CZECH TECHNICAL UNIVERSITY in PRAGUE
Faculty of Electrical Engineering

ČVUT FEL katedra počítačů

Diploma Thesis

Pattern Recognition for Real World Data

Miroslav Čepek

Supervisor: Ing. Pavel Kordík

Study programme: Informatika a výpocetní technika

January 2006

Prohlášení

Prohlašuji, že jsem svou diplomovou práci vypracoval samostatně a použil jsem pouze podklady uvedené v přiloženém seznamu.

Nemám závažný důvod proti užití tohoto školního díla ve smyslu §60 Zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon).

V Úvalech dne 26. 1. 2005 ....

Abstract

A long time sleep EEG recording is very important in diagnosis of many diseases and sleep difficulties. For easier diagnosis it is necessary to know so called sleep stages. Because manual Sleep stages scoring is a time consuming work, there is a constant desire to automate the process. The automation is discussed in many works, but the most of them uses only Fourier transform to extract relevant features from the signal. In this work I will try to find out if some new features do not describe the Sleep stages better. Among many feature extraction methods the Wavelet transform, Statistic and few others will be used. After the feature extraction, usefull features will be selected with various data mining methods and then each selection will be evaluated with several classifiers.

Abstrakt

Dlouhodobé záznamy spánkového EEG jsou velmi důležité při diagnóze mnoha závažných onemocnění a poruch spánku. Pro ulehčení určování diagózy je důležité znát takzvané spánkové fáze. Protože manuální ohodnocování EEG záznamu je dlouhá a monotóní práce, je dlouhodobě zájem o automatizaci této úlohy. Automatizace je tématem mnoha prací, ale většina používá pouze fourierovu transformaci pro extrakci příznaků ze signálu. V této práci jsm si vytyčil úkol prozkoumat, zda nějaké jiné příznaky nepopisují spánkové fáze lépe. Při extrakci příznaků použiji Waveletovou transformaci, statistiku a několik dalších metod. Po extrakci příznaků použiji několik metod pro výběr příznaků, abych zkusil najít významné příznaky. Každou podmnožinu vybraných příznaků zkusím vyhodnotit několika klasifikačními metodami.

Contents

List of figures xii

List of tables xiv

1 Introduction and Motivation 1

I Theoretical part 3

2 Electroencefalogram (EEG) 5
2.1 A Brief History 5
2.2 Origin of EEG 5
2.3 Measuring of EEG 5
2.4 Description of EEG 5
2.5 Sleep Scoring 6

3 Features 9
3.1 Introduction 9
3.2 Fourier Transform 9
3.2.1 Introduction 9
3.2.2 Continuous Fourier Transform 9
3.2.3 Discrete Fourier Transform 10
3.2.4 Windowed Fourier Transform 10
3.3 Wavelets 11
3.3.1 Introduction 11
3.3.2 Continuous Wavelet Transformation 12
3.3.3 Discrete Wavelet Transformation 13
3.4 Hjorth's Parameters 13
3.4.1 Activity, Mobility, Complexity 14
3.5 Important Volume Changes 14
3.6 Function Approximation with Rational and Polynomial Functions 15
3.6.1 Polynomial Approximation 15
3.6.2 Rational Approximation 15
3.7 Statistic 16
3.7.1 Basics 16
3.7.2 Statistic parameters 17
3.7.3 Normal Probability Distribution 19
3.7.4 Interval Parameter Estimations 19

4 Clustering Methods 21
4.1 Vector Quantisation 21
4.2 K-Means 22

5 Classification 23
5.1 Naive Bayesian classificator 23
5.2 Bayesian Net 23

5.3 Decision Tables . . . 24
5.4 Simple Logistic . . . 24
5.5 Decision Tree . . . 25
5.6 GAME . . . 25

6 Feature Selection and Feature Ranking 27
6.1 Introduction . . . 27
6.2 GAME . . . 27
6.3 Weka . . . 27
6.3.1 Feature Search Methods . . . 28
6.3.2 Feature Subset Evaluators . . . 28

II Practical Part 29

7 Goals to Achieve and Proposed Concept 31

8 List of Tested Features 33
8.1 Note on Symbols . . . 33
8.2 Features in Josef Rieger's work . . . 33
8.3 My Features . . . 34
8.3.1 Continuous Wavelet Transformation Statistics . . . 34
8.3.2 Fourier Coefficients Statistics . . . 35
8.3.3 Fourier Frequency Band Spectrum Energy . . . 35
8.3.4 Spectral Centroid . . . 36
8.3.5 Time Domain Features – Ratio Values of Time Signal . . . 36
8.3.6 Hjorth Parameters . . . 36
8.3.7 Important Volume Changes Detector . . . 36
8.3.8 Polynomial Interpolation Parameters . . . 36
8.3.9 Rational Interpolation Parameters . . . 37

9 Evaluation of Proposed Concept on Snoring Data 39
9.1 Introduction . . . 39
9.2 Experimental Setup . . . 39
9.3 Data Acquisition . . . 40
9.4 Breathing Stages Assignment . . . 40
9.5 Feature Extraction . . . 41
9.6 Feature Ranking, Feature Selection and its Results . . . 41
9.6.1 Introduction . . . 41
9.6.2 Class: Not Breathing . . . 43
9.6.3 Class: Breathing Irregularly . . . 44
9.6.4 Class: Breathing Regularly . . . 45
9.6.5 Class: Snoring . . . 45
9.6.6 Feature Selection Conclusions . . . 46
9.7 Classificator Building and Their Results . . . 47
9.7.1 Classification Results – Class "Not Breathing" . . . 47
9.7.2 Classification Results – Class "Breathing Irregularly" . . . 50
9.7.3 Classification Results – Class "Breathing Regularly" . . . 52
9.7.4 Classification Results – Class "Snoring" . . . 54
9.7.5 Results Discussion and Confusion Matrix . . . 56
9.8 Evaluation of Proposed Concept – Conclusions . . . 56

10 Application of the Concept on MIT Sleep Data 59
10.1 Description . . . 59
10.2 Experiment Setup . . . 59
10.3 Data Acquisition . . . 59
10.4 Features Extraction . . . 60
10.5 Sleep Stages Balancing . . . 60

10.6 Feature Ranking, Feature Selection and Its Results . . . 60
10.6.1 Introduction . . . 60
10.6.2 Sleep Stage Wake . . . 61
10.6.3 Sleep Stage 1 . . . 62
10.6.4 Sleep Stage 2 . . . 64
10.6.5 Sleep Stage 3 . . . 65
10.6.6 Sleep Stage 4 . . . 66
10.6.7 Sleep Stage REM . . . 68
10.6.8 Feature Selection Conclusions . . . 69
10.7 Classifiers Building and Their Results . . . 70
10.7.1 Patient Wake – Results . . . 70
10.7.2 Sleep Stage 1 – Results . . . 71
10.7.3 Sleep Stage 2 – Results . . . 71
10.7.4 Sleep Stage 3 – Results . . . 72
10.7.5 Sleep Stage 4 – Results . . . 73
10.7.6 Sleep Stage REM – Results . . . 74
10.8 Sleep Stages Results and Conclusions . . . 74

11 Future work 79

12 Conclusion 81

A Brief Overview of Used Programs 83

B Complete list of features 87
B.1 Note on Symbols . . . 87
B.2 Features in original work . . . 87
B.3 New Tested Features . . . 88
B.3.1 Continuous Wavelet Transformation Statistics . . . 88
B.3.2 Fourier Coefitients Statistics . . . 89
B.3.3 Fourier Frequency Band Spectrum Energy . . . 90
B.3.4 Spectral Centroid . . . 92
B.3.5 Time Domain Features – Ratio Values of Time Signal . . . 92
B.3.6 Hjorth Parameters . . . 92
B.3.7 Important Volume Changes Detector . . . 92
B.3.8 Polynomial Interpolation Parameters . . . 93
B.3.9 Rational Interpolation Parameters . . . 93

C Evaluation of Proposed Concept – List of Selected Features 95
C.1 Class: Not Breathing . . . 95
C.2 Class: Breathing Irregulary . . . 98
C.3 Class: Breathing Regulary . . . 101
C.4 Class: Snoring . . . 104

D Sleep EEG Data – List of Selected Features 107
D.1 Sleep Stage Wake . . . 107
D.2 Sleep Stage 1 . . . 109
D.3 Sleep Stage 2 . . . 111
D.4 Sleep Stage 3 . . . 115
D.5 Sleep Stage 4 . . . 119
D.6 Sleep Stage REM . . . 122

E Enclosed CD-ROM content 127

List of Figures

2.1 Placing electrodes in system 10/20 . . . 6
2.2 EEG frequency bands illustration on real data . . . 6
2.3 Illustration of Sleep stages in frequency domain . . . 7

3.1 Square (left) and Bartlett's window . . . 11
3.2 Hamming's (left) and Blackman's window . . . 11
3.3 Explaining frequency and time resolution . . . 12
3.4 Discrete wavelet transformation scheme . . . 13
3.5 Characteristic changes of curve shape, illustrating the dependence of individual parameters. 14
3.6 Normal distribution density . . . 19
3.7 Probabilities of Normal distribution ranges . . . 20
3.8 Examples of probable and non probable distributions . . . 20
3.9 Comparison of Normal distribution and Student's distribution with three degrees of freedom 20

4.1 Example of clustering . . . 21

5.1 Example of Bayes network . . . 24
5.2 Example of decision tree . . . 25
5.3 a structure of sample GAME neural network . . . 26
5.4 Less common data mining tasks the GAME provides. . . . 26

7.1 Concept illustration – number of section refers to application of the concept on MIT sleep data. . . . 32

9.1 Evaluation of proposed concept flowchart . . . 39
9.2 Importing the data to Class Assignment Tool . . . 40
9.3 Class assigning procedure – Example of "Not Breathing" Class . . . 41
9.4 Class assigning procedure – Example of "Breathing Irregularly" Class . . . 41
9.5 Class assigning procedure – Example of "Breathing Regularly" Class . . . 42
9.6 Class assigning procedure – Example of "Snoring" Class . . . 42
9.7 Feature extraction in Class Assignment Tool . . . 43
9.8 Error rate achieved by classificators for class "Not Breathing" . . . 48
9.9 Example of error – instance misclassified in preprocessing . . . 49
9.10 Example of error – example of border instance . . . 49
9.11 Error rate achieved by classificators for class "Breathing Irregularly" . . . 50
9.12 Example of error – example of bad learned classificator . . . 51
9.13 Example of error – instance misclassified in preprocessing or classificator error. . . . 51
9.14 Error rate achieved by classificators for class "Breathing Regularly" . . . 52
9.15 Example of error – error caused by border instance . . . 53
9.16 Example of error – poorly learned classifier – the middle row, GAME with all features . 53
9.17 Error rate achieved by classificators for class "Snore" . . . 54
9.18 Example of error – error in class assignment during preprocessing . . . 55
9.19 Example of error – bad learned classifiers (according to my opinion) . . . 55

10.1 Experimental setup for Sleep EEG experiment . . . 59
10.2 Error rate achieved by classificators for Sleep Stage Wake . . . 70

10.3 Error rate achieved by classificators for Sleep Stage 1 71
10.4 Error rate achieved by classificators for Sleep Stage 2 72
10.5 Error rate achieved by classificators for Sleep Stage 3 73
10.6 Error rate achieved by classificators for Sleep Stage 4 73
10.7 Error rate achieved by classificators for Sleep Stage REM 74
10.8 Example of confusing instances in electrode EOG for sleep stage REM 75
10.9 Example of confusing instances for sleep stage REM 76
10.10Example of confusing instances for sleep stage 2 and 3 76

11.1 Sleep recording in frequency domain with sleep stages 79

A.1 GAME Screenshot 83
A.2 Weka Screenshot 84
A.3 Matlab screenshot 84
A.4 CATool screenshot 85

List of Tables

2.1 EEG frequency bands 8

9.1 Results on feature groups significance for stage ”Not Breathing” 44
9.2 Results on feature groups significance for stage ”Breathing Irregularly” 44
9.3 Results on feature groups significance for stage ”Breathing Regularly” 45
9.4 Results on feature groups significance for stage ”Snoring” 46
9.5 Results on feature groups significance for all stages – abbreviations meaning is in the text on page 46 46
9.6 Error rate achieved by classificators for class ”Not Breathing” 47
9.7 Error rate achieved by classificators for class ”Breathing Irregularly” 50
9.8 Confusion table for Simple Logistic classificator classifying all classes at once 56
9.9 Results on feature groups significance for all stages (In percents of all features) 56
9.10 Best results achieved by classifiers for each class (Error rate and Accuracy are in percentage of misclassified or correctly classified instances.) 57

