HYBRID OF OPTIMIZED RANDOM FOREST AND EXTREME GRADIENT BOOSTING FOR ONLINE LEARNING STYLE CLASSIFICATION

by

HAZIQAH SHAMSUDIN

Thesis submitted in fulfilment of the requirements
for the degree of
Master of Science

March 2019

ACKNOWLEDGEMENT

"In the name of Allah, most Gracious, most Compassionate"

First and foremost, I am most grateful to my supervisors Dr. Umi Kalsom Yusof, School of Computer Science in Universiti Sains Malaysia, for her unconditional support, encauragement, and unequivocal time, energy and opinion that have been given throughout the duration of my study in accomplishing this thesis.

I would like to express my gratitude to my co-supervisor Pn Maziani Sabudin for providing beneficial assistant and support throughout my research journey. I would also like to express my appreciation to most lecturers in School of Computer Science that have never tired of giving valuable knowledge and insight along this period of study. Personally, I would like to thank my senior Mohd Nor Akmal bin Khalid for his guidance throughout this research process, his guidance in helping me with giving a lot of tips. Additionally, I would also like to thank my friends Nur Izzati Abd Kader, Nur Aqilah Paskhal and Nurfarahin Mohd Noor which is working on a Master together and helping out each other. I would also like to express my gratitude to all my course mate and thank them for helping to stay strong while giving me valuable opinion and knowledge throughout these challenging years.

Above all, none of this would have been possible without the love and patience of my family who has been a constant source of love, concern, support and strength for all these years. Lastly, I thank my parents for their unconditional support and for always give me encouragement to stay patience and to continue with this fight.

ACKNOWLEDGEMENT

"In the name of Allah, most Gracious, most Compassionate"

First and foremost, I am most grateful to my supervisors Dr. Umi Kalsom Yusof, School of Computer Science in Universiti Sains Malaysia, for her unconditional support, encauragement, and unequivocal time, energy and opinion that have been given throughout the duration of my study in accomplishing this thesis.

I would like to express my gratitude to my co-supervisor Pn Maziani Sabudin for providing beneficial assistant and support throughout my research journey. I would also like to express my appreciation to most lecturers in School of Computer Science that have never tired of giving valuable knowledge and insight along this period of study. Personally, I would like to thank my senior Mohd Nor Akmal bin Khalid for his guidance throughout this research process, his guidance in helping me with giving a lot of tips. Additionally, I would also like to thank my friends Nur Izzati Abd Kader, Nur Aqilah Paskhal and Nurfarahin Mohd Noor which is working on a Master together and helping out each other. I would also like to express my gratitude to all my course mate and thank them for helping to stay strong while giving me valuable opinion and knowledge throughout these challenging years.

Above all, none of this would have been possible without the love and patience of my family who has been a constant source of love, concern, support and strength for all these years. Lastly, I thank my parents for their unconditional support and for always give me encouragement to stay patience and to continue with this fight.

TABLE OF CONTENTS

ACKNOWLEDGEMENT ii

TABLE OF CONTENTS iii

LIST OF TABLES viii

LIST OF FIGURES x

LIST OF ABBREVIATIONS xii

ABSTRAK xiv

ABSTRACT xvi

CHAPTER 1 – INTRODUCTION

1.1 Preliminaries 1

1.2 Motivation 5

1.3 Problem Statement 6

1.4 Goal and Objectives of the study 8

1.5 Study Scope and Limitations 8

1.6 Outline of Thesis 9

CHAPTER 2 – LITERATURE REVIEWS

2.1 Introduction 12

2.2 Educational Data Mining 12

2.2.1 Traditional Learning 14

2.2.2 Online Learning 14

2.3 Learning Style 18

2.4 Automated Approaches in Learning Style Detection 21

2.4.1 Literature-based 22

2.4.2 Data-driven 23

2.5 Techniques Used in Automated Approaches of Learning Style Detection.......... 25

2.5.1 Feature Selection 25

2.5.1(a) Filter based Approach in Attribute Selection.................... 25

2.5.1(b) Wrapper based Approach in Attribute Selection 26

2.5.1(c) Embedded feature selection in attribute selection.............. 26

2.5.2 Classification... 27

2.5.2(a) Single classification algorithms in learning style detection..... 27

2.5.2(b) Hybrid Algorithms in learning style detection 32

2.5.2(c) Hyperparameter Optimization 33

2.6 Model Evaluation Performance 36

2.7 Trends and Future Directions 39

2.8 Summary 41

CHAPTER 3 – RESEARCH METHODOLOGY

3.1 Introduction 43

3.2 Research Methodology 43

3.3 Datasets.. 47

3.4 Instrumentation and Result Analysis... 47

3.4.1 Hardware and Software Requirement 49

3.4.2 Experimentations and Analysis 49

3.5 Performance Metrices 50

3.5.1 Notations and Parameter 50

3.5.2 Confusion Matrix 51

3.5.3 Sampling Method 53

3.5.4 Constraints 54

3.6 Summary 55

CHAPTER 4 – DETERMINING THE CONTRIBUTING ATTRIBUTES IN LEARNING STYLE PREDICTION USING RANK BY IMPORTANCE AND RANDOM FOREST TECHNIQUES

4.1 Introduction 56

4.2 Feature Selection Method 56

4.2.1 Rank by Importance 58

4.2.1(a) Overall Steps in Rank by Importance Method 59

4.2.2 Correlation based Feature Selection 60

4.2.2(a) Overall Steps in Correlation based Feature Selection Method . 61

4.2.3 Recursive Feature Elimination 62

4.2.3(a) Overall Steps in Recursive Feature Elimination Method 62

4.3 Classification method using Tree-based Algorithms Approach 63

4.3.1 Random Forest 64

4.3.1(a) Initialization 65

4.3.1(b) Build of Tree in Random Forest Model 65

4.3.2 Extreme Gradient Boosting 66

4.3.2(a) Initialization 66

4.3.2(b) Boosting method 67

4.4 Experimental Results and Discussion 67

4.4.1 Experimental Results for Feature Selection 68

4.4.1(a) Result of using Rank by Importance Method 68

4.4.1(b) Result of using Correlation based Method 69

4.4.1(c) Result of using Recursive Feature Elimination Method 69

4.4.2 Experimental Results for classification using Tree-based Algorithms 70

4.4.2(a) Random Forest 70

4.4.2(b) Extreme Gradient Boosting 71

4.4.3 Computational Result of the Feature Selection Methods and Tree-based Algorithms 72

4.5 Summary 77

CHAPTER 5 – INCORPORATING HYPERPARAMETER OPTIMIZATION TO IMPROVE THE PERFORMANCE OF RANDOM FOREST AND EXTREME GRADIENT BOOSTING

5.1 Introduction 79

5.2 Hyperparameter Optimization 79

5.3 Incorporating Hyperparameter Optimization in the Tree-based Algorithms 80

5.3.1 Random Forest with Hyperparameter Optimization 83

5.3.2 Extreme Gradient Boosting with Hyperparameter Optimization 85

5.4 Computational Result of Tree-based Algorithms with hyerparameter optimization 87

5.4.1 Random Forest 87

5.4.2 Extreme Gradient Boosting 93

5.4.3 Overall comparison Hyperparameter Optimization in Tree-based Algorithms 97

5.5 Summary 102

CHAPTER 6 – PROPOSED HYBRID OF RANDOM FOREST AND EXTREME GRADIENT BOOSTING TO FURTHER IMPROVE THE PERFORMANCE OF TREE-BASED ALGORITHMS

