Classification of Sinhala Songs based on Emotions

A dissertation submitted for the Degree of Master of Science in Computer Science

K.M.H.B. Abeyratne
University of Colombo School of Computing
2019

Declaration

The thesis is my original work and has not been submitted previously for a degree at this or any other university/institute.

To the best of my knowledge it does not contain any material published or written by another person, except as acknowledged in the text.

Student Name: K.M.Harshi Abeyratne

Registration Number: 2016mcs002

Index Number: 16440025

Signature: Date:

This is to certify that this thesis is based on the work of Ms. K.M.Harshi Abeyratne under my supervision. The thesis has been prepared according to the format stipulated and is of acceptable standard.

Certified by:

Supervisor Name: Dr. K.L. Jayaratne

Signature: Date:

Abstract

Classification of Sinhala Songs has received less attention from researchers in the Sri Lankan context. The purpose of the current study is to shed light on emotion based classification based on Music Emotion Recognition (MER) which is a subdomain of Music Information Retrieval (MIR). In achieving this the authors followed three steps namely; selection of an appropriate emotion model, extraction and selection of low level music feature and training of the classifier utilizing supervised machine learning algorithms. 18 experiments were conducted to analyze the level of accuracy in each feature selection algorithm in combination with the supervised machine learning algorithms. Results of the current study suggest that the combination of ReliefF based feature selection algorithm and Random Forest supervised machine learning algorithm yields highest accuracy classification of Sinhala songs based emotions. The highest accuracy level of 91.98% in the current study is identified as the highest result achieved within the domain of Sinhala songs classification according to literature review conducted.

Acknowledgement

I would first like to thank my supervisor Dr. K. L. Jayaratne of the University Of Colombo School Of Computing. The door to Dr. Jayaratne's room was always open whenever I ran into a trouble spot or had a question about my research or writing. He consistently steered me in to the right direction whenever he thought I needed it. I would like to express my sincere gratitude for his patience, motivation, enthusiasm, and immense knowledge.

I would also like to thank the music experts, specially to the Music Director Mr. Suresh Maliyadde who were involved in the emotions labeling for this study. Without their passionate participation and input, emotion labeling of songs could not have been successfully conducted. I thank Mr. Nandasara, Lecturer, University of Colombo School of Computing for the help he provided in finding Sinhala songs.

I would also like to acknowledge Dr. Manjusri Wickramasinghe, Senior Lecturer, University of Colombo School of Computing, Mr. Roshan Rajapaksha Lecturer, University of Colombo School of Computing and Mr.Rangana Amaraweera Lecturer, University of Colombo School of Computing, for their very valuable comments on this research.

My sincere thanks also goes to the all the staff members of library, University of Colombo School of Computing.

Finally, I must express my profound gratitude to my parents and to my husband for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AMMMFCCOA	Area Method of Moments of MFCCs Overall Average
AMMMFCCOSD	Area Method of Moments of MFCCs Overall Standard Deviation
BSOA	Beat Sum Overall Average
CBS	Correlation Based Selection
COA	Compactness Overall Average
COSD	Compactness Overall Standard Deviation
FOLEWOA	Fraction Of Low Energy Windows Overall Average
FOLEWOSD	Fraction Of Low Energy Windows Overall Standard Deviation
HP	High Arousal and Positive Valence
IGBS	Information Gain Based Selection
LAPV	Low Arousal and Positive Valence
LN	Low Arousal and Negative Valence
LPC	Linear Predictive Coefficients
LPCOSD	LPC Overall Standard Deviation
MER	Music Emotion Recognition
MFCC	Mel-frequency Cepstral Coefficients
MIDI	Musical Instrument Digital Interface
MIR	Music Information Retrieval
MLkNN	Multi-Label k-NearestNeighbors
MMOA	Method of Moments Overall Average
MMOSD	Method of Moments Overall Standard Deviation
RBS	ReliefF Based Selection
RMSOA	Root Mean Square Overall Average
RMSOSD	Root Mean Square Overall Standard Deviation
RMS	Root-Mean-Square
SBOA	Strongest Beat Overall Average
SBOSD	Strongest Beat Overall Standard Deviation
SCOA	Spectral Centroid Overall Average
SCOSD	Spectral Centroid Overall Standard Deviation
SFOA	Spectral Flux Overall Average
SFOSD	Spectral Flux Overall Standard Deviation
SMO	Sequential Minimal Optimization
SOSBOA	Strength Of Strongest Beat Overall Average
SOSBOSD	Strength Of Strongest Beat Overall Standard Deviation
SRPOA	Spectral Rolloff Point Overall Average
SRPOSD	Spectral Rolloff Point Overall Standard Deviation
SVOA	Spectral Variability Overall Average
SVOSD	Spectral Variability Overall Standard Deviation

ZCOA	Zero Crossings Overall Average
ZCOSD	Zero Crossings Overall Standard Deviation
SVM	Support Vector Machine
WEKA	Waikato Environment for Knowledge Analysis

Chapter 1

Introduction

1.1 Background

Music plays a major role in human life from ancient history to current digital era. It is universally accepted that human emotions and music has a strong bond. When considering the emotions music can be classified into different categories based on genres, cultural backgrounds, rhythm etc. [1]. In this digital era music listeners have access to millions of songs containing different emotions. Accessibility to music improves day by day. Hence, it is essential to have a more efficient organization and search methods to search for a song. Due to the subjective nature of human insights, it is a huge challenge to classify songs based on emotions.

However, Music Emotion Recognition (MER) has gained increasing attention in the field of Music Information Retrieval (MIR) during past years. There are different types of emotion models MIR researchers use for their studies. Researchers around the world follow different approaches to classify songs based on emotions. From those approaches, machine learning techniques are famous among researchers for classification and training [2].

All humans pose universal emotions like happy, sad or angry, but based on the context or the culture they grew up this music classification might differ from country to country. Even though music information retrieval MIR is a hot research topic among researchers it is rear to find an emotion classification study conducted for the Sri Lankan music. As per the Sri Lankan music experts, it is not advisable to directly apply the findings of other researches (Outside the Sri Lankan context) findings to Sinhala songs since the nature of the Sri Lankan music and the music standards followed by the Sri Lankan musicians are different from world standard [1].

Even though other researchers around the world have gained acceptable results in the field of music emotion recognition, the capability of founding an acceptable approach to classify the Sinhala songs based on emotions is problematic. Through this study the author has taken a considerable effort to address the above mentioned problem using machine learning algorithms.

1.2 Motivation

Music Information Retrieval (MIR) is a hot topic around the world. This field is widely studied in psychology, signal processing, neuroscience, musicology, machine learning and in many other domains [4].However, only a limited number of studies have been done for the Sri Lankan music. Sri Lankan music is different from that of other countries, especially compared with western music [1]. According to the Sri Lankan musicians, they cannot directly adopt the findings of published studies to Sri Lankan music [1].

It was identified that there is a gap in Music Information Retrieval (MIR) in the Sri Lankan context. As a step of filling that gap, the present study attempted to address one question in the category of Music Emotion Retrieval (MER) . During the past 10 years, Music Emotions Recognition has attracted an increasing attention to the field of MIR. Although MIR and music emotions are widely studied worldwide it is still at an early stage. When the Sri Lankan context is considered, it is very rare to find studies based on Emotions Recognition.

By listening to a song, human mind categorizes it according to the emotion contained in that song. This emotion classification is based on music and the lyrics of the song. By considering a scenario, the composer receives lyrics of a song. By reading the words of the song, he can understand the emotion that contains. Composers spend a significant time to create a melody for a song. After composing a melody, is there a way to compare it with the emotion implied by lyrics? It is essential for the melody created by the composer to be matched with the emotion of lyrics. For a composer to make sure that the melody created by him matches with the required emotions, it is necessary to have a method for identifying the emotions possessed by the melody. Another scenario is that if we want to automatically recommend a song for a particular person, based on their emotions, we have to capture the emotions and it is essential to classify the songs based on the emotions.

We are living in a digital world where people always seek solutions through digital technologies. This is the same among the song listeners. Typical song listeners have access to thousands of songs locally via their smart phones or any other music players. Simultaneously, song listeners have access to billions of songs across the Internet with the development of technologies like cloud based music services. Due to the wide availability of the songs, the song listeners have a huge collection of songs. Sri Lankan songs (Sinhala songs) are categorized mainly based on the song title or the singer. If a particular person is in a happy mood and if he/she needs to hear a happy song, then he/she has to manually go through each song to find out whether a particular song matches with his/her emotions. It is hard to handle songs manually as it consumes massive time. Therefore, searching for a particular song can be a complicated task.

Above descriptions are few application level requirements that motivated us to continue this study. A well trained automated classification system which classifies the songs based on the emotions will address the above requirement. Apart from the above requirements, this study will address other application level requirements such as music sorting and provide direction for future enhancement of MIR, especially in Sri Lankan Sinhala music context.

1.3 Aims and Objectives

Aim of this study is to identify the possibility of applying machine learning techniques for automatic classification of Sinhala songs based on emotions.

To accomplish this aim present study focused on following objectives,

1. Identify the most suitable emotional model to adopt in the context of Sri Lankan music (Sinhala songs).
2. Identify the most suitable music features to be extracted from songs with respect to Sinhala songs.
3. Identify the most suitable training model to automatically classify Sinhala songs based on emotions.

1.4 Research Questions

The primary research problem addressed in this research is the classification of Sinhala songs based on emotions. Following three research questions were addressed with the aim of a solution to the primary research question;

01. What is the emotional model/emotions that is most suitable to adopt in the context of Sri Lankan music(Sinhala songs)?

Since 1936 several emotional models have been suggested by researchers. These models can be categorized into 2 approaches as categorical and dimensional models of emotion. By analyzing those models, need to find most suitable emotion model/emotions that matches with the context of Sri Lanka [4], [5], [6], [7].

02. The second question, which was necessary to be addressed was, what are the feature selection and the extraction methods and significant features that are relevant to emotion classification with respect to Sinhala songs.

Since this study focused only on Music Information retrieval, lyrics of songs were not considered for this study. According to previous studies, it is clear that it is difficult to obtain a satisfactory result by extracting one feature [7].To continue the research, it was essential to identify the music features that are more relevant to Sri Lankan music. Most importantly, those features should be capable of classifying the emotions that the author selected for the first question.

03. The third question was, what is the classification approach that automatically classifies Sinhala songs based on emotions?

Over time, many machine learning algorithms were applied to study the relationship between music features and emotion labels, such as support vector machine [8] Neural networks [9] and k-nearest neighbor [10]. Here, the author analyzed the applicability of available methods and identified a method that is most suitable for the data set.

1.5 Scope

Since, this study was conducted in the Sri Lankan context, the applicability for other context is not promising. According to author's knowledge, there was no data set that matched the requirements of the present study. This led to the creation of a new data set from scratch. When creating the data set, emotion classification will take more time and labor cost. According to previous studies, there are 2 ways of classify emotions.

1. Classify the emotions of songs using music experts' knowledge [1].
2. Classify the emotions of songs using a survey [4].

From above mentioned types it was decided to use the experts' knowledge in emotion classification as it allows to complete the study within the given time period.

Even though this study selected the Sri Lanka context, this study was limited to Sinhala songs. Author did not consider Tamil songs or songs of any other languages in Sri Lanka.

1.6 Research Contribution

A considerable number of emotion classifications based on songs have been conducted considering the music of the song. Most of them are based on English/western songs. According to the experts in the domain of Sri Lankan music, their findings cannot be applied to Sinhala songs directly, because of the difference between western music and Sri Lankan music. Researchers have also suggested that it is more effective to create models according to the context of target users [5]. But according to author's knowledge, there is only one study that has been carried out in the context of Sri Lankan music related to emotion classification. According to that study, the best accuracy that they have obtained was not adequate to act as an evidence for a good emotion classification system [7]. The objective of the present study was to find out a new approach that is capable of classifying Sinhala songs based on emotions.

1.7 Organization of the Dissertation

Subsequent chapters of the dissertation are organized as follows. The second chapter is about the previous research conducted on music emotion classification. That chapter discusses the different approaches that are used by researchers around the world for music emotion classification especially using machine learning techniques. Then in the same chapter, the author discusses current state of the automatic music emotion classification in Sri Lankan context. That chapter covers both the strengths and weaknesses of those studies.

Third Chapter of the dissertation describes research methodology that the author used to address the research questions. Chapter Four includes an overview of the experimental setup and chapter five has been dedicated to explain the result achieved and discussion based on the result. The concluding chapter (Chapter six) discusses the further improvements and the potential future work.

Chapter 2

Literature Review

2.1 Introduction

This literature review pertains to the research topic “Classification of Sinhala songs based on emotions. The review first explains the different emotion models proposed by several researchers and how music emotion retrieval studies use those emotion models. Then the literature review discusses feature selection and extraction with reference to previous studies. In this section, the author discusses both feature selection and extraction methods. Third section of the review is model training approaches of different studies. Both supervised and unsupervised approaches will be explored in this section. The chapter concludes with a review of different approaches which were used in studies conducted within the Sri Lankan context.

2.2 Music Emotions

Hevner, [3] conducted a study on the psychological relationship between music and emotions and since then, lots of studies have been done on Music and emotions. Music emotions have been widely studied in many areas including psychology, signal processing, neuroscience, musicology and machine learning. Researchers in the field of music emotion have highlighted three common issues suggesting that it is still at an early stage of development.

Those issues are namely; (i). Collection of ground truth data, (ii). Choice of emotion model, and (iii).the relationship between emotion and individual acoustic features. The literate review revealed that previous studies on music information have followed two approaches namely; categorical and dimensional models of emotion [5].

2.2.1Categorical Model

Researchers state that humans experience their emotions as categories. The nature of these categories are that its difference to each other. This approach suggests that there are a limited number of primary emotion classes which are applicable universally. Among those classes are emotions such as happiness, sadness, anger, fear, disgust, and surprise. These primary emotions are prevalent in all cultures and from these secondary emotions can be derived from these primary emotions [2].

