international diabetes federation. *Diabetic Medicine*, *23*(5), 469-480.

Anderson, P., Critchley, J., Chan, J., Cockram, C., Lee, Z., Thomas, G., & Tomlinson, B. (2001). Factor analysis of the metabolic syndrome: obesity vs insulin resistance as the central abnormality. *International Journal of Obesity*, *25*(12), 1782.

Ayubi, E., Khalili, D., Delpisheh, A., Hadaegh, F., & Azizi, F. (2015). Factor analysis of metabolic syndrome components and predicting type 2 diabetes: Results of 10-year follow-up in a middle eastern population. *Journal of Diabetes*, *7*(6), 830-838.

Babyak, M. A., & Green, S. B. (2010). Confirmatory factor analysis: an introduction for psychosomatic medicine researchers. *Psychosomatic Medicine*, *72*(6), 587–597.

Balkau, B., & Charles, M.-A. (1999). Comment on the provisional report from the who consultation. *Diabetic Medicine*, *16*(5), 442-443.

Balkau, B., Deanfield, J. E., Després, J.-P., Bassand, J.-P., Fox, K. A., Smith, S. C., ... others (2007). International day for the evaluation of abdominal obesity (idea): a study of waist circumference, cardiovascular disease, and diabetes mellitus in 168 000 primary care patients in 63 countries. *Circulation*, *116*(17), 1942–1951.

Bateman, L. A., Slentz, C. A., Willis, L. H., Shields, A. T., Piner, L. W., Bales, C. W., ... Kraus, W. E. (2011). Comparison of aerobic versus resistance exercise training effects on metabolic syndrome (from the studies of a targeted risk reduction intervention through defined exercise - strride-at/rt). *American Journal of Cardiology*, *108*(6), 838-44.

Batey, L. S., Goff, D. C., Tortolero, S. R., Nichaman, M. Z., Chan, W., Chan, F. A., ... Labarthe, D. R. (1997). Summary measures of the insulin resistance syndrome are adverse among mexican-american versus non-hispanic white children. *Circulation*, *96*(12), 4319-4325.

Baumann, P., Hochbaum, D., & Yang, Y. (2018). A comparative study of the leading machine learning techniques and two new optimization algorithms. *European Journal of Operational Research*. doi: https://doi.org/10.1016/j.ejor.2018.07.009

Boyce, D. (2017). Chapter 17 - evaluation of medical laboratory tests. In J. D. Placzek & D. A. Boyce (Eds.), *Orthopaedic physical therapy secrets (third edition)* (Third Edition ed., pp. 125–134). Elsevier. doi: https://doi.org/10.1016/B978-0-323-28683-1.00017-5

Bradley, A. P. (1997). The use of the area under the roc curve in the evaluation of machine learning algorithms. *Pattern Recognition*, *30*(7), 1145–1159.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32.

Brenner, H., & Gefeller, O. (1997). Variation of sensitivity, specificity, likelihood ratios and predictive values with disease prevalence. *Statistics in Medicine*, *16*(9), 981–991.

Brozova, J., Cechurova, D., & Lacigova, S. (2016). [metabolic syndrome in patients with diabetes mellitus type 1, prevalence, impact on morbidity and mortality]. *Vnitrni Lekarstvi*, *62*(Suppl 4), 8-14.

Cani, P. D., & Hul, M. V. (2015). Novel opportunities for next-generation probiotics targeting metabolic syndrome. *Current Opinion in Biotechnology*, *32*, 21 - 27. doi: https://doi.org/10.1016/j.copbio.2014.10.006

Carpenter, G. A. (1989). Neural network models for pattern recognition and associative memory. *Neural Networks*, 3.

Carpenter, G. A., & Grossberg, S. (1988). The art of adaptive pattern recognition by a self-organizing neural network. *IEEE Computer*, *21*(3), 77-88.

Carpenter, G. A., Grossberg, S., Markuzon, N., Reynolds, J. H., & Rosen, D. B. (1992). Fuzzy artmap: A neural network architecture for incremental supervised learning of analog multidimensional maps. *IEEE Transactions on Neural Networks*, *3*(5), 698-713.

Carpenter, G. A., Grossberg, S., & Reynolds, J. H. (1991). Artmap: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network. *Neural Networks*, *4*(5), 565-588.

Carpenter, G. A., Grossberg, S., & Rosen, D. B. (1991). Fuzzy art: Fast stable learning and categorization of analog patterns by an adaptive resonance system. *Neural Networks*, *4*(6), 759-771.

Carpenter, G. A., & Markuzon, N. (1998). Artmap-ic and medical diagnosis: Instance counting and inconsistent cases. *Neural Networks*, *11*(2), 323-336.

Carpenter, G. A., & Ross, W. D. (1995). Art-emap: A neural network architecture for object recognition by evidence accumulation. *IEEE Transactions on Neural Networks*, *6*(4), 805-818.

Carroll, S., Cooke, C., & Butterly, R. (2000). Metabolic clustering, physical activity and fitness in nonsmoking, middle-aged men. *Medicine & Science in Sports & Exercise*, *32*(12), 2079–2086.

Carroll, S. J., Paquet, C., Howard, N. J., Adams, R. J., Taylor, A. W., & Daniel, M. (2014). Validation of continuous clinical indices of cardiometabolic risk in a cohort of australian adults. *BMC Cardiovascular Disorder*, *14*, 27.

Cervantes, L., Lee, J.-s., & Lee, J. (2007). Agent-based approach to distributed ensemble learning of fuzzy artmap classifiers. In N. T. Nguyen, A. Grzech, R. J. Howlett, & L. C. Jain (Eds.), *Agent and multi-agent systems: Technologies and applications: First kes international symposium, kes-amsta 2007, wroclaw, poland, may 31â€“ june 1, 2007. proceedings* (p. 805-814). Berlin, Heidelberg: Springer Berlin Heidelberg.

Chang, C.-H., Yen, C.-H., & Chen, M.-Y. (2010). Using principal component analysis to develop a single-parameter screening tool for metabolic syndrome. *BMC public health*, *10*(1), 708.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). Smote: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, *16*, 321–357.

Chawla, N. V., Japkowicz, N., & Kotcz, A. (2004). Special issue on learning from imbalanced data sets. *ACM Sigkdd Explorations Newsletter*, *6*(1), 1–6.

Chen, H., Xiong, S., & Ren, X. (2014). Evaluating the risk of metabolic syndrome based on an artificial intelligence model. *Abstract and Applied Analysis*, *2014*.

Chen, J., Muntner, P., Hamm, L. L., Jones, D. W., Batuman, V., Fonseca, V., ... He, J. (2004). The metabolic syndrome and chronic kidney disease in us adults. *Annals of Internal Medicine*, *140*(3), 167-174.

Coley, D. A. (1999). *An introduction to genetic algorithms for scientists and engineers*. World Scientific Publishing Company.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297.

Craig, C. L., Marshall, A. L., Sjöström, M., Bauman, A. E., Booth, M. L., Ainsworth, B. E., ... others (2003). International physical activity questionnaire: 12-country reliability and validity. *Medicine & Science in Sports & Exercise*, *35*(8), 1381–1395.