6.1 Introduction 104

6.2 The Hybrid of optimized Random Forest and Extreme Gradient Boosting RF(XGB) 104

6.2.1 Incorporating the Extreme Gradient Boosting function in the Random Forest 104

6.2.2 The Overall Flow of The Proposed Hybrid of Random Forest and Extreme Gradient Boosting 105

6.3 Comparative Analysis on the Performance of Algorithms 109

6.3.1 Accuracy Value 109

6.3.2 Receiver Operating Curve and Area Under Curve Value 111

6.3.3 Overall Comparison of result 113

6.4 Summary 114

CHAPTER 7 – CONCLUSION AND OUTLOOK

7.1 Introduction 116

7.2 Research Contributions 118

7.3 Future Works 118

REFERENCES 120

APPENDICES

LIST OF TABLES

		Page
Table 2.1	Table of comparison between data-driven and literature-based approach	22
Table 2.2	A summary of the techniques adopted by previous researchers in the learning style detection	34
Table 3.1	Summary of dataset	47
Table 3.2	Attributes match to dimension of Felder Silverman Learning Style Model (FSLSM)	48
Table 4.1	List of attributes with importance value for each of the dimension using Rank by Importance method	68
Table 4.2	Attributes selected for each of the dimension	70
Table 4.3	Accuracy value of feature selection against RF algorithm	71
Table 4.4	Overall value of accuracy feature selection against classifier	72
Table 4.5	List of selected attributes using different feature selection method	73
Table 4.6	Overall value of accuracy feature selection against classifier	74
Table 4.7	Average value of accuracy feature selection against classifier	77
Table 4.8	Comparison of average accuracy results with literature	77
Table 5.1	Parameter setup for RF	84
Table 5.2	Parameter setup for Xgb	87
Table 5.3	Optimal *mtry* value for input dimension	88
Table 5.4	Optimal *mtry* value for processing dimension	89
Table 5.5	Optimal *mtry* value for perception dimension	90
Table 5.6	Optimal *mtry* value for understanding dimension	91
Table 5.7	Optimal *mtry* value for understanding dimension	92
Table 5.8	Overall result of comparison for RF algorithm with and without hyperparameter optimization	93

Table 5.9 Most optimal parameter value of *max_depth*, *col_sample_bytree*, and *nrounds* for all dimension 96

Table 5.10 Overall result of comparison for Xgb algorithm with and without hyperparameter optimization 96

Table 5.11 Overall result of comparison for tree-based algorithms with and without hyperparameter optimization 99

Table 5.12 Comparison of average accuracy results with literature 100

Table 6.1 Parameter setting for RF(Xgb) 109

Table 6.2 Percentage of accuracy for RF(Xgb) 110

Table 6.3 Result of comparison between the tree-based algorithm and RF(Xgb) 114

Table 6.4 Comparison of average accuracy results with literature 114

Table 7.1 Mapping research objectives to the research contributions 118

LIST OF FIGURES

		Page
Figure 1.1	Structure of the thesis	11
Figure 2.1	The content structure of chapter 2	13
Figure 2.2	Illustration diagram of naive bayes algorithm (Padmanaban, 2014)	28
Figure 2.3	Illustration diagram of decision tree (DT) algorithm	30
Figure 2.4	Illustration diagram of artificial neural network (ANN) algorithm	31
Figure 3.1	Research methodology	44
Figure 3.2	Confusion matrix	51
Figure 4.1	Four general basic steps of feature selection	57
Figure 4.2	Illustration diagram of random forest algorithm	65
Figure 4.3	Illustration diagram of extreme gradient boosting algorithm	66
Figure 4.4	Percentage of accuracy based on attributes selected for respective dimension	76
Figure 4.4(a)	Processing dimension	76
Figure 4.4(b)	Perception dimension	76
Figure 4.4(c)	Input dimension	76
Figure 5.1	Grid search diagram	80
Figure 5.2	Working flow of the tree based algorithm in learning style detection with hyperparameter optimization	81
Figure 5.3	Step 1 of the working flow of the tree based algorithm in learning style detection with hyperparameter optimization	81
Figure 5.4	Step 2 of the working flow of the tree based algorithm in learning style detection with hyperparameter optimization	82
Figure 5.5	Step 3 of the working flow of the tree based algorithm in learning style detection with hyperparameter optimization	82
Figure 5.6	Optimized *ntree* for input dimension	88

Figure 5.7 Optimized *ntree* for processing dimension 89

Figure 5.8 Optimized *ntree* for perception dimension 90

Figure 5.9 Optimized *ntree* for understanding dimension 91

Figure 5.10 Parameter value with the highest accuracy for input dimension 93

Figure 5.11 Parameter value with the highest accuracy for processing dimension 94

Figure 5.12 Parameter value with the highest accuracy for perception dimension 95

Figure 5.13 Parameter value with the highest accuracy for understanding dimension 95

Figure 6.1 Illustration diagram of parameter in Xgb 105

Figure 6.2 The working flow diagram RF(Xgb) 107

Figure 6.3 The combination of RF(Xgb) from the individual algorithm of RF and Xgb 108

Figure 6.4 ROC value for FSLSM dimension of hybrid RF(Xgb) 113

Figure 6.4(a) Input dimension 113

Figure 6.4(b) Processing dimension 113

Figure 6.4(c) Perception dimension 113

Figure 6.4(d) Understanding dimension 113

LIST OF ABBREVIATIONS

AH	Adaptive Hypermedia
AHS	Adaptive Hypermedia System
AI	Artificial Intelligence
AIWBES	Adaptive Intelligence Web Based Educational System
ANN	Artificial Neural Network
CART	Classification and Regression Tree
CFS	Correlation based Feature Selection
Corre	Correlation
CV	cross validation
DT	Decision Tree
DM	Data Mining
DEEDS	Digital Electronics Education and Design Suite
EDM	Educational Data Mining
FCBF	fast correlated-based filter
FSLSM	Felder Silverman Learning Style Model
ILS	Index of Learning Style
ITS	Intelligent Tutoring System
NN	Neural Network
NB	Naive Bayes
RbI	Rank by Importance
RF	Random Forest

RFE	Recursive Feature Elimination
SRL	self-regulated learning
TP	True Positive
TN	True Negative
UM	User Model
Xgb	Extreme Gradient Boosting