Henver [3] discussed about eight clusters of affective adjectives (Fig 2.1). Each cluster contains several emotions with similar characteristics. Henver's model [3] is a circular model, where neighboring clusters varies in a cumulative way [11].This model is called Henver's "mood clock". Farnsworth refined and regrouped this model into 10 groups. He points out that Hevner's clock model does not describe internally consistent mood patterns [12]. The major drawback of the categorical approach to Music Emotion Recognition (MER) is that the small number of primary emotion classes are not sufficient when compared with the richness of music emotion perceived by humans. Moreover, using a large number of emotion classes could overwhelm the subjects, so it is also not considered practical for psychological studies. For example, for the affective terms *calm/peaceful*, *carefree, laid-back/mellow*, *relaxed*, and *soft*, we cannot simply quantify their similarity as zero just because they are different words or as one because they are synonyms. This *ambiguity* and *granularity* issue of emotion description gives rise to the proposal of the dimensional approach to MER [2].

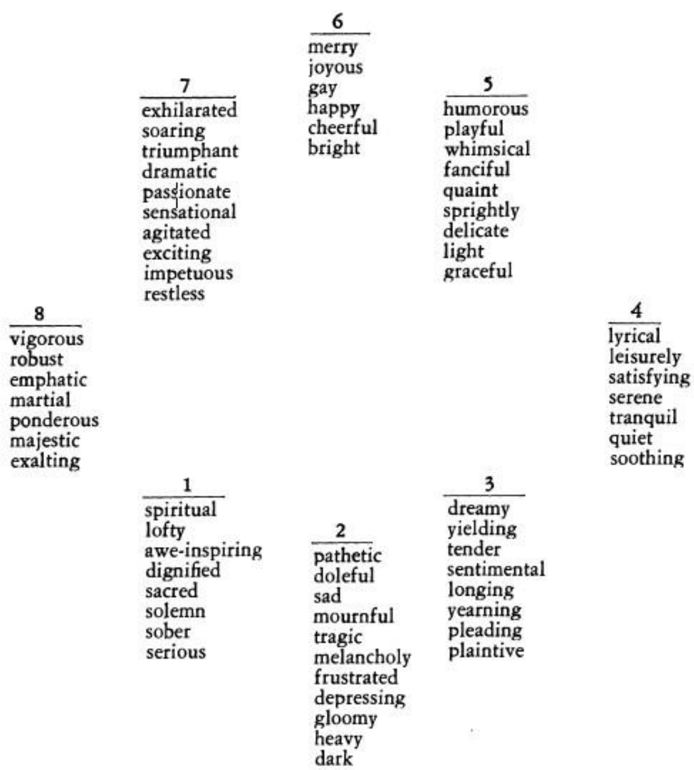

Fig 2.1: Hevner's eight clusters of affective terms

2.2.2 Dimensional Model

Categorical approach always tries to distinguish emotions from one another [5] Russell [13] stated that whilst having a cognitive representation of categories like happy, excited, or sad may be worthwhile, it is imperative to identify the cognitive representation for affect resembles a scientific theory in other ways[13]. Dimensional approach focuses on identifying the emotions based on the placement of emotions. For emotion placement researchers suggest small number of emotion dimensions with name axes. These emotion dimensions have been found through correlation analysis between affective terms [2]. The dimensional model considers all affective terms arising from independent neurophysiologic systems: valence (negative to positive) and arousal (calm to exciting) [4].

Russel's model [13] covers a wide range of situations that include experience of one's own emotional states and judging of another's emotional state (Fig 2.2). In his research Russell [13] uses a number of rating scales of affective terms. Using these scales Russell [13] describes the emotion of music stimulus. Then using factor analysis and correlations between scales, he obtained a small number of fundamental dimensions [13]. These dimensions include emotions: valence (or pleasantness; positive and negative affective states), arousal (or activation; energy and stimulation level), and potency (or dominance; a sense of control or freedom to act) [2]

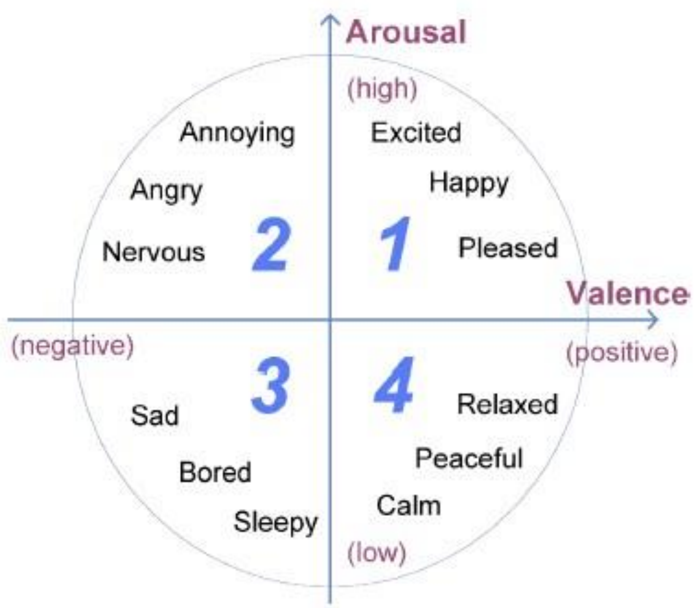

Fig 2.2: The 2D valence-arousal emotion space [Russell 1980]

Through the review of previous studies, it could be revealed that most of the music information retrieval studies did not directly adopt models from either categorical or dimensional approaches. Laurier, Grivolla and Herrera [14] in their study used the emotions; Angry, Happy, Sad and Relax. They have stated that they used a categorical approach to represent the emotions. However, they have highlighted that, with the above mentioned emotions they cover four parts of the 2D representation suggested by Russell [13] with valence and arousal dimensions. Laurier et al. [14] further explained "Happy" and "Relaxed" are positive valence with high and low arousal respectively. "Angry" and "sad" are negative valence with high and low arousal respectively [14]. Wagner, Ki and, Andr´e [15] used a 2D- emotional model Johannes Wagner et al. [15] stated that, there is no common agreement on the basic emotions suggested by theorists, they used the above model targeting four emotion classes, joy, anger, sadness, and pleasure. They argued that, this model provides a more simplified representation of human emotions using two dimensions; arousal and valence [15].

With reference to the Sri Lankan context, author of this study found only one previous study. As per the authors of that study Lakshitha and Jayaratne [1], along with the knowledge of experts in Sri Lankan Music domain, they selected five emotion categories which are most likely to be visible in Sri Lankan Sinhala music. Those are Happy, Excited, Sad, Calm and Heroic. Even though in their study Lakshitha and Jayaratne [1] discussed on emotion models used by others studies, the Lakshitha and Jayaratne [1] study has not stated whether they followed the categorical approach or dimensional approach [1]. However, the applicability of these models into real world scenarios are problematic since all these psychological models have been proposed in laboratory settings where social context is loosely applied [7]. There is still no consensus on which emotional model or how many emotion categories should be used [2]. The present study attempts to identify best emotional model and emotions that are more applicable for the Sri Lankan context.

2.3Music Feature Extraction

Conventionally, MER categorizes emotions in to a number of classes such as happy, angry, sad, and relaxed. Subsequently machine-learning techniques are applied in training a classifier. To represent the acoustic property of a music piece, features such as timbre, rhythm, and harmony are extracted. [2]. Music listening experience is a multidimensional one. The emotion perception of music is associated with different patterns of acoustic cues. Examples such as, while arousal is related to tempo (fast/slow), pitch (high/low), loudness level (high/low), and timbre (bright/soft), valence is related to mode (major/minor) and harmony (consonant/dissonant) [2].

According to the Henver [3] emotion perception is rarely dependent on a single type of music. There is interrelationship, which is prevalent between the judgements of the listeners and mood and musical parameters such as tempo, dynamics, rhythm, timbre, articulation, pitch, mode, tone attacks, and harmony. Early experiments showed that the most important music element for excitement was swift tempo; modality was important for sadness and happiness but useless for excitement and calm; and melody played a very small part in producing a given affective state [5]. As per Wieczorkowska [10] most of the researchers extract features such RMS features, root mean square of amplitude of the signal, zero-crossing rate, loudness, pitch, harmonicity, bandwidth, spectral statistical moments, spectral flux, Mel-frequency cepstral coefficients, linear prediction coefficients and spectral roll-off for their studies.

As cited by Kim, Schmidt, Migneco, Morton ,Richardson,Scott, Speck, and Turnbull [16] MacDorman has examined the capability of applying multiple acoustic features to predict emotions. They found out that when combined all five features could be used to obtain better predictions. Moreover they point out that Eerola et al developed open-source feature extraction code, called MIR tool box which have developed a specific subset of informative audio features for emotion including dynamics, timbre, harmony, register, rhythm, and articulation [16].
Song, Dixon and Pearce [4] used two type of features namely mean and standard deviation with 55 features. They extracted these features using the MIR toolbox. These 55 features had been further categorized into four main features: dynamics, rhythm, spectral, and harmony.

The study of Yang, Lin, Su, and Chen [17] to extract features used the spectral contrast algorithm, DWCH algorithm, and two computer programs called PsySound and Marsyas. They constructed a 114-dimension feature space using these methods. Using PsySounds Yang et al. [17] modeled parameters of auditory sensation based on some psychoacoustic models. They have measured the output of loudness, level, dissonance, and pitch. Loudness measured include loudness, sharpness (sound brightness), and tumbrel width PsySound. As level measures they used sound pressure level, background noise level etc. [17].

Daniel McEnnis, Cory McKay, Ichiro Fujinaga and Philippe Depalle proposed method to extract music features. Which is called Jaudio [18].Researchers around the world used this method to extract music features for their studies. Specially the researchers who conducted MIR researchers for Indian songs achieved high result with this feature extraction method [19]. Using this method 72 low level features, which belong to high level features like rhythm and timber, can be extracted from a song [18].

Feng et al. used another 3 features namely relative tempo, the mean and standard deviation of average silence ratio. They used a classifier called BP neural network for this feature identification [9] In the study of Wieczorkowska [10], he suggests to use signal parameters based on the time domain, spectrum and evolution of sound features. That study used MIDI files to extract features. From audio signal the author has extracted, mean, median, standard deviation, minimum, maximum, and range of voiced pitch signal, pitch for the first and last voiced frame, slope of pitch contour, pitch derivative statistics, speaking rate, i.e. average number of syllables per second, intervals, maximum and minimum pitch position, regression coefficients, and mean square error for regression coefficient and a few more features [10].

When the Sri Lankan context is considered, author of this research found only one study [1]. In that study only the melody was extracted by previous researchers. Lakshika and Jayaratne [1] have concluded that melody in isolation is not capable of differentiating between an array of emotions conveyed by songs. They suggest to use other acoustic features together with melody to get a better result.

2.4 Music Feature Selection

Once music features are extracted next step is feature selection [20]. MIR researchers use different method to select features. Purpose of feature selection is to reduce the noise. The creation of noise can be attributed to the imperfection of technologies that are used to extract data or the inherent limitation of the data itself. Researchers state the presence of a large number of features causes the learning model to over fit resulting in performance degeneration [20].

Based on the data set these feature selection methods can be categorized in to two types. If the training set is labelled supervised feature selection methods can be used [21]. Unsupervised feature selection methods can be used for unlabeled data sets [21].Filter models, wrapper models and embedded models can be taken as three main supervised feature selection methods. Out of these, filter models distinguish feature selection from classifier learning. Due to this the bias of a learning algorithm does not interact with that of the bias of a feature selection algorithm. Within filter models, the literature review suggests that Relief method [23] Fisher score method and Information Gain based method [24], [25] are popular among researchers.

The only research performed in the Sri Lankan context used the InfoGainAttributeEval method for feature selection. But the author of that study has stated that the chosen feature selection method resulted in a low accuracy level. Therefore that study eliminate the feature selection stage and conducted the research with full feature set [1].

2.5 Model Training

The author has identified that, after selecting a suitable emotion model and feature extraction, researchers have tried try deferent model training mechanisms to classify song based on emotions [2]. Over the time, different researchers have used different machine learning algorithms to learn about relationship between music features and emotion labels [2]. This includes both supervised and unsupervised approaches [1], [6].

Researchers have used algorithms such as support vector machines [8], neural networks [9] and k-nearest neighbor [10]. Researchers have different viewpoints in Music information retrieval, and therefore MIR works can be categorized into three approaches. The categorical approach to MER categorizes emotions into a number of discrete classes and applies machine learning techniques to train a classifier. The predicted emotion labels can be incorporated into a text- based or metadata-based music retrieval system. The dimensional approach to MER defines emotions as numerical values over a number of emotion dimensions. A regression model is trained to predict the emotion values that represent the affective content of a song, thereby representing the song as a point in an emotion space [2].

By considering the fact that a song may express more than one emotion, multi label classification algorithms, such as multi label SVM and MLkNN , have been applied to assign multiple emotion labels to a song. MotivatedF states that the emotion perception is influenced by personal factors such as cultural background, generation, sex, and personality [2]. Yang, proposed a fuzzy approach that measures the strength of each emotion in association with the song under the classification and provides a more objective measurement of emotion [26]. However, according to researchers, human emotional response to music depends on the interplay between musical, personal, and situational factors [2]. They propose in developing music mood classification techniques for today's music and users, MIR researchers should extend classical mood models according to the context of targeted users and music listening reality [5].

As per Lakshitha and Jayaratne [1] have attempted utilizing different classification algorithms together with different classifier combination methodologies. The best accuracy level attained by (authors) had been inadequate to recognize as an effective emotion classification system. This has led to their conclusion that melody in isolation is incapable in differentiating emotions conveyed by songs [1].

2.6 Chapter Summary

The above literature review highlights the following key points. In Music information retrieval a predominant number of previous studies follow the same process. It could be observed that they select the emotions based on their research requirement. Then music features were selected and extract those features were extracted using different methods. Finally, train the selected model was trained using selected emotions and extracted features. Feature selection depends on emotion selection while the training model depends on the extracted features. Previous studies around the world state that they have achieved significant results by following this process.

However, previous studies pertaining the Sri Lankan context state that they had not been able to achieve a significant result in classifying Sinhala songs based on emotions. Therefore the present study focuses on finding appropriate emotions related to the Sri Lankan context. Selecting and extracting suitable features from Sri Lankan Context. Finally, this study will propose a training model that is applicable for the Sri Lankan context.

Chapter 3

Methodology

3.1 Introduction

This chapter of the thesis presents the author's approach to address the research problem described in the first chapter. Objective of this chapter is to elaborate the research methodology of this study, which would ultimately lead to the solution for identified research problem. This chapter is inclusive of Fundamental approach key, Design considerations and the overall architecture diagram of proposed methodology.