Dai, L., Gonçalves, C. M. V., Lin, Z., Huang, J., Lu, H., Yi, L., ... An, D. (2015). Exploring metabolic syndrome serum free fatty acid profiles based on gc–sim–ms combined with random forests and canonical correlation analysis [Journal Article]. *Talanta*, *135*, 108-114.

DeBoer, M. D., Filipp, S. L., & Gurka, M. J. (2018). Use of a metabolic syndrome severity z score to track risk during treatment of prediabetes: An analysis of the diabetes prevention program [Journal Article]. *Diabetes Care*, *41*(11), 2421-2430.

DeBoer, M. D., Gurka, M. J., Woo, J. G., & Morrison, J. A. (2015). Severity of metabolic syndrome as a predictor of cardiovascular disease between childhood and adulthood:

the princeton lipid research cohort study. *Journal of the American College of Cardiology*, *66*(6), 755–757.

DeBoer, M. D., Gurka, M. J., Woo, J. G., & Morrison, J. A. (2015). Severity of the metabolic syndrome as a predictor of type 2 diabetes between childhood and adulthood: the princeton lipid research cohort study. *Diabetologia*, *58*(12), 2745-52.

de Carvalho Vidigal, F., Bressan, J., Babio, N., & Salas-Salvadó, J. (2013). Prevalence of metabolic syndrome in brazilian adults: a systematic review. *BMC Public Health*, *13*(1), 1198.

de Edelenyi, F. S., Goumidi, L., Bertrais, S., Phillips, C., MacManus, R., Roche, H., . . . Lairon, D. (2008). Prediction of the metabolic syndrome status based on dietary and genetic parameters, using random forest. *Genes & Nutrition*, *3*(3), 173.

Dekker, J. M., Girman, C., Rhodes, T., Nijpels, G., Stehouwer, C. D., Bouter, L. M., & Heine, R. J. (2005). Metabolic syndrome and 10-year cardiovascular disease risk in the hoorn study. *Circulation*, *112*(5), 666–673.

De Luca, C., & Olefsky, J. M. (2008). Inflammation and insulin resistance. *FEBS Letters*, *582*(1), 97–105.

De Maesschalck, R., Jouan-Rimbaud, D., & Massart, D. L. (2000). The mahalanobis distance. *Chemometrics and Intelligent Laboratory Systems*, *50*(1), 1–18.

Demir, S., Artim-Esen, B., Sahinkaya, Y., Pehlivan, O., Alpay-Kanitez, N., Omma, A., . . . Inanc, M. (2016). Metabolic syndrome is not only a risk factor for cardiovascular diseases in systemic lupus erythematosus but is also associated with cumulative organ damage: a cross-sectional analysis of 311 patients. *Lupus*, *25*(2), 177-84.

Dheeru, D., & Karra Taniskidou, E. (2017). *UCI machine learning repository.*

Dietterich, T. G. (1999). Machine learning. In R. A. Wilson & F. C. Keil (Eds.), *The mit encyclopedia of the cognitive sciences.* MIT Press.

Direk, K., Cecelja, M., Astle, W., Chowienczyk, P., Spector, T. D., Falchi, M., & Andrew, T. (2013). The relationship between dxa-based and anthropometric measures of visceral fat and morbidity in women. *BMC Cardiovascular Disorders*, *13*(1), 25.

Domingos, F., & Serra, A. (2014). Metabolic syndrome: a multifaceted risk factor for kidney stones. *Scandinavian Journal of Urology*, *48*(5), 414-9.

Drager, L. F., Togeiro, S. M., Polotsky, V. Y., & Lorenzi-Filho, G. (2013). Obstructive sleep apnea: a cardiometabolic risk in obesity and the metabolic syndrome. *Journal of the American College of Cardiology*, *62*(7), 569–576.

Dragsbæk, K., Neergaard, J. S., Laursen, J. M., Hansen, H. B., Christiansen, C., Beck-Nielsen, H., . . . Henriksen, K. (2016). Metabolic syndrome and subsequent risk of type 2 diabetes and cardiovascular disease in elderly women: Challenging the current definition. *Medicine*, *95*(36).

Drehmer, M., Pereira, M. A., Schmidt, M. I., Alvim, S., Lotufo, P. A., Luft, V. C., & Duncan, B. B. (2016). Total and full-fat, but not low-fat, dairy product intakes are inversely associated with metabolic syndrome in adults. *Journal of Nutrition*, *146*(1), 81-9.

DuBose, K. D., & McKune, A. J. (2014). The relationship between objectively measured physical activity, salivary cortisol, and the metabolic syndrome score in girls. *Pediatric Exercise Science*, *26*(3), 221-30.

Dusseault-Belanger, F., Cohen, A. A., Hivert, M. F., Courteau, J., & Vanasse, A. (2013). Validating metabolic syndrome through principal component analysis in a medically diverse, realistic cohort. *Metabolic Syndrome and Related Disorders*, *11*(1), 21-28.

Ebbert, J. O., & Jensen, M. D. (2013). Fat depots, free fatty acids, and dyslipidemia. *Nutrients*, *5*(2), 498–508.

Eckel, R. H., Grundy, S. M., & Zimmet, P. Z. (2005). The metabolic syndrome. *The Lancet*, *365*(9468), 1415–1428.

Eiben, A. E., Smith, J. E., et al. (2003). *Introduction to evolutionary computing* (Vol. 53). Springer.

Eisenmann, J. C. (2008). On the use of a continuous metabolic syndrome score in pediatric research. *Cardiovascular Diabetology*, *7*.

Eisenmann, J. C., Wickel, E. E., Welk, G. J., & Blair, S. N. (2005). Relationship between adolescent fitness and fatness and cardiovascular disease risk factors in adulthood:

the aerobics center longitudinal study (acls). *American Heart Journal*, *149*(1), 46-53.

Ekelund, U., Brage, S., Froberg, K., Harro, M., Anderssen, S. A., Sardinha, L. B., ... Andersen, L. B. (2006). Tv viewing and physical activity are independently associated with metabolic risk in children: the european youth heart study. *PLoS Medicine*, *3*(12), e488.

Esposito, K., Pontillo, A., Di Palo, C., Giugliano, G., Masella, M., Marfella, R., & Giugliano, D. (2003). Effect of weight loss and lifestyle changes on vascular inflammatory markers in obese women: a randomized trial. *Jama*, *289*(14), 1799–1804.

Expert Panel on Detection, E. (2001). Executive summary of the third report of the national cholesterol education program (ncep) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (adult treatment panel iii). *Jama*, *285*(19), 2486-2497.

Falkner, B., & Cossrow, N. D. (2014). Prevalence of metabolic syndrome and obesity-associated hypertension in the racial ethnic minorities of the united states. *Current Hypertension Reports*, *16*(7), 449.

Faria, G., Pestana, D., Aral, M., Preto, J., Norberto, S., Calhau, C., ... Taveira-Gomes, A. (2014). Metabolic score insights on the development and prediction of remission of metabolic syndrome after gastric bypass [Journal Article]. *Annals of Surgery*, *260*(2), 279-286.