HIBRID HUTAN RAWAK DAN EXTREME GRADIENT BOOSTING YANG DIOPTIMUM UNTUK MENGKLASIFIKASI GAYA BELAJAR ATAS TALIAN

ABSTRAK

Perlombongan data pendidikan telah menarik banyak perhatian dalam kalangan penyelidik sejak beberapa dekad yang lalu. Perlombongan data pendidikan digunakan untuk memperoleh pemahaman yang lebih berkaitan tingkah laku pelajar dengan membina model berdasarkan data yang dikumpul melalui alatan pembelajaran yang digunakan untuk memperbaiki sistem pembelajaran supaya lebih bersifat peribadi dan boleh suai. Gaya belajar setiap individu di dalam pembelajaran atas talian ditentukan melalui interaksi dan tingkah laku mereka terhadap sistem tersebut. Model gaya belajar *Felder-Silverman's* merupakan teori model atas talian paling kerap digunakan untuk menentukan gaya belajar individu. Di peringkat awal, dalam menentukan gaya belajar, pengguna diminta untuk mengisi soal selidik yang direka untuk menentukan gaya belajar individu di akhir pembelajaran. Walaubagaimanapun, kaedah ini mengambil masa dan hasil dapatan tidak memuaskan disebabkan ia dipengaruhi oleh faktor tingkah laku manusia. Oleh itu, penyelidik mula mengkaji gaya belajar dengan menggunakan pendekatan automatik di mana fail log aktiviti dikumpul untuk memahami interaksi tingkah laku pengguna dengan sistem. Kandungan log fail terdiri daripada beberapa atribut yang dipadankan dengan sistem seperti bilangan lawatan, urutan tindakan dan istilah carian terpilih dan masa yang diperuntukkan. Ia juga termasuk penjejakan aktiviti di dalam sistem seperti carian, mengambil peperiksaan, kuiz, ujian penilaian kendiri, penggunaan forum, penghantaran email dan ruang perbincangan termasuk membaca atau memuat turun bahan daripada sistem. Disebabkan oleh banyak atribut yang terlibat dalam meramalkan gaya belajar, prestasi dan kecekapan klasifikasi ramalan masih rendah di antara 58%-89% (peratusan tertinggi adalah algoritma J48). Tambahan lagi kebanyakan penyelidik tidak mengoptimumkan parameter pilihan di mana ia turut menyumbang kepada prestasi matriks yang rendah. Penggunaan algoritma hibrid juga kurang diberi penekanan dalam bidang penentuan gaya belajar secara automatik di mana pertimbangan untuk menggunakan kaedah ini boleh membantu memperbaiki prestasi dalam meramalkan gaya belajar. Kajian ini dijalankan untuk menentukan atribut yang paling relevan dalam meramalkan gaya belajar. Pertama, tiga kaedah berbeza digunakan untuk memilih bilangan atribut yang paling relevan (ciri); iaitu *Rank by Importance* (*RbI*), penghapusan ciri rekursif (*RFE*) dan kolerasi. Seterusnya, algoritma berasaskan pokok digunakan untuk menentukan gaya belajar pengguna. Algoritma berasaskan pokok yang digunakan dalam kajian ini adalah hutan rawak (*RF*) dan (*Extreme Gradient Boosting*) (*Xgb*). Untuk mengoptimumkan prestasi matriks, parameter bagi algoritma berasaskan pokok dioptimumkan dengan menggunakan pengoptimuman hyperparameter berasaskan teknik carian grid. Daripada eksperimen tersebut (*RF*) dan (*Xgb*) kemudian dihibrid untuk meningkatkan lagi peratusan ketepatan dalam meramalkan gaya belajar. Hasil menunjukkan, memilih atribut yang paling relevan membantu dalam meningkatkan peratus ketepatan dengan peningkatan sebanyak (0.05%). Dengan menggabungkan pengoptimuman hyperparameter kepada algoritma berasakan pokok, peratus ketepatan meningkat secara positif sebanyak (0.03%). Akhir sekali, dengan melakukan hibrid kepada algoritma tersebut, peratus ketepatan meningkat sebanyak (0.05%) dengan nilai ketepatan sebanyak 96%. Ini membantu dalam mencari teknik yang paling berkesan dalam menentukan gaya belajar pelajar.

HYBRID OF OPTIMIZED RANDOM FOREST AND EXTREME GRADIENT BOOSTING FOR ONLINE LEARNING STYLE CLASSIFICATION

ABSTRACT

Educational Data Mining (EDM) have raised a lot of attention among researchers since the last few decades. EDM is used to gain more insight into the behavior of learners by building models based on data collected from learning tools which result in improving learning system to be more personalized and adaptive. Learning style of specific users in the online learning system is determined based on their interaction and behaviour towards the system. Felder-Silverman's learning style model is the most common online learning theory used in determining the learning style. Initially, in determining the users' learning styles, users are asked to fill in the questionnaires which is designed to learn their learning style at the end of the learning sessions. However, this method is time consuming and the result are not reliable due to the human factors behavior. Thus, the researchers started to study the learning style by using an automated approach in which the activity log files are collected in order to understand the interactivity behaviour of the users with the system. The content of the log files consists of several related attributes matched to the system such as the number of visits, sequences of actions and selected search terms, and time spent. It also includes activity tracking such as the searching, enroll in exam, quiz, self-assessment test, using forum, sending email and discussion board including reading or downloading materials from the system. Due to many attributes in predicting the learning style, the performance and efficiency of the classification and prediction are still poor which is between 58%-89% (highest was J48 algorithm. In addition, most of the researchers have not optimized the parameters selections which also contribute in the low performance matrices. The usage of hybrid algorithms is also less highlighted in the area of automated learning style detection where the consideration of this method may results in better performance in the prediction of learning style. This research is conducted to determine the most relevant attributes in predicting the learning style. First, three different methods are used to select the most relevant number of attributes (features); which are Rank by Importance (RbI), Recursive Feature Elimination (RFE) and Correlation. Next, tree based algorithm is being used to detect the learning style of the user. The tree based algorithms used in this study are the Random Forest (RF) and Extreme Gradient Boosting (Xgb). In order to optimize the results of the performance matrices, the parameters of the tree-based algorithms is being optimized by using the grid search hyperparameter optimization. From the experiments, RF and Xgb is then hybrid RF (Xgb) to improve the percentage of accuracy in predicting the learning style. Results shown, selecting the most relevant attributes helps in increasing the percentage of accuracy in learning style detection with an increment of 0.05%. By incorporating the hyperparameter optimization to the tree-based algorithms results in a positive increment in terms of the percentage of accuracy (0.03%). Finally, by doing a hybrid on the tree-based algorithms the percentage of accuracy is further improved by (0.05%) with accuracy value of 96%. The results help to find the most effective approach in determining the learning style of the student.

CHAPTER 1

INTRODUCTION

1.1 Preliminaries

Educational Data Mining (EDM) is a field that make full use of statistical, machine learning, and Data Mining (DM) algorithms over the various types of educational data. EDM is concerned with developing methods to explore the unique types of data in educational settings and using these methods, to better understand the students and settings in which they learn (Baker, 2010). The EDM process converts raw data from educational systems into useful information that could potentially have a great impact on educational research and practice. EDM emerges as a paradigm oriented to design models, tasks, methods, and algorithms for exploring data from educational settings (Luan, 2002). The aim of EDM is to enhance the understanding of various stakeholder about the way people learn in a data-driven way which can result in the improvement of the online learning to be more personalized and adaptive, and learning can be optimized as one of the main goals (Vahdat *et al.*, 2015).

In the past years, the research works aim to designate online learning environment based on learning style increased substantially (Feldman *et al.*, 2015). This is because, by determining the students learning style, an adaptive learning environment can be obtained. Based on the previous research, it is stated that, the first step to achieve an adaptive learning environment is by identifying students' learning style (Chang *et al.*, 2009; Dağ and Geçer, 2009; Feldman *et al.*, 2015).

Learning style is known as learning ways or preferences which are widely used by the learners on how materials are presented, how to work with it and how to internalize information (Litzinger *et al.*, 2005). It is also known as the learning choices and differences of an individual. Identifying students' learning style has several benefits of making students aware of their strength and weaknesses when it comes to learning and the possibility to personalize their learning environment to learning style (Bernard *et al.*, 2017). This learning style identifies capability of each students, how they handle certain courses, subject and the system. Learning style is meant to determine the learning preferences of each students either in a traditional classroom or online learning based.