3.2 Fundamental Approach

Since the ultimate objective of the study is to address the research problem specified in the introduction chapter, it is necessary to address the research questions. Methodology devises a new approach to classify Sinhala songs based on emotions and evaluate the capability of different supervised learning methods. As the initial step, proposed emotion models were evaluated. Based on comments by pervious researchers and with the support of music experts the emotion model for the study was finalized and will be explain further in the next chapter. Since there is no data set available for this research based on the selected emotions a new data set had to be created, containing 162 labeled songs. Newly created data set was used to extract music features from the songs.

Based on studies of previous researchers and suggestions of Sri Lankan music experts, features to be extracted were selected. Then with several extraction methods features were extracted and extracted features were saved in a CSV file format. Different feature selection algorithms were used as a dimension reduction tool to remove insignificant features from the data set. That filtered data set was fed to WEKA with a CSV format. Then researchers tried several supervised algorithms and analyzed those algorithms to get a significant classification of Sinhala songs based on their emotions.

3.3 Design Concerns

It was identified as important to consider about several factors and the need to address them in a proper way in order to achieve the goal of this research.

1. The data set: Since there is no dataset available for this study, creating a database with wider representation of Sinhala songs is one key concern.
2. Quality of the selected songs is another concern. Each song can be in different file formats. Amount of features extracted and the quality depend on the file format.
3. Select emotions that could be identified in Sinhala songs.
4. Select Features to be extracted from the songs. Some features might not be available in Sinhala songs. Therefore, selecting relevant features is important.
5. Identify the best classifier for Sinhala Songs

3.4 Overall Architecture

Figure 3.1 presents the overall architecture diagram of proposed methodology. This architecture has been modeled to address the research questions effectively.

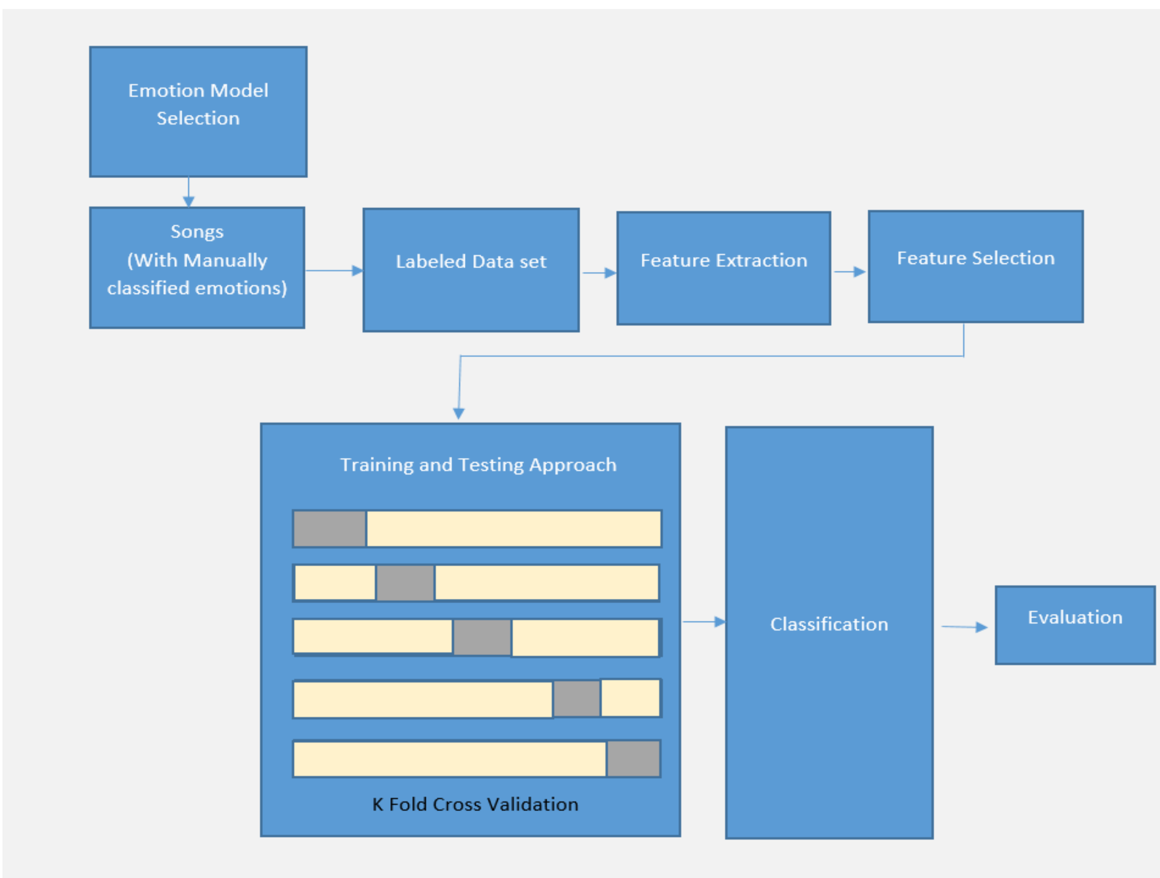

Fig 3.1: Overall architecture diagram of the methodology

3.5. Chapter Summary

This chapter was dedicated to provide an idea of the fundamental approach used by the researcher of this study. Furthermore, this chapter covers Design Concerns and the Overall Architecture of this study. Since this chapter provide an abstract level idea of this study, the steps for deploying this architecture would be broadly elaborated in the next Chapter.

Chapter 4

Research Design

4.1 Introduction

The pervious chapter covered the approach or the methodology we used in this study. Research would continue according to the architectural diagram presented in section 3.4.This chapter is dedicated to describe the experimental set up used by the author to follow the steps in proposed methodology. This chapter covers the methods, tools and other resources used in the study for completion of subsequent phases. Entire experiment was conducted using freely available tools in combination with the self-implemented methods.

4.2 Emotion Model Selection

MIR researchers use different emotions for their studies. Most of the time selection of emotions is based on available data set. However, the researchers who used their own data sets selected emotions with the help of emotion models. These emotion models are suggested by the researchers who engaged in human and emotion studies. Different emotion models were analyzed to identify the most appropriate emotions for the present study. Since these emotions needed to be extracted from Sinhala songs it is essential to match these emotions for the Sri Lankan context.

During the survey of the literature it was observed that most of the studies on music information retrieval have used Happy and Sad emotions [14], [2], [15].Happy and Sad emotions are primary emotions that are included in both categorical and dimensional emotion models. The study conducted by Lakshitha and Jayaratne which is the only study found in this field in the Sri Lankan context, and it has used Happy and Sad as emotions. Other than those, Lakshitha and Jayaratne have used other emotions such as calm, excited and heroic. However, without identifying distinct emotions, we decided to use a well established emotional model. After analyzing both categorical and dimensional emotion models with the expert knowledge author decided to select emotions from Russel’s 2D valence-arousal emotion space (Fig 2.2).Author used several methods to identify most suitable emotion model and emotions, which will be described in detail later in this chapter.

4.2.1 Methods used to finalize the emotions

Following are the methods author used to finalize the emotions for the study.

Previous Studies

The present study identified that Russel's 2D valence – arousal is used quite frequently by researchers in the in the field of MIR. This model covers a wide range of emotions. Instead of using the entire model, researchers decided to use selected emotions from this model. It was observed that there were many successful attempts for song classification based on emotions with this Russel's model [13]. However, as per the author's knowledge this is the first attempt of using Russel's model on Sinhala songs.

Domain Expert Knowledge

Musicians in the Sinhala domain can be considered as the experts who have a clear idea about the Sinhala songs. To identify the most suitable emotions for Sri Lankan context we used the expertise of such musicians. We used their insight to select Russel's 2D valence – arousal emotion model [13]. Russel's 2D valence – arousal emotion model introduces the 12 emotions of humans. It was clear that we could not classify songs according to all these emotions due to several constrains.

1. For a classification, it is essential to have significant amount of songs for each category. Since author does not have a freely available data set and has to create it from scratch creating a data set that is, sufficient to classify 12 emotions is not practical because of time constrains.
2. It is impossible to guaranty that the finding a significant amount of Sinhala songs for each emotion is possible.
3. Since there is no successful attempt in Sinhala song classification based on emotions author's ultimate target of this study is to classify Sinhala songs for at least a few emotions.

With the guidance of the experts at the first stage author decided to use entire range of each category of the Russel's 2D valence – arousal emotion model. Then label the range with a unique code.

High Arousal and Positive Valence - HP

Low Arousal and Negative Valence - LN

Low Arousal and Positive Valence – LP

From the High arousal and Negative valence category it was not practical to select any emotion. Experts state that it is impossible to find a significant amount of Sinhala songs for this category. At the second stage of this study, author focused on classifying emotions separately instead of the emotion range. Therefore two emotions from each category were selected for classification.

Following are the emotions selected for each category.
High Arousal and Positive Valence - Excited and Happy
Low Arousal and Negative Valence - Sad and Sleepy
Low Arousal and Positive Valence - Relax and Calm

4.3 Data Set

Since the data has a high impact on the outcome of this study the author identified, that a proper data set is a high priority of the research. Data set plays a major role in any study because in order to produce high accuracy it is essential to have a rich data set. This is a common factor regardless of the study field or scope. In a classification system, it is necessary to have an adequate amount of data. Since the present study is on song classification, it is clear that having adequate amount of music files for each category is important because classifier needs to be trained using them.

This data should labeled. On the other hand, these data should provide ground truth. Since the emotion is a subjective factor, there is no universal ground truth. Because of that, researchers describe that creating a ground truth data set for music emotions have a serious practical problem. Author of this study identified that it is essential to create a new data set for this study. Since studies on song classification based on emotions are less in the Sri Lankan context and the only available data set does not match with the requirements of the present study. Due to the above mentioned reason, a fresh data set was created from scratch.

It was identified that it is essential to find an answer to the first research question before the implementation of the data set. The author first focused on the first research question and finalized an emotion model that is called Russel’s 2D valence – arousal emotion model. 200 Sinhala songs were selected randomly. Then the music expert labeled those 200 songs based on the Russel’s 2D valence – arousal emotion model.

Following are the results of manual labeling,
High Arousal and Positive Valence - HP (Excited - 29 songs,Happy-26 songs, Pleased- 6 songs)
Low Arousal and Negative Valence - LN (Sad - 28 songs, Bored- 4 songs, Sleepy- 24 songs)
Low Arousal and Positive Valence - LP(Relax- 27 songs, Peaceful-9 songs, Calm-28 songs)

From the randomly selected 200 songs, music experts state that 20 songs were not suitable for any category of the model. Therefore to be in line with the Russel's 2D valence – arousal emotion model we decided to remove those 20 songs from the data set.

Author identified that some emotions do not have significant amount of songs to train a classifier. Therefore it was decided to remove following songs and emotions from the data set
High Arousal and Positive Valence - HP (Pleased- 6 songs)
Low Arousal and Negative Valence - LN (Bored- 4 songs)
Low Arousal and Positive Valence - LP (Peaceful-9 songs)
With the above mentioned changes refined data set was as Table 4.1

Description	Label	Number of songs
High Arousal and Positive Valence	HP	55
Low Arousal and Negative Valence	LN	52
Low Arousal and Positive Valence	LP	55
Total		162

Table 4.1:First data set details

Here both excited and happy represent the High Arousal and Positive Valence emotions therefore it was decided to label both emotions with the code 'HP' which contains 55 songs.

52 songs belong to sad and sleepy emotions, which represent the Low Arousal and Negative Valence category. We labeled it with code 'LN'

Low Arousal and Positive Valence category contain 55 songs which has both relax and calm emotions and labeled with code ‘LP’
All together, we created a data set with 162 Sinhala songs.

At the second stage of the study we again labeled the same song set based on a specific emotion. Here the target was to examine the capability of classifying the songs based on a specific emotion. Author decided to implement this second stage of the study because of the significant accuracy achieved at the first stage (accuracy details will be discussed in next chapter)
Table 4.2 is the dataset we created for the second stage.

Emotion	**Label**	**Song**
Excited	E	29
Happy	H	26
Sad	S	28
Sleepy	SL	24
Relax	R	27
Calm	C	28
Total		**162**

Table 4.2: Second data set details

For this task, the author received help from Sri Lankan music experts. They provided the songs and these songs were categorized using their expertise. These songs were in WAV file format. The first ninety (90) seconds of each song was identified and extracted without making any changes to the original file format WAV.

Duration of songs was decided as 90 seconds because the only previous study conducted in the Sri Lankan context [1] has considered that 90 sec is the adequate length for Sinhala songs as it covers the chorus, intermediate music between the chorus and verse, and a verse [1]. Music experts also agree with this 90 sec length. Author decided to keep the WAW audio file format as it is because it is a lossless file format that allows more data in the audio clip.

4.4 Music Feature Extraction

To train the classifier it is essential to identify music features that are unique to each emotion category. This section of the study focuses on finding the answer for the second research question of the study. That is to identify what are most relevant music features that need to be extracted from the Sinhala songs and identify methods to extract these features. Therefore, the next step of this study was extracting features from each song.

Music Feature extraction is also done in two stages, for the first stage author used the following dataset containing 162 songs with data labeled according to the emotion rangers. Which contained 162 Sinhala songs.

High Arousal and Positive Valence – HP 55

Low Arousal and Negative Valence - LN 52

Low Arousal and Positive Valence – LP 55

Then at the second stage, we used the following data set with labeled emotions

High Arousal and Positive Valence - HP (Excited - 29 songs, Happy-26 songs)

Low Arousal and Negative Valence - LN (Sad - 28 songs, Sleepy- 24 songs)

Low Arousal and Positive Valence - LP(Relax- 27 songs, Calm-28 songs)

MIR researchers extracted different types of features to train their classifiers. Author of the present study focused on the feature extraction method proposed by Daniel McEnnis et al. [18].In their study they proposed a framework to extract features of songs. They state their framework is capable to handle multidimensional features and handle dependencies among features. This dependencies handling helps to avoid duplication of calculations [18].

Based on the result attained by the previous studies and the music experts' ideas we examined following music features to be extracted [5], [2], [10], [16].

4.4.1 Feature Description

1. Zero Crossing

   The extraction technique used in the feature is to count the number of times it crosses the zero within a given window [18], [27].

   Following is the definition of zero crossing.

(xn−1< 0 and xn> 0) or (xn−1 > 0 and xn< 0) or (xn−1 6= 0 and xn = 0) .

2. RMS

   RMS is identified as a good measure of the power of a signal and it extracts the Root Mean Square from a set of samples. The calculation technique involves taking the summation of the squares of each sample, dividing this by the number of samples in the window, and finding the square root of the result [18] [28].