Fawcett, T. (2006). An introduction to roc analysis. *Pattern Recognition Letters*, *27*(8), 861–874.

Ferrannini, E., & Cushman, W. C. (2012). Diabetes and hypertension: the bad companions. *The Lancet*, *380*(9841), 601–610.

Ferreira, I., Boreham, C. A., Twisk, J. W., Gallagher, A. M., Young, I. S., Murray, L. J., & Stehouwer, C. D. (2007). Clustering of metabolic syndrome risk factors and arterial stiffness in young adults: the northern ireland young hearts project. *Journal of Hypertension*, *25*(5), 1009-1020.

Ford, E. S., & Li, C. (2008). Defining the metabolic syndrome in children and adolescents: will the real definition please stand up? *The Journal of Pediatrics*, *152*(2), 160-164.

e13.

Franks, P. W., Ekelund, U., Brage, S., Wong, M.-Y., & Wareham, N. J. (2004). Does the association of habitual physical activity with the metabolic syndrome differ by level of cardiorespiratory fitness? *Diabetes Care*, *27*(5), 1187–1193.

Franz, R. (2006). Representations for genetic and evolutionary algorithms. In *Representations for genetic and evolutionary algorithms* (pp. 6–32). Springer, USA.

French, R. M. (1992). Semi-distributed representations and catastrophic forgetting in connectionist networks. *Connection Science*, *4*(3-4), 365–377.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 1189–1232.

Gaio, V., Nunes, B., Fernandes, A., Mendonca, F., Horta Correia, F., Beleza, A., ... Barreto da Silva, M. (2014). Genetic variation at the cyp2c19 gene associated with metabolic syndrome susceptibility in a south portuguese population: results from the pilot study of the european health examination survey in portugal. *Diabetology & Metabolic Syndrome*, *6*(1), 23.

Galassi, A., Reynolds, K., & He, J. (2006). Metabolic syndrome and risk of cardiovascular disease: a meta-analysis. *The American Journal of Medicine*, *119*(10), 812-819.

Gami, A. S., Witt, B. J., Howard, D. E., Erwin, P. J., Gami, L. A., Somers, V. K., & Montori, V. M. (2007). Metabolic syndrome and risk of incident cardiovascular events and death: a systematic review and meta-analysis of longitudinal studies. *Journal of the American College of Cardiology*, *49*(4), 403–414.

Gandek, B., Ware, J. E., Aaronson, N. K., Apolone, G., Bjorner, J. B., Brazier, J. E., ... others (1998). Cross-validation of item selection and scoring for the sf-12 health survey in nine countries: results from the iqola project. *Journal of Clinical Epidemiology*, *51*(11), 1171–1178.

Ginsberg, H. N., & MacCallum, P. R. (2009). The obesity, metabolic syndrome, and type 2 diabetes mellitus pandemic: Ii. therapeutic management of atherogenic dyslipidemia. *The Journal of Clinical Hypertension*, *11*(9), 520–527.

Goldberg, D. E., & Lingle, R., Jr. (1985). Allelesloci and the traveling salesman problem. In *Proceedings of the 1st international conference on genetic algorithms* (pp. 154–159). Hillsdale, NJ, USA: L. Erlbaum Associates Inc.

Gomez-Marcos, M. A., Patino-Alonso, M. C., Recio-Rodriguez, J. I., Anton-Alvarez, J., Cabrejas-Sanchez, A., Fernandez-Alonso, C., ... Garcia-Ortiz, L. (2013). Confirmatory factor analysis to assess the measure of adiposity that best fits the diagnosis of metabolic syndrome and relationship to physical activity in adults. *European Journal of Nutrition*, *52*(5), 1451-1459.

Granger, E., Henniges, P., Oliveira, L. S., & Sabourin, R. (2006). Particle swarm optimization of fuzzy artmap parameters. In *The 2006 ieee international joint conference on neural network proceedings* (p. 2060-2067). IEEE.

Graziano, F., Grassi, M., Sacco, S., Concas, M. P., Vaccargiu, S., Pirastu, M., & Biino, G. (2015). Probing the factor structure of metabolic syndrome in sardinian genetic isolates [Journal Article]. *Nutrition, Metabolism and Cardiovascular Diseases*, *25*(6), 548-555.

Grefenstette, J. J. (1986). Optimization of control parameters for genetic algorithms. *IEEE Transactions on Systems, Man, and Cybernetics*, *16*(1), 122-128.

Grossberg, S. (1976a). Adaptive pattern classification and universal recoding: Ii. feedback, expectation, olfaction, illusions. *Biological cybernetics*, *23*(4), 187-202.

Grossberg, S. (1976b). Adaptive pattern classification and universal recoding: I. parallel development and coding of neural feature detectors. *Biological Cybernetics*, *23*(3), 121-134.

Grossberg, S. (1982). How does a brain build a cognitive code? In *Studies of mind and brain* (p. 1-52). Springer.

Grossberg, S. (1987a). *The adaptive brain i: Cognition, learning, reinforcement, and rhythm*. Elsevier.

Grossberg, S. (1987b). *The adaptive brain ii: Vision, speech, language, and motor controlam-sterdam.* Elsevier/North-Holland.

Grossberg, S. (1987c). Competitive learning: From interactive activation to adaptive

resonance. *Cognitive Science*, *11*(1), 23-63.

Gurka, M. J., Golden, S. H., Musani, S. K., Sims, M., Vishnu, A., Guo, Y., . . . DeBoer, M. D. (2017). Independent associations between a metabolic syndrome severity score and future diabetes by sex and race: the atherosclerosis risk in communities study and jackson heart study. *Diabetologia*, *60*(7), 1261-1270.

Gurka, M. J., Lilly, C. L., Oliver, M. N., & DeBoer, M. D. (2014). An examination of sex and racial/ethnic differences in the metabolic syndrome among adults: A confirmatory factor analysis and a resulting continuous severity score. *Metabolism*, *63*(2), 218-225.

Hadi, A. A., Naing, N. N., Daud, A., & Nordin, R. (2006). Reliability and construct validity of the malay version of the job content questionnaire (jcq) among secondary school teachers in kota bharu, kelantan, malaysia. *Southeast Asian Journal of Tropical Medicine and Public Health*, *37*(6), 1254.

Haffner, S. M., Valdez, R. A., Hazuda, H. P., Mitchell, B. D., Morales, P. A., & Stern, M. P. (1992). Prospective analysis of the insulin-resistance syndrome (syndrome x). *Diabetes*, *41*(6), 715–722.

Hastie, T., Tibshirani, R., Friedman, J., & Franklin, J. (2005). The elements of statistical learning: data mining, inference and prediction. *The Mathematical Intelligencer*, *27*(2), 83-85.

Haupt, R. L., & Haupt, S. E. (2004). Practical genetic algorithms. hoboken, new jersey: A john wiley & sons. *Inc., Publication*.

He, H., & Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, *21*(9), 1263–1284.