10.1 Results on feature groups significance for sleep stage Wake 61
10.2 Results on electrode significance for sleep stage Wake 62
10.3 Results on feature groups significance for sleep stage 1 63
10.4 Results on electrode significance for sleep stage 1 63
10.5 Results on feature groups significance for sleep stage 2 64
10.6 Results on electrode significance for sleep stage 2 65
10.7 Results on feature groups significance for sleep stage 3 66
10.8 Results on electrode significance for sleep stage 3 66
10.9 Results on feature groups significance for sleep stage 4 67
10.10 Results on electrode significance for sleep stage 4 67
10.11 Results on feature groups significance for sleep stage REM 68
10.12 Results on electrode significance for sleep stage REM 69
10.13 Feature groups significance for all sleep stages 69
10.14 Significance of electrodes for all sleep stages 69
10.15 Error rate achieved by classificators for stage Wake 70
10.16 Error rate achieved by classificators for sleep stage 1 71
10.17 Error rate achieved by classificators for sleep stage 2 72
10.18 Error rate achieved by classificators for sleep stage 3 72
10.19 Error rate achieved by classificators for sleep stage 4 73
10.20 Error rate achieved by classificators for sleep stage REM 74
10.21 Best results achieved by classifiers for each class (Error rate and Accuracy are in percentage of misclassified or correctly classified instances.) 75
10.22 Comparison my error rate with other works. Empty spaces in the table means that the work did not concern on that sleep stage. Note that in work [19] the results are presented only in form of charts and there are no exact values. 77

C.1 Selected features – Class: Not Breathing – CfsSubsetEvaluator and BestFirst, CfsSubsetEvaluator and GeneticSearch 95
C.2 Selected features – Class: Not Breathing – GAME selected features, Ranker and ChiSquaredEvaluator 96
C.3 Selected features – Class: Not Breathing – Ranker and GainRatio, Ranker and CInfoGain 97

C.4 Selected features – Class: Breathing Irregulary – CfsSubsetEvaluator and BestFirst, CfsSubsetEvaluator and GeneticSearch . . . 98
C.5 Selected features – Class: Breathing Irregulary – GAME selected features, Ranker and ChiSquaredEvaluator . . . 99
C.6 Selected features – Class: Breathing Irregulary – Ranker and GainRatio, Ranker and CInfoGain . . . 100
C.7 Selected features – Class: Breathing Regulary – CfsSubsetEvaluator and BestFirst, CfsSubsetEvaluator and GeneticSearch . . . 101
C.8 Selected features – Class: Breathing Regulary – GAME selected features, Ranker and ChiSquaredEvaluator . . . 102
C.9 Selected features – Class: Breathing Regulary – Ranker and GainRatio, Ranker and CInfoGain . . . 103
C.10 Selected features – Class: Snoring – CfsSubsetEvaluator and BestFirst, CfsSubsetEvaluator and GeneticSearch . . . 104
C.11 Selected features – Class: Snoring – GAME selected features, Ranker and ChiSquaredEvaluator . . . 105
C.12 Selected features – Class: Snoring – Ranker and GainRatio, Ranker and CInfoGain . . . 106

D.1 Selected features – Wake – GAME selected features . . . 108
D.2 Selected features – Sleep Stage 1 – GAME selected features . . . 109
D.3 Selected features – Sleep stage 2 – CfsSubsetEvaluator and BestFirst, CfsSubsetEvaluator and GeneticSearch . . . 111
D.4 Selected features – Sleep Stage 2 – GAME selected features, Ranker and ChiSquaredEvaluator . . . 112
D.5 Selected features – Sleep Stage 2 – Ranker and GainRatio, Ranker and InfoGain . . . 114
D.6 Selected features – Sleep stage 3 – CfsSubsetEvaluator and BestFirst, CfsSubsetEvaluator and GeneticSearch . . . 115
D.7 Selected features – Sleep Stage 3 – GAME selected features, Ranker and ChiSquaredEvaluator . . . 116
D.8 Selected features – Sleep Stage 3 – Ranker and GainRatio, Ranker and InfoGain – please note, that sets of selected features are the same for both methods . . . 118
D.9 Selected features – Sleep stage 4 – CfsSubsetEvaluator and BestFirst, CfsSubsetEvaluator and GeneticSearch . . . 119
D.10 Selected features – Sleep Stage 4 – GAME selected features, Ranker and ChiSquaredEvaluator . . . 120
D.11 Selected features – Sleep Stage 4 – Ranker and GainRatio, Ranker and InfoGain . . . 121
D.12 Selected features – Sleep stage REM – CfsSubsetEvaluator and BestFirst, CfsSubsetEvaluator and GeneticSearch . . . 123
D.13 Selected features – Sleep Stage REM – GAME selected features, Ranker and ChiSquaredEvaluator . . . 124
D.14 Selected features – Sleep Stage REM – Ranker and GainRatio, Ranker and InfoGain . . . 125

E.1 Enclosed CD-ROM content . . . 127

Chapter 1

Introduction and Motivation

According to profesor Matoušek the complaints of poor night sleep can be found in about one third of the population and in 13% occur the symptoms with such an intensity that they negatively influence the well-being. The therapy of these sleep difficulties is based on whole-night polygraphic examination, consisting of multichannel electroencephalogram (EEG) and many others signals. For the purpose of the diagnosis, the 8 hours´ recordings is divided in segments of 30s length. Each of these segments is separately evaluated and the corresponding sleep stage is assigned. Sleep stages were introduced in 1967 by Association for the Psychophysiological Study of Sleep (APSS) and this organisation also introduced patterns, which identifies each sleep stage. These patterns were published in the Manual for Visual Sleep Scoring [20].

Through the patterns are clear and deciding on each segment is not too difficult, the recording consists from several thousands of multichannel segments. Hence the procedure is time-consuming and requires experienced neurophysiologist. Consequently, the method of a visual scoring represents a substantial limitation, decreasing the capacity of the sleep laboratories and reducing the access to the diagnostics and therapy of sleep difficulties.

In the literature, several methods of computerised scoring of the sleep records have been suggested. However, none of those methods has been broadly used in the clinical work. The reason is that the automatic assessment fails to correctly classify the sleep stages in almost one third of cases, and the typical diagnostic procedure still is based on the visual inspection of the recording.

Many of these unsuccessful methods are based on frequency band powers, which are the most natural, because APSS defined the Sleep stages also by frequencies powers in the signal.

In this work I will try answer following question: *"Are there any better signal features which may lead to more accurate classification?"*

The examined features will be extracted using Wavelet transform, Statistic, also Fourier transform and few other methods. Among these extracted features I will rank and select usefull features with several feature selection methods. Feature selection methods were taken from commonly available data mining tools. After feature selection, several classifiers will be used to evaluate selected features. And than, based on results of classifiers, I will try to find an answer to my question.

In my work I will often refer to diploma thesis and some other work done by Josef Rieger who graduated in 2004. His work was not exactly on the same field as my work (the Sleep EEG) but it is a little bit more general – the long term EEG. In this work I will also use some of his features.

I will divide my work into two parts. The first part will cover basic theory for methods used in feature extraction in chapter 3, feature selection in chapter 6, classification in chapter 5.

In the second part I will describe all experiments I made. Based on many past experimental results it was proved that for each data set different signal features are significant. Therefore I propose a concept leading to the utilisation of proper signal features. Using significant features gives very often better classification accuracy. This concept can be applied general to any time-series data. The concept is proposed in chapter 7. Later in chapter 8 I will list all signal features I will use in my work. And finally in chapter 9 I will evaluate proposed concept on simple data. In chapter 10 I will use the concept on sleep EEG data where I gather enough information which allows me to answer my question.

Part I

Theoretical part

Chapter 2

Electroencefalogram (EEG)

2.1 A Brief History

The human interest in EEG begins in 1875 when English physician Richard Caton discovered that if two electrodes are applied to the surface of animal head a sensitive instrument can show continuous fluctuation of electric potential between those two electrodes. Later same activity was found on human head. The scientist who discovered this was German physiatrist Hans Berger. First between 1902 and 1910 he tried to record EEG of various animals and he was in general unsuccessful. But in 1924 he turned to recording human EEG and he succeeded. His first article was published in 1929 in Archiv forschung Psychiatrie. He found rhythmic changes (brain waves) varied with the individual's state of consciousness. He found two basic frequency bands of EEG and gave them their names (alpha and beta).

Rest of the world began to interest in human EEG in 1933 when two English physicians, Edgar Adrian and Brain Matthews, were able to repeat Berger's experiments. These distinguished English investigators gained for Berger's discovery the scientific support it needed.

Since then the EEG is used in many medicine fields.

2.2 Origin of EEG

A basic stone of neural system is a neural cell (neuron). This neuron consists of a body (soma), many "inputs" which connects it to its neighbours and gets informations from them. But the biggest part of neuron is called axon. This axon (it may be called an "output" of neuron cell) transfers small electrical charges from the body to other neighbouring neurons. This charge is impossible to measure on living animals or humans. But if we take a large group of tens of thousands cell we are able to measure its activity even through skull. And if we record this electrical activity, we got EEG.

2.3 Measuring of EEG

Recording of EEG is provided by large number of electrodes placed on mans head. International standard for placing electrodes on the head was developed is 1960's and is called 10/20. The name have origin in distances between electrodes. The distance is 10 or 20 percent of mans head as shown on the figure 2.1.

In total its used 21 electrodes. 19 of them are signal and are placed on the head. Remaining two electrodes are referential.and are placed on ear lobes. For easier placing are often used so-called EEG cap or EEG belts with build-in electrodes. To eliminate contact resistance a conducting gel is applied.

2.4 Description of EEG

In general there are four EEG frequency bands called alpha, beta, theta and delta. Each band is connected to some mental activity as described in table 2.1 below.

To get better idea of EEG frequency bands see figure 2.2. On this figure the real Sleep EEG data with some activity are shown. The original EEG time domain data were converted into frequency domain by

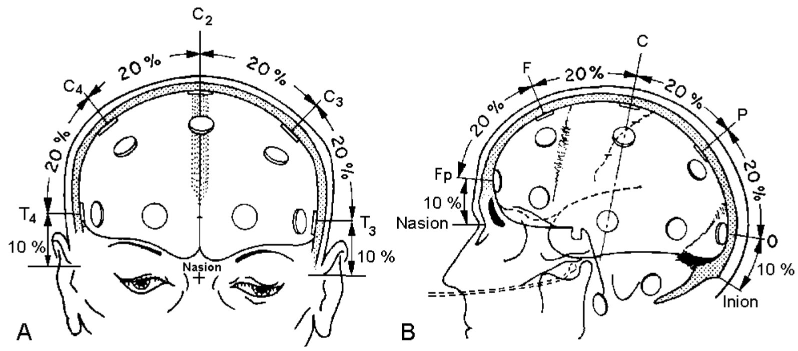

Figure 2.1: Placing electrodes in system 10/20

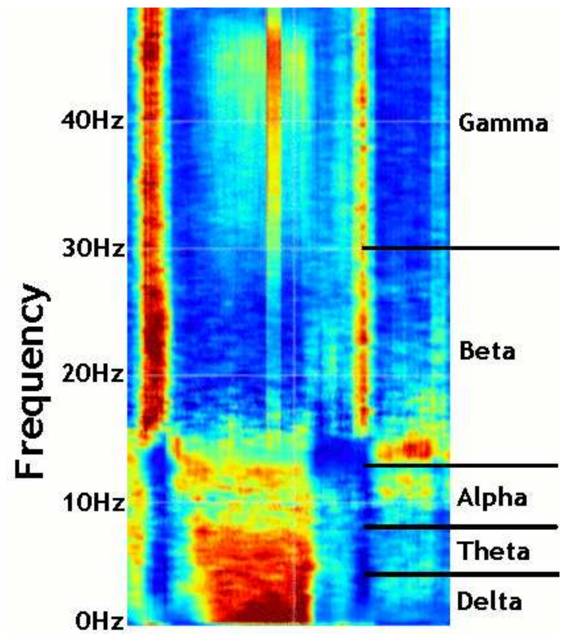

Figure 2.2: EEG frequency bands illustration on real data

Fourier transform and visualised. This picture comes from diploma thesis of Josef Rieger [4].

This division is mainly historical and have no logical sense. Alpha and beta bands were discovered and named by Hans Berger in 1929. Later in 1936 Grey Walter discovered and named delta and theta waves. Gamma frequency band was created by Jasper and Andrews in 1938. They divided early beta described by Berger as frequencies higher than 13 Hz to nowadays beta and gamma.

2.5 Sleep Scoring

In 1967 the Association for Psychophysiological Study of Sleep chartered a committee of sleep researchers to establish a standard system for visually scoring stages of sleep. The terminology and scoring system produced by this committee were quickly adopted after its publishing in 1968 and are used to nowadays.

It describes main characteristics of each stage. I will describe only basic characteristics of each sleep stage. For detailed insight please see [20] or [16].

Motion time – This is an artifact stage – it occurs when patient moves.

Stage Wake – This stage occures when patient is awake. It is characterised by alpha activity and/or a low voltage, mixed frequency EEG. Certain subjects (alpha producers) may have a virtually continuous alpha record; other subjects may show little or no alpha activity in the waking record. This stage is usually, but not necessarily, accompanied by a relatively high tonic EMG, and often REMs and eye blinks are present in the EOG tracing.

Stage 1 – This stage is the transition from wakefulness to the other sleep stages or following body movements during sleep. During nocturnal sleep, Stage 1 tends to be relatively short, lasting from about 1 to 7 minutes.

Stage 2 – This stage is defined by the presence of sleep spindles and/or K complexes and the absence of sufficient high amplitude, slow activity to define the presence of Stages 3 and 4 (see below).

The sleep spindle are defined as trains of 12-16 Hz waves of 10 μV or higher amplitude. It must contains at least six consecutive waves or must last longer than 0.5 s.