3. Spectral Centroid

   Spectral Centroid can be identified as a measure of the “Centre of Mass” of the power spectrum which is derived by calculating the mean bin of the power spectrum. The result which is a number between 0 to 1 provides a representation on the fraction of the total number of bins [18] [29].

4. Spectral Rolloff Point

   Spectral Rolloff Point identifies the amount of the right-skewedness of the power spectrum as the measure. The spectral Rolloff Point is identified as the fraction of bins in the power spectrum at which 85% of the power is at lower frequencies [18] [29].

5. Spectral Flux

   Spectral Flux is considered as a good measure of the amount of spectral change of a signal. The extraction technique is to extract the spectral flux from a window of samples and the preceding window. To spectral flux, first the difference between the current value of each magnitude spectrum bin in the current window from the corresponding value of the magnitude spectrum of the previous window should be calculated. Then these differences are squared and the result is derived by addition of the squared values [18] [29].

6. Compactness

   Compactness is related to Spectral Smoothness as per the definition by McAdams [27]. Rather than summing over partials, compactness sums over frequency bins of an FFT, which is the difference. This provides an indication of the noisiness of the signal [18].

7. Fraction Of Low Energy window

Fraction of Low Energy window is calculated by taking the mean value of the RMS of the last 100 windows and to find what fraction of these 100 windows are below the mean. This extracts the Fraction of Low Energy Windows form one window to another. Hence this is a good measure of how much of a signal is quiet relative to the rest of a signal [18].

8. Strongest Beat

   Strongest Beat is a technique to extract the strongest beat in a signal, in beats per minute. This is identified by finding the highest bin in the beat histogram [18].

9. Beat Sum

   Beat Sum is a measure to identify how important a role regular beats play in a music piece. This is identified using the sum of all values in the beat histogram [18].

10. Strength Of Strongest Beat

    This is a measure to identify the strength of the strongest beat in comparison to other possible beats. The strength of strongest beat is calculated by finding the entry in the beat histogram corresponding to the strongest beat and dividing it by the sum of all entries in the beat histogram [18].

11. LPC

    LPC calculates the linear predictive coefficients of a signal [18].

12. Method of Moments

    This is a feature which consists of the first five statistical moments on the spectrograph. There are several important aspects in this namely; the area (zeroth order), mean (first order), Power Spectrum Density (second order), Spectral Skew (third order), and Spectral Kurtosis (fourth order) [18].

13. Area Method of Moments

    Area Method of Moments considers a series of frames of spectral data as a two dimensional image. Then it will be analyzed using two-dimensional method of moments (Fujinaga1997). This is capable of giving a description of the spectrograph, including its changes, over a relatively short time frame [18].

14. Mel-frequency cepstral coefficients (MFCC)

    This is a leading feature used for speech recognition. The Mel frequency cepstrum (MFC) is a representation of the short term power spectrum of a sound, based on a linear cosine transform of a log power spectrum on a nonlinear mel scale of frequency. The MFC has been proven to approximate the human auditory system's response more closely than the normal cepstrum, and MFCCs are coefficients that collectively make up an MFC. This is useful for describing a spectrum window [18] [29].

From above features, overall average and standard deviation was calculated as suggested by Daniel McEnnis at el. [18].Using the overall average and standard deviation 72 features were extracted from each and every song. Table No 4.3 explain those 72 features.

Extracted Feature Name	Number of Extracted Features per song
Spectral Centroid Overall Standard Deviation	1
Spectral Rolloff Point Overall Standard Deviation	1
Spectral Flux Overall Standard Deviation	1
Compactness Overall Standard Deviation	1
Spectral Variability Overall Standard Deviation	1
Root Mean Square Overall Standard Deviation	1
Fraction Of Low Energy Windows Overall Standard Deviation	1
Zero Crossings Overall Standard Deviation	1
Strongest Beat Overall Standard Deviation	1
Beat Sum Overall Standard Deviation	1
Strength Of Strongest Beat Overall Standard Deviation	1
LPC Overall Standard Deviation	10
Method of Moments Overall Standard Deviation	5

Area Method of Moments of MFCCs Overall Standard Deviation	10
Spectral Centroid Overall Average	1
Spectral Rolloff Point Overall Average	1
Spectral Flux Overall Average	1
Compactness Overall Average	1
Spectral Variability Overall Average	1
Root Mean Square Overall Average	1
Fraction Of Low Energy Windows Overall Average	1
Zero Crossings Overall Average	1
Strongest Beat Overall Average	1
Beat Sum Overall Average	1
Strength Of Strongest Beat Overall Average	1
LPC Overall Average	10
Method of Moments Overall Average	5
Area Method of Moments of MFCCs Overall Average	10
Total Number of Extracted Features per song	72

Table 4.3: Extracted features from Sinha songs

4.5 Music Feature Selection

Generally these type of studies conducted are characterized with high level of noise. There are several reason for this. Two of the major reasons are imperfections in the technologies used to collect data and the source of the data itself. Dimensionality reduction is one of most popular techniques that is being used to remove noise and redundant features. There are two main categories of dimensionality reduction namely, feature extraction and feature selection. The feature selection approach is capable of selecting a small subset of features, which results in minimizing redundancy and maximizing relevance to the target such as the class labels in classification. On the other hand representative feature selection techniques include techniques such as Information Gain, Relief, Fisher Score and Lasso.

Some approaches, project features into a new feature space with lower dimensionality and the newly constructed features can usually be identified as combinations of original features. Principle Component Analysis, Linear Discriminant Analysis (LDA) and Canonical Correlation Analysis (CCA) are some of the examples for that kind of techniques. Author of this study identified these kind of methods as not suitable for present study because they pose poses a problematic situation for further analysis of new features, because there is no physical meaning for the transformed features obtained from those methods.

For the present study author used supervised feature selection techniques. Supervised feature selection techniques were utilized for the purpose of the present study. They can be classified in to three areas namely; filter models, wrapper models and embedded models. The filter model separates feature selection from classifier learning. This results in bias of a learning algorithm not interacting with the bias of a feature selection algorithm. Supervised feature selection depends on general characteristics of the training data set. These characteristics are distance, consistency, dependency, information and correlation [25] are among the most representative algorithms of the filter model [20]. MIR researchers have mentioned that it is difficult to identify the most suitable feature selection method for a certain data set without checking the result of feature selection algorithm [20]. Present study investigates the Filter methods to select features. They are the Relief method [23] information gain based method [24], [25] and Correlation-based Feature Selection [31].

4.5.1 Feature selection methods

Relief based selection: This evaluates the worth of an attribute by repeatedly sampling an instance and taking the value of the given attribute for the closest instance of the same and different class [23].

Correlation-based selection: This method calculates the correlation between each attribute and the output variable. Only those attributes that have a moderate to high positive or negative correlation will be selected and attributes with a low correlation (value close to zero) will be drooped [31].

Information gain based selection: In this method the attributes that contribute more information will be selected as they have a higher information gain value. Attributes that do not add much information will have lower score and can be removed [26].

4.5.2 Feature selection using Weka

In the Weka environment, feature selection is divided into two parts;

- Attribute Evaluator
- Search Method

The attribute evaluator is the technique where each attribute in the dataset is evaluated in the context of the output variable. On the other hand search method is the technique used to try or navigate different combinations of attributes in the dataset in order to derive at a short list of chosen features. Some attribute evaluator techniques require the use of specific search methods. For an example; CorrelationAttributeEval technique mentioned in the next section of this study can only be complemented with a Ranker Search Method which evaluates each attribute and lists the results in a rank order.

For the feature selection table 4.4 attribute evaluators and search methods were used.

Attribute Evaluator	**Search Method**
ReliefF based selection	Ranker
Correlation based selection	Ranker
Information Gain based selection	Ranker

Table 4.4: Features selection methods used for the present study

All 72 features were ranked using the table 4.4 methods. Each of the Attribute Evaluator gave a rank list which is different from each other. These ranks were used in selecting the most significant features to train the classifier. (Next chapter 5.3 describes the selected features)

4.6 Model Training

It was identified that after selecting suitable features researchers try different model training mechanisms to classify songs based on emotions. Over the time, these researchers used different approaches. This study conducted the model training in 2 phases. At the first stage, it used the dataset, which was labeled according to the arousal and valence values. (Table 4.1) At the second stage it used the dataset with specific emotions which created as the second dataset (Table 4.2). Based on the viewpoints of researchers MIR projects can be catergorized in to three approaches. The categorical approach to MER characterized with categorizing emotions into a number of discrete classes and applying machine-learnig techniques to train a classifier. The predicted emotion labels can be linked in to a text-based or metadata-based music retrieval system. The dimensional approach defines emotions as numerical values over a number of emotion dimensions. A regression model is trained for the purpose of predicting emotion values that represent the affective content of a song, to represent the song as a point in an emotional space [2].

Present study used the approach that categorize emotions in to number of discrete classes and applies machine-learning techniques to train a classifier. For this classification, researchers used several algorithms. Since the data set is labeled and follow supervised learning approach, author focused only on supervised learning algorithms.

4.6.1 Algorithms used to train the classifier

SMO

In 1998, John C. Platt proposed the SMO algorithm which became the fastest quadratic programming optimization algorithm, especially for linear SVM and sparse data performance. According to John C. Platt, for the purpose of training a support vector machine, it requires the solution of a very large quadratic programming (QP) optimization problem. SMO breaks a large QP problem into a series of smallest possible QP problems. These small QP problems are solved analytically, which avoids using a numerical QP optimization as an inner loop which is a time consuming. Time taken for SMO’s computation is dominated by SVM evaluation; hence SMO is fastest for linear SVMs and sparse data sets. On real-world sparse data sets, SMO can be more than 1000 times faster than the chunking algorithm [32].

Random Forests

Random forests are an ensemble learning method for classification, regression and other tasks that operates by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. Random decision forests correct form decision trees’ habit of over fitting to their training set [33].

Naïve Bayes

This is a technique which greatly simplifies the learning by assuming that features are independent given class. In practice Naïve bayes often competes well with more sophisticated classifiers. Naïve bayes has proved to be effective in many practical applications. Naïve bayes has shown to be optimal for some important classes of concepts that have a high degree of feature dependencies [34].

4.6.2 Model Training with selected algorithms.

Using the mentioned data set, feature selection methods and classification algorithms in this chapter author conducted a series of experiments. These experiments are mainly divided in to two phases based on the data set. Inside each phase, experiments are subdivided based on feature selection methods and classification algorithms

4.6.3 First phase of the model training

For the first phase author used the first data set which contain arousal and valence values (Table 4.1) Using that data set author conducted table 4.5 experiment series.

Experiment Number	Feature selection method	Classification Algorithm
Experiment 1.1.2	ReliefF based selection	SMO
Experiment 1.1.3	ReliefF based selection	Random Forests
Experiment 1.1.4	ReliefF based selection	Naïve Bayes
Experiment 1.2.2	Correlation based selection	SMO
Experiment 1.2.3	Correlation based selection	Random Forests
Experiment 1.2.4	Correlation based selection	Naïve Bayes
Experiment 1.3.2	Information Gain based method	SMO

Experiment 1.3.3	Information Gain based method	Random Forests
Experiment 1.3.4	Information Gain based method	Naïve Bayes

Table 4.5: Experiment list with first data set

4.6.4 Second phase of the model training

For the second phase author used the second data set which contain 6 emotions (Table 4.2)Using that data set author conducted Table 4.6 experiment series.

Experiment Number	Feature selection method	Classification Algorithm
Experiment 2.1.1	ReliefF based selection	SMO
Experiment 2.1.2	ReliefF based selection	Random Forests
Experiment 2.1.3	ReliefF based selection	Naïve Bayes
Experiment 2.2.1	Correlation based selection	SMO
Experiment 2.2.2	Correlation based selection	Random Forests
Experiment 2.2.3	Correlation based selection	Naïve Bayes
Experiment 2.3.1	Information Gain based method	SMO
Experiment 2.3.2	Information Gain based method	Random Forests
Experiment 2.3.3	Information Gain based method	Naïve Bayes

Table 4.6:Experiment list with second data set

By conducting table 4.6 experiment, author analysis the accuracy of each combination. These experiments were done in the Weka environment. Accuracy level and the result of the confusion matrix were used to evaluate the result. This will be explained in the next chapter. Result of the above experiments were used to propose an approach that can classify Sinhala songs based on the emotion with a high level of accuracy.

4.7 Chapter Summery

Design chapter of the thesis covered the experimental set up used in the study. It elaborates each step of the classification. Emotion model selection covered the methods author used to finalize an emotion model. Then the next section explains the data set implementation process. Feature extraction section described techniques that are used to extract data and the type of features extracted from the songs. Feature selection section was dedicated to elaborate the feature selection models used in the study and how it is implemented in Weka environment. Final section of the chapter explained the algorithms used to train the classifier and the series of experiments conducted to find the best accuracy combination.

Chapter 5

Results and Discussion

5.1 Introduction

This chapter presents the evaluation and discussion of the findings of the present study. Chapter consists of main sections namely evaluation of results and discussion based on results. All the experiments relating to the study were carried out in a Weka environment. Here the entire experiment was done in two phases. In the first phase author used the data set, which labeled according to Arousal and valance ranges. During the second phase data set was labeled according to the emotions of songs.

5.2 Music Feature selection result

At the music feature extraction stage 72 features were extracted from each Sinhala song. Those extracted data were fed to Weka environment for feature selection. Feature selection was done in two phases,

Phase 1: feature selection of labeled arousal and valance data set (First data set)

Phase 2: feature selection with labeled emotion data set (Second data set)

First and second data set, which contained the songs with labeled emotions were fed to Weka. Appendix A to F shows the result based on feature selection method.

In both data sets author noticed top ranking features are almost common to all three feature selection methods. Their ranking positions were different from one method to the other. However, to derive at a conclusion on significant features it was essential to check the accuracy with machine learning algorithms which will be explained later in this chapter.

5.3 Model Training result

K-fold cross validation method was used to evaluate the results of the present study. Here the emotion classification can be identified as a single labeled classification. Because it was assumed that each song can covey only one emotion from the selected set of emotions.

5.3.1 K-fold cross validation

K-fold cross validation is a popular validation method especially in classification studies.