Heagerty, P. J., Lumley, T., & Pepe, M. S. (2000). Time-dependent roc curves for censored survival data and a diagnostic marker. *Biometrics*, *56*(2), 337–344.

Hert, M., Schreurs, V., Vancampfort, D., & Winkel, R. (2009). Metabolic syndrome in people with schizophrenia: a review. *World Psychiatry*, *8*(1), 15-22.

Heshmat, R., Shafiee, G., Kelishadi, R., Babaki, A. E. S., Motlagh, M. E., Arefirad, T., . . . Qorbani, M. (2015). Is the association of continuous metabolic syndrome risk

score with body mass index independent of physical activity? the caspian-iii study. *Nutrition Research and Practice*, *9*(4), 404-410.

Hillier, T. A., Rousseau, A., Lange, C., Lepinay, P., Cailleau, M., Novak, M., . . . Balkau, B. (2006). Practical way to assess metabolic syndrome using a continuous score obtained from principal components analysis. *Diabetologia*, *49*(7), 1528-1535.

Hirose, H., Takayama, T., Hozawa, S., Hibi, T., & Saito, I. (2011). Prediction of metabolic syndrome using artificial neural network system based on clinical data including insulin resistance index and serum adiponectin. *Computers in Biology and Medicine*, *41*(11), 1051-1056.

Holland, J. H. (1992). *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. MIT press.

Hosseini, M., Sarrafzadegan, N., Kelishadi, R., Monajemi, M., Asgary, S., & Vardanjani, H. M. (2014). Population-based metabolic syndrome risk score and its determinants: The isfahan healthy heart program. *Journal of Research in Medical Sciences*, *19*(12), 1167-74.

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55.

Huang, Y. C. (2013). The application of data mining to explore association rules between metabolic syndrome and lifestyles. *Health Information Management Journal*, *42*(3), 29–36.

Hunt, K. J., Resendez, R. G., Williams, K., Haffner, S. M., & Stern, M. P. (2004). National cholesterol education program versus world health organization metabolic syndrome in relation to all-cause and cardiovascular mortality in the san antonio heart study. *Circulation*, *110*(10), 1251–1257.

Huo, D., Wang, W., Li, X., Gao, Q., Wu, L. J., Luo, Y. X., . . . Guo, X. H. (2013). Evaluation of two single-factor models of metabolic syndrome: a confirmatory factor analysis for an adult population in beijing. *Lipids in Health and Disease*, *12*(1), 61.

Hwu, C.-M., Hsiung, C. A., Wu, K.-D., Lee, W.-J., Shih, K.-C., Grove, J., . . . Group, S. S.

(2008). Diagnosis of insulin resistance in hypertensive patients by the metabolic syndrome: Aha vs. idf definitions. *International Journal of Clinical Practice*, *62*(9), 1441–1446.

Isomaa, B., Almgren, P., Tuomi, T., Forsen, B., Lahti, K., Nissen, M., . . . Groop, L. (2001). Cardiovascular morbidity and mortality associated with the metabolic syndrome. *Diabetes Care*, *24*(4), 683-9.

Ivanovic, D., Kupusinac, A., Stokic, E., Doroslovacki, R., & Ivetic, D. (2016). Ann prediction of metabolic syndrome: a complex puzzle that will be completed. *Journal of Medical Systems*, *40*(12), 264.

Janghorbani, M., & Amini, M. (2016). Utility of continuous metabolic syndrome score in assessing risk of type 2 diabetes: The isfahan diabetes prevention study. *Annals of Nutrition and Metabolism*, *68*(1), 19-25.

Japkowicz, N., & Stephen, S. (2002). The class imbalance problem: A systematic study. *Intelligent Data Analysis*, *6*(5), 429–449.

Je, Y., Kim, Y., & Park, T. (2017). Development of a self-assessment score for metabolic syndrome risk in non-obese korean adults [Journal Article]. *Asia Pacific Journal of Clinical Nutrition*, *26*(2), 220-226.

Jeong, S., Jo, Y. M., Shim, S.-O., Choi, Y.-J., & Youn, C.-H. (2014). A novel model for metabolic syndrome risk quantification based on areal similarity degree. *IEEE Transactions on Biomedical Engineering*, *61*(3), 665–679.

Jia, X. X., Chen, Q. C., Wu, P. P., Liu, M., Chen, X. X., Xiao, J., . . . Wang, S. M. (2018). Dynamic development of metabolic syndrome and its risk prediction in chinese population: a longitudinal study using markov model [Journal Article]. *Diabetology & Metabolic Syndrome*, *10*(1), 24.

Jiang, B., Li, B., Wang, Y., Han, B., Wang, N., Li, Q., . . . Lu, Y. (2016). The nine-year changes of the incidence and characteristics of metabolic syndrome in china: longitudinal comparisons of the two cross-sectional surveys in a newly formed urban community. *Cardiovascular Diabetology*, *15*(1), 84.

Johnson, J. L., Slentz, C. A., Houmard, J. A., Samsa, G. P., Duscha, B. D., Aiken, L. B., . . . Kraus, W. E. (2007). Exercise training amount and intensity effects on metabolic syndrome (from studies of a targeted risk reduction intervention through

defined exercise). *The American Journal of Cardiology*, *100*(12), 1759 - 1766. doi: http://dx.doi.org/10.1016/j.amjcard.2007.07.027

Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A*, *374*(2065), 20150202.

Jung, K. J., Jee, Y. H., & Jee, S. H. (2017). Metabolic risk score and vascular mortality among korean adults: The korean metabolic syndrome mortality study [Journal Article]. *Asia-Pacific Journal of Public Health*, *29*(2), 122-131.

Jung, U. J., & Choi, M.-S. (2014). Obesity and its metabolic complications: the role of adipokines and the relationship between obesity, inflammation, insulin resistance, dyslipidemia and nonalcoholic fatty liver disease. *International Lournal of Molecular Sciences*, *15*(4), 6184–6223.

Kahn, R. (2007). Metabolic syndrome is it a syndrome? does it matter? *Circulation*, *115*(13), 1806-1811.

Kahn, R., Buse, J., Ferrannini, E., & Stern, M. (2005). The metabolic syndrome: time for a critical appraisal. *Diabetologia*, *48*(9), 1684–1699.

Kakudi, H. A., Loo, C. K., & Moy, F. M. (2017). Predicting metabolic syndrome using risk quantification and ensemble methods. In *Computational intelligence (ssci), 2017 ieee symposium series on* (pp. 1–8).

Kakudi, H. A., Loo, C. K., & Pasupa, K. (2017). Risk quantification of metabolic syndrome with quantum particle swarm optimisation. In *Proceedings of the 26th international conference on world wide web companion* (pp. 1141–1147).

Kang, G. D., Guo, L., Guo, Z. R., Hu, X. S., Wu, M., & Yang, H. T. (2012). Continuous metabolic syndrome risk score for predicting cardiovascular disease in the chinese population. *Asia Pacific Journal of Clinical Nutrition*, *21*(1), 88-96.