K complex is a waveform beginning with negative wave immediately followed by a high positive slow wave. The duration must be longer than 0.5s and its peak-to-peak amplitude must be bigger than 200 μV.

Stage 3 – This stage is identified if frequencies of deep delta band (below 2Hz) occupies between 20% and 50% of an epoch.

Stage 4 – This stage is identified if frequencies of deep delta band (below 2Hz) occupies more than 50% of an epoch.

Stage REM – This stage contains the REM waveforms which must appear as a saccadic eye movement, with rapid changes in angular velocity. This means that position of eyes must be also recorded.

For better illustration of above presented stage descriptions see figure 2.3. It was obtained from Josef Rieger. It presents one electrode recording of sleep EEG recording in frequency domain.

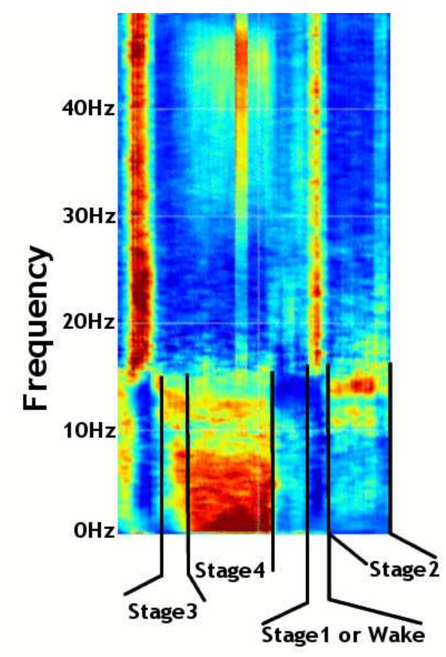

Figure 2.3: Illustration of Sleep stages in frequency domain

Band name	Frequency [Hz]	Description
delta	0 – 4 Hz	EEG of adult person with this frequency occurs in very deep sleep, hypnosis or in coma. If is present in wake state, it is always pathological phenomenon. The higher amplitude or better locality mean more serious problems. But this frequency band is normal for very young children.
theta	4 – 8 Hz	This frequency band is, for healthy people, specific in deep sleep and is often connected to dreaming. This waves occurs in central, temporal and parietal parts of head. Pathological occurrences are if amplitude of theta is as twice as high as alpha waves or is higher than 15 μV.
alpha	8 – 13 Hz	Alpha waves can be typically found in back side of head and are characteristic for awake state without mental or physical activity. The highest amplitude achieves in closed–eyes state and fades with opening eyes and growing activity.
beta	13 – 30 Hz	Are connected to the front part of the head and occurs with high mental activity like concentration, thinking or emotions.
gamma	above 30 Hz	This frequency band is specific to humans and is only band found in every part of the brain. Occurs when brain needs to simultaneously process information from different areas. A good memory is associated with gamma waves activity and deficiency creates learning disabilities.

Table 2.1: EEG frequency bands

Chapter 3

Features

3.1 Introduction

In general a feature may be almost anything – for example colour, taste, smell, temperature, consistency, mean value and many others. It may be said that a feature is some property of examined object. The process of gaining features from the object is called feature extraction.

The problem with the features is that it is often not known which are the best to describe behavior of the classificator (or any other task). The idea to extract as many features as possible is not good because many features will be misguiding and classification methods (or any others) may become lost in quantity of data. The solution is to use a process called feature selection. Feature selection is a process which tries to identify useful features and discart the others. Methods of feature selection is described later in this work.

I selected many features to be extracted from the signal. Some features were taken from diploma thesis of Josef Rieger [4] and I was interested how good they are. The frequency band powers features were exteracted because they were mentioned in Sleep scoring manual [20]. Wavelet and Fourier transorm statistic features were extracted becouse I thought they may be significant. The other features were extracted because I was interested how significant they are for describing such complex signal. Basics of mathematical methods I used for feature extraction in my work are described below in this chapter.

3.2 Fourier Transform

3.2.1 Introduction

Fourier transform is the one of the most widely used transform in many science fields. It was developed in early 19th century by Joseph Fourier.

3.2.2 Continuous Fourier Transform

If function f meets following criteria – $f(x)$ and $\frac{df(x)}{dx}$ are piecewise continuous for $\forall x \in \mathcal{R}$ and

$$\int_{-\infty}^{+\infty} f(x)dx < +\infty$$

then function $\mathcal{F}(i\omega)$ computed by equation 3.1 is Fourier image of $f(x)$ in frequency space.

$$\mathcal{F}(i\omega) = \int_{-\infty}^{+\infty} x(t)e^{-i\omega t}dt \tag{3.1}$$

where
ω angular frequency.
i imaginary unit $i^2 = -1$.
Reverse transformation is defined by following formula :

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathcal{F}(i\omega)e^{i\omega t}d\omega \tag{3.2}$$

3.2.3 Discrete Fourier Transform

For discrete real world signals we can't use above formula 3.1 because the signal is not continuous and we can't integrate it and yet another modification must take place because real world signal always have finite length. There is no need to despair because the definition formula 3.1 can be changed to fit the new situation. First, the integral is substituted with a sum. The second is then in theory solved by repeating of the signal to both positive and negative infinity.

This modification is called Discrete Fourier transform.

$$\mathcal{X}\left(\frac{k}{NT}\right) = T\sum_{n=0}^{N-1} x(nT)e^{-\frac{2i\pi}{N}nk} \tag{3.3}$$

where
- N number of samples in the signal.
- T length of signal in time.
- k frequency coefficient number.
- i imaginary unit $i^2 = -1$.

There is one very important note about frequency spectrum. In continuous Fourier transform the transformed function was also continuous and you get information about every frequency in the continuous function. But in discrete case you get only set of points (or sequence) in frequency space so you get information only about few frequencies.

The number of frequencies in spectrum is equal to the number of samples in signal. Frequency values in the spectrum begins with $f_0 = 0$ Hz and continues with uniform step $\Delta f = \frac{1}{NT} = \frac{f_{Sample}}{N}$. In other words frequency spectrum value k is $f_k = \frac{k}{NT} = k\frac{f_{Sample}}{N}$. Symbol f_S stands for sampling frequency of the signal and meaning of other symbols is the same as above.

3.2.4 Windowed Fourier Transform

Above described transforms have one big disadvantage for non-stationary signals. The result of the transform lacks any time information. This mean that Fourier transform can tell maximum about frequencies present in a signal but can't tell when these frequencies occurs. This shows a problem in all time $\rightarrow$ frequency transformations. You can get good frequency resolution and poor time resolution or vice versa, but not both of them at the same time. The time and frequency information in such transformation is constant and you can get more from one and less from the other. This is sometimes compared to Heisenberg uncertainty principle known from quantum physics. To address this issue with Fourier transformation the modification called the windowed (or short time) Fourier transform (WFT) was developed other approach to this problem shows wavelet transform described later in this work.

Basic idea of WFT is quite simple. If the signal is not stationary we can try to split the signal to parts that can be assumed to be stationary. We lost some frequency information but we got some time information. Exactly as Heisenberg uncertainty principle tells.

The part of the signal which we will be interested in is selected with so called window. With this window the new signal is produced with following formula :

$$x_w(k) = x(k) * w(k+\tau) \tag{3.4}$$

where
- $x(k)$ k-th value in signal.
- $w(k)$ k-th value in window function.
- τ translation of the window.
- $x_w(k)$ final signal.

And this new signal x_w we pass through Fourier transform. This new signal differs from the original one but its frequency spectrum is more equal to real spectrum.

There are several windows to choose. The choice

Basic types of windows are shown below: (Symbol N below means length of the window.)

Square window

$$w(k) = \begin{cases} 1, & \forall k \in \langle 0, \cdots, N \rangle \\ 0, & elsewhere \end{cases} \tag{3.5}$$

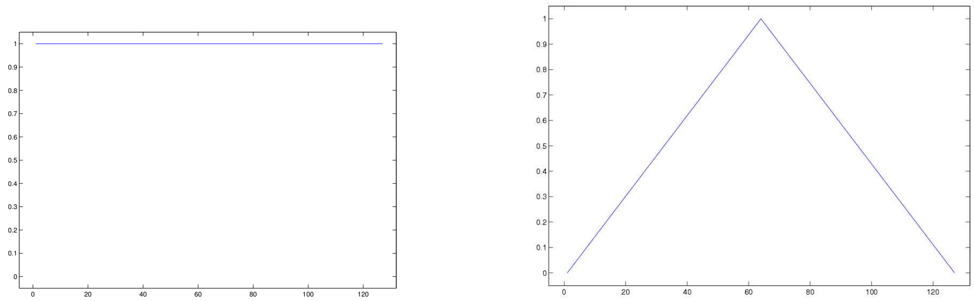

Figure 3.1: Square (left) and Bartlett's window

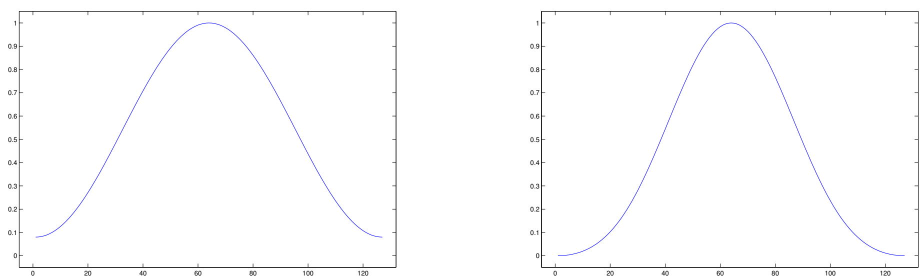

Figure 3.2: Hamming's (left) and Blackman's window

Bartlett's window

$$w(k) = \begin{cases} \frac{k}{N/2}, & 0 \le k \le N/2 \\ 2 - \frac{2k}{N/2}, & N/2 \le k \le N \\ 0, & elsewhere \end{cases} \tag{3.6}$$

Hamming's window

$$w(k) = \begin{cases} 0.54 - 0.46\cos\left(\frac{2\pi n}{N}\right), & \forall k \in \langle 0, \cdots, N \rangle \\ 0 & elsewhere \end{cases} \tag{3.7}$$

Blackman's window

$$w(k) = \begin{cases} 0.42 - 0.5\cos\left(2\pi\frac{k}{N}\right) + 0.08\cos\left(4\pi\frac{k}{N}\right), & \forall k \in \langle 0, \cdots, N \rangle \\ 0 & elsewhere \end{cases} \tag{3.8}$$

3.3 Wavelets

3.3.1 Introduction

As mentioned above in 3.2.4 there is a problem with frequency/time resolution in any time $\rightarrow$ frequency transform. One way for solving this problem is windowed Fourier transform. Another is shown in this section.

Wavelet transformation is quite young although its root goes deep in history. The very foundation of wavelet transformation was laid by Joseph Fourier in early 19th century when he discovered the Fourier transformation. The first mention of what we now call wavelet comes from Alfred Haar from 1909. The concept of wavelets in its present theoretical form comes from Jean Morlet and the team at the Marseilles Theoretical Physics Centre. The methods of wavelet analysis have been developed mainly by

Y. Meyer and his colleagues, who have ensured the methods' dissemination. The main algorithm dates back to the work of Stephane Mallat in 1988. Now wavelets are in focus of many researchers.

3.3.2 Continuous Wavelet Transformation

Continuous wavelet transformation (also CWT) is defined by following formula taken from [2].

$$CWT_x^{\tau}(s,\tau) = \frac{1}{\sqrt{s}} \int_{-\infty}^{+\infty} f(t)\Phi^* \left(\frac{t-\tau}{s}\right) dt \tag{3.9}$$

where

$f(t)$	is transformed function.
Φ^*	is so called Mother wavelet.
$s \in \mathcal{R}$	is scale.
$\tau \in \mathcal{R}$	is translation in time.

As you can see in the definition the Continuous wavelet transformation is an equation of two variables – scale and translation. The term wavelet means a small wave. The smallness refers to the condition that this (window) function is of finite length (or compactly supported). The wave refers to the condition that this function is oscillatory. The term mother wavelet implies that the functions with different region of support that are used in the transformation process are derived from this main function, or the mother wavelet. In other words, the mother wavelet is a prototype for generating the other window functions.

The term translation is used in same meaning as in Windowed Fourier transformation. It locates beginning of the mother wavelet.

The parameter scale in the wavelet theory is similar to the scale used in maps. As in the case of maps, high scales correspond to a non-detailed global view (of the signal), and low scales correspond to a detailed view. Similarly, in terms of frequency, low frequencies (and high scales) correspond to a global information of a signal, whereas high frequencies (low scales) correspond to a detailed information of a hidden pattern in the signal.