This method has a single parameter call "K". K refers to the number of groups that given a sample will be split into. Author used the k values as 10 for this study. Therefore in this study data were split in to 10 equal sized sets. Then the first set was used as the test set and remaining 9 sets were used to train the model. In the K-fold cross validation this process is repeated until each set is used as a test set. This evaluation is recommended for the limited sample data set because the advantage is every data point gets to be in a test set exactly one time and every data point is in the training set 9 times.Weka Environment was used to do these experiments as mentioned before. Next section of the chapter explains the gain result with the confusion Matrix for each experiment.

5.3.2 Model training experiments and results with first data set

In the first phase of experiment (with first Data set) nine experiments were conducted combining feature selection methods and supervised training algorithms. To find the highest accuracy with most significant features of each combination, feature selection was varied. Therefore following results and confusion matrix shows only the highest accuracy levels of each combination.

Experiment 1.1.1: This experiment used the ReliefF based selection ranking result to train the classifier with SMO algorithm (Fig 5.1).Following is the highest accuracy gained with respect to these 2combinations.With all features, classification achieved 87.037 %accuracy, with top 17 features it increased the accuracy until 90.7407 %.That is the highest accuracy gained in the combination of ReliefF feature selection and SMO algorithm.

a = HP, b = LP, c = LN

a	b	c	
51	4	0	a
5	50	0	b
0	6	46	c

Fig 5.1:Confusion matrix of Experiment 1.1.1

Experiment 1.1.2: The Combination of ReliefF based selection and Random forests algorithm was used in this experiment (Fig 5.2). With all 72 features it achieved 95.679 %.with the reduction of features the values changed but not exceed 95.679 % .However the same 95.679 % accuracy was gained with the top ranked 15 features which was the highest accuracy author achieved for this combination.

a = HP, b = LP, c = LN

a	b	c	
54	1	0	a
4	50	1	b
0	1	51	c

Fig 5.2:Confusion matrix of Experiment 1.1.2

Experiment 1.1.3: ReliefF based feature selection was used with Naïve bayes algorithm in this experiment (Fig 5.3). With all 72 features, accuracy was 81.4815 % .With the top 45 features it was increased to 85.1852 % .That was the highest result author gained for this combination. With top 20 features, accuracy was 74.6914 %.

a = HP, b = LP, c = LN

a	b	c	
46	9	0	a
9	43	3	b
2	1	49	c

Fig 5.3: Confusion matrix of Experiment 1.1.3

Experiment1.2.1: Correlation based selection method was used as the feature selection method and SMO algorithm was used to train the classifier (Fig 5.4) .With all extracted 72 features the accuracy was 87.037 %.However using the Correlation based selection method rank author reduced the features and found highest accuracy as 90.7407% which was gained using the top 24 features.

a = HP, b = LP, c = LN

a	b	c	
51	4	0	a
4	51	0	b
0	7	45	c

Fig 5.4: Confusion matrix of Experiment 1.2.1

Experiment 1.2.2: This experiment used the Correlation based selection method and Random forests algorithm (Fig 5.5). With all 72 features, the accuracy was 95.679 %.It was identified that even with top 19 features this95.679 % accuracy can be achieved. However, this was the highest accuracy obtained for this experiment.

a = HP, b = LP, c = LN

a	b	c	
54	1	0	a
3	51	1	b
0	2	50	c

Fig 5.5: Confusion matrix of Experiment 1.2.2

Experiment 1.2.3: Naïve bayes algorithm was used as the model-training algorithm and Correlation based selection was the feature selection method (Fig 5.6). With all 72 features, the accuracy was 81.4815 %.It increased till 85.1852 % with the top 42 features. That was the highest accuracy for this experiment.

a = HP, b = LP, c = LN

a	b	c	
47	8	0	a
8	44	3	b
1	3	48	c

Fig 5.6: Confusion matrix of Experiment 1.2.3

Experiment 1.3.1: This experiment used the Information Gain based method to select features (Fig 5.7). SMO algorithm was used to train the classifier. With all 72 features, accuracy was 87.037 % and it marked the highest accuracy of 90.1235 % with top 24 features.

a = HP, b = LP, c = LN

a	b	c	
49	6	0	a
3	51	1	b
0	6	46	c

Fig 5.7: Confusion matrix of Experiment 1.3.1

Experiment 1.3.2: Information Gain based method used together with Random forests algorithm in this experiment (Fig 5.8).95.679% was the highest accuracy for this experiment. This 95.679 % of accuracy was gained with the all 72 features. However, author identified that using the top 18 features it is possible to achieve the 95.679 % of accuracy.

a = HP, b = LP, c = LN

a	b	c	
54	1	0	a
3	51	1	b
0	2	50	c

Fig 5.8: Confusion matrix of Experiment 1.3.2

Experiment 1.3.3: For this experiment, Naïve bayes was used as the classifier-training algorithm and Information Gain based method was used to select features (Fig 5.9). With all 72 features accuracy was 81.4815 % it was increased till 84.5679 % with top 45 features. That was the highest accuracy for this experiment.

a = HP, b = LP, c = LN

a	b	c	
47	8	0	a
10	42	3	b
2	2	48	c

Fig 5.9: Confusion matrix of Experiment 1.3.3

The Table 5.1 presents the summary of the results obtained from above experiments.

Experiment Number	Feature selection method	Classification Algorithm	Accuracy With all features	Highest accuracy	Number of features for Highest Accuracy
Experiment 1.1.1	ReliefF based selection	SMO	87.037 %	90.7407%	Top Ranked 17
Experiment 1.1.2	ReliefF based selection	Random Forests	95.679 %	95.679 %	Top Ranked 15
Experiment 1.1.3	ReliefF based selection	Naïve Bayes	81.4815 %	85.1852%	Top Ranked 45
Experiment 1.2.1	Correlation based selection	SMO	87.037 %	90.7407 %	Top Ranked 24
Experiment 1.2.2	Correlation based selection	Random Forests	95.679 %	95.679 %	Top Ranked 19
Experiment 1.2.3	Correlation based selection	Naïve Bayes	81.4815 %	85.8025 %	Top Ranked 42
Experiment 1.3.1	Information Gain based selection	SMO	87.037 %	90.1235%	Top Ranked 24
Experiment 1.3.2	Information Gain based selection	Random Forests	95.679 %	95.679 %	Top Ranked 18
Experiment 1.3.3	Information Gain based selection	Naïve Bayes	81.4815 %	84.5679 %	Top Ranked 45

Table 5.1: Summary of classification accuracies (First data set).

The result gained from the Table 5.1 is graphically represent in figure 5.10.

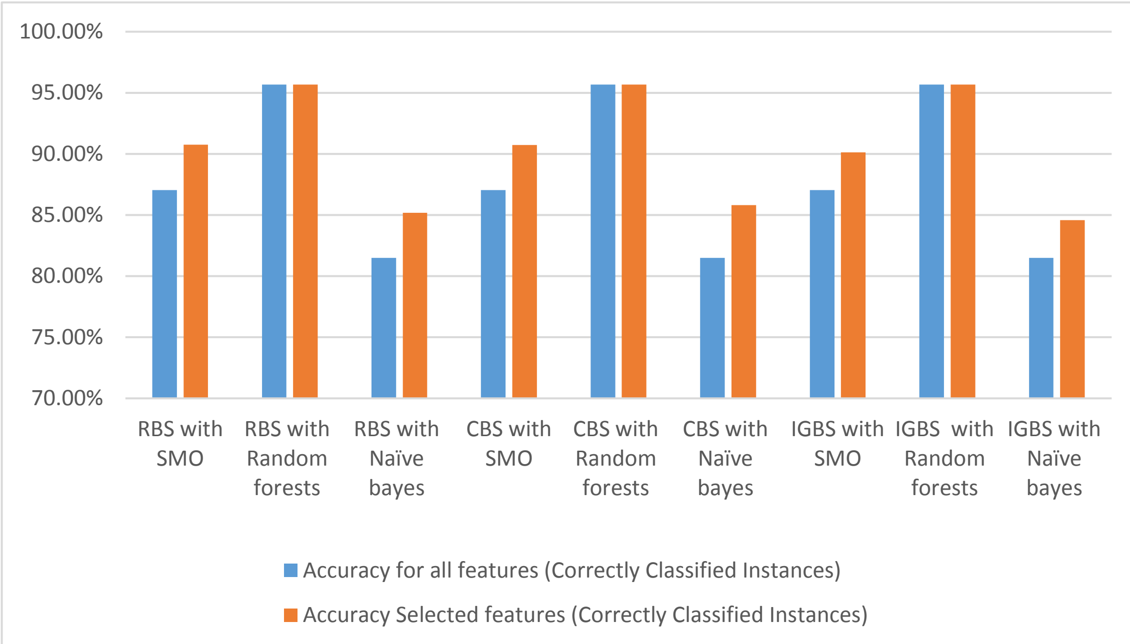

Fig 5.10: Comparison of Emotion classification accuracies (First data set).

5.3.3. Model training experiments and results with second data set

In the second phase of the experiments author used the second data set, which were labeled according to the selected emotions. In this stage, 09 experiments were conducted using three feature selection methods and three supervised learning algorithms. Following are the results and confusion matrix obtained from those experiments.

Experiment 2.1.1: This experiment used the ReliefF based selection as feature selection method and SMO algorithm to train the classifier (Fig 5.11). With all 72 features accuracy was 82.716 %. It was increased till 90.7407 % by selecting up to 22 features.

a=E, b=C, c=H, d= R, e=S f=SL

a	**b**	**c**	**d**	**e**	**f**	
28	**1**	**0**	**0**	**0**	**0**	**a**
2	**26**	**0**	**0**	**0**	**0**	**b**
1	**0**	**25**	**0**	**0**	**0**	**c**
1	**1**	**1**	**24**	**0**	**0**	**d**
0	**2**	**1**	**4**	**20**	**1**	**e**
0	**0**	**0**	**0**	**0**	**24**	**f**

Fig 5.11: Confusion matrix of Experiment 2.1.1

Experiment 2.1.2: This experiment used the ReliefF based selection as feature selection method and Random forests algorithm to train the classifier (Fig 5.12). With all 72 features accuracy was 90.7407 %. It was increased up to 91.9753 % upon selecting 12 top ranked features.

a=E, b=C, c=H, d= R, e=S f=SL

a	b	c	d	e	f	
27	1	1	0	0	0	a
2	25	1	0	0	0	b
1	0	25	0	0	0	c
1	0	1	23	2	0	d
0	2	0	0	25	1	e
0	0	0	0	0	24	f

Fig 5.12: Confusion matrix of Experiment 2.1.2

Experiment 2.1.3: This experiment used the ReliefF based selection as feature selection method and Naïve bayes algorithm to train the classifier (Fig 5.13). With all 72 features, accuracy was 80.2469 %. It was increased up to 88.8889 % upon selecting 12 top ranked features.

a=E, b=C, c=H, d= R, e=S f=SL

a	b	c	d	e	f	
27	2	0	0	0	0	a
3	25	0	0	0	0	b
1	1	21	3	0	0	c
1	1	0	24	1	0	d
0	3	0	1	22	2	e
0	0	0	0	0	24	f

Fig 5.13: Confusion matrix of Experiment 2.1.3

Experiment 2.2.1: This experiment used the Correlation based selection as feature selection method and SMO algorithm to train the classifier (Fig 5.14). With all 72 features, accuracy was 82.716%. It was increased up to 90.1235 % upon selecting 22 top ranked features.

a=E, b=C, c=H, d= R, e=S f=SL

a	b	c	d	e	f	
28	1	0	0	0	0	a
2	26	0	0	0	0	b
0	1	23	2	0	0	c
1	1	1	24	0	0	d
0	2	1	3	21	1	e
0	0	0	0	0	24	f

Fig 5.14: Confusion matrix of Experiment 2.2.1

Experiment 2.2.2: This experiment used the Correlation based selection as feature selection method and Random forests algorithm to train the classifier (Fig 5.15). With all 72 features accuracy was 90.7407 %. It was increased up to 91.9753 % upon selecting 15 top ranked features.

a=E, b=C, c=H, d= R, e=S f=SL

a	b	c	d	e	f	
27	1	1	0	0	0	a
3	24	1	0	0	0	b
1	0	25	0	0	0	c
1	0	1	23	2	0	d
0	0	1	0	26	1	e
0	0	0	0	0	24	f

Fig 5.15: Confusion matrix of Experiment 2.2.2

Experiment 2.2.3: This experiment used the Correlation based selection as feature selection method and Naïve bayes algorithm to train the classifier (Fig 5.16). With all 72 features, accuracy was 80.2469 %. It was increased up to 89.5062 % upon selecting 26 top ranked features.

a=E, b=C, c=H, d= R, e=S f=SL

a	b	c	d	e	f	
26	3	0	0	0	0	a
3	25	0	0	0	0	b
1	0	23	2	0	0	c
1	1	0	25	0	0	d
0	3	0	2	22	1	e
0	0	0	0	0	24	f

Fig 5.16: Confusion matrix of Experiment 2.2.3

Experiment 2.3.1: This experiment used the Information Gain based selection as feature selection method and SMO algorithm to train the classifier (Fig 5.17). With all 72 features, accuracy was 82.716%. It was increased up to 91.358 % upon selecting 16 top ranked features.

a=E, b=C, c=H, d= R, e=S f=SL

a	b	c	d	e	f	
28	1	0	0	0	0	a
2	26	0	0	0	0	b
0	1	25	0	0	0	c
1	1	1	24	0	0	d
1	1	1	3	21	1	e
0	0	0	0	0	24	f

Fig 5.17: Confusion matrix of Experiment 2.3.1

Experiment 2.3.2: This experiment used the Information Gain based selection as feature selection method and Random forests algorithm to train the classifier (Fig 5.18). With all 72 features, accuracy was 90.7407 %. It was increased up to 91.358 % upon selecting 11 top ranked features.

a=E, b=C, c=H, d= R, e=S f=SL

a	b	c	d	e	f	
26	1	1	1	0	0	a
3	24	1	0	0	0	b
1	0	24	1	0	0	c
1	0	1	23	2	0	d
0	2	0	0	25	0	e
0	0	0	0	0	24	f

Fig 5.18: Confusion matrix of Experiment 2.3.2

Experiment 2.3.3: This experiment used the Information Gain based selection as feature selection method and Naïve bayes algorithm to train the classifier (Fig 5.19). With all 72 features, accuracy was 80.2469 %. It was increased up to 88.2716 % upon selecting 23 top ranked features.

a=E, b=C, c=H, d= R, e=S f=SL

a	b	c	D	e	f	
26	3	0	0	0	0	a
3	25	0	0	0	0	b
1	0	23	2	0	0	c
1	1	0	25	0	0	d
0	2	0	4	20	2	e
0	0	0	0	0	24	f

Fig 5.19: Confusion matrix of Experiment 2.3.3

The Table 5.2 presents the summary of experiments conducted in phase two using the second data set.