Karimi-Alavijeh, F., Jalili, S., & Sadeghi, M. (2016). Predicting metabolic syndrome using decision tree and support vector machine methods. *ARYA Atherosclerosis*, *12*(3), 146-152.

Kass, G. V. (1980). An exploratory technique for investigating large quantities of categorical

data. *Applied Statistics*, 119–127.

Kaur, J. (2014). A comprehensive review on metabolic syndrome. *Cardiology Research and Practice*, *2014*.

Klein, B. E., Klein, R., & Lee, K. E. (2002). Components of the metabolic syndrome and risk of cardiovascular disease and diabetes in beaver dam. *Diabetes Care*, *25*(10), 1790–1794.

Koufakou, A., Georgiopoulos, M., Anagnostopoulos, G., & Kasparis, T. (2001). Cross-validation in fuzzy artmap for large databases. *Neural Networks*, *14*(9), 1279-1291.

Kubat, M., Holte, R. C., & Matwin, S. (1998). Machine learning for the detection of oil spills in satellite radar images. *Machine Learning*, *30*(2-3), 195–215.

Laaksonen, D. E., Lakka, H.-M., Niskanen, L. K., Kaplan, G. A., Salonen, J. T., & Lakka, T. A. (2002). Metabolic syndrome and development of diabetes mellitus: application and validation of recently suggested definitions of the metabolic syndrome in a prospective cohort study. *American Journal of Epidemiology*, *156*(11), 1070–1077.

Lakka, H. M., Laaksonen, D. E., Lakka, T. A., Niskanen, L. K., Kumpusalo, E., Tuomilehto, J., & Salonen, J. T. (2002). The metabolic syndrome and total and cardiovascular disease mortality in middle-aged men. *Jama*, *288*(21), 2709-2716.

Landsberg, L. (2001). Insulin-mediated sympathetic stimulation: role in the pathogenesis of obesity-related hypertension (or, how insulin affects blood pressure, and why). *Journal of Hypertension*, *19*(3), 523–528.

Lao, X. Q., Zhang, Y. H., Wong, M. C. S., Xu, Y. J., Xu, H. F., Nie, S. P., . . . Yu, I. T. S. (2012). The prevalence of metabolic syndrome and cardiovascular risk factors in adults in southern china. *BMC Public Health*, *12*(1), 64.

Liese, A. D., Mayer-Davis, E. J., Tyroler, H. A., Davis, C., Keil, U., Duncan, B. B., & Heiss, G. (1997). Development of the multiple metabolic syndrome in the aric cohort: joint contribution of insulin, bmi, and whr. *Annals of Epidemiology*, *7*(6), 407–416.

Liew, S. M., Doust, J., & Glasziou, P. (2011). Cardiovascular risk scores do not account

for the effect of treatment: a review. *Heart*, *97*(9), 689–697.

Liew, W. S., Seera, M., & Loo, C. K. (2016). Hierarchical parallel genetic optimization fuzzy artmap ensemble. *Neural Processing Letters*, *44*(2), 451–470.

Lin, C. C., Bai, Y.-M., Chen, J.-Y., Hwang, T.-J., Chen, T.-T., Chiu, H.-W., . . . others (2010). Easy and low-cost identification of metabolic syndrome in patients treated with second-generation antipsychotics: artificial neural network and logistic regression models. *Journal of Clinical Psychiatry*, *71*(3), 225.

Lin, Z., Gonçalves, C. M. V., Dai, L., Lu, H.-m., Huang, J.-h., Ji, H., . . . Liang, Y.-z. (2014). Exploring metabolic syndrome serum profiling based on gas chromatography mass spectrometry and random forest models. *Analytica Chimica Acta*, *827*, 22–27.

Loh, W.-Y. (2014). Fifty years of classification and regression trees. *International Statistical Review*, *82*(3), 329–348.

Loo, C. K., Liew, W. S., Seera, M., & Lim, E. (2015). Probabilistic ensemble fuzzy artmap optimization using hierarchical parallel genetic algorithms. *Neural Computing and Applications*, *26*(2), 263–276.

Lopez-Candales, A. (2001). Metabolic syndrome x: a comprehensive review of the pathophysiology and recommended therapy. *Journal of Medicine*, *32*(5-6), 283-300.

Magnussen, C. G., Cheriyan, S., Sabin, M. A., Juonala, M., Koskinen, J., Thomson, R., . . . Raitakari, O. T. (2016). Continuous and dichotomous metabolic syndrome definitions in youth predict adult type 2 diabetes and carotid artery intima media thickness: The cardiovascular risk in young finns study. *Journal of Pediatrics*, *171*, 97-+.

Maison, P., Byrne, C. D., Hales, C. N., Day, N. E., & Wareham, N. J. (2001). Do different dimensions of the metabolic syndrome change together over time? *Diabetes Care*, *24*(10), 1758–1763.

Malhotra, R. (2015). A systematic review of machine learning techniques for software fault prediction. *Applied Soft Computing*, *27*, 504 - 518.

Malik, S., Wong, N. D., Franklin, S. S., Kamath, T. V., Gilbert, J., Pio, J. R., & Williams,

G. R. (2004). Impact of the metabolic syndrome on mortality from coronary heart disease, cardiovascular disease, and all causes in united states adults. *Circulation*, *110*(10), 1245–1250.

Marriott, S., & Harrison, R. F. (1995). A modified fuzzy artmap architecture for the approximation of noisy mappings. *Neural Networks*, *8*(4), 619-641.

Masuyama, N., Loo, C. K., & Dawood, F. (2018). Kernel bayesian art and artmap. *Neural Networks*, *98*, 76 - 86. doi: https://doi.org/10.1016/j.neunet.2017.11.003

McCloskey, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. In *Psychology of learning and motivation* (Vol. 24, pp. 109–165). Elsevier.

McNeill, A. M., Rosamond, W. D., Girman, C. J., Golden, S. H., Schmidt, M. I., East, H. E., . . . Heiss, G. (2005). The metabolic syndrome and 11-year risk of incident cardiovascular disease in the atherosclerosis risk in communities study. *Diabetes Care*, *28*(2), 385–390.

Meigs, J. B., Wilson, P. W., Nathan, D. M., Da'Agostino, R. B., Williams, K., & Haffner, S. M. (2003). Prevalence and characteristics of the metabolic syndrome in the san antonio heart and framingham offspring studies. *Diabetes*, *52*(8), 2160–2167.

Melillo, P., De Luca, N., Bracale, M., & Pecchia, L. (2013). Classification tree for risk assessment in patients suffering from congestive heart failure via long-term heart rate variability. *IEEE Journal of Biomedical and Health Informatics*, *17*(3), 727–733.

Meng, C., Zeleznik, O. A., Thallinger, G. G., Kuster, B., Gholami, A. M., & Culhane, A. C. (2016). Dimension reduction techniques for the integrative analysis of multi-omics data. *Briefings in Bioinformatics*, *17*(4), 628–641.

Miller, B., Fridline, M., Liu, P. Y., & Marino, D. (2014). Use of chaid decision trees to formulate pathways for the early detection of metabolic syndrome in young adults. *Computational and Mathematical Methods in Medicine*.