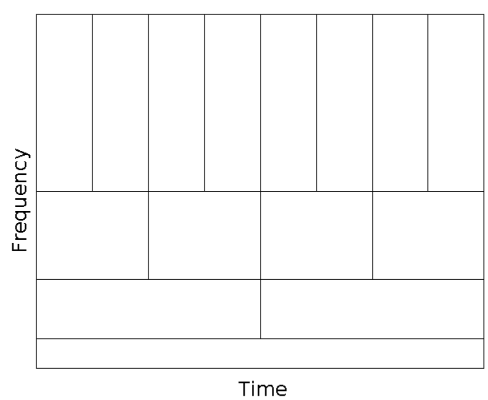

Figure 3.3: Explaining frequency and time resolution

The figure 3.3 is commonly used to explain how the time and frequency resolution should be interpreted. Each box in the picture presents wavelet transform value in time–frequency plane. Width of each box shows values of time represented by this box (or value of wavelet transform). Same height of each box shows values of frequency presented by this box. As may be seen, at low frequencies (or high scales), the height of the boxes are shorter, but their widths are longer. This corresponds to better frequency resolutions, since there is less ambiguity regarding the value of the exact frequency and there is poor time resolution, since more ambiguity regarding the value of the exact time.

Regardless of the dimensions of the boxes, the areas of all boxes are the same and determined by Heisenberg's inequality. As a summary, the area of a box is fixed for each mother wavelet, whereas different mother wavelets can result in different areas. However, all areas are lower bounded by $\frac{1}{4}\pi$. That is, we cannot reduce the areas of the boxes as much as we want due to the Heisenberg's uncertainty principle. On the other hand, for a given mother wavelet the dimensions of the boxes can be changed, while keeping the area the same. And so the time and frequency resolution may be controlled.

If the continuous wavelet transform is applied to discrete signals the definition formula 3.9 must be redefined in appropriate way :

$$CWT_s^\tau(n) = \frac{1}{\sqrt{s}} \sum_{n \in \mathcal{Z}} f(n) \Psi^* \left(\frac{n - \tau}{s} \right) \tag{3.10}$$

The meaning of all symbols is the same as in the definition 3.9. Just translation τ can't be now a real number, but must be multiply of whole number (you can't shift less than one sample).

3.3.3 Discrete Wavelet Transformation

For time discrete signals the computation of continuous wavelet transform with formula 3.10 is highly ineffective. So another method called discrete wavelet transform was developed. This method tries to minimise computational time with preserving maximum time–frequency information.

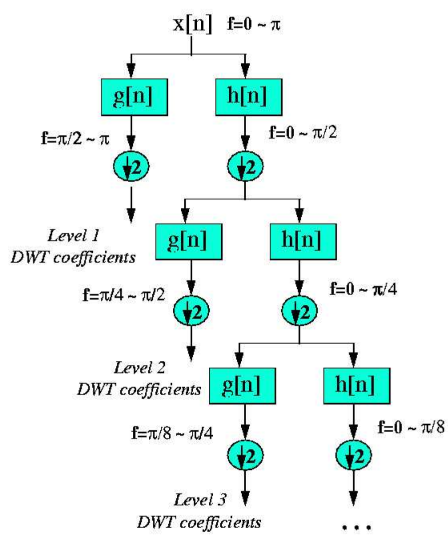

Figure 3.4: Discrete wavelet transformation scheme

The computational scheme is easy to understand. It is drawn on figure 3.4. In every level following procedure is performed.

1. The signal pass through half band lowpass and highpass filter. So you got two signals one with higher half of frequency spectrum and the other with lower half.
2. These signals are subsampled by 2. This mean every second sample from both signals is dropped.
3. The signal with lower part of frequencies is input to next step of discrete wavelet transformation, while signal with higher frequencies are discrete wavelet transformation coefficients on this level.

This procedure is applied until enough levels are computed or until there is a signal to work on. It is easy to understand that in every level it is doubled frequency precision while time accuracy is halved.

3.4 Hjorth's Parameters

In this section I will briefly describe three parameters introduced by Bo Hjorth in [12] in 1970. He tries to find new parameters which describe the EEG signal. These parameters are usually used for small slices of signal and are easy to compute. Both these reasons allows Hjorth's parameters to be computed on-line. These parameters are defined through momentum of random variable and are named Activity, Mobility and Complexity.

3.4.1 Activity, Mobility, Complexity

Here is the short description of each parameter:

Activity giving a measure of squared standard deviation of amplitude (in statistics is called variance or mean power).

Mobility giving a measure of standard deviation of slope with reference to the standard deviation of the amplitude. It is expressed as a ratio per time unit and may be conceived also as mean frequency.

Complexity giving a measure of excessive details, with reference to the softness of curve shape. It is expressed as the number of standard slopes actually generated during the average time required for generation of one standard amplitude as given by mobility.

If we want a mathematical description we must first define a spectral moment. A spectral moment m of order n is defined as

$$m_n = \int_{-\infty}^{+\infty} \omega^n S(\omega) d\omega$$

Where
ω is frequency.
$S(\omega)$ is power spectrum.

Also it is shown in [12] that m_0 is equal to variance of the signal σ_a^2. And further spectral moment of even order are equal to the variance of higher signal derivations ($m_2 = \sigma_{x'}^2$, $m_4 = \sigma_{x''}^2$, etc...).

Activity is simply equal to zero spectral moment m_0 or signal variance in time domain σ_x^2. Better $activity = m_0 = \sigma_x^2$.

Mobility is a square root of division m_2 and m_0/ In other words $mobility = \sqrt{\frac{m_2}{m_0}} = \frac{\sigma_{x'}}{\sigma_x}$.

Complexity is a square root of division m_4, m_0 and m_2. Better $complexity = \sqrt{\frac{m_4/m_2}{m_2/m_0}} = \frac{\sigma_{x''}/\sigma_{x'}}{\sigma_{x'}/\sigma_x}$.

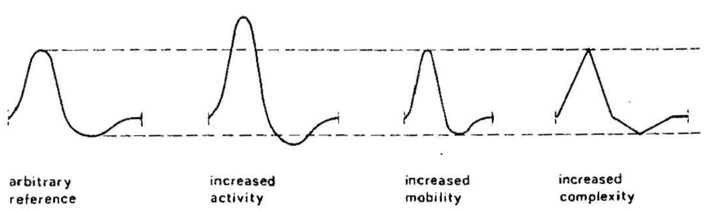

Figure 3.5: Characteristic changes of curve shape, illustrating the dependence of individual parameters.

More information on this area can be found in original article [12].

3.5 Important Volume Changes

This was introduced by A. Sporka in [14].

Lets define a small range of amplitudes and refer it as a window. The window will be shifted along the signal and it will generate events as described in following algorithm:

1. Place the centre of the window the first sample in the signal.
2. Move the window by one sample right.
3. If the sample is out of the window (its amplitude is out of the range covered by the window), then generate an event and shift the centre the window to be equal to the amplitude of current sample.
4. Continue by step 2 until the end of the signal.

When the algorithm is over, the number of events is counted and acts as an feature. For details on this see [14].

3.6 Function Approximation with Rational and Polynomial Functions

Sometimes it is good idea to replace one function with another. The reason for doing this may be that the original function is too complex to compute or some other characteristics are needed which the original function does not have (high order derivations, integral) or the original function is not known, but is given as a list of points in space. Than the interpolation or approximation is need. Difference between interpolation and approximation is that interpolation curve goes right through given points. In case of approximation it is not necessary.

The approximation function is denoted as $\phi(x)$ and it is a function with many parameters. The task of approximation is to set parameters to values which produce a minimum of error function which is:

$$err = \sum \|f(x_i) - \phi(x_i)\|$$

There are many methods how to find and set up approximation function ϕ. I will outline ideas of two approaches I used in my work.

3.6.1 Polynomial Approximation

In this case the approximation function ϕ is a polynomial of n-th degree – $\phi = p_0 + xp_1 + \cdots x^n p_n$. And m denotes number of known points of original function.

The error function to minimise is

$$err = \sum_{i=1}^{m} \mid f(x_i) - \sum_{j=0}^{n} x_i^j p_j \mid$$

Now we want to find a minimum of this err function. It may be done with mathematical analysis – we look for point $(p_0, p_1, \cdots, p_n)$ where

$$\frac{\partial err}{\partial p_k} = 0$$

for all $k = 0, \cdots, n$.

Now the minimum condition can be rewritten as

$$\sum_{i=0}^{n} y_i x_i^k = \sum_{i=0}^{m} \sum_{j=0}^{n} x_i^{j+k} p_j$$

for all $k = 0, \cdots, n$.

More informations may be found in [28].

3.6.2 Rational Approximation

In previous paragraph I outlined how to approximate a function with a polynomial. But for some functions and some intervals, the optimal rational function approximation is able to achieve higher accuracy than the optimal polynomial approximation with the same number of coefficients.

Let the final rational function be a fraction of two polynomials with numerator of degree m and denominator of degree k :

$$R(x) \equiv \frac{p_0 + xp_1 + \cdots + x^m p_m}{1 + xq_1 + \cdots + x^k q_k}$$

And the task is to set parameters p_0, $p_1 \cdots p_m$, $q_1 \cdots q_k$ to values which ensure minimum error of approximation:

$$err = R(x) - f(x)$$

where $f(x)$ is approximated function.

One way how to find coefficients p_i and q_j is to find minimum of following formula

$$p_0 + x_i p_1 + \cdots + x_i{}^m p_m = [f(x_i) \pm err](1 + x_i q_1 + \cdots + x_i{}^k q_k) i = 1, 2, \cdots m + k + 2$$

And above formula is searched iteratively for local extrema with this algorithm:

1. Find an initial rational function with $m + k + 2$ extrema x_i.
2. Solve above formula for new rational coefficients p_i, q_j and r.
3. Evaluate the resulting R(x) to find its actual extrema (which will not be the same as the guessed values).
4. Replace each guessed value with the nearest actual extreme of the same sign.
5. Go back to step 2 and iterate to convergence.

And desired result of this algorithm is global minimum of error err (and resulting best approximation of $f(x)$). Under a broad set of assumptions presented in [29], this method will converge.
More informations may also be found in [28].

3.7 Statistic

Here I will introduce some statistical parameters I will use for describing the signal. Parts of statistical theory I will also use for feature selection and some parts will be used in other parts of theory.

3.7.1 Basics

Random Variable

Random variable is mapping $X : \Omega \rightarrow \mathcal{R}$ which for $\forall t \in \mathcal{R}$ holds

$$X^{-1}((-\infty, t\rangle) = \omega \in \Omega | X(\omega) \leq t \in \mathcal{A}$$

Where $\mathcal{A}$ is set of events that can occur.

Probability

First lets define what is a relative frequency of an event. It is defined in very natural way. Let N be number of trials and n_i number of occurrences of event ω_i. Then relative frequency of event: ω_i is

$$P(\omega_i) = \frac{n_i}{N}$$

If the number of trials N rise to positive infinity then the relative frequency converge to a probability of event occurrence.

$$p(\omega_i) = \lim_{n \to +\infty} \frac{n_i}{N}$$

Distribution Function and Density Function

Let X be a random variable. Then a real function $F(t)$ defined for $\forall t \in \mathcal{R}$ by formula

$$F(t) = P\left[X \in (-\infty, t\rangle\right] = P[X \leq t]$$

we call a *Distribution function* of random variable X.
This formula show how probable is are events between "most left-one" and the selected. How the "most left-one" and selected events are determined and which lies events between is on will of the user.
The distribution function must hold following properties:

- $0 \leq F(t) \leq 1$.
- $F(t1) \leq F(t2)$, if $t1 \leq t2$.

- $\lim_{t \to -\infty} F(t) = 0$ and $\lim_{t \to +\infty} F(t) = 1$.
- F is upper semi-continuous function.
- F have at most countable number of point of discontinuity.

The probability distributions can be divided into two groups. This groups are – continuous and discrete. The definition of discrete probability distribution is as follows:

Distribution of random variable X is called discrete if exists a finite or countable set of real numbers $\{t_1, t_2, ...\}$ and for all t_i from such set is $P[X = t_i] = p_i > 0$ and distribution function F of random variable X is

$$F(t) = \sum_{i: t_i \leq t} p_i$$

Function $P(t) = P[X = t]$ defined $\forall t \in \mathcal{R}$ is called *Probability function* of random variable X.

And all probability distributions that don't fit this definition are continuous. And for continuous functions we can define a density function as follows:

Let the $F(t)$ be a distribution function of random variable X. Then if exists positive real function $f(t)$ for which in $\forall x \in \mathcal{R}$ is valid statement

$$F(x) = \int_{-\infty}^{x} f(t) dt$$

Then $f(t)$ is called a density function of random variable X.

Above I have spoken only about probability of events from range $(-\infty, x\rangle$ but of course one can compute a probability of some interval range of events. The formula is rather simple

$$P[u \leq X \leq v] = (1 - P[v \leq X]) - P[u \leq X]$$

Just small note on what is a sample and a population. The population is a set of all objects or people we want to study or describe. It may be infinite. The sample is a (selected) subset of a population.

3.7.2 Statistic parameters

Mean

The term *mean* denotes two things, which are slightly different

- The first it is the arithmetic average value. It is also called *sample mean* or expected value.
- The second is the expected value of a random variable, which is also called the *population mean* .

In all following text I will use the term *mean* in the first meaning (as the average or sample mean).

This sample mean can be computed by very well known formula

$$avg(X) = \bar{X} = \frac{1}{N} \sum_{i=1}^{N} x_i$$

.

Just for completeness I will show the formula for the population mean. The formula assumes that the distribution function is known. For discrete distribution functions the formula is

$$EX = \sum_{i} p_i x_i$$

where p_i is probability of event i and x_i is the value of the event.