Experiment Number	Feature selection method	Classification Algorithm	Accuracy With all features	Highest accuracy	Number of features for Highest Accuracy
Experiment 2.1.1	ReliefF based selection	SMO	82.716 %	90.7407 %	Top ranked 22
Experiment 2.1.2	ReliefF based selection	Random Forests	90.7407 %	91.9753 %	Top ranked 12
Experiment 2.1.3	ReliefF based selection	Naïve Bayes	80.2469 %	88.8889 %	Top ranked12
Experiment 2.2.1	Correlation based selection	SMO	82.716 %	90.1235 %	Top ranked 22
Experiment 2.2.2	Correlation based selection	Random Forests	90.7407 %	91.9753%	Top ranked 15
Experiment 2.2.3	Correlation based selection	Naïve Bayes	80.2469 %	89.5062 %	Top ranked 26
Experiment 2.3.1	Information Gain based selection	SMO	82.716 %	91.358 %	Top ranked 16
Experiment 2.3.2	Information Gain based selection	Random Forests	90.7407 %	91.358 %	Top ranked 11

Experiment 2.3.3	Information Gain based selection	Naïve Bayes	80.2469 %	88.2716 %	Top ranked 23

Table 5.2: Summary of classification accuracies (Second data set).

The result gained from the Table 5.2 is graphically represent in figure 5.20.

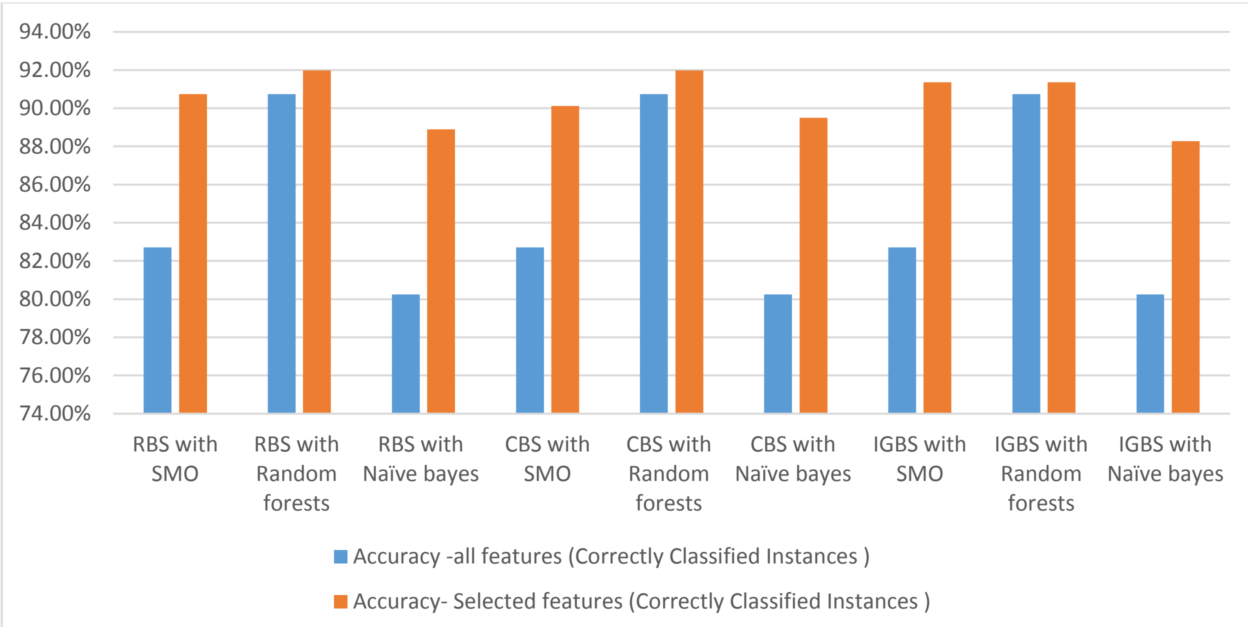

Fig 5.20: Comparison of Emotion classification accuracies (second data set).

5.4 Most Significant features for Sinhala songs/Music

By analyzing the results in table 5.2, 21 most significant features were identified. These features were common to all three feature selection methods. For the common feature selection author used top ranking 26 features in every feature selection method. Because 26 is the highest number of features author used to achieve high accuracy.

	Name of the Feature
1	AMMMFCCOA 6
2	AMMMFCCOA 9
3	MMMFCCOA 10
4	AMMMFCCOA 8
5	AMMMFCCOA 5
6	AMMMFCCOA 2
7	AMMMFCCOSD 4

8	AMMMFCCOA 4
9	AMMMFCCOSD 7
10	AMMMFCCOA 3
11	AMMMFCCOA 7
12	FOLEWOSD
13	AMMMFCCOSD 10
14	AMMMFCCOSD 5
15	AMMMFCCOSD 8
16	AMMMFCCOSD 6
17	AMMMFCCOSD 9
18	COA
19	SFOA
20	MMOA 1
21	MMOSD 1

Table 5.3: Selected top 21 common features methods.

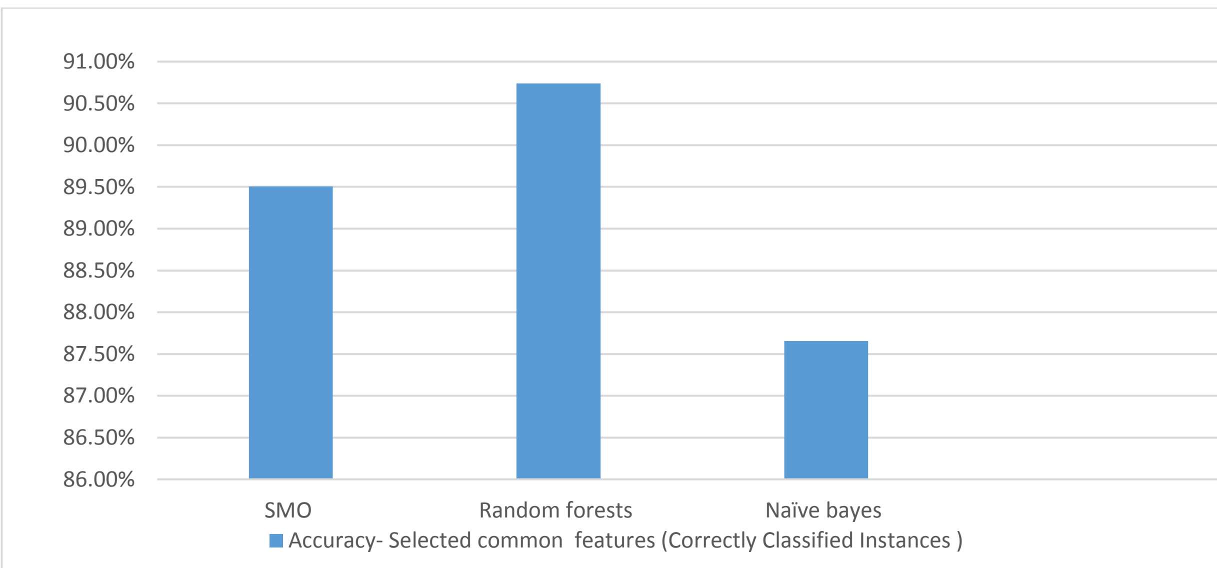

Fig 5.21: Emotion classification accuracies with common 21 features

5.5 Discussion

Present study was focused on finding an approach that can classify Sinhala songs based on emotions elicit. The results obtained in the experiments provide sufficient evidence to suggest that the present study achieved significant results in addressing the research problem.

First question of the study was to identifying the most suitable emotion model/emotions relevant to Sinhala songs. The first phase of the study was conducted using data labeled with arousal and valance instead of specific emotions, because it was identified in previous studies which were conducted with respect to Sri Lankan context was not significant in terms of accuracy levels. Researchers of these studies stated that they were unable to achieve significant results in their studies.

Due to the above situation, present study first focused on achieving a significant accuracy on high-level classification. However, the result or the accuracy level gained was significantly high with arousal and valance labeling. All algorithms that were used to train the classifier resulted in more than 85% of accuracy. It declared the possibility of applying a data set, which is labeled with emotions. Therefore, author used the second data set and conducted experiments in second phase. The accuracy gained from the second data set was more than 88% for all selected algorithms. This provides guidance to all MER researches who expect to conduct studies on Sinhala songs that, it is not necessary to do a higher level classification. With Sinhala songs, it is possible to do more meaningful emotion classification.

The second research question was the selection and the extraction of features that are relevant to the emotion classification with respect to Sinhala songs. After comparing the results of both first and second data set, it was identified that it is unnecessary to analyze the result of first dataset, as the second dataset was capable of providing a significant level of accuracy.

An important observation was the possibility of increasing the accuracy level by reducing the least significant features. As depicted in the figure 5.10 and figure 5.20 classifier was able to improve their classification abilities when the selected feature sets are given as input to them. These features were selected using three feature selection methods (Table 5.3). The aim of the feature selection was to reduce the computational time for model training by reducing the number of features.

According to the observation author came across, from the extracted 72 features most of the features were least significant in feature classification. In Experiment, 2.2.3 author used 26 features to achieve the highest level of accuracy, which was the highest number of features used.

It is important to highlight that in some experiments the mentioned highest levels of accuracy were achieved even with high number of features. Here the author states only the highest accuracy with least features. This observation gives an idea that their might be situations where the combination of some high ranked features and one or two features form least significant features providing better results. However, this study did not focus on that type of combinations. This study was focused on the feature ranks produced by the feature selection methods and these rankings were used to reduce features.

Other observation made by the author was ReliefF based selection, Correlation based selection and Information gained methods top ranked features were almost the same (Appendix D to F) even though the ranking methods were different. It was identified that from the top ranked 26 features of all three methods 21 features were common for these three feature selection methods (Table 5.3).Which were the most significant features that produced the high level of accuracy. The numbering of table (Table 5.3) not indicate any ranking order. These common 21 features were able to produce more than 87% accuracy in all three classification algorithms (Figure 5.21) But those accuracy levels did not exceed the highest accuracy levels of selected features with respect to feature selection methods and supervised learning algorithms (Figure 5.20)

Another highlighted point was from these 21 features, (Table 5.3) some features were instance of one feature. For example these common feature set contain 9 instances of Area Method of Moments of MFCCs Overall Average (AMMMFCCOA) feature. The total number of AMMMFCCOA in the 72 feature set was 10. From that 10 instance 9 were selected as top common features. It gives an indication AMMMFCCOA has a key relationship with Music Emotion recognition in Sinhala songs.

Area Method of Moments of MFCCs Overall Standard Deviation (AMMMFCCOSD) feature demonstrated the second highest contribution for the common feature set as it represented 7 features from 21 common features (Table 5.3).

Here the important aspect is that both AMMMFCCOA and AMMMFCCOSD are the different calculations of Area Method of Moments of MFCC.Since the 16 of 21 features represented the Area Method of Moments of MFCC it is clear that, Area Method of Moments of MFCC can make a huge impact on accuracy of MER.

Mel-frequency cepstral coefficient (MFCC) is a commonly used timbre feature among MIR researchers [2], [14],[17], [35]. This MFCC together with Area Method of Moments created the

AMMMFCC which includes sequence of random variables each independent of the previous ones, mean, Power Spectrum Density, Spectral Skew and Spectral Kurtosis [18].

It was observed that most of the features in common feature set (Table 5.3) represent the "timber" feature in music. These results indicate that timber plays an important role in music perception [30]. MIR researchers agree that timber play major role in larger scale movement of tension and relaxation, thus contribute to the expressions inherent in musical form [30]. The result of the present study is consistent with previous studies.

With the result of the feature extraction and selection author observed that feature extraction methods by Daniel McEnnis et. al. [18] and features selected using the three feature selection methods (Table 4.4) can produce significant accuracy in emotion classification relevant to the Sinhala songs.

Final question of the present study was finding an approach that automatically classifies the Sinhala songs based on emotions. The approach of the present study was classification of emotions using supervised learning algorithms. With the results achieved in this study it can be highlighted that supervised learning algorithms perform well in emotion classification of Sinhala songs.

Random forest algorithm demonstrate highest results with all feature selection methods (Table 5.2)The reason for this high level of accuracy might be the nature of this algorithm, which easily measures the relative importance of each feature on the prediction. This algorithm reduces impurity across all trees in the forest [33]. Random forest algorithm provided the highest accuracy even with the selected common feature set (Fig 5.21).
However, the Sequential minimal optimization (SMO) algorithm and the Naïve bayes algorithm also gave a significant accuracy.

After analyzing results with the common feature set and results with respect to specific feature selection method with supervised algorithm, author came to a conclusion that, specific feature selection method with supervised algorithm gave high accuracy than the common feature set. From all conducted experiments ReliefF based selection with Random forest algorithm gave the highest accuracy of 91.9753 % with 12 features. That is the best result author was able to achieve from this study.

Therefore the observation of author is, that even though most experiments gave the significant accuracy for emotion classification of Sinhala songs ReliefF based selection with Random forest algorithm is the most suitable approach to classify Sinhala songs based on their emotions.

Based on the above discussion it is clear the present study can propose an approach that can classify Sinhala songs based on emotions. For that, the emotions and the features used in the present study can be applied widely. This study explains methods to emotion model selection, feature extraction and feature selection for Sinhala songs, which led to significant levels of accuracy in emotion classification based on supervised machine learning algorithms.

5.5 Chapter Summery

This chapter presented the results gained in different levels through different experiments and the discussion based on those results. Experiment results were presented in different ways to maintain clarity of the study. Discussion section explained significant observations of the author and how it provided the answers for the three research questions. The last section of the chapter explained how results and observations provided an approach that can classify Sinhala songs based on their emotions.

Chapter 6

Conclusion and Future works

6.1 Conclusion

There is a significant amount of research studies conducted in the areas of Music Information Retrieval (MIR) and Music Emotion Retrieval (MER). Given the contextual differences (i.e. subjectivity of emotions) in music, it calls for more and more context based studies specifically on music emotion retrieval. Due to this reason and limited existing knowledge on music emotion retrieval of Sinhala songs, the author identified the need to conduct a study focusing on this. Enhancing the existing level of accuracy in predicting the music information retrieval was also an intention behind conducting the present study. The findings of the study were able to extend the knowledge of music emotion retrieval with a higher level of accuracy compared to previous studies as intended.