Mitchell, M. (1998). *An introduction to genetic algorithms*. MIT press.

Miyauchi, Y., & Nishimura, H. (2018). Bayesian network modeling for specific health

checkups on metabolic syndrome. In D. E. Holmes & L. C. Jain (Eds.), *Advances in biomedical informatics.* Springer International Publishing.

Mochizuki, K., Miyauchi, R., Misaki, Y., Ichikawa, Y., & Goda, T. (2013). Principal component 1 score calculated from metabolic syndrome diagnostic parameters is a possible marker for the development of metabolic syndrome in middle-aged japanese men without treatment for metabolic diseases. *European Journal of Nutrition*, *52*(1), 67-74.

Mohamud, W. N. W., Ismail, A., Khir, A. S., Ismail, I. S., Mu, K. I., Kadir, K. A., ... Wan, M. W. B. (2012). Prevalence of metabolic syndrome and its risk factors in adult malaysians: results of a nationwide survey. *Diabetes Research and Clinical Practice*, *96*(1), 91–97.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the prisma statement. *PLoS Medicine*, *6*(7), e1000097.

Montague, C. T., & O'rahilly, S. (2000). The perils of portliness: causes and consequences of visceral adiposity. *Diabetes*, *49*(6), 883–888.

Morris, A. (2018). *Angptl4—the link binding obesity and glucose intolerance.* Nature Publishing Group 75 VARICK ST, 9TH FLR, NEW YORK, NY 10013-1917 USA.

Morse, S. A., Zhang, R., Thakur, V., & Reisin, E. (2005). Hypertension and the metabolic syndrome. *The American Journal of the Medical Sciences*, *330*(6), 303–310.

Mottillo, S., Filion, K. B., Genest, J., Joseph, L., Pilote, L., Poirier, P., ... Eisenberg, M. J. (2010). The metabolic syndrome and cardiovascular risk: a systematic review and meta-analysis. *Journal of the American College of Cardiology*, *56*(14), 1113–1132.

Moy, F. M., Hoe, V. C. W., Hairi, N. N., Buckley, B., Wark, P. A., Koh, D., ... Bulgiba, A. M. (2014). Cohort study on clustering of lifestyle risk factors and understanding its association with stress on health and wellbeing among school teachers in malaysia (cluster): a study protocol. *BMC Public Health*, *14*(1), 611.

Murguia-Romero, M., Jimenez-Flores, R., Mendez-Cruz, A. R., & Villalobos-Molina, R. (2013). Predicting metabolic syndrome with neural networks. In *Mexican international conference on artificial intelligence* (p. 464-472). Springer.

Murthy, S. K. (1998). Automatic construction of decision trees from data: A multi-disciplinary survey. *Data Mining and Knowledge Discovery*, *2*(4), 345–389.

Musani, S. K., Martin, L. J., Woo, J. G., Olivier, M., Gurka, M. J., & DeBoer, M. D. (2017). Heritability of the severity of the metabolic syndrome in whites and blacks in 3 large cohorts. *Circulation: Cardiovascular Genetics*, *10*(2), e001621.

Negnevitsky, M. (2005). *Artificial intelligence: a guide to intelligent systems*. Pearson Education.

Neto, A. S., de Campos, W., Dos Santos, G. C., & Junior, O. M. (2014). Metabolic syndrome risk score and time expended in moderate to vigorous physical activity in adolescents. *BMC Pediatrics*, *14*(1), 42.

Obokata, M., Negishi, K., Ohyama, Y., Okada, H., Imai, K., & Kurabayashi, M. (2015). A risk score with additional four independent factors to predict the incidence and recovery from metabolic syndrome: Development and validation in large japanese cohorts. *PloS One*, *10*(7), e0133884.

Oda, E., Abe, M., Veeraveedu, P. T., & Watanabe, K. (2007). Considerable disagreement among definitions of metabolic syndrome for japanese. *Circulation Journal*, *71*(8), 1239–1243.

Okosun, I. S., Lyn, R., Davis-Smith, M., Eriksen, M., & Seale, P. (2010). Validity of a continuous metabolic risk score as an index for modeling metabolic syndrome in adolescents. *Annals of Epidemiology*, *20*(11), 843-851.

Olza, J., Aguilera, C. M., Gil-Campos, M., Leis, R., Bueno, G., Valle, M., ... Gil, A. (2015). A continuous metabolic syndrome score is associated with specific biomarkers of inflammation and cvd risk in prepubertal children. *Annals of Nutrition and Metabolism*, *66*(2-3), 72-79.

Palaniappan, L., Carnethon, M. R., Wang, Y., Hanley, A. J., Fortmann, S. P., Haffner, S. M., & Wagenknecht, L. (2004). Predictors of the incident metabolic syndrome in adults: the insulin resistance atherosclerosis study. *Diabetes Care*, *27*(3), 788–793.

Palaniappan, R., & Eswaran, C. (2009). Using genetic algorithm to select the presentation order of training patterns that improves simplified fuzzy artmap classification performance. *Applied Soft Computing*, *9*(1), 100-106.

Palomo, I., Alarcon, M., Moore-Carrasco, R., & Argiles, J. M. (2006). Hemostasis alterations in metabolic syndrome (review). *International Journal of Molecular Medicine*, *18*(5), 969-974.

Park, H. S., & Cho, S. B. (2012). Evolutionary attribute ordering in bayesian networks for predicting the metabolic syndrome. *Expert Systems with Applications*, *39*(4), 4240-4249.

Paul, T. K., Ueno, K., Iwata, K., Hayashi, T., & Honda, N. (2008). Genetic algorithm based methods for identification of health risk factors aimed at preventing metabolic syndrome. *Simulated Evolution and Learning, Proceedings*, *5361*, 210-+.

Pencheva, T., Atanassov, K., & Shannon, A. (2009). Modelling of a roulette wheel selection operator in genetic algorithms using generalized nets. *International Journal of Bioautomation*, *13*(4), 257–264.

Pepe, M. S. (2003). *The statistical evaluation of medical tests for classification and prediction*. Oxford University Press, USA.

Perveen, S., Shahbaz, M., Keshavjee, K., & Guergachi, A. (2019). Metabolic syndrome and development of diabetes mellitus: Predictive modeling based on machine learning techniques [Journal Article]. *IEEE Access*, *7*, 1365-1375.

Potteiger, J. A., Claytor, R. P., Hulver, M. W., Hughes, M. R., Carper, M. J., Richmond, S., & Thyfault, J. P. (2012). Resistance exercise and aerobic exercise when paired with dietary energy restriction both reduce the clinical components of metabolic syndrome in previously physically inactive males. *European Journal of Applied Physiology*, *112*(6), 2035–2044.

Povel, C. M., Beulens, J. W., van der Schouw, Y. T., Dolle, M. E. T., Spijkerman, A. M. W., Verschuren, W. M. M., ... Boer, J. M. A. (2012). Metabolic syndrome model definitions predicting type 2 diabetes and cardiovascular disease. *Diabetes Care*, DC112546.