There are many learning style models available in the last 30 years, where over 70 theories were developed (Coffield *et al.*, 2004). Some of the widely used learning style models include Kolb learning theory, Felder Silverman Learning Style Model (FSLSM), Vark learning theory and Dunn and Dunn learning style model. Some of the model can overlap to each other. The most commonly used learning style model is the FSLSM. It has four different dimension which are processing, understanding, perception and input. FSLSM also incorporated different elements from different learning style model such as Kolb, Pask and Myers-Briggs (Mokhtar *et al.*, 2010). There are several reasons mentioned by previous researchers as on why FSLSM is the most preferable model in the area of automated learning style detection. One of the reasons is because the validity and reliability of the Index of Learning Style (ILS) scale used in FSLSM are already in a mature state as mentioned by Felder and Spurlin (2005). FSLSM is also suggested as the most preferable model for adaptive e-learning systems to provide learning styles based adaption. In addition, this model is also the most used model in the literature (Ciloglugil, 2016).

Initially, in determining the learning style of users in an online learning, researchers ask the students to fill in the questionnaire at the end of the learning session (Truong, 2016). The output obtained from the questionnaire is input back to the system to be used further to personalize the online learning system (Bernard *et al.*, 2017). While these instruments present good reliability and validity, they have been subjected to some criticism. Firstly, filling out a questionnaire is a boring task that requires an additional amount of work from the students, given that some questionnaires have more than 100 items. Secondly, students tend to choose answers arbitrarily if they are not aware of the importance or the future uses of the questionnaire. Thirdly, students can be influenced by the way questionnaire is formulated, which lead them to give more appropriate answers. Fourthly, questionnaires are prepared to the assumption that students are aware of their learning preferences, but this is not always the case. Finally, learning styles can vary over time. A questionnaire is a static approach, as soon as the learning style changes, the results of the questionnaire are no longer valid (Mokhtar *et al.*, 2010; Feldman *et al.*, 2015).

Therefore, researcher come out with an alternative, where they determine the learning style automatically (Truong, 2016). This is done by collecting the log files on the interacting behaviour between the user and the system. The content of the log files consists of several related attributes matched to the system such as the number of visits, characteristics and types of objects chosen, sequences of actions and selected search terms, number of visits, time spent and performance. It also includes the activities tracked such as the searching, enroll in exam, quiz, self-assessment test, using forum, sending email and discussion board including reading or downloading of materials from the system (Truong, 2016; Ciloglugil, 2016). These attributes were then matched with the learning style model. Then, the result is further analyzed using machine learning algorithms until the learning style of the user is determined.

Unfortunately, there are still weaknesses in this method where researchers are not certain on which attributes are relevant in determining the learning style. This is because, the current percentage of accuracy of determining the learning style is still low which is between 50%-85% (Bernard *et al.*, 2017; Maaliw III, 2016). One of the application in machine learning which address the problem of reducing irrelevant and redundant attributes is feature selection method.

Feature selection is a term commonly used in data mining to describe the techniques available for reducing inputs to a manageable size for process and analysis. It helps in understanding data, reducing computation requirement, reducing the effect of the curse of dimensionality and improving the predictor performance (Kim *et al.*, 2003). There are three general classes of feature selection namely as filters, wrappers and embedded (Miao and Niu, 2016).

Different classification algorithms have been used in order to have an automated learning style identification. One of the most popular methods is rule-based in which researchers "translated" different learning styles based on the learning style model (Graf *et al.*, 2009). In a learning environment the learning styles of a student is a decisive factor. In many cases there is a mismatch between personal learning styles and the learning demands of different disciplines.

Researchers have applied some widely used techniques such as Artificial Neural Network (ANN), Naive Bayes (NB) and Decision Tree (DT). However, there are still gap within the usage of the stated algorithm in terms of the accuracy of the result obtained. One of the ways to enhance the performance of the algorithms, is by doing hyperparameter optimization in the selected algorithms. On the other hand, some researchers also proposed to hybrid the machine learning algorithms. However, in the domain area of the automated detection of learning style there are very few proposed hybrid algorithms and the result is in the average percentage of 68% (Özpolat and Akar, 2009; Cha *et al.*, 2006; Hung *et al.*, 2016).

In machine learning, hyperparameter optimization is the problem of choosing a set of optimal hyperparameters for a learning algorithm. The problem of identifying a good value for hyperparameters (λ) where $\lambda = parameter$ is called the problem of hyperparameter optimization (Bergstra and Bengio, 2012). The critical step in hyperparameter optimization is to choose the set of trials $\lambda^1 ... \lambda^s$. The most commonly used technique in hyperparameter optimization is by doing a grid search technique. Grid search technique is simple to implement and parallelization is trivial. Other than that, it is also reliable in low dimensional spaces. It is required to choose a set of values for each parameters. Then, the set of trials is formed by assembling every possible combination of this values which leads in determining the most possible optimal value of the parameters.

Lastly, one of the crucial steps to further improve the performance of the algorithms is by doing a hybrid. Numerous methods have been suggested for the creation of hybrid of classifiers (Dietterich, 2000). Although many methods of hybrid have been proposed, yet there is no clear picture of which method is the best (Vilalta and Drissi, 2002). Thus, an active area of research in supervised learning is the study of methods for the construction of good hybrid algorithms.

1.2 Motivation

Learning style have raised a lot of attention among researchers in the last decade. It is used to gain more insight into the behavior of learners by building models based on collected data from learning tools which result in improving learning system to be more personalized and adaptive. Adaptive is a situation of having an ability to change to suit the different conditions. Learning style of a specific user is determined based on their interaction and behavior towards the system. By determining an accurate learning style of the user it will lead to a better adaptivity of the online learning system which will then increased the user performance.

In the past years, researchers had started to determine learning style automatically by using the activity log files from the online learning system along with machine learning algorithms. But, the accuracy value is still poor due to many attributes involved and other factors that were not considered when applied the machine learning algorithms such as hyperparameter optimization and hybridization of algorithms. This resulted in imperfect learning style detection which might cause poor adaptability of the online learning system. With that reason, this study used different methods of machine learning approaches such as the tree-based algorithms and feature selection methods. Detail discussions on the problem and objectives are discussed further in the next section.

1.3 Problem Statement

From previous research, determining an accurate learning style's models play an important role in increasing the adaptivity of the online learning system. Theoretically, learning style varies from time to time, due to the change of learner's behaviour and the insincerity of the user when answering the questionnaire, so the validity of the learning style obtained using questionnaire is doubted (Mokhtar *et al.*, 2010).

Hence, an automated way of determining the learning style was introduced by using the log files from the online learning system. However, the log files contain many attributes which were used in predicting the learning style and this result in lower percentage of accuracy in the prediction model. Therefore, to improve the percentage of accuracy, this research focus on finding the most effective algorithms in determining the learning style. This is because, from the previous research, the percentage of accuracy obtained in the learning style detection is still low which are between the range of 75%-85% (Bernard *et al.*, 2017; Maaliw III, 2016; Özpolat and Akar, 2009).