For the emotion identification of the present study, Russel's 2D Valence – Arousal Emotion Model was used. In conducting the study it was decided to consider only six emotion categories of the said model. This was imperative due to unavailability of Sinhala songs to suit certain emotion categories. Due to the importance of this decision Sri Lankan music experts were also consulted in the process of selecting emotions as well as songs.

Due to the unavailability of a suitable dataset of Sinhala Songs, the author developed a new dataset of Sinhala songs. This data set was created in two stages namely; labeling songs based on valence – arousal range and using six categories of emotions selected. In labeling songs, the expertise of a Sri Lankan musician was obtained. Subsequently 72 features were extracted from each song from the data set. Feature extraction was done using the method proposed by Daniel McEnnis,et al[19].Extracted features were selected using three feature selection methods, which are popular among MIR researchers for music feature selection. These feature selection methods were ReliefF based selection, Correlation based selection and Information Gain based method. Selected features were used to train the classifier. For the model training three supervised algorithms were used. Those are Random forest algorithm, Sequential minimal optimization (SMO) algorithm and the Naïve bayes algorithm. Based on the results gained from the feature selection and classifier training 21 common features were selected and trained using the Random forest, Sequential minimal optimization (SMO) and Naïve bayes algorithms. Result obtained from the common features and different feature selection methods were compared to evaluate the possibility of proposing a common features set for Sri Lankan context. Common feature set accuracies were lower than feature selection methods results.Even though the common features set was able to achive more than 87% of accuracy for all algorithms, ReliefF based selection with Random forest algorithm gave the highest accuracy of 91.9753 % with 12 features.

Based on the all experiments and results the author of the present study propose to use Russel's 2D Valence – Arousal Emotion Model with modifications for emotion identification. Daniel McEnnis,et al [18] model for feature extraction of the song. ReliefF based selection is for feature selection and finally Random forest algorithm for classifier training. The above proposed approach promises a significant level of accuracy in emotions classification for Sinhala songs based on emotions.

6.2 Future Works

The results of the study showed that a better classification of emotions based on Sinhala songs was able to be achieve as intended at the beginning of the study. However, there are several directions that future researchers who have the interest of carrying out studies in this particular area of study can focus on. First, areas are to engage in subjective data collection using the feedback from music listeners. Since there are studies that have been conducted using similar methods in the international arena, it would be interesting to see how it applies in to the Sri Lankan context. The second area is to combine both music excerpts as well as lyrics in future studies. The present study focused only on music excerpts of songs. Combination of music excerpts of songs and lyrics, in the future will enable more meaningful classification. The third area is the possibility of applying deep leaning methods, in place of supervised machine leaning algorithms which was used in the present study.

Apart from the above future research directions there is one more important aspect that future researchers can consider which involves mainly psychoanalysis of the listeners. Psychological studies suggest that our emotional response to music that we listen depends on interplay between musical, personal and situational factors. The present study focused only on musical factors, hence future studies can focus on considering personal and situational factors for the study as well. Under the influence of situational factors such as listening mood and listening environment, a person's emotion perception of the same song could vary significantly. For an example, when we are in a sad mood, a happy song may be not so happy to us. Therefore, it would be great if the Music

Information Retrieval system could detect the listing mood by using techniques like prosodic cues, body movements, or physiological signals.

References

[1] M. Lakshitha and K. Jayaratne, "Melody Analysis for Prediction of the Emotions Conveyed by Sinhala Songs", in *2016 IEEE International Conference on Information and Automation for Sustainability (ICIAfS)*, Galle, Sri Lanka, 2016.

[2]Y.-H. Yang and Homer H. Chen, "Machine Recognition of Music Emotion: A Review," ACM Transactions on Intelligent Systems and Technology, vol. Vol. 3, no. No. 3,, p. Article 40, May 2012.

[3]K. Hevner, "Experimental Studies of the Elements of Expression in Music," The American Journal of Psychology, vol. Vol. 48, no. No. 2, pp. pp. 246-268, 1936.

[4]Y. Song, S. Dixon and M. Pearce, "Evaluation Of Musical Features For Emotion," in International Society for Music Information Retrieval, Porto, Portugal, 2012.

[5]X. H. J, S. Downie and A. F. Ehmann, "Lyric Text Mining In Music Mood Classification," in 10th International Society for Music Information Retrieval Conference, Kobe, Japan, 2009.

[6]R. Sarkar, S. Dutta, A. Roy and S. K. Saha, "Emotion Based Categorization of Music Using Low Level Features and Agglomerative Clustering," in National Conference on Computer Vision, Pattern Recognition, Image Processing, and Graphics, Mandi, 2018.

[7]X. Hu, "Music and Mood: Where Theory and Reality Meet," in iConference, Urbana-Champaign, IL, USA, 2010.

[8]K. Bischoff, C. S. Firan, C. Laurier and M. Sordo, "Music Mood And Theme Classification - A Hybrid Approach," in 10th International Society for Music Information Retrieval Conference (ISMIR 2009), Kobe, Japan, 2009.

[9]Y. Feng, Y. Zhuang and Y. Pan, "Popular Music Retrieval by Detecting Mood," in SIGIR'03, Toronto, Canada., July 28-August 1, 2003.

[10]A. A. Wieczorkowska, "Towards Extracting Emotions from Music," Intelligent Media Technology for Communicative Intelligence , vol. vol 3490, pp. pp 228-238, 2005.

[11]P. Ekman, "An argument for basic emotions", *Cognition and Emotion*, vol. 6, no. 3-4, pp. 169-200, 1992.

[12]P. Farnsworth, "A Study of the Hevner Adjective List", *The Journal of Aesthetics and Art Criticism*, vol. 13, no. 1, p. 97, 1954.

[13]Russell, "A circumplex model of affect.", *Journal of Personality and Social Psychology*, vol. 39, no. 6, pp. 1161-1178, 1980.

[14]C. Laurier, J. Grivolla and P. Herrera, "Multimodal Music Mood Classification Using Audio and Lyrics", in *2008 Seventh International Conference on Machine Learning and Applications*, San Diego, CA, USA, 2008.

[15]Wagner, J. Kim and E. Andr´e, "From Physiological Signals To Emotions: Implementing And Comparing Selected Methods For Feature Extraction And Classification", Institute of Computer Science, University of Augsburg, 2005.

[16]Y. E. Kim, E. Schmidt, R. Migneco, B. G. Morton, P. Richardson, J. Scott, J. A. Speck and D. Turnbull, "Music Emotion Recognition: A State Of The Art Review", in *11th International Society for Music Information Retrieval Conference (ISMIR 2010)*, 2010.

[17]Yang, Y. Lin, Y. Su and H. Chen, "A Regression Approach to Music Emotion Recognition", *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 16, no. 2, pp. 448-457, 2008.

[18]D. McEnnis, C. McKay, I. Fujinaga and P. Depalle, "Jaudio: A Feature Extraction Library", Queen Mary, University of London, 2005.

[19]B. Patra, D. Das and S. Bandyopadhyay, "Unsupervised Approach to Hindi Music Mood Classification", *Mining Intelligence and Knowledge Exploration*, pp. 62-69, 2013.

[20] J. Tang, S. Alelyani and H. Liu, "Feature Selection for Classification: A Review", *In Data Classification: Algorithms and Applications*, pp. pp. 37-64, 2014.

[21]J. Weston, A. Elisseeff and B. Scholkopf, "Use of the Zero-Norm with Linear Models and Kernel Methods", *Journal of Machine Learning Research*, pp. 1439-1461, 2003.

[22]P. Mitra, C. Murthy and S. Pal, "Unsupervised feature selection using feature similarity", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 3, pp. 301-312, 2002.

[23]M. Robnik and I. Kononenko, "Theoretical and Empirical Analysis of ReliefF and RReliefF", *Machine Learning Journal*, vol. 53, pp. 23-69, 2003.

[24]B. Azhagusundari and A. Thanamani, "Feature Selection based on Information Gain", *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, vol. -2, no. 2278-3075-2, 2013.

[25]Hanchuan Peng, Fuhui Long and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1226-1238, 2005.

[26]Y.-H. Yang, C.-C. Liu and a. H. H. Chen, "Music emotion classification: A fuzzy approach," in 14th ACM international conference on Multimedia, Santa Barbara, California, USA, January 2006.

[27]S. Pouyanfar and H. Sameti, “Music Emotion Recognition Using Two Level Classification,” 2014 Iranian Conference on Intelligent Systems (ICIS), Jan. 2014.

[28]X. L. Huang, L. F. Yang, Y. Xu, and S. T. Sun, “Feature selection and feature learning in arousal dimension of music emotion by using shrinkage methods,” Multimedia Systems, Oct. 2015.

[29]J. L. Zhang, x. I. A. O. Hu, “improving music mood classification using lyrics, audio and social tags,” thesis, 2010.

[30]S. McAdams, "Perspectives on the Contribution of Timbre to Musical Structure", *Computer Music Journal*, vol. 23, no. 3, pp. 85-102, 1999.

[31]M. Hall, "Correlation-based Feature Selection for Discrete and Numeric Class Machine Learning", in *Seventeenth International Conference on Machine Learning*, 2000, pp. 359-366.

[32]J. Platt, "Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines", in *Technical Report MSR-TR-98-14*, 1998.

[33]T. Ho, "Random decision forests", in *Proceedings of 3rd International Conference on Document Analysis and Recognition*, 1995.

[34]I. Rish, "An empirical study of the naive Bayes classifier", T.J. Watson Research Center, Hawthorne, NY, 2001.

[35]P. Saari, T. Eerola and O. Lartillot, "Generalizability and Simplicity as Criteria in Feature Selection: Application to Mood Classification in Music", in *IEEE Transactions on Audio Speech and Language Processing*, 2011, pp. 1802 – 1812.

Appendix

Appendix A :The feature Rank according to ReliefF based selection (First data set)

Rank	Name of the Feature	Rank	Name of the Feature
1	0.263416 72 AMMMFCCOA 10	37	0.014453 53 LPCOA 6
2	0.254448 68 AMMMFCCOA 6	38	0.014039 41 SVOA
3	0.253341 71 AMMMFCCOA 9	39	0.013742 18 LPCOSD 7
4	0.229948 70 AMMMFCCOA 8	40	0.013667 8 ZCOSD
5	0.229851 67 AMMMFCCOA 5	41	0.012233 37 SCOA
6	0.210626 30 AMMMFCCOSD 4	42	0.011697 14 LPCOSD 3
7	0.204327 65 AMMMFCCOA 3	43	0.011628 20 LPCOSD 9
8	0.202076 33 AMMMFCCOSD 7	44	0.010936 50 LPCOA 3
9	0.199615 66 AMMMFCCOA 4	45	0.010606 48 LPCOA 1
10	0.189199 69 AMMMFCCOA 7	46	0.00984 23 MMOSD 2
11	0.188319 64 AMMMFCCOA 2	47	0.009584 27 AMMMFCCOSD 1
12	0.080505 7 FOLEWOSD	48	0.009511 26 MMOSD 5
13	0.05163 36 AMMMFCCOSD 10	49	0.009348 10 BSOSD
14	0.045789 28 AMMMFCCOSD 2	50	0.008866 24 MMOSD 3
15	0.045548 39 SFOA	51	0.008011 29 AMMMFCCOSD 3
16	0.03889 40 COA	52	0.00788 9 SBOSD
17	0.035671 34 AMMMFCCOSD 8	53	0.00787 62 MMOA 5
18	0.035573 31 AMMMFCCOSD 5	54	0.006332 11 SOSBOSD
19	0.033 58 MMOA 1	55	0.005228 16 LPCOSD 5
20	0.028372 35 AMMMFCCOSD 9	56	0.005111 51 LPCOA 4
21	0.028279 32 AMMMFCCOSD 6	57	0.00496 47 SOSBOA
22	0.027227 22 MMOSD 1	58	0.004241 61 MMOA 4
23	0.025775 45 SBOA	59	0.00411 19 LPCOSD 8
24	0.023717 60 MMOA 3	60	0.003255 55 LPCOA 8
25	0.022945 59 MMOA 2	61	0.003068 5 SVOSD
26	0.021481 3 SFOSD	62	0.003028 15 LPCOSD 4
27	0.02044 2 SRPOSD	63	0.002793 6 RMSOSD
28	0.018968 49 LPCOA 2	64	0.002639 13 LPCOSD 2
29	0.018175 63 AMMMFCCOA 1	65	0.002135 25 MMOSD 4
30	0.01816 4 COSD	66	0.001768 17 LPCOSD 6
31	0.017006 1 SCOSD	67	0.001665 54 LPCOA 7
32	0.016968 42 RMSOA	68	0.000974 43 FOLEWOA
33	0.016955 38 SRPOA	69	0 21 LPCOSD 10
34	0.016863 44 ZCOA	70	0 57 LPCOA 10
35	0.015778 12 LPCOSD 1	71	-0.002123 56 LPCOA 9
36	0.01572 46 BSOA	72	-0.004456 52 LPCOA 5

Appendix B: The feature Rank according to Correlation based selection. (First data set)