Prasad, D., Kabir, Z., Dash, A., & Das, B. (2012). Prevalence and risk factors for metabolic syndrome in asian indians: A community study from urban eastern india. *Journal of Cardiovascular Disease Research*, *3*(3), 204.

Ragland, D. R. (1992). Dichotomizing continuous outcome variables: dependence of the magnitude of association and statistical power on the cutpoint. *Epidemiology*, *3*(5),

434-440.

Rampal, S., Mahadeva, S., Guallar, E., Bulgiba, A., Mohamed, R., Rahmat, R., . . . Rampal, L. (2012). Ethnic differences in the prevalence of metabolic syndrome: results from a multi-ethnic population-based survey in malaysia. *PloS One*, *7*(9), e46365.

Rao, R. B., Krishnan, S., & Niculescu, R. S. (2006). Data mining for cmproved cardiac care. *ACM SIGKDD Explorations Newsletter*, *8*(1), 3–10.

Ratcliff, R. (1990). Connectionist models of recognition memory: constraints imposed by learning and forgetting functions. *Psychological Review*, *97*(2), 285.

Reaven, G. M. (1988). Role of insulin resistance in human disease. *Diabetes*, *37*(12), 1595-607.

Richardson, F. M., & Thomas, M. S. (2008). Critical periods and catastrophic interference effects in the development of self-organizing feature maps. *Developmental Science*, *11*(3), 371–389.

Romero-Saldana, M., Fuentes-Jimenez, F. J., Vaquero-Abellan, M., Alvarez-Fernandez, C., Molina-Recio, G., & Lopez-Miranda, J. (2016). New non-invasive method for early detection of metabolic syndrome in the working population. *European Journal of Cardiovascular Nursing*, *15*(7), 549-558.

Rosen, C. A., Lee, A. S., Osborne, J., Zullo, T., & Murry, T. (2004). Development and validation of the voice handicap index-10. *The Laryngoscope*, *114*(9), 1549–1556.

Saeed, A., Ab Hamid, S. H., & Mustafa, M. B. (2016). The experimental applications of search-based techniques for model-based testing: Taxonomy and systematic literature review. *Applied Soft Computing*, *49*, 1094–1117.

Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the roc plot when evaluating binary classifiers on imbalanced datasets. *PloS One*, *10*(3), e0118432.

Sakawa, M. (2012). *Genetic algorithms and fuzzy multiobjective optimization*. Springer Science & Business Media.

Sattar, N., Gaw, A., Scherbakova, O., Ford, I., O'Reilly, D. S. J., Haffner, S. M., ... others (2003). Metabolic syndrome with and without c-reactive protein as a predictor of coronary heart disease and diabetes in the west of scotland coronary prevention study. *Circulation*, *108*(4), 414–419.

Seidell, J. C., Cigolini, M., Charzewska, J., Ellsinger, B., Deslypere, J., & Cruz, A. (1992). Fat distribution in european men: A comparison of anthropometric measurements in relation to cardiovascular risk factors. *International Journal of Obesity and Related Metabolic Disorders: Journal of the International Association for the Study of Obesity*, *16*(1), 17–22.

Sheehan, J. P. (2004). Fasting hyperglycemia: etiology, diagnosis, and treatment. *Diabetes Technology & therapeutics*, *6*(4), 525–533.

Sheehan, M. T., & Jensen, M. D. (2000). Metabolic complications of obesity: Pathophysiologic considerations. *Medical Clinics*, *84*(2), 363–385.

Shimoda, A., Ichikawa, D., & Oyama, H. (2018). Prediction models to identify individuals at risk of metabolic syndrome who are unlikely to participate in a health intervention program. *International Journal of Medical Informatics*, *111*, 90–99.

Simmons, R., Alberti, K., Gale, E., Colagiuri, S., Tuomilehto, J., Qiao, Q., ... others (2010). The metabolic syndrome: useful concept or clinical tool? report of a who expert consultation. *Diabetologia*, *53*(4), 600–605.

Sivanandam, S., & Deepa, S. N. (2008). *Introduction to genetic algorithms*. Springer-Verlag Berlin Heidelberg.

Smits, M. M., Woudstra, P., Utzschneider, K. M., Tong, J., Gerchman, F., Faulenbach, M., ... Kahn, S. E. (2013). Adipocytokines as features of the metabolic syndrome determined using confirmatory factor analysis. *Annals of Epidemiology*, *23*(7), 415-421.

Soldatovic, I., Vukovic, R., Culafic, D., Gajic, M., & Dimitrijevic-Sreckovic, V. (2016). sims score: Simple method for quantifying metabolic syndrome. *Plos One*, *11*(1), 10.

Steinberg, G. B., Church, B. W., McCall, C. J., Scott, A. B., & Kalis, B. P. (2014). Novel predictive models for metabolic syndrome risk: A "big data" analytic approach. *The American Journal of Managed Care*, *20*(6), e221-8.

Strobl, C., Boulesteix, A.-L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics*, *8*(1), 25.

Sullivan, L. M., Massaro, J. M., & D'Agostino Sr, R. B. (2004). Presentation of multivariate data for clinical use: The framingham study risk score functions. *Statistics in medicine*, *23*(10), 1631–1660.

Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science*, *240*(4857), 1285.

Tan, C., Sasagawa, Y., Kamo, K. I., Kukitsu, T., Noda, S., Ishikawa, K., ... Kokai, Y. (2016). Evaluation of the japanese metabolic syndrome risk score (jamrisc): a newly developed questionnaire used as a screening tool for diagnosing metabolic syndrome and insulin resistance in japan. *Environmental Health and Preventive Medicine*, *21*(6), 470.

Thivel, D., Malina, R. M., Isacco, L., Aucouturier, J., Meyer, M., & Duche, P. (2009). Metabolic syndrome in obese children and adolescents: Dichotomous or continuous? *Metabolic Syndrome and Related Disorders*, *7*(6), 549-555.

Timar, O., Sestier, F., & Levy, E. (2000). Metabolic syndrome x: a review. *The Canadian Journal of Cardiology*, *16*(6), 779-789.

Tomek, I. (1976). Two modifications of cnn. *IEEE Transactions on Systems, Man and Cybernetics*, *6*, 769–772.

Tsou, M. T., Chang, B. C. C., Huang, W. H., & Hsu, C. P. (2014). Prevalence of metabolic syndrome and risk factor analysis among urban elderly in one medical center in northern taiwan. *International Journal of Gerontology*, *8*(3), 127-132.

Tu, J. V. (1996). Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *Journal of Clinical Epidemiology*, *49*(11), 1225–1231.

Turpin, S. M., Nicholls, H. T., Willmes, D. M., Mourier, A., Brodesser, S., Wunderlich, C. M., ... others (2014). Obesity-induced cers6-dependent c 16: 0 ceramide production promotes weight gain and glucose intolerance. *Cell Metabolism*, *20*(4), 678–686.

Ushida, Y., Kato, R., Niwa, K., Tanimura, D., Izawa, H., Yasui, K., . . . Honda, H. (2012). Combinational risk factors of metabolic syndrome identified by fuzzy neural network analysis of health-check data. *BMC Medical Informatics and Decision Making*, *12*(1), 80.