This research focuses on increasing the accuracy in predicting a learning style. It concerns with improving current techniques to explore the unique types of data in educational settings, to better understand students and their learning environment (Baker, 2010). Selecting the most relevant attributes in learning style prediction plays a crucial role in increasing the percentage of accuracy. The performance of the learning algorithms over the test set, often motivates feature selection, which consists of detecting the relevant features and discarding the irrelevant ones (Sánchez-Maroño *et al.*, 2007). Determining the relevant attributes through feature selection has several advantages which are first, improving the performance of the machine learning algorithm. Additionally, it can aid in data understanding, and gaining knowledge about the process (Guyon and Elisseeff, 2006).

Improving the performance of the selected algorithms plays an important role in increasing the learning style prediction. Several factors which can lead to a better prediction of accuracy are including by doing a hyperparameter optimization and hybrid of algorithms. Hyperparameter optimization is the process of identifying a good value for hyperparameters; i.e (λ) where $\lambda = parameter$. It is the minimization of parameters over a subset of parameter. While hybrid is the result of combination of two or more algorithms for efficiently solving the problem. It is commonly designed to yield a better performance of the algorithms compared to individual algorithm.

With that, this research was conducted to determine the most relevant attributes and the most effective algorithm in predicting the learning style of the user.

The following are the three research questions:

1. What is the most relevant attribute that can be used to determine the learning style,
2. What is the most effective algorithm that can be used to determine the learning style,
3. How to improve the performance of algorithm.

1.4 Goal and Objectives of the study

The main goal of this research is to find the most effective approach in determining the learning style of the student. The effectiveness of the result can be further evaluate by comparing and contrasting the result obtain from literature review.

The objectives of this research are:

1. To determine the contributing attributes of the user's learning style model using feature selection methods and tree-based algorithms,
2. To improve the tree-based algorithms using hyperparameter optimization,
3. To design and evaluate the hybrid of Random Forest RF and Extreme Gradient Boosting Xgb to improve the accuracy of learning style prediction.

1.5 Study Scope and Limitations

In determining the most effective algorithm used in increasing the adaptiveness of the adaptive learning system, there are many rules need to be taken into consideration. Therefore, various parameters, constraints and uncertainties need to be clarified in order to make the study more achievable. With that, some scopes and limitations have been clarified to complete the study. The scopes and limitations of this research are given as follows:

(i) This study uses available dataset obtained from the previous researcher (Maaliw III, 2016). Therefore, the attribute of the datasets and the variable used is fixed as per requirement from paper.

(ii) The comparison of proposed algorithms is based on the obtained result with different literature review.

(iii) This research did not taking into consideration any demographic value such as gender, country and age group.

(iv) For the computational part, the focus was on using machine learning techniques. This study focused on the feature selection techniques and classification task.

1.6 Outline of Thesis

This thesis is organized into seven chapters. Figure 1.1 shows the structure of the thesis.

Brief descriptions of the content of each chapter are given as follows:

(i) The thesis begins with Chapter 1 which consist of problem statement, the goal and objectives of the research and also the scope and significance of the research.

(ii) Next, Chapter 2 reviews in detail the related work of domain problem and at the same time emphasis on the selected domain which would help in understanding the overall context of the thesis.

(iii) Chapter 3 describes the research methodology of this thesis which consist of research framework, data sources, instrumentation, problem description, performance measure, and experimentation and analysis.

(iv) Chapter 4 discusses the detail on selecting the most relevant attributes. It first involves in determining the most relevant attribute by using three different feature selection methods which are RbI, RFE, and Correlation. Then, the methods are evaluated by using the selected tree-based algorithms which are RF and Xgb.

(v) Chapter 5 provides the result of predicting the learning style using the tree based algorithms by incorporating hyperparameter optimization. Other than that, the performances of the algorithms are compared with the current literature result.

(vi) Chapter 6 provides the results of the learning style prediction by enhancing the tree based algorithms. Then, the performances of the algorithms are compared with current literature result.

(vii) Chapter 7 discussed the conclusion of the thesis along with the future works to be done as the study outcome.

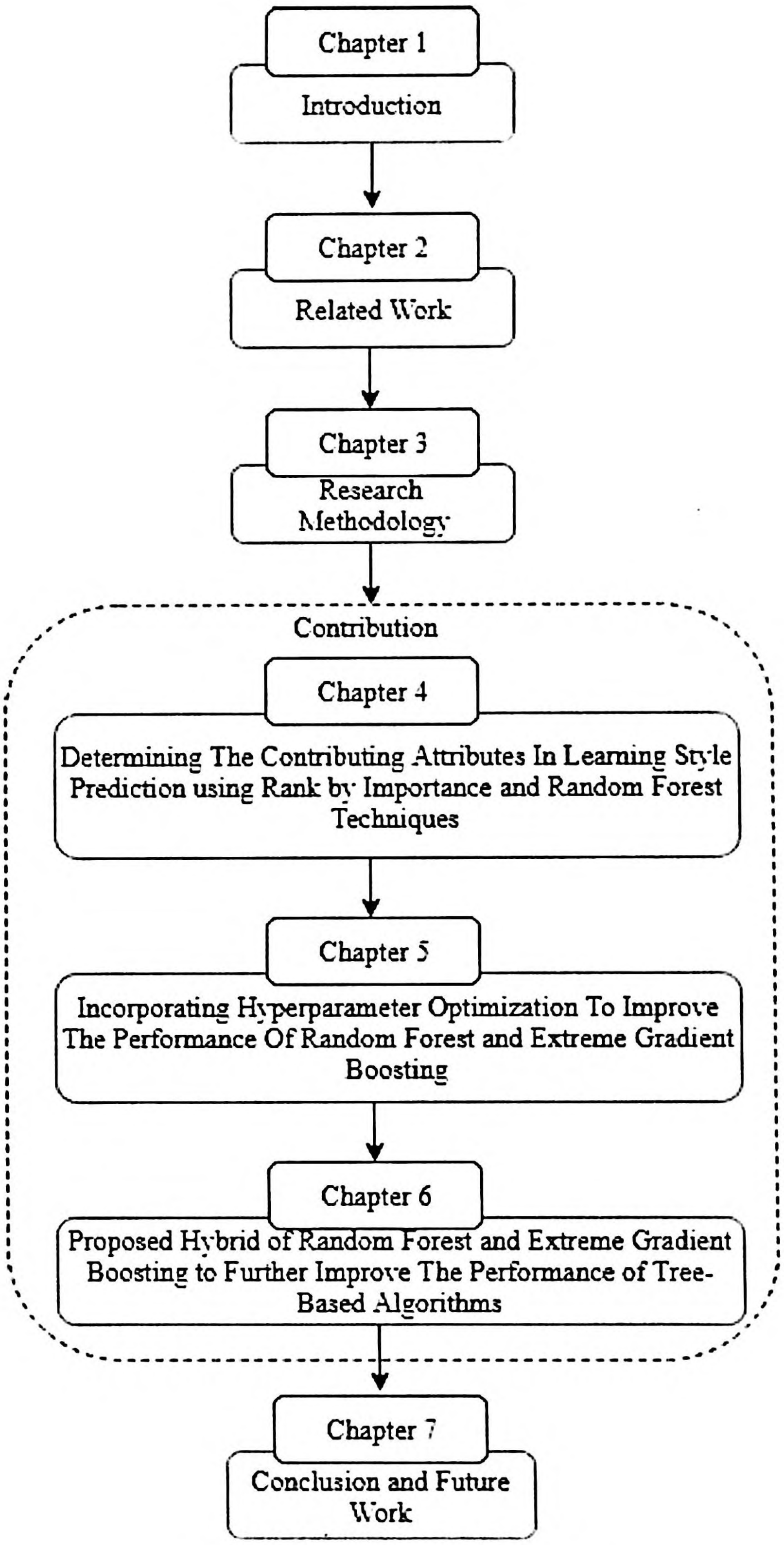
Chapter 1
Introduction
Chapter 2
Related Work
Chapter 3
Research Methodology
Contribution
Chapter 4
Determining The Contributing Attributes In Learning Style Prediction using Rank by Importance and Random Forest Techniques
Chapter 5
Incorporating Hyperparameter Optimization To Improve The Performance Of Random Forest and Extreme Gradient Boosting
Chapter 6
Proposed Hybrid of Random Forest and Extreme Gradient Boosting to Further Improve The Performance of Tree-Based Algorithms
Chapter 7
Conclusion and Future Work