For continuous distribution functions the population mean can be computed as follows

$$EX = \int_{-\infty}^{+\infty} x f(x) dx$$

Where $f(x)$ is a density function.

Variance

For sample variance the formula is

$$\sigma_X^2 = var(X) = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \bar{X})^2$$

In plain language, it can be expressed as "The average of the square of the distance of each data point from the mean".

And the population variance is defined this way:

$$\sigma_X^2 = var(X) = E((X - \mu)^2)$$

Standard deviation

For the sample standard deviation the formula is as follows:

$$\sigma_X = \sqrt{var(X)} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(X_i - \bar{X})^2}$$

It is just square root of variance.

Median

Let $f(x)$ be a distribution density function, then a median value m is defined this way

$$P(X \leq m) = \int_{-\infty}^{m} f(x)dx = 0.5$$

In other words the median is a value which divides the population (or sample) into two equal sized parts.

r-th central moment

R-th central moment is a arithmetic mean of r-th powered differences between events and sample mean. In formula

$$M_r^{'} = \frac{1}{N}\sum_{i=1}^{N}(X_i - \bar{X})^r, forr = 2, 3, 4, ...$$

.

Kurtosis

Kurtosis is a measure of the wideness of the probability distribution. A high kurtosis distribution has a sharper "peak" and fatter "tails", while a low kurtosis distribution has a more rounded peak with wider "shoulders".

In mathematical expression it is

$$\nu_2 = \frac{m_4}{m_2^2}$$

Where μ_4 is fourth central moment and m_2^2 is variance.

Skewness

Skewness is a measure of the asymmetry of the probability distribution of a real-valued random variable. Roughly speaking, a distribution has positive skew (right-skewed) if the higher tail is longer and negative skew (left-skewed) if the lower tail is longer (confusing the two is a common error).

Written in mathematical language the skewness is defined as

$$\nu_1 = \frac{\mu_3}{\sigma^3}$$

Where μ_3 is third central moment and σ is standard deviation.

3.7.3 Normal Probability Distribution

There are many probability distributions but one is very common in the nature. It is so called Normal distribution.

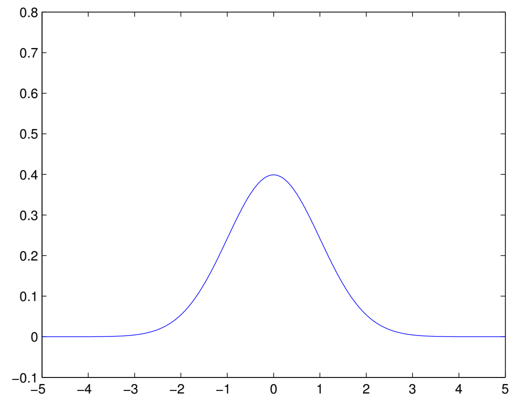

Figure 3.6: Normal distribution density

Its probability denstition function is on figure 3.6. This shape is often called Gaussian. The mathematical formula for density function is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \forall x \in \mathcal{R}$$

The distribution function is

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$

Note that $f(x)$ have no primitive function and $F(x)$ is not Riemann integrable and its values can be obtained only by numerical integration.

The normal distribution has two parameters μ and σ. The μ parameter is a mean and it is controls the position of the maximum of the density function. The σ is standard deviation and controls the wideness of the Gaussian.

3.7.4 Interval Parameter Estimations

Very often the data are known but the parameters of the distribution function is unknown, but there is a desire to know them for further work – for example for hypothesis testing. If the whole population is known then there is no problem to obtain exact value, but if a population is not known or is impossible to obtain, some estimation is needed. Because of random nature of obtained values the exact value of the parameter may vary. Very often the values are assumed to be from Normal distribution and the estimated parameter is a mean (or μ) of the normal distribution. Now, if the interval centred around the mean is taken then it covers a range of values which occurs with some probability (see the figure 3.7). This range is called *Confidence interval* . The range grows in length with rising probability it covers. The confidence interval is often chosen to cover 95% of values. The value range also depends on number of values in the sample and more values means narrower interval because it is less probable to obtain many values with low probability of occurrence. The example is on the figure 3.8 below. The green distribution is far less probable then the blue.

Yet the probability that the data are from the green distribution (see figure 3.8) is not equal to zero although it is very near to it. This means that the data may (with almost zero probability) be from the green distribution but almost certainly are not.

Now I have talked enough and its time for some mathematics. I will use only normal distribution mean estimation in my work, so I will put in formulas only for estimation of mean values. First the formula for population mean interval estimation is usefull in situation when the variance of the data is known

$$P(\mu - z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \bar{X} < \mu + z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}) = 1 - \alpha$$

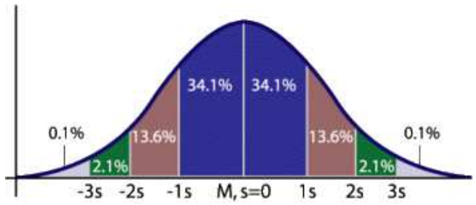

Figure 3.7: Probabilities of Normal distribution ranges

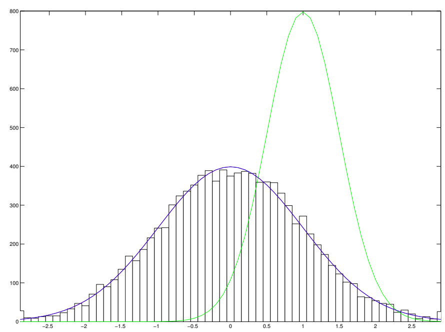

Figure 3.8: Examples of probable and non probable distributions

where

μ	is population mean.
$\bar{X}$	is sample mean.
n	is number of samples.
α	is significance level.
$1-\alpha$	is confidence coefficient.
$z_{1-\frac{\alpha}{2}}$	is $1-\frac{\alpha}{2}$ fractile of normal distribution.

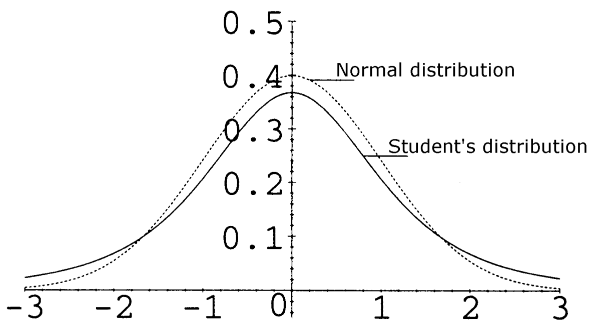

Figure 3.9: Comparison of Normal distribution and Student's distribution with three degrees of freedom

But more often situation is that the population variance is also unknown and is replaced by sample variance. But then instead of normal distribution the Student's distribution must be used. It have almost similar shape to normal distribution but is a little wider (see figure 3.9). It also introduce new parameter called degrees of freedom. The formula for this case is

$$\mu = \bar{X} \pm t_{n-1;\left(1-\frac{\alpha}{2}\right)} \frac{s}{\sqrt{n}}$$

where $t_{n-1;\left(1-\frac{\alpha}{2}\right)}$ is $1-\frac{\alpha}{2}$ fractile of Student's distribution with $n-1$ degrees of freedom. All other symbols are the same as above.

Chapter 4

Clustering Methods

The clustering is, in my work, unsupervised learning problem and its goal is to find a structure in a set of unlabelled data. A loose definition of clustering could be "the process of organising objects into groups whose members are similar in some way". A cluster is therefore a collection of objects which are somehow "similar" and are "dissimilar" to the objects belonging to other clusters. On figure 4.1 you can see a little example.

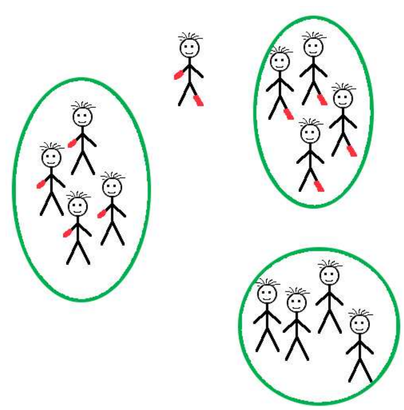

Figure 4.1: Example of clustering

The object is assigned to a cluster according to some clustering criterion. It can be shown that there is no absolute "best" clustering criterion (for example distance in space, broken bone on image above, value of some function, etc.) which would be independent of the final aim of the clustering. Consequently, it is the user which must supply this clustering criterion, in such a way that the result of the clustering will suit his needs.

4.1 Vector Quantisation

The goal of Vector quantisation is to approximate propability density $p(\mathbf{x})$ of input vectors $\mathbf{x} \in \mathcal{R}^n$ with finite set of so called code vectors (sometimes codeword or code book vector) $\mathbf{w_i} \in \mathcal{R}^n$ $i \in 1, \cdots, n$. If we have set of code vectors we assign to every $\mathbf{x}$ it's code vector $\mathbf{w_c}$ which satisfies following formula:

$$c = \arg \min_{i=1,\cdots,n} \{\|\mathbf{x} - \mathbf{w_i}\|\}.$$

Where $\|\mathbf{x}\|$ is suitable norm (most propably Euclidean norm).

$$E = \int \|\mathbf{x} - \mathbf{w_c}\|^2 p(\mathbf{x}) d\mathbf{x} \tag{4.1}$$

One way to find the placing of code vectors is to minimise error of vector quantisation. Error is defined by formula 4.1.

The $\mathbf{w_c}$ is code vector for $\mathbf{x}$. But in real world the propability density is often unknown and the vector quantisation problem is defined by learning set $T = \{\mathbf{x}^{(t)}, t = 1, \cdots, k\}$. In this case the error is defined with following formula :

$$E = \frac{1}{k} \sum_{i=1}^{k} \|\mathbf{x}^{(t)} - \mathbf{w_c}\|^2 \tag{4.2}$$

4.2 K-Means

One of the methods which solve vector quantisation problem is called K-Means algorithm. This algorithm was published in 1967 by MacQueen [5]. As every vector quantisation problem solver it tries to locate code vectors that presents every cluster in data. The optimality of the solution is measured with error function 4.2.

The algorithm in its simplest form consists from following steps:

1. *Place K points into space represented by the objects that are being clustered. These vectors are initial code vectors ($\mathbf{w_c}$).*
2. *Assign each object $\mathbf{x_i}$ to the group presented by the nearest code vector.*
3. *When all objects are in some group. Recalculate positions of code vectors to decrease the error function.*
4. *Repeat step 2 and 3 until error function decreases.*

Although that the algorithm stops in finite time, it can stop in suboptimal solution and the solution often depends on initial solution. To reduce this effect it is advised to run this algorithm several times with different initial configuration.

Chapter 5

Classification

The task of Classification is on first sight very similar to Clustering. But the goal is different. In clustering tries to find similarity in unlabelled object and group them together. The goal of classification is to find patterns which describe the apriori known group. This pattern can help to understand already known data and to predict behaviour of new instances.

The text about classifiers implemented in Weka is inspired by [27] and [11]. Part about GAME is taken from [9].

5.1 Naive Bayesian classificator

Naive Bayes classifier provides a simple approach and clear semantic to representing, using and learning probabilistic knowledge.

This classifier make two important assumptions about the data. First it assumes that the predictive attributes are conditionally independent given class. The second is that no hidden or latent attributes influence the classification process. These assumptions allow very efficient learning and classification algorithms.

The classification process is quite simple. After the vector of attributes values $\mathbf{x}$ is made, the Bayes rule is computed to obtain probability of each class. The Bayes rule is

$$p(C = c \mid \mathbf{X} = \mathbf{x}) = \frac{p(C = c)p(\mathbf{X} = \mathbf{x} \mid C = c)}{p(\mathbf{X} = \mathbf{x})}$$

where :

$\mathbf{X}$	is vector of random variables denoting the observed values.
$\mathbf{x}$	is vector of observed values.
C	is a random variable denoting the class.
c	is label for one particular class.

After that, the most probable class is classified.

The training of this classifier is in finding the probabilities and parameters (mean and variance) of normal distribution. To obtain these parameters the maximum likelihood is often used.

More informations are in [27].

5.2 Bayesian Net

In many cases the Naive Bayes classifier as presented above is too simple and some more powerful model is needed. Often the probabilities are not dependent on only one fact, but may be modified by many facts. The influences may be presented in form of acyclic graph. An example of such graph is on figure 5.1. The nodes of this graph are components (or facts) of the system and in system on the figure 5.1 takes discrete values. The nodes are labelled as **A**, **B**, ... and their corresponding variables in lowercase letters. Each link in the net is directional and joins two nodes – it presents influence of one node upon another. Thus in net on figure 5.1 node **A** directly influences **D**, **B** influences **C**, ...

Suppose some Bayes network is available, then through a direct application of Bayes rule, the probability of any configuration of variables can be computed. To perform this computation we need so called

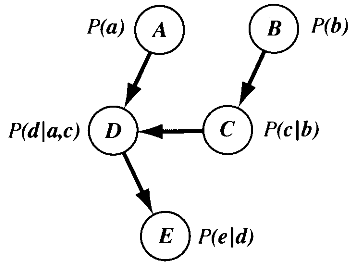

Figure 5.1: Example of Bayes network

conditional probability tables, which gives a probability of any variable at a node for each conditioning vent.

The learning and creating of the network goes beyond scope of my work. Further information may be found in [27] or [11].