Van Schependom, J., Yu, W., Gielen, J., Laton, J., De Keyser, J., De Hert, M., & Nagels, G. (2015). Do advanced statistical techniques really help in the diagnosis of the metabolic syndrome in patients treated with second-generation antipsychotics? *Journal of Clinical Psychiatry*, *76*(10), e1292-9.

Vapnik, V. (2013). *The nature of statistical learning theory*. Springer science & business media.

Veganzones, D., & Séverin, E. (2018). An investigation of bankruptcy prediction in imbalanced datasets. *Decision Support Systems*, *112*, 111 - 124. doi: https://doi.org/10.1016/j.dss.2018.06.011

Vigdor, B., & Lerner, B. (2007). The bayesian artmap. *IEEE Transactions on Neural Networks*, *18*(6), 1628-1644.

Vikram, N. K., Pandey, R. M., Misra, A., Goel, K., & Gupta, N. (2009). Factor analysis of the metabolic syndrome components in urban asian indian adolescents. *Asia Pacific Journal of Clinical Nutrition*, *18*(2), 293-300.

Vishnu, A., Gurka, M. J., & DeBoer, M. D. (2015). The severity of the metabolic syndrome increases over time within individuals, independent of baseline metabolic syndrome status and medication use: The atherosclerosis risk in communities study. *Atherosclerosis*, *243*(1), 278–285.

Walker, S., Gurka, M., Oliver, M., Johns, D., & DeBoer, M. (2012). Racial/ethnic discrepancies in the metabolic syndrome begin in childhood and persist after adjustment for environmental factors. *Nutrition, Metabolism and Cardiovascular Diseases*, *22*(2), 141–148.

Ware Jr, J. E., Kosinski, M., & Keller, S. D. (1996). A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. *Medical Care*, *34*(3), 220–233.

Whitley, D. (1994). A genetic algorithm tutorial. *Statistics and Computing*, *4*(2), 65–85.

WHO. (1999). *World health organization. definition, diagnosis and classification of diabetes mellitus and its complications: report of a who consultation. part 1: diagnosis and classification of diabetes mellitus. geneva, switzerland* (Tech. Rep.). Geneva: World health organization. Retrieved from http://apps.who.int/iris/bitstream/10665/66040/1/WHO_NCD_NCS_99.2.pdf

WHO. (2018). *Global health observatory (gho) data: Raised blood pressure.* Retrieved from http://www.who.int/gho/ncd/risk_factors/blood_pressure_prevalence_text/en/

Wijndaele, K., Beunen, G., Duvigneaud, N., Matton, L., Duquet, W., Thomis, M., ... Philippaerts, R. M. (2006). A continuous metabolic syndrome risk score - utility for epidemiological analyses. *Diabetes Care*, *29*(10), 2329-2329.

Wijndaele, K., Duvigneaud, N., Matton, L., Duquet, W., Delecluse, C., Thomis, M., ... Philippaerts, R. M. (2009). Sedentary behaviour, physical activity and a continuous metabolic syndrome risk score in adults. *European Journal of Clinical Nutrition*, *63*(3), 421-429.

Wildman, R. P., Gu, D., Reynolds, K., Duan, X., Wu, X., & He, J. (2005). Are waist circumference and body mass index independently associated with cardiovascular disease risk in chinese adults?–. *The American Journal of Clinical Nutrition*, *82*(6), 1195–1202.

Wiley, J. F., & Carrington, M. J. (2016). A metabolic syndrome severity score: A tool to quantify cardio-metabolic risk factors. *Preventive Medicine*, *88*, 189-95.

Williamson, J. R. (1996). Gaussian artmap: A neural network for fast incremental learning of noisy multidimensional maps. *Neural Networks*, *9*(5), 881-897.

Wilson, D. L. (1972). Asymptotic properties of nearest neighbor rules using edited data. *IEEE Transactions on Systems, Man, and Cybernetics*, *2*(3), 408–421.

Wilson, P. W., D'Agostino, R. B., Levy, D., Belanger, A. M., Silbershatz, H., & Kannel, W. B. (1998). Prediction of coronary heart disease using risk factor categories. *Circulation*, *97*(18), 1837–1847.

Wong, Y., Cook, P., Roderick, P., & Somani, B. K. (2016). Metabolic syndrome and kidney stone disease: A systematic review of literature. *Journal of Endourology*, *30*(3), 246-53.

Worachartcheewan, A., Nantasenamat, C., Isarankura-Na-Ayudhya, C., & Prachayasittikul, V. (2013). Quantitative population-health relationship (qphr) for assessing metabolic syndrome. *EXCLI Journal*, *12*, 569-583.

Worachartcheewan, A., Shoombuatong, W., Pidetcha, P., Nopnithipat, W., Prachayasittikul, V., & Nantasenamat, C. (2015). Predicting metabolic syndrome using the random forest method. *The Scientific World Journal*, *2015*, 581501.

Xue, F., & Michels, K. B. (2007). Diabetes, metabolic syndrome, and breast cancer: a review of the current evidence. *The American Journal of Clinical Nutrition*, *86*(3), 823S-835S.

Yousefzadeh, G., Shokoohi, M., Najafipour, H., & Shadkamfarokhi, M. (2015). Applying the framingham risk score for prediction of metabolic syndrome: The kerman coronary artery disease risk study, iran. *ARYA Atherosclerosis*, *11*(3), 179.

Zadeh, L. A. (1965). Fuzzy logic and its applications. *New York, NY, USA*.

Zadeh-Vakili, A., Tehrani, F. R., & Hosseinpanah, F. (2011). Waist circumference and insulin resistance: a community based cross sectional study on reproductive aged iranian women. *Diabetology & Metabolic Syndrome*, *3*(1), 18.

Zhao, X., Xu, K., Shi, H., Cheng, J., Ma, J., Gao, Y., . . . Zhou, L. (2014). Application of the back-error propagation artificial neural network (bpann) on genetic variants in the ppar-gamma and rxr-alpha gene and risk of metabolic syndrome in a chinese han population. *Journal of Biomedical Research*, *28*(2), 114.

Zhou, X.-H., McClish, D. K., & Obuchowski, N. A. (2009). *Statistical methods in diagnostic medicine* (Vol. 569). John Wiley & Sons.

Zimmet, P. Z., Shaw, J. E., & Alberti, K. G. M. (2005). Mainstreaming the metabolic syndrome: a definitive definition. *The Medical Journal of Australia*, *183*(4), 175–176.

Zou, T. T., Zhou, Y. J., Zhou, X. D., Liu, W. Y., Van Poucke, S., Wu, W. J., . . . Pan, X. Y. (2018). Mets risk score: A clear scoring model to predict a 3-year risk for metabolic syndrome [Journal Article]. *Hormone and Metabolic Research*, *50*(9), 683-689.

Zweiri, Y., Whidborne, J., & Seneviratne, L. (2003). A three-term backpropagation algorithm. *Neurocomputing*, *50*, 305 - 318.