Figure 1.1: Structure of the thesis

CHAPTER 2

LITERATURE REVIEWS

2.1 Introduction

This chapter reviewed related works in the area of the learning style detection. In this chapter, the motivation and inspiration of adopting the domain problem as well as the chosen methodology for the research study is emphasized. Additionally, the potential trend and direction derived from the review are discussed within the scope of this research study. The organization of this chapter is given in Figure 2.1.

2.2 Educational Data Mining

Educational data mining (EDM) can be defined as an emerging discipline, concerned with developing methods for exploring the unique types of data that come from educational settings, and using those methods to better understand students, and the settings which they learn in (Dutt *et al.*, 2017). Romero and Ventura (2007) in their paper list the primary applications of EDM which are analysing and visualizing of data, providing feedback for students, and recommendations for students. In addition, the applications also included predicting students performance, students' modeling, detecting undesirable students' behaviors, social network analysis, developing concept maps, constructing courseware and planning and scheduling (B.Namratha, 2016).

EDM emerges as a paradigm oriented to design models, tasks, methods, and algorithms for exploring data from educational settings (Peña-Ayala, 2014). EDM methods can be used to identify an operationalize specific behaviors, theorized to represent motivation or metacognition (Winne and Baker, 2013). Just as early efforts to understand online behaviors,

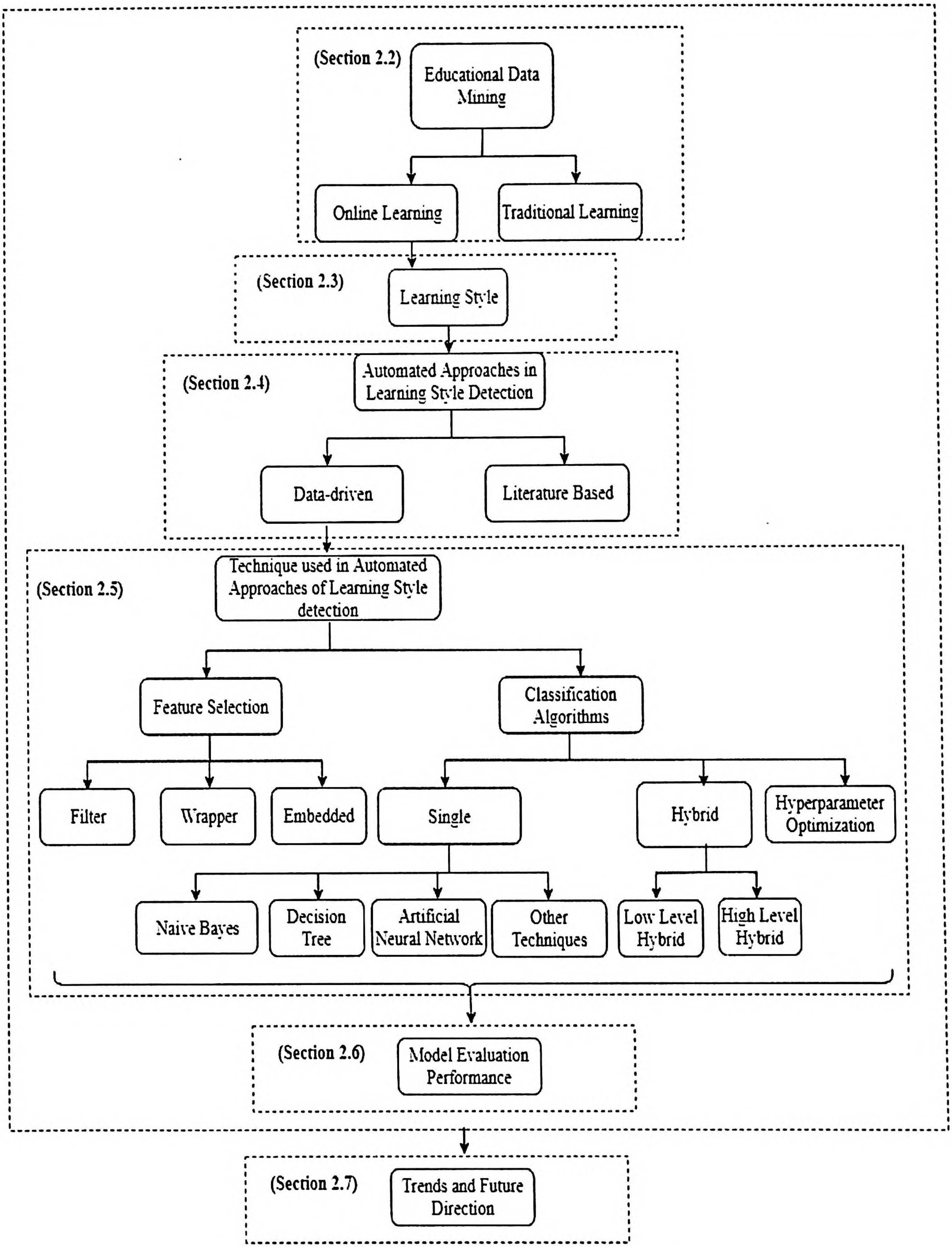

Figure 2.1: The content structure of chapter 2

early efforts at EDM involved in mining website log data (Amershi and Conati, 2009), but now it is integrated, instrumented, and sophisticated kind of learning systems which provide more data. EDM generally emphasizes in reducing learning into small components that can be analyzed and then influenced by a software which is adapted to the students (Arnold, 2010).

In the last few years, EDM methods have enable considerable expansion in the sophistication of student learning models. In particular, EDM methods enable researchers to make higher-level inferences about students' behavior, such as when a student is gaming the system, when a student has "slipped" (making an error despite knowing a skill), and when a student is engaging in self-explanation (Shih *et al.*, 2011). EDM emerged as a research area in recent years for researchers all over the world from different and related research areas, which are traditional learning and online learning (Romero and Ventura, 2010).

2.2.1 Traditional Learning

In traditional learning, usually teacher talks more than the student. In terms of the subject matter, in traditional learning, the teacher conducts the lesson according to the study program and the existing curriculum.

To emphasis in the learning process, the students learn "what" instead of "how", the students and the teachers are busy completing the required subject matter quota, the students are not involved in inquiry-based education and in problem solving, but rather in tasks set by the teachers. In terms of motivation of the students', it will be differ from time to time but there is no option of having a break and they still need to proceed with the lesson and the subject matter is "distant" from them (Irwin-Zarecka, 2017).