5.3 Decision Tables

The decision table is very simple representation of hypotheses. The decision table consists of two components: a schema and a body. The schema is a set of features that are included in the table. The body contains labelled instances from the space defined by features from the scheme.

The new unknown instance is classified as follows – let $\mathcal{I}$ be the set of labelled instances in the table exactly matching given instance I. Only features in the scheme are required for match. If $\mathcal{I}$ is empty, then default (most probable) class in the table is returned. Elsewere the class with the most instances in $\mathcal{I}$ is returned.

The estimation of average error of the decision table is computed by following formula:

$$\widehat{err}(h,\tau) = \frac{1}{\|\tau\|} \sum_{(x_i,y_i)\in\tau} L(h(x_i), y_i)$$

where:

x_i	is an instance.
y_i	is actual class of instance x_i.
h	is hypothesis function (class assigned be the table).
$L(u,v)$	is a penalty function – penalty for deciding to classify the instance to class u while it belongs to class v.
τ	is testing set.

In learning process we try to find subset of features which minimises this error function

$$A^* = arg \min_{A^\dagger \subseteq A} err(DTM(A^\dagger, S), f)$$

where:

A	is set of features.
$A^\dagger$	is a current schema.
S	is a learning set.
$DTM(A^\dagger, S)$	is a decision table build over $A^\dagger$ and S.
f	is actual class assignment to instances.

The strategies how to find the optimal subset goes beyond this chapter and are discused below in chapter 6.

5.4 Simple Logistic

The linear logistic regression perform a least-squares fit of a set of parameters $\beta_0, \cdots, \beta_k$ to a numeric target variable to form a model

$$f(x) = \beta_0 + \sum_{i=1}^{k} x_i \beta_i$$

When we obtain the value $f(x)$ it says nothing about class the x belongs to. One method is to use posterior class probabilities $P(G = j \mid X = x)$ for the J classes via linear functions in x while at the same time ensuring they sum to one and remain in $< 0, 1 >$. The probabilities are computed as

$$P(G = j \mid X = x) = \frac{e^{F_j(x)}}{\sum_{k=1}^{J} e^{F_k(x)}} \text{where} \sum_{k=1}^{J} F_k(x) = 0$$

Where $F_j(X) = \beta_j^T * x$ and it is usually fit by maximum likehood estimates for the parameters β_j. And the assigned class is is the one with the highest posterior probability.

One way to find these estimates is based on the LogitBoost algorithm. It performs forward stage-wise fitting of additive logistic regression models. More for information please see [27].

5.5 Decision Tree

This technique is based on natural and intuitive idea to classify objects through a sequence of question in which the next question depends on answer to current question. This very useful for categorical and ordinal data, but can be extended also to contnuous variables. Such sequence of questions can be displayed as a direct decision tree. Root node and inner nodes are the questions and leaf nodes are classes to which evaluated object belongs. The links are answers to the these questions.

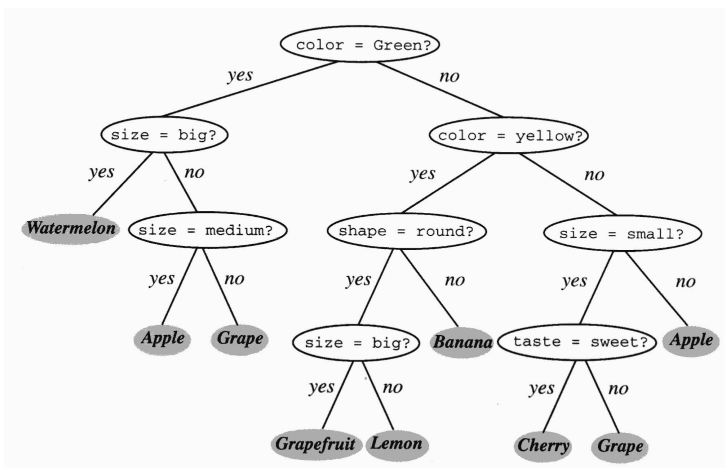

Figure 5.2: Example of decision tree

On the picture 5.5 (taken from [11]) an example of such tree can be seen. Things to decide about in the example are pieces of fruit. The features in this case are colour, size, shape and taste.

The building of this tree contain only one question – how to pick the question asked in node. The most popular approach is to select features to ask according to their information entropy or information gain. For more details see [27] or [11].

5.6 GAME

The GAME combines two of three main methods of softcomputing – the neural network and the genetic algorithms [1]. During the learning process the layers are added iteratively until the output of the network shows too high error rate. In each layer a lot of neural cells are generated (with help of genetic algorithm) and the best cells are selected to the network.

[1]For basic informations on neural network see [6] or [8], for informations on genetic algorithms see [7].

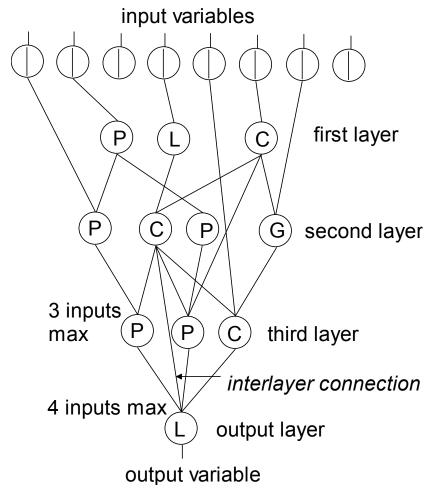

Figure 5.3: a structure of sample GAME neural network

The figure 5.3 shows a structure of sample neural network. On picture you may see a network with several hidden layers created by selecting successful units of different kind.

The GAME network build this way can provide several data mining tasks. The less common are shown on figure 5.4. I will use one shown there – feature ranking and one not shown – classification.

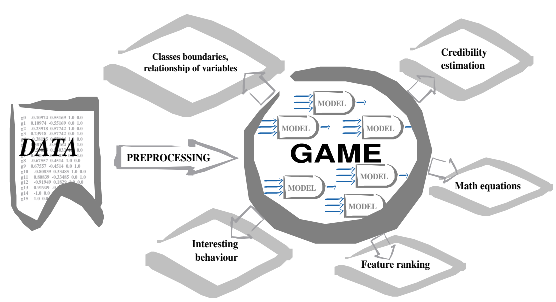

Figure 5.4: Less common data mining tasks the GAME provides.

The recall of the network is the same as in case of any other feed forward neural network. See [8] for details on artificial neural networks. For more information on GAME see [9].

The network can work as an classifier in way all others neural networks do. It learns to react with some output for each instance in learning set, then in classification stage the unknown instance is given to the network it again produce some output and the output is decoded as some class the instance belongs to.

Chapter 6

Feature Selection and Feature Ranking

6.1 Introduction

Above in chapter 3 we discuss the problem how to extract features from a signal. But if too many features are provided to the classification (or any other algorithm) the phenomenon called curse of dimensionality can occur. On the other hand if too there are too little features they may not carry enough information for softcomputing methods to decide correctly. The task to find proper number of features is a task for feature selection algorithms.

In practical applications, it is impossible to obtain complete set of relevant variables. Therefore, the modelled system is open system, and all important variables that are not included in the data set (for what reason ever) are summarised as noise [31].

In fact, even if theoretically more variables should provide one with better modelling accuracy, in real cases it has been observed many times that this is not the case. Successful models seem to be good balance of model complexity and available information. In fact, variables selection tends to produce models that are simpler, clearer, computationally less expensive and, moreover, providing often better prediction accuracy [27].

6.2 GAME

The GAME uses genetic algorithm to build a neural network. It is possible to say that while the genetic algorithm is building the network, it performs a feature ranking. The feature ranking (according to [22]) works this way – in the initial population of the first layer units are randomly generated. Connection to certain input variable is represented as in corresponding gene locus. Numbers of ones in locus are therefore uniformly distributed at the beginning of genetic algorithm. After several epochs of genetic algorithm, numbers of ones in gene loci representing more important input variables increases whereas numbers in loci of least significant variables decreases.

This fact can be used for the estimation of variable's significance. In each layer of the network, after the last epoch of genetic algorithm, before the best gene from each niche is selected, we count how many genes (units) are connected to each input variable. This number is accumulated for each input variable and when divided by sum of accumulated numbers for all input variables, we get the proportional significance of each input variable.

6.3 Weka

In Weka the feature selection process is divided into two separate processes. One is feature search method and the second is feature subset evaluator. Here is their brief description of used methods. It is taken from [27].

6.3.1 Feature Search Methods

Search methods traverse the attribute space to find a good subset. Quality is measured by the chosen feature subset evaluator described later.

BestFirst

BestFirst performs greedy hill climbing with backtracking. The adjustable parameter is how many consecutive nonimproving nodes must be encountered before the system backtracks. The method can search forward from the empty set of attributes, backward from the full set, or start at an intermediate point (specified by a list of features) and search in both directions by considering all possible single-feature additions and deletions.

GeneticSearch

GeneticSearch uses a simple genetic algorithm [30]. Parameters include population size, number of generations, and probabilities of crossover and mutation. The list of features can be specified as the starting point, which becomes a member of the initial population.

Ranker

Ranker is not a search method, in meaning of above two methods, but is a ranking scheme for individual attributes. It sorts attributes by their individual evaluations and can be used only with some selected attribute evaluators.

Ranker not only ranks attributes but also performs attribute selection by removing the lower-ranking ones. The cutoff threshold may be specified below which attributes are discarded, or number of retained features may be specified.

6.3.2 Feature Subset Evaluators

Subset evaluators take a subset of features and return a number which measure a quality of the subset and guides the further search.

CfsSubsetEval

CfsSubsetEval assesses the predictive ability of each feature individually and the degree of redundancy among them. It preferres sets of features that are highly correlated with the class but are not correlated with another featrues. An option iteratively adds attributes that have the highest correlation with the class, provided that the set does not already contain an attribute whose correlation with the attribute in question is even higher.

Following feature evaluators are used with Ranker search method to generate a ranked list from which Ranker discards a given number of features.

InfoGainAttributeEval

InfoGainAttributeEval evaluates features by measuring their information gain with respect to the class. It discretises numeric features first. This method, along with the next three, can treat missing as a separate value or distribute the counts among other values in proportion to their frequency.

ChiSquaredAttributeEval

ChiSquaredAttributeEval evaluates features by computing the chi-squared statistic with respect to the class.

GainRatioAttributeEval

GainRatioAttributeEval evaluates features by measuring their gain ratio with respect to the class.

Part II

Practical Part

Chapter 7

Goals to Achieve and Proposed Concept

The Sleep EEG is a very important medicine tool for many clinical applications. To take full advantage of sleep EEG it is necessary to know so called Sleep stages. These Sleep stages was defined in 1967 by Association for the Psychophysiological Study of Sleep and they guide for scoring the sleep called Manual for Visual Sleep Scoring [20] is used ever since. Sleep scoring by a human expert is a very time consuming task and normally could require hours to classify a whole night recording. In past 20 years many works tried to automate the sleep scoring. Most of these works uses only Fourier transformation to extract features from data.

*In this work I will try to find if features extracted with another feature extraction methods can not produce features which distinguish sleep stages better.*To answer this problem I propose a concept consisting of following steps:

- Data acquisition
- Preprocessing
- Feature extraction
- Feature ranking and feature selection with several methods
- Evaluation of selected features with several classifiers
- Conclusions

For illustration the concept is showed on figure 7.1. It shows section numbers of each step in sleep stages classification experiment.

This concept is inspired by CRISP Data mining industry standard. The concept presented here covers the data part of the CRISP. The business and data understanding was the part of prior education to this work and naturally are not covered by the concept. Deployment, which is also not covered, is planed in the future in Faculty of medicine in Hradec Králové.

Before I use this concept on sleep EEG data, I will evaluate it on a bit simpler data. The goal of this evaluation is to find weak places, properties and problems of my concept. The simpler data are recordings of patient breathing during his/her sleep.

After that if the concept will prove functional I will process the Sleep EEG data. Based on results I will try to find the answer on the question I asked in the introduction of my work.