Rank	Name of the Feature	Rank	Name of the Feature
1	0.5747 72 AMMMFCCOA 10	37	0.1653 23 MMOSD 2
2	0.56 71 AMMMFCCOA 9	38	0.1614 51 LPCOA 4
3	0.56 68 AMMMFCCOA 6	39	0.1557 41 SVOA
4	0.5499 67 AMMMFCCOA 5	40	0.154 48 LPCOA 1
5	0.5498 70 AMMMFCCOA 8	41	0.1469 8 ZCOSD
6	0.5486 69 AMMMFCCOA 7	42	0.1387 62 MMOA 5
7	0.5481 33 AMMMFCCOSD 7	43	0.1375 20 LPCOSD 9
8	0.5365 66 AMMMFCCOA 4	44	0.1306 6 RMSOSD
9	0.5206 30 AMMMFCCOSD 4	45	0.1256 15 LPCOSD 4
10	0.5204 65 AMMMFCCOA 3	46	0.1254 5 SVOSD
11	0.5003 64 AMMMFCCOA 2	47	0.1228 17 LPCOSD 6
12	0.4401 36 AMMMFCCOSD 10	48	0.1227 18 LPCOSD 7
13	0.4044 7 FOLEWOSD	49	0.1109 27 AMMMFCCOSD 1
14	0.3784 40 COA	50	0.1054 47 SOSBOA
15	0.3717 39 SFOA	51	0.1053 19 LPCOSD 8
16	0.3389 32 AMMMFCCOSD 6	52	0.1031 25 MMOSD 4
17	0.3388 35 AMMMFCCOSD 9	53	0.1012 13 LPCOSD 2
18	0.3336 22 MMOSD 1	54	0.0941 11 SOSBOSD
19	0.3306 58 MMOA 1	55	0.0871 24 MMOSD 3
20	0.3294 34 AMMMFCCOSD 8	56	0.0852 16 LPCOSD 5
21	0.3293 31 AMMMFCCOSD 5	57	0.0799 4 COSD
22	0.2662 3 SFOSD	58	0.0784 53 LPCOA 6
23	0.245 59 MMOA 2	59	0.0687 56 LPCOA 9
24	0.2331 60 MMOA 3	60	0.0624 10 BSOSD
25	0.2257 45 SBOA	61	0.0599 9 SBOSD
26	0.2073 49 LPCOA 2	62	0.0562 28 AMMMFCCOSD 2
27	0.205 44 ZCOA	63	0.049 61 MMOA 4
28	0.1989 63 AMMMFCCOA 1	64	0.0468 52 LPCOA 5
29	0.1979 46 BSOA	65	0.0461 55 LPCOA 8
30	0.1967 2 SRPOSD	66	0.0455 54 LPCOA 7
31	0.1963 42 RMSOA	67	0.0431 26 MMOSD 5
32	0.1954 38 SRPOA	68	0.032 29 AMMMFCCOSD 3
33	0.1834 1 SCOSD	69	0.027 50 LPCOA 3
34	0.18 14 LPCOSD 3	70	0.0158 43 FOLEWOA
35	0.1672 37 SCOA	71	0 57 LPCOA 10
36	0.1658 12 LPCOSD 1	72	0 21 LPCOSD 10

Appendix C:The feature Rank according to Information Gain based selection. (First data set)

Rank	Name of the Feature	Rank	Name of the Feature
1	1.3717 64 AMMMFCCOA 2	37	0 14 LPCOSD 3
2	1.3717 66 AMMMFCCOA 4	38	0 15 LPCOSD 4
3	1.3717 69 AMMMFCCOA 7	39	0 16 LPCOSD 5
4	1.3717 70 AMMMFCCOA 8	40	0 17 LPCOSD 6
5	1.334 67 AMMMFCCOA 5	41	0 18 LPCOSD 7
6	1.3254 30 AMMMFCCOSD 4	42	0 11 SOSBOSD
7	1.2587 33 AMMMFCCOSD 7	43	0 10 BSOSD
8	0.9486 65 AMMMFCCOA 3	44	0 9 SBOSD
9	0.9486 71 AMMMFCCOA 9	45	0 4 COSD
10	0.9486 68 AMMMFCCOA 6	46	0 2 SRPOSD
11	0.9486 72 AMMMFCCOA 10	47	0 57 LPCOA 10
12	0.471 36 AMMMFCCOSD 10	48	0 5 SVOSD
13	0.4563 7 FOLEWOSD	49	0 8 ZCOSD
14	0.3898 28 AMMMFCCOSD 2	50	0 6 RMSOSD
15	0.313 32 AMMMFCCOSD 6	51	0 56 LPCOA 9
16	0.313 35 AMMMFCCOSD 9	52	0 19 LPCOSD 8
17	0.2438 40 COA	53	0 20 LPCOSD 9
18	0.2295 39 SFOA	54	0 21 LPCOSD 10
19	0.2039 58 MMOA 1	55	0 47 SOSBOA
20	0.1952 22 MMOSD 1	56	0 62 MMOA 5
21	0.1903 31 AMMMFCCOSD 5	57	0 50 LPCOA 3
22	0.1903 34 AMMMFCCOSD 8	58	0 61 MMOA 4
23	0.1795 45 SBOA	59	0 52 LPCOA 5
24	0.1444 3 SFOSD	60	0 46 BSOA
25	0.1288 59 MMOA 2	61	0 43 FOLEWOA
26	0.1181 37 SCOA	62	0 51 LPCOA 4
27	0.1173 48 LPCOA 1	63	0 53 LPCOA 6
28	0.11 44 ZCOA	64	0 55 LPCOA 8
29	0.1025 60 MMOA 3	65	0 25 MMOSD 4
30	0.0999 49 LPCOA 2	66	0 23 MMOSD 2
31	0.0944 38 SRPOA	67	0 24 MMOSD 3
32	0.0916 63 AMMMFCCOA 1	68	0 26 MMOSD 5
33	0 41 SVOA	69	0 29 AMMMFCCOSD 3
34	0 13 LPCOSD 2	70	0 27 AMMMFCCOSD 1
35	0 12 LPCOSD 1	71	0 54 LPCOA 7
36	0 42 RMSOA	72	0 1 SCOSD

Appendix D:The feature Rank according to ReliefF based selection (Second data set)

Rank	Name of the Feature	Rank	Name of the Feature
1	0.323092 68 AMMMFCCOA 6	37	0.012083 63 AMMMFCCOA 1
2	0.321817 71 AMMMFCCOA 9	38	0.011765 23 MMOSD 2
3	0.321231 72 AMMMFCCOA 10	39	0.011275 53 LPCOA 6
4	0.314969 70 AMMMFCCOA 8	40	0.011125 37 SCOA
5	0.314863 67 AMMMFCCOA 5	41	0.009491 8 ZCOSD
6	0.299752 64 AMMMFCCOA 2	42	0.009418 19 LPCOSD 8
7	0.295758 30 AMMMFCCOSD 4	43	0.009341 42 RMSOA
8	0.295266 66 AMMMFCCOA 4	44	0.009175 14 LPCOSD 3
9	0.293452 33 AMMMFCCOSD 7	45	0.009158 46 BSOA
10	0.280159 65 AMMMFCCOA 3	46	0.008624 50 LPCOA 3
11	0.270509 69 AMMMFCCOA 7	47	0.008555 48 LPCOA 1
12	0.07295 7 FOLEWOSD	48	0.008555 29 AMMMFCCOSD 3
13	0.06658 36 AMMMFCCOSD 10	49	0.008527 27 AMMMFCCOSD 1
14	0.042089 31 AMMMFCCOSD 5	50	0.006906 15 LPCOSD 4
15	0.042077 34 AMMMFCCOSD 8	51	0.006735 43 FOLEWOA
16	0.037278 32 AMMMFCCOSD 6	52	0.006181 11 SOSBOSD
17	0.037274 35 AMMMFCCOSD 9	53	0.006075 26 MMOSD 5
18	0.036667 28 AMMMFCCOSD 2	54	0.005597 41 SVOA
19	0.036029 40 COA	55	0.005436 51 LPCOA 4
20	0.035864 39 SFOA	56	0.005055 47 SOSBOA
21	0.030172 58 MMOA 1	57	0.004061 10 BSOSD
22	0.023234 22 MMOSD 1	58	0.003991 25 MMOSD 4
23	0.023086 45 SBOA	59	0.003123 17 LPCOSD 6
24	0.018433 3 SFOSD	60	0.002959 54 LPCOA 7
25	0.018266 59 MMOA 2	61	0.002891 24 MMOSD 3
26	0.018153 4 COSD	62	0.002554 55 LPCOA 8
27	0.017454 2 SRPOSD	63	0.00133 56 LPCOA 9
28	0.015708 20 LPCOSD 9	64	0.001311 62 MMOA 5
29	0.015269 60 MMOA 3	65	0.000697 13 LPCOSD 2
30	0.014748 44 ZCOA	66	0.000269 61 MMOA 4
31	0.014253 38 SRPOA	67	0.000136 5 SVOSD
32	0.013625 1 SCOSD	68	0 57 LPCOA 10
33	0.013535 18 LPCOSD 7	69	0 21 LPCOSD 10
34	0.013121 12 LPCOSD 1	70	-0.000219 16 LPCOSD 5
35	0.012688 9 SBOSD	71	-0.000683 6 RMSOSD
36	0.012107 49 LPCOA 2	72	-0.003974 52 LPCOA 5

Appendix E :The feature Rank according to Correlation based selection. (Second data set)

Rank	Name of the Feature	Rank	Name of the Feature
1	0.323092 68 AMMMFCCOA 6	37	0.012083 63 AMMMFCCOA 1
2	0.321817 71 AMMMFCCOA 9	38	0.011765 23 MMOSD 2
3	0.321231 72 AMMMFCCOA 10	39	0.011275 53 LPCOA 6
4	0.314969 70 AMMMFCCOA 8	40	0.011125 37 SCOA
5	0.314863 67 AMMMFCCOA 5	41	0.009491 8 ZCOSD
6	0.299752 64 AMMMFCCOA 2	42	0.009418 19 LPCOSD 8
7	0.295758 30 AMMMFCCOSD 4	43	0.009341 42 RMSOA
8	0.295266 66 AMMMFCCOA 4	44	0.009175 14 LPCOSD 3
9	0.293452 33 AMMMFCCOSD 7	45	0.009158 46 BSOA
10	0.280159 65 AMMMFCCOA 3	46	0.008624 50 LPCOA 3
11	0.270509 69 AMMMFCCOA 7	47	0.008555 48 LPCOA 1
12	0.07295 7 FOLEWOSD	48	0.008555 29 AMMMFCCOSD 3
13	0.06658 36 AMMMFCCOSD 10	49	0.008527 27 AMMMFCCOSD 1
14	0.042089 31 AMMMFCCOSD 5	50	0.006906 15 LPCOSD 4
15	0.042077 34 AMMMFCCOSD 8	51	0.006735 43 FOLEWOA
16	0.037278 32 AMMMFCCOSD 6	52	0.006181 11 SOSBOSD
17	0.037274 35 AMMMFCCOSD 9	53	0.006075 26 MMOSD 5
18	0.036667 28 AMMMFCCOSD 2	54	0.005597 41 SVOA
19	0.036029 40 COA	55	0.005436 51 LPCOA 4
20	0.035864 39 SFOA	56	0.005055 47 SOSBOA
21	0.030172 58 MMOA 1	57	0.004061 10 BSOSD
22	0.023234 22 MMOSD 1	58	0.003991 25 MMOSD 4
23	0.023086 45 SBOA	59	0.003123 17 LPCOSD 6
24	0.018433 3 SFOSD	60	0.002959 54 LPCOA 7
25	0.018266 59 MMOA 2	61	0.002891 24 MMOSD 3
26	0.018153 4 COSD	62	0.002554 55 LPCOA 8
27	0.017454 2 SRPOSD	63	0.00133 56 LPCOA 9
28	0.015708 20 LPCOSD 9	64	0.001311 62 MMOA 5
29	0.015269 60 MMOA 3	65	0.000697 13 LPCOSD 2
30	0.014748 44 ZCOA	66	0.000269 61 MMOA 4
31	0.014253 38 SRPOA	67	0.000136 5 SVOSD
32	0.013625 1 SCOSD	68	0 57 LPCOA 10
33	0.013535 18 LPCOSD 7	69	0 21 LPCOSD 10
34	0.013121 12 LPCOSD 1	70	-0.000219 16 LPCOSD 5
35	0.012688 9 SBOSD	71	-0.000683 6 RMSOSD
36	0.012107 49 LPCOA 2	72	-0.003974 52 LPCOA 5

Appendix F:The feature Rank according to Information Gain based selection. (Second data set)

Rank	Name of the Feature	Rank	Name of the Feature
1	2.3005 64 AMMMFCCOA 2	37	0 2 SRPOSD
2	2.3005 69 AMMMFCCOA 7	38	0 6 RMSOSD
3	2.3005 70 AMMMFCCOA 8	39	0 9 SBOSD
4	2.3005 66 AMMMFCCOA 4	40	0 15 LPCOSD 4
5	2.2628 67 AMMMFCCOA 5	41	0 13 LPCOSD 2
6	2.2125 33 AMMMFCCOSD 7	42	0 14 LPCOSD 3
7	2.2041 30 AMMMFCCOSD 4	43	0 12 LPCOSD 1
8	2.0319 71 AMMMFCCOA 9	44	0 10 BSOSD
9	2.0319 72 AMMMFCCOA 10	45	0 11 SOSBOSD
10	2.0319 65 AMMMFCCOA 3	46	0 24 MMOSD 3
11	2.0253 68 AMMMFCCOA 6	47	0 63 AMMMFCCOA 1
12	0.5871 36 AMMMFCCOSD 10	48	0 26 MMOSD 5
13	0.5223 7 FOLEWOSD	49	0 47 SOSBOA
14	0.3375 28 AMMMFCCOSD 2	50	0 51 LPCOA 4
15	0.255 40 COA	51	0 50 LPCOA 3
16	0.2547 39 SFOA	52	0 49 LPCOA 2
17	0.2472 32 AMMMFCCOSD 6	53	0 52 LPCOA 5
18	0.2472 35 AMMMFCCOSD 9	54	0 53 LPCOA 6
19	0.2375 22 MMOSD 1	55	0 54 LPCOA 7
20	0.2317 31 AMMMFCCOSD 5	56	0 57 LPCOA 10
21	0.2317 34 AMMMFCCOSD 8	57	0 56 LPCOA 9
22	0.2253 58 MMOA 1	58	0 55 LPCOA 8
23	0.185 37 SCOA	59	0 48 LPCOA 1
24	0 60 MMOA 3	60	0 46 BSOA
25	0 19 LPCOSD 8	61	0 27 AMMMFCCOSD 1
26	0 21 LPCOSD 10	62	0 45 SBOA
27	0 20 LPCOSD 9	63	0 62 MMOA 5
28	0 17 LPCOSD 6	64	0 59 MMOA 2
29	0 23 MMOSD 2	65	0 29 AMMMFCCOSD 3
30	0 18 LPCOSD 7	66	0 61 MMOA 4
31	0 16 LPCOSD 5	67	0 38 SRPOA
32	0 25 MMOSD 4	68	0 41 SVOA
33	0 4 COSD	69	0 44 ZCOA
34	0 5 SVOSD	70	0 43 FOLEWOA
35	0 3 SFOSD	71	0 42 RMSOA
36	0 8 ZCOSD	72	0 1 SCOSD