2.2.2 Online Learning

Online learning system is an online courses or learning software or interactive learning environments that use intelligent tutoring systems, virtual labs, or simulations. There are plenty of online elarning environment available nowdays. Online courses offered through a learning or course management systems such as blackboard, moodle or sakai. Examples of learning software and interactive learning environment are those from Kaplan, Khan Academy, and Agile mind. When online learning systems use data to change the response towards the students' performance they become adaptive learning environments (Oxman *et al.*, 2014).

The second-generation tutors are usually called Intelligent Tutoring System (ITS) (VanLehn *et al.*, 2007). ITS is a computer software designed to simulate a human instructor's skills behavior and guidance. Such ITS systems are capable in interpreting complex students' responses and can be learnt as they work, they are able to distinguish where and why a student understanding has gone amiss and to give hints to aid the student to understand the material given. Intelligent tutoring system delivers numerous benefits of a human instructor to very large numbers of students.

Intelligent tutoring systems can also deliver real-time data to instructors and developers who want to improve their teaching styles (Aleven and Koedinger, 2002). Various ITS deliver step-by-step monitoring of the student's solution as it is being produced (Naser, 2016), while others offer feedback on the final solution (Aleven and Koedinger, 2002). An ITS provides customized computer-based instruction to students (Conati *et al.*, 2002; Naser, 2016).

An intelligent tutoring system is a system designed similar to the teachers' behavior in teaching. It can help students studying set of subjects by series of lessons, many questions in each lesson, and offering specific instruction with feedback. It can explain complex student responses and learn what they want. The system makes a profile for each student and estimate the student's degree of skill. This type of system can change its tutoring behavior in real time. The aim is not merely to know a whether a response is correct or not, but to understand which of the students' response is incorrect (Al-Bastami and Naser, 2017).

Studies on adaptive educational hypermedia become popular in the recent years, due to the expanded use of e-learning around the world. However, the adoption of these adaptive systems in actual e-learning platforms is still not widespread. One of the main reasons is that current educational systems do not offer personalization options has to do with a variety of challenging issues. These challenges are all discussed in the recent field literature and can be grouped into four main thematic categories (Brusilovsky and Peylo, 2003). These categories are inter-operability, open corpus knowledge, usage across a variety of delivery devices, and the design of meta adaptive systems (Somyürek, 2015).

A hypermedia application offers its learners much freedom to navigate through a large hyperspace. AH offers its learners personalized content, presentation, and navigation support. The researchers presented a survey of AH architecture, defined a new taxonomy of adaptation techniques and also introduced a set of requirements and a modular structure that can be used to update the first generic AH model adaptive hypermedia application model (AHAM) that was introduced 10 years ago (Knutov *et al.*, 2009). AEHS systems offer an alternative to the non-individualized instruction approach, by providing various services adapted to the learner profile. These systems are based on user models which characterise each individual and can use these models to offer learners educational experiences which fit their needs. To achieve this, AEHS are comprised of several sub-components which have their own distinct behaviours and properties. On the other hand, the hierarchy of the evaluation task is shown by the evaluation-wide system (Mulwa *et al.*, 2010).

Web based educational systems have changed the paradigm of teaching and learning. They have made the teaching-learning depends on time and space. Materials and information stored and maintained at one place can be accessed globally at any time on demand.

Integrating intelligence, interactivity and simulation in the web based courseware makes it not only attractive but also essential for effective learning process. Computer use in education started with the implementation of textbook as a model for developing learning support tools (Brusilovsky, 1996). With the advancement of technology, the concept,started as electronic copy of the hardcopy of the text book, which undergo many modifications. Additional like audio, video, animation etc. were also introduced. Birth of internet introduced global access along with hypertext technology. The passive text books were made more vibrant and functionally effective by integrating with ITS technology.

E-learning scenario today is dominated by Learning Management Systems (LMS) like blackboard (Bradford *et al.*, 2007). These systems support several activities performed by instructors and learners during e-learning process. While the learners use the system for learning, communicating and collaboration teachers use it to develop web-based course notes and quizzes, communicate with learners , and evaluate and monitor learner's evolution.

However, all the LMS-based courses offer the same educational material for all learners irrespective of their individual differences in skill and knowledge levels, goals and interests of the learners. This approach is changed by Adaptive web based educational systems (AWBES). These systems are found to perform better than LMS (Brusilovsky, 2004). Adaptive quizzes, intelligent solution analyzers, adaptive class monitoring systems and adaptive collaborative systems perform their respective functions more efficiently.

The web based educational system incorporates adaptivity and intelligence by building a model of individual learner's goals, preferences and knowledge. The system is then adapted to the learner's needs by using the model through the interactivity with the learner. Intelligence is brought to play in the system by imitating human teacher through tutoring functionality. This includes functions like coaching, helping in problem solving. There are some systems which are intelligent (Mitrovic, 2003), and adaptive (Yano *et al.*, 2002). Some systems incorporate both adaptive and intelligent. The first intelligent and adaptive web based educational systems were developed in 1995-1996 (Brusilovsky and Peylo, 2003). These were then followed by many interesting ones. Adaptive and intelligent technologies are the different ways of adding adaptive or intelligent functionalities to education systems. Most of the technologies are derived from ITS and adaptive hypermedia systems. These methods are found to exist before internet era and termed classic technologies.

In developing the Adaptive Intelligence Web Based Educational System (AIWBES) there are four models which need to be taken into consideration which are the user model, network model, knowledge model and environment model. User model gives an indication of the user's current knowledge, interests, or activity including the learning style of the user. Learning styles play an important role in adaptive e-learning systems as it refer to students' preferred ways of learning. Different students have different preferred ways of learning. Some students prefer learning through texts and readings while some understand better through images (Truong, 2016). Learning styles can be differentiated by the way students process information. If the students is active, they do not perform well in a classroom situation but they learn effectively through interaction with other students (Felder and Silverman, 1988). Section 2.3 discusses in detail about the learning style.

2.3 Learning Style

A learning style is a student's consistent way of responding to and using stimuli in the context of learning (Felder and Silverman, 1988). This is because some students may prefer to learn by doing hands-on activities while others prefer to read and reflect about it. Other than that, it is also observed that there are some students who prefer working alone. It can be described further as the entirety of the preferences a student has for low learning material, how they process information and how they internalize the information (Truong, 2016).

Learning style of a specific user is determined based on their interaction and behavior towards the system. By determining an accurate learning style of the user will lead to a better adaptivity of the online learning which will then increase the users' performance. Different students have different preferred ways of learning. Some may deal with theories, others may learn through experiments and examples (Bernard *et al.*, 2017).

With that, researchers come out with e-learning system, which allow researchers to observe student's behaviors throughout the learning process. Data mining and computerized algorithms, is used to quickly identify and analyze trends in big data set which opens opportunities to develop a new framework to observe and measure learning style through online behaviors. Learning styles are also useful sources to develop an adaptive e-learning system that effectively personalizes learning resources to individuals' learning needs (Bernard *et al.*, 2017). A survey paper by Dutt *et al.* (2017) even suggested that learning styles models were the most useful framework for adaptive system development among other sources such as previous knowledge and student background.

Learning styles play an important role in adaptive e-learning systems as it refers to students' preferred ways of learning. Different students have different preferred ways to learn.Some students may prefer learning through texts and readings while some may understand better through images (Truong, 2016). Learning styles can be differentiated between the way students process information. If the students is an active students, they do not perform well in a classroom situation but they learn effectively through interaction with other students (Felder and Silverman, 1988).