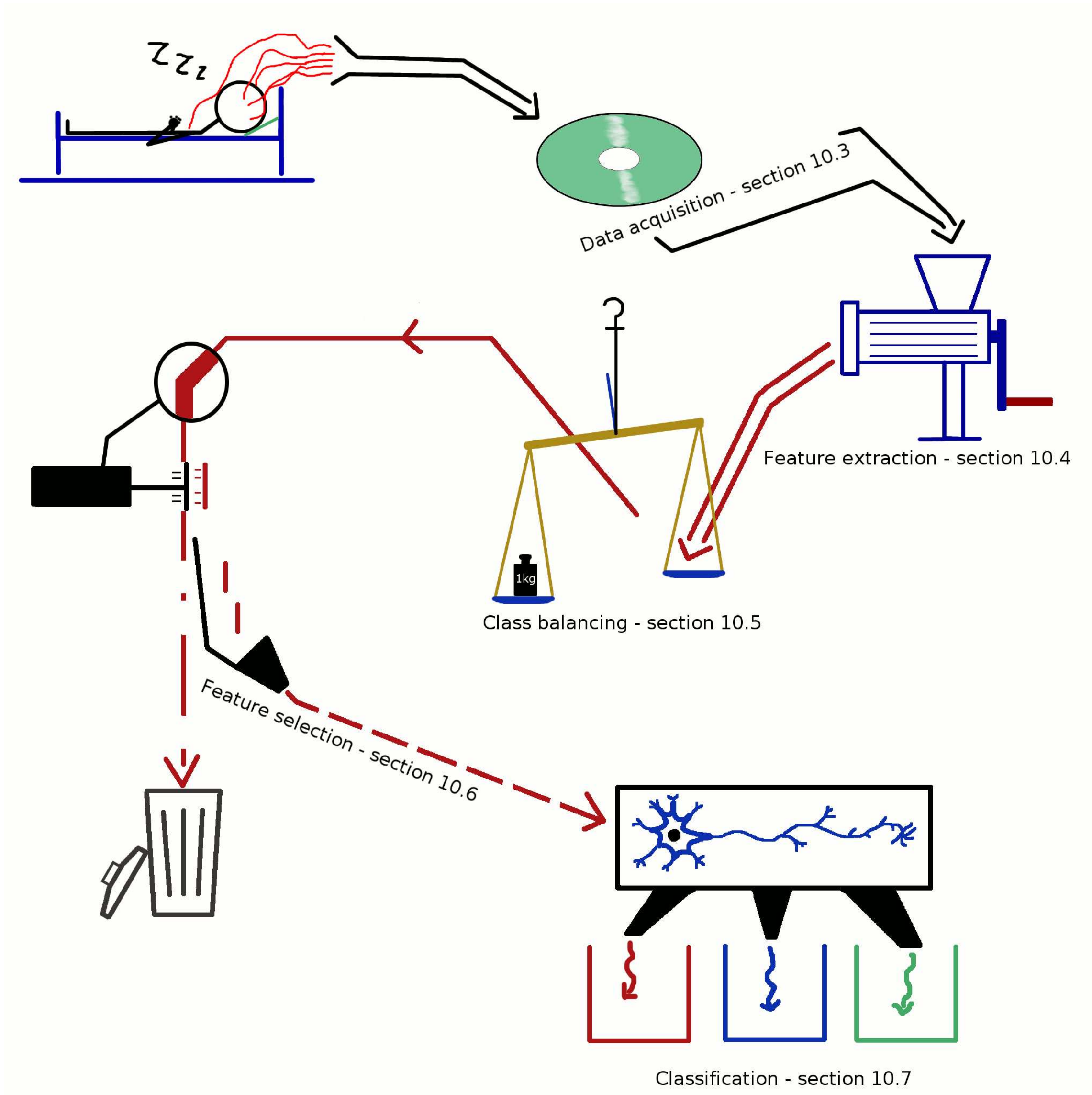

Figure 7.1: Concept illustration – number of section refers to application of the concept on MIT sleep data.

Chapter 8

List of Tested Features

In this chapter I will show list of formulas I will use to extract features. The list of formulas presented here is shortened a little. To see complete list of features see appendix B on page 87.

A short summary of theory needed to compute these equations is in above chapters in Theoretical part of my thesis.

8.1 Note on Symbols

In following sections I will use some common symbols. Here I will describe its meaning.

N	Number of samples in currently processed window.
x	all samples in time domain window.
x'	first derivation of signal x.
x''	second derivation of signal x.
x_i	i-th sample in time domain signal.
$CWT(x)_i$	i-th value in continuous wavelet transformation of signal x.
$FFT(x)_i$	i-th value in Fourier transformation of signal x.

8.2 Features in Josef Rieger's work

Here is list of features used by Ing. Josef Rieger in his diploma thesis [4].

Average absolute amplitude

$$A_{mean\ abs} = \frac{1}{N}\sum_{j=1}^{N}|x_j|$$

Average amplitude

$$A_{mean} = \frac{1}{N}\sum_{j=1}^{N}x_j$$

Maximum positive amplitude

$$A_{max+} = \max_{j=1,\cdots,N} x_i$$

Maximum negative amplitude

$$A_{max-} = -\max_{j=1,\cdots,N}(-x_j)$$

Variance

$$varX = \frac{1}{N}\sum_{j=1}^{N}(x_j - \bar{x})^2$$

Skewness

$$\gamma_1 = \frac{\mu_3}{\sigma^3} = \frac{\mu_3}{(\mu_2)^{\frac{3}{2}}}$$

Kurtosis

$$\gamma_2 = \frac{\mu_4}{\sigma^4} - 3 = \frac{\mu_4}{(\mu_2)^2} - 3$$

Average absolute first derivation

$$D^{(1)}_{mean\ abs} = \frac{1}{T_{vz}(N-1)} \sum_{j=2}^{N} |x_j - x_{j-1}|$$

Absolute maximum of first derivation

$$D^{(1)}_{max\ abs} = \frac{1}{T_{vz}} \max_{j=2,\cdots,N} |x_j - x_{j-1}|$$

Absolute maximum of second derivation

$$D^{(2)}_{max\ abs} = \frac{1}{T_{vz}} \max_{j=2,\cdots,N} |x'_j - x'_{j-1}|$$

8.3 My Features

8.3.1 Continuous Wavelet Transformation Statistics

These features are selected statistics properties of continuous wavelet transform.

CWT_mean

$$CWT_{mean}(x) = \frac{1}{N} \sum_{j=1}^{N} CWT(x)_j$$

CWT_std_dev

$$CWT_{std\_dev}(x) = \sigma_{CWT} = \sqrt{\frac{1}{N-1} \sum_{j=1}^{N} \left(CWT(x)_j - \overline{CWT(x)} \right)^2}$$

CWT_variance

$$CWT_{var}(x) = \frac{1}{N-1} \sum_{j=1}^{N} \left(CWT(x)_j - \overline{CWT(x)} \right)^2$$

CWT_median

$$\mathcal{P}\left(\mathcal{X} \geq CWT_{median}\right) = \int_{-\infty}^{CWT_{median}} f(x)dx = 0.5$$

CWT_max

$$CWT_{max}(x) = \max_{j=1\cdots N} (CWT(x)_j)$$

CWT_min

$$CWT_{min}(x) = \min_{j=1\cdots N} (CWT(x)_j)$$

CWT_rms

$$CWT_{RMS}(x) = \frac{1}{N} \sum_{j=1}^{N} (CWT(x)_j^2)$$

CWT_skewness

$$\gamma_1 = \frac{\mu_3}{\sigma^3} = \frac{\mu_3}{(\mu_2)^{\frac{3}{2}}}$$

CWT_kurtosis

$$\gamma_2 = \frac{\mu_4}{\sigma^4} - 3 = \frac{\mu_4}{(\mu_2)^2} - 3$$

CWT_max2rms

$$CWT_{max2rms}(x) = \frac{CWT_{max}(x)}{CWT_{rms}(x)}$$

CWT_max2mean

$$CWT_{max2mean}(x) = \frac{CWT_{max}(x)}{CWT_{mean}(x)}$$

CWT_max2median

$$CWT_{max2median}(x) = \frac{CWT_{max}(x)}{CWT_{median}(x)}$$

8.3.2 Fourier Coefficients Statistics

These features are selected statistics properties of Fourier transform.

FFT_mean

$$FFT_{mean}(x) = \frac{1}{N}\sum_{j=1}^{N} FFT(x)_j$$

FFT_std_dev

$$FFT_{std\_dev}(x) = \sigma_{FFT} = \sqrt{\frac{1}{N-1}\sum_{j=1}^{N}\left(FFT(x)_j - \overline{FFT(x)}\right)^2}$$

FFT_variance

$$FFT_{var}(x) = \frac{1}{N-1}\sum_{j=1}^{N}\left(FFT(x)_j - \overline{FFT(x)}\right)^2$$

And so on as in the case of Wavelet transform. The complete list can be found in section B.3.2 on page 89.

8.3.3 Fourier Frequency Band Spectrum Energy

These features represents band powers of selected frequency ranges. The frequency ranges corresponding to the EEG frequency band division are specified in the beginning of my work.

FFT_band_0-2

$$FFT_{band\_0-2}(x) = \sum_{j \in 0...2\ Hz\ FFT\ coefitients} |fft(x)|$$

Formulas for power of bands 2-4 Hz, 4-6 Hz, 6-8 Hz, 8-10 Hz, 10-13 Hz, 13-19 Hz, 19-24 Hz, 24-30 Hz, 30-50 Hz can be found in section B.3.3 on page 90.

FFT_bandalpha

$$FFT_{band\_aplha}(x) = \sum_{j \in alpha\ Hz\ FFT\ coefitients} |fft(x)|$$

Formulas for power of regular EEG bands Beta, Gamma, Theta, Delta can be found in section B.3.3 on page 90.

FFT_band_delta2alpha_ratio

$$FFT_{band\_delta2theta\_ratio} = \frac{FFT_{band\_delta}(x)}{FFT_{band\_alpha}(x)}$$

Formulas for band power ratios delta/theta, delta/alpha, delta/beta, delta/gamma, theta/alpha, theta/beta, theta/gamma, alpha/beta, alpha/gamma, beta/gamma is in section B.3.3 on page 91.

8.3.4 Spectral Centroid

This term means the frequency which divides a spectrum power into two equal halves.

spectralcentroidfreq

$$\sum_{i \in 0 \ldots Spectral\_centroid\ Hz\ FT\ coefitients} |fft(x)| = \sum_{i \in Spectral\_centroid \ldots f_max\ Hz\ FT\ coefitients} |fft(x)|$$

8.3.5 Time Domain Features – Ratio Values of Time Signal

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1} N(x_i^2)}$$

Max2rms_ratio

$$max2rms\_ratio = \frac{\max(x)}{RMS}$$

Max2mean_ratio

$$max2rms\_ratio = \frac{\max(x)}{mean(x)}$$

Max2median_ratio

$$max2rms\_ratio = \frac{\max(x)}{median(x)}$$

8.3.6 Hjorth Parameters

hjorth_activity

$$activity = m_0 = \sigma_x^2$$

hjorth_mobility

$$mobility = \frac{\sigma_{x'}}{\sigma_x}$$

hjorth_complexity

$$complexity = \frac{\sigma_{x''}/\sigma_{x'}}{\sigma_{x'}/\sigma_x}$$

8.3.7 Important Volume Changes Detector

See 3.5 for details.

number_of_changes

$$number\_of\_changes = number\_of\_generated\_events$$

8.3.8 Polynomial Interpolation Parameters

Let P be a polynomial of 3th degree. $P(x) = \alpha_1 x^3 + \alpha_2 x^2 + \alpha_3 x + \alpha_4$. The polynomial will be fitted to the signal and α_1, α_2, α_3, α_4 will act as features this way.

interp_poly_p1

$$interp\_poly\_p1 = \alpha_1$$

interp_poly_p2

$$interp\_poly\_p2 = \alpha_2$$

Formulas for other interpolation coefitients is in section B.3.8 on page 93.

interp_poly_rms

$$interp\_poly\_rms = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - P(x_i))^2}$$

8.3.9 Rational Interpolation Parameters

Let R be a fraction of two polynomials of 3th degree. $R(x) = \frac{\alpha_1 x^3 + \alpha_2 x^2 + \alpha_3 x + \alpha_4}{x^3 + \beta_1 x^2 + \beta_2 x + \beta_3}$. The fraction of polynomials will be fitted to the signal and α_1, α_2, α_3, α_4, β_1, β_2, β_3 will act as features this way.

interp_rational_p1

$$interp\_rational\_p1 = \alpha_1$$

interp_rational_q1

$$interp\_rational\_q1 = \beta_1$$

Formulas for other interpolation coefitients is in section B.3.9 on page 93.

interp_poly_rms

$$interp\_poly\_rms = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - R(x_i))^2}$$

Chapter 9

Evaluation of Proposed Concept on Snoring Data

9.1 Introduction

In this chapter I will introduce an experiment with simple and easy to understand data. This experiment will be used to demonstrate the capabilities and problems of my concept. The proposed concept as I figured out in 7 consists of following steps : **Data acquisition** , **Preprocessing** , **Feature extraction** , **Feature ranking and feature selection with several methods** , **Evaluation of selected features with several classifiers** and **Evaluation Conclusion** . In each step of the concept I use several different methods which may contains specific errors and problems which, in combination with complex EEG data, may be more difficult to identify (or even unidentifiable). Also I want to gather some more experience with this concept and in case of problems, adjust it.

Data used in this experiment comes from Faculty of Medicine of Charles University in Hradec Králové and was recorded in Sleep laboratory on one patient. The experiment goal will be to distinguish between breathing states of a patient. The recording I use in this experiment We have the complete recording from one such examination, It consists from 8 hours of record and includes EEG, sleep position of the patient, microphone, breathing, etc... . For my experiment I will use just two signals from above mentioned and it will be Air flow and Microphone signals. The Air flow signal is measured in front of the patients mouth and shows direction and amount of air breathed in and out by patient and the Microphone signal which shows intensity of the noise in a patient sleep room. The breathing states I try to distinguish will be "Not-Breathing", "Breathing Regularly", "Breathing Irregularly" and "Snoring". The recording is not scored for this experiment so I will have to assign above mentioned classes to the data.

In concept step **Building Multiple Classifiers** I will not build classifiers which tries to identify all classes at once. Instead I will try to build separate classifiers which try to recognise one particular class. Other classes will be recognised as "others". Maybe a little better – I will build classifier which recognise two states – "Patient is snoring" (for example) and "Patient is not snoring".