Learning style concept by Dunn (1984) stated that learning style of each student is different and personalized in a unique way, while they proceed to learn and memorize a new knowledge. Then, Kolb has developed research on experiential learning theory. Thereafter, Kolb has been developed researches on "Experiential Learning Theory". He made these assistive researches to support that concept. Learning style is a signed indicators that how the students perceive, interact and response to the learning environments (Özerem and Akkoyunlu, 2015). Different researcher states that learning style composes of distinguish behaviors which shows how the student learns knowledge from the environment and adopt these knowledge to himself (Bhagat *et al.*, 2015). Stone (2014) expressed learning style as unchanged personal process group which guides us while we receive knowledge from our environment.

There are many learning style model available in this area as mentioned by Truong (2016) in the last 30 years, over 70 theories were developed. One of the most commonly used is the Felder-Silverman's model, which differentiates learning styles through 4 different dimension which are:-

- Perception (Sensory/Intuitive)
- Input (Visual/Verbal)
- Processing (Active/Reflective) and
- Understanding (Sequential/Global)

This theory is by far the most widely used in adaptive learning system (accounted for 70.6% of all papers in the survey conducted by Truong (2016). FSLSM can described the students' learning style in great detail. For the perception dimension which consists of sensing and intuitive learning style it describes a preference for processing information. In this dimension, learners with sensing learning styles prefer to learn facts and concrete materials, using their sensory experiences of particular instances as a primary source. On the other hand, intuitive learners prefer to learn abstract learning material, such as theories and their underlying meanings, with general principles rather than concrete instances being a preferred source of information.

The input dimension consists of visual and verbal learning style. Visual learners prefer materials such as graphs, charts or videos, while verbal learners prefer words either written or spoken. In the third dimension which is the processing dimension consist of active and reflective learning style. Active learners prefer to learn by doing, experimentation and collaborative while reflective learners prefer to think the information and absorb it alone or in small groups.

Lastly, for the understanding dimension it consists of sequential and global learning style, where sequential learner prefer information to be provided in a linear (serial) fashion and tend to make small steps through learning material while global learner tend to make larger leaps from non-understanding to understanding and tend to require seeing the "big picture" before understanding a topic.

Several ways proposed by the researchers in detecting the learning style. One of the ways is by using an automated approach where, the use of machine learning algorithms is incorporated in it. Section 2.4 discusses in details the previous approach used and trend.

2.4 Automated Approaches in Learning Style Detection

Automatic detection of learning styles in an online learning systems is an important problem in which many researchers have proven their interest because it enhances learning efficiency by implementing adaptation dynamically (Dung and Florea, 2012a). Many studies aim at the automatic detection of learning styles to avoid intentional or unintentional wrong answers, and to save student's time on answering a questionnaire. There are two main approaches in the automated detection of learning style, namely as the data-driven and literature-based approaches. Table 2.1 shows the comparison from the previous literature on the data-driven and literature based approach. The details of both approaches are discussed in subsection 2.4.1 and 2.4.2.

Table 2.1: Table of comparison between data-driven and literature-based approach

Author	Literature-Based	Data-Driven
Dung and Florea (2012b)	/	
Graf *et al.* (2009)	/	
Latham *et al.* (2012)	/	
Popescu (2009)	/	
Sangineto *et al.* (2008)	/	
Alkhuraiji *et al.* (2011)		/
Carmona *et al.* (2008)		/
García *et al.* (2007)		/
Ahmad and Shamsuddin (2010)		/
Kelly and Tangney (2006)		/
Alkhuraiji *et al.* (2011)		/
Cha *et al.* (2006)		/
Crockett *et al.* (2017)		/
Özpolat and Akar (2009)		/
Georgiou and Makry (2004)		/
Kolekar *et al.* (2010)		/
Lo and Shu (2005)		/
Stathacopoulou *et al.* (2005)		/
Cabada *et al.* (2009)		/
Villaverde *et al.* (2006)		/
Yannibelli *et al.* (2006)		/
Chang *et al.* (2009)		/
Maaliw III (2016)		/
Bernard *et al.* (2017)		/
Gilbert and Han (1999)		/

2.4.1 Literature-based

The idea of the literature-based approach is to use the behaviours of students in order to get hints about their learning style preferences and then apply a simple rule-based method to calculate learning styles from the number of matching hints. The behaviours refer to the interaction of the user with the system such as the number of forum post and other. This approach is similar to the method used for calculating learning styles in the index of learning style questionnaire (ILS) and has the advantage to be generic and applicable for the data gathered from any course, due to the fact that FSLSM is developed for learning.

The Index of Learning Styles is an on-line survey instrument used to assess preferences on four dimensions (active/reflective, sensing/intuitive, visual/verbal, and sequential/global) of a learning style model formulated by Felder and Silverman (1988). However, the approach might have problems in estimating the importance of the different hints used for calculating the learning styles. A method of using this approach was proposed by Graf *et al.* (2007). The authors analysed the behaviours of 127 learners during an object oriented modeling course in LMS Moodle.

The literature-based approach apply a simple rule-based method to calculate the learning style of the user by using the matching hints of the interaction between the user and the system. This approach is similar to the method used for calculating learning style preferences by the learning style instruments. The literature-based approach has the advantage that it is generic and applicable to data gathered from any learning course because learning style models are developed for learning in general. However, the approach might have problems in estimating the importance of the different hints used for calculating the learning style preferences (Feldman *et al.*, 2015).

2.4.2 Data-driven

The automatic detection of learning styles in data-driven approaches is carried out by an Artificial Intelligence. (AI) classification algorithm that takes the user model as input and returns the students' learning style preferences as output. This approach has the advantage as it uses real data to classify the user, so it can be very accurate (Graf *et al.*, 2007; Truong, 2016; Feldman *et al.*, 2014). However, the approach strictly depends on the available data and therefore, a representative dataset is crucial to build an accurate classifier. This classifier able to identify learning styles from data of the same learning course and at the same time identify learning styles from data of any other course.

The data-driven approach is more commonly used than the literature-based approach because the latter requires of having some knowledge of psychology and cognitive science to correctly estimate the importance of the hints. In contrast, data-driven approaches are more familiar to computer science researchers because it requires to gather relevant information for the model's user and then use an AI classification algorithm to automatically detect the learning style preferences. The works which employ a data-driven approach apply several AI classification algorithms to automatically detect learning styles (Bernard *et al.*, 2017; Maaliw III, 2016).

In addition, data-driven approach uses sample data in order to build a model that imitates the ILS questionnaire to identify learning styles from the behaviours of learners. The advantage of the approach is that the model can be very accurate due to the use of real data. However, the approach strictly depends on the available data (Dung and Florea, 2012a). Therefore, it may be difficult to have a good data set used for detecting learning styles because the data are scattered on different courses. One of the studies in this approach is conducted by García *et al.* (2007). The authors observed the behaviours of learners during an online course in the SAVER system and performed two experiments to show the effectiveness of bayesian networks to identify learning styles based on the behavior of students. The approach consider the active/reflective, sensing/intuitive, and the sequential/global dimension of FSLSM. The result showed that the bayesian network obtains good results for the sensing/intuitive dimension and can detect the active/reflective and sequential/global dimension provided that the students have some learning experiences in web-based courses and that they are encouraged to communicate with each other via communication tools. Section 2.5, discusses the previous techniques used in the detection