9.2 Experimental Setup

The picture 9.1 shows the setup of my experiment.

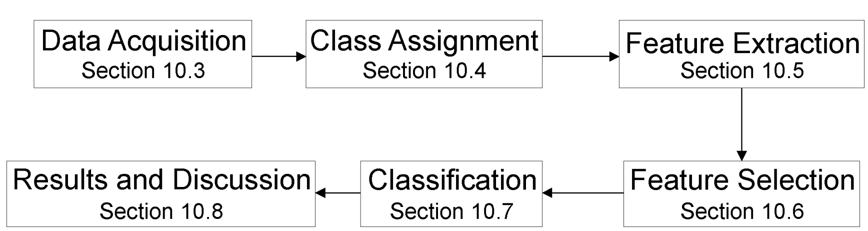

Figure 9.1: Evaluation of proposed concept flowchart

At first I will use my Class Assignment Tool described above to extract features from the signals and to assign classes to the signal windows. For the assignment I will not use any special method. I will assign the classes on my own will according to events in the time-domain signal. In the second step above described features will extracted from a signal. The third stage will be a feature ranking. I will try to determine which features are important for classifying each class. To do this I will use methods implemented in Weka tool and the GMDH and genetic methods implemented in the GAME (for details on both programs see the previous sections). The next step will be the building different models with selected features and I will again use Weka and GAME tools. Then I will evaluate these models using of number of missclassified classes. And finally some some conclusions of this experiment will be done.

9.3 Data Acquisition

As in the most cases the first step of experiment is to acquire the data. The data come to me in binary format called European Data Format (or EDF). The definition of this format can be found in [13]. This format is widely used in EEG recording systems. The task so was to import the breathing data from this EDF to my Matlab processing program – CATool (see A for details). Part of this program of mine is an import routine which reads data in EDF to CATool.

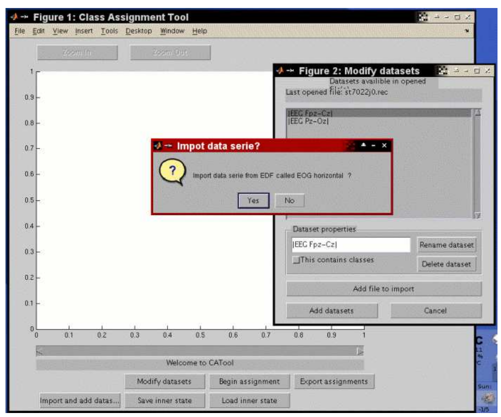

Figure 9.2: Importing the data to Class Assignment Tool

9.4 Breathing Stages Assignment

Here I will describe how I classified the signal. As usual I used my Class Assignment Tool (or CATool in short) wrote in Matlab for this task.

I have manually assigned the classes to the signal windows (see figures 9.3,9.4,9.5,9.6). Because I have no video nor any other information about the sleep, I used only time signals. I tried to hold the following simple rules for breathing stages assignment:

- Class "Not Breathing" – when in signal window is a little of air flow and small amount of microphone noise. Example signal shape is shown on figure 9.3
- Class "Breathing Irregularly" – when in signal window is some air flow, but it have no repeating shape. And amplitude of microphone is low. Example signal shape is shown on figure 9.4
- Class "Breathing Regularly" – when in signal window is some air flow and it have a repeating (sinusoid) shape. And amplitude of microphone is low. Example signal shape is shown on figure 9.5
- Class "Snoring" – when there is a lot of noise in the window. Example signal shape is shown on figure 9.6

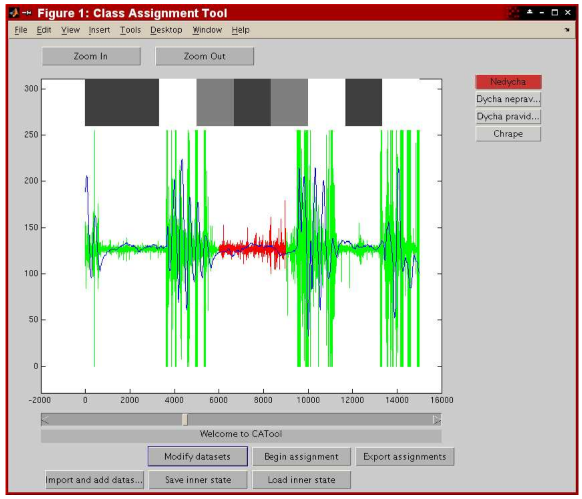

Figure 9.3: Class assigning procedure – Example of "Not Breathing" Class

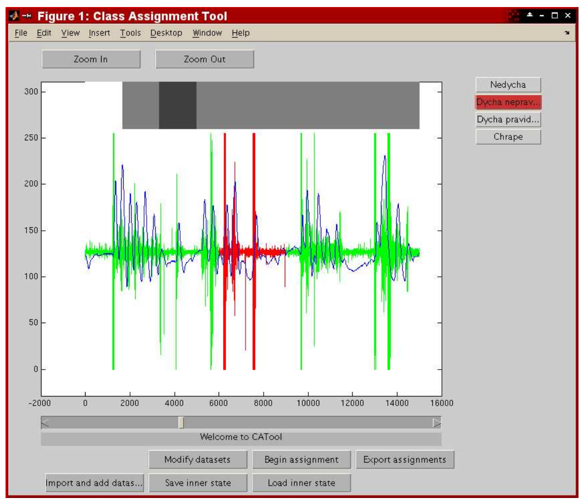

Figure 9.4: Class assigning procedure – Example of "Breathing Irregularly" Class

9.5 Feature Extraction

When breathing stages were assigned I extracted the features from the signals. This was done in my CATool which calls Matlab build-in functions. The list of extracted features for each signal is exactly same as described in 8.2. The picture 9.7 illustrates the feature extraction.

After I had done the assignments and feature extraction I have exported and saved the feature values and assignments to file formats suitable for using in Weka and GAME knowledge mining programs.

9.6 Feature Ranking, Feature Selection and its Results

9.6.1 Introduction

As I wrote above in section 6 reducing the dimension of input data is very helpfull for correct data classification. It improves the speed of classification, reduces learning time and even can improve accuracy of the classification, because it leaves out unimportant or even confusing features. I also want to find which features are more important for distinguishing breathing classes. Let me recall that I will try to build classifiers which will distinguish one particular class from the others.

For feature selection I used above mentioned data mining tools called Weka and GAME. In Weka there are many feature selection and feature ranking methods. Each method is a combnation of feature

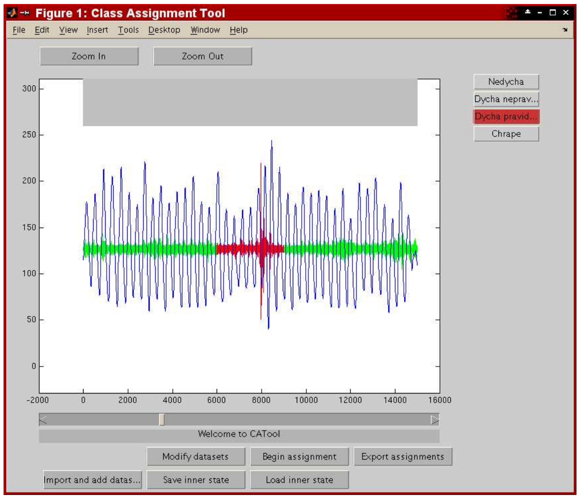

Figure 9.5: Class assigning procedure – Example of "Breathing Regularly" Class

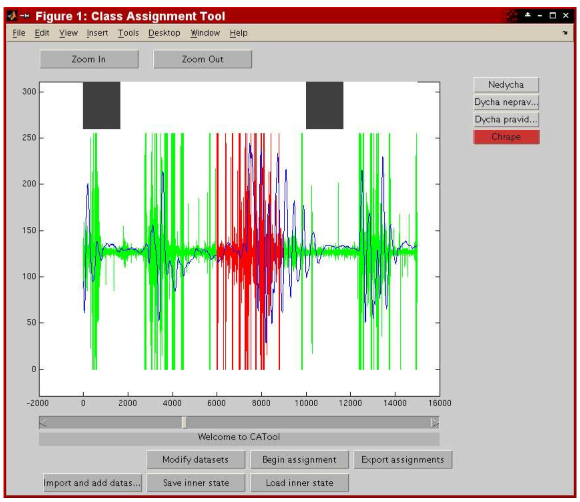

Figure 9.6: Class assigning procedure – Example of "Snoring" Class

search method and evaluator of currently selected features. I choose five combination and here is the list of combinations I choose with names as can be found in Weka.

Evaluator	**Feature search method**
CfsSubsetEval	BestFirst
CfsSubsetEval	GeneticSearch
ChiSquaredAttributeEval	Ranker
GainRatioAttributeEval	Ranker
InfoGainAttributeEval	Ranker

These combinations can be divided into two general groups the first group have the Ranker as the feature search method and the other have some other search method. The result of the first group is the sorted list of features according to their significance for distinguishing one class from the others. In this case I took 30 most important features and I will try to build suitable models from them. The result of the second group is just list of selected features without any hint how significant the features are. In this case I took all selected features.

One of the results of building a model in the GAME is a ranked list of features. Let me remind that the ranking is a percentage of neural inputs connected to particular features. The higher importance for distinguishing the class from the others implies the higher number of neurons connected to it. Because of the random nature of genetic algorithms, which is the GAME based on, I decide to use statistic

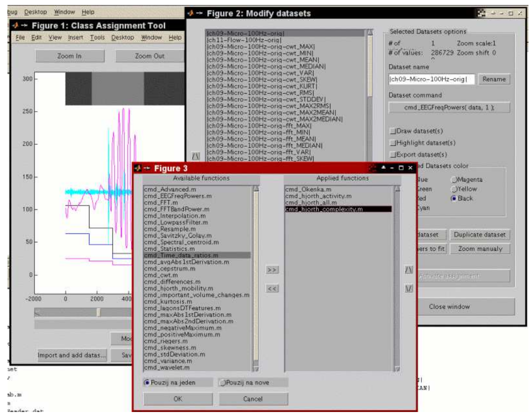

Figure 9.7: Feature extraction in Class Assignment Tool

methods to decide which features to use. I assumed that ranking gained from the GAME is normally distributed and I used a interval parameter estimation method to determine a 0.95 confidence interval of the mean value (see theoretical part for details). Then I selected the features which have a chance to get better importance than 1 percent. In other words I selected features which in their 0.95-confidence interval contains values 0.01 and higher.

Now is the time to show you the selected features. I will group the features according to the classes they should distinguish.

9.6.2 Class: Not Breathing

Features are listed is tables C.1 on page 95, C.2 on page 96, C.3 on page 97. I have done statistic on selected features by all methods. I have counted types of features and I found following facts:

In following list are numbers of selected features in each feature group:

- There are 42 Wavelet transform statistic based features. That is 25.6%.
- There are 48 Fourier transform based statistic features. That is 29.3%.
- There are 32 Fourier transform based band powers features. That is 19.5%.
- There are 24 Time-domain features selected by Josef Rieger. That is 14.6%.
- There are 6 Hjorth time domain features. That is 3.7%.
- There are 9 others time domain features. That is 5.5%.

In this list is the same statistics on all extracted features I got following numbers:

- There are 24 Wavelet transform statistic based features. That is 16.5%.
- There are 24 Fourier transform based statistic features. That is 16.5%.
- There are 42 Fourier transform based band powers features. That is 28.8%.
- There are 20 Time-domain features selected by Josef Rieger. That is 13.7%.
- There are 6 Hjorth time-domain features. That is 4.1%.
- There are 30 others time-domain features. That is 20.5%.

Now the feature set which gain higher percentage in selected features than it was its percentage a priory in extracted features, the feature set is important for the correct classification of the class. At least the feature selection methods think so.

Finally as the result I may say this – for distinguishing "Not Breathing" class from the others the Fourier transform based statistic features are most important. Second in the queue are the Wavelet transform based statistics features. The others hold their significance or even lost the percentage importance in the selected features.

	Significance in	
Feature group	**extracted features**	**selected features**
Wavelet transform	16.5%	25.6%
Fourier transform statistics	16.5%	29.3%
Fourier transform based band power	28.8%	19.5%
Time-domain features by Josef Rieger	13.7%	14.6%
Time-domain Hjorth's features	4.1%	3.7 %
Other time-domain features	20.5%	5.5%

Table 9.1: Results on feature groups significance for stage "Not Breathing"

9.6.3 Class: Breathing Irregularly

Features are listed is tables C.4 on page 98, C.5 on page 99, C.6 on page 100. I have done statistic on selected features by all methods. I have counted types of features and I found following facts:

In following list are numbers of selected features in each feature group:

- There are 42 Wavelet transform statistic based features. That is 26.1%.
- There are 40 Fourier transform based statistic features. That is 24.8%.
- There are 25 Fourier transform based band powers features. That is 15.5%.
- There are 31 Time-domain features selected by Josef Rieger. That is 19.3%.
- There are 10 Hjorth time domain features. That is 6%.
- There are 13 others time domain features (like polynomial approximation). That is 8.1%.

In this list is the same statistics on all extracted features I got following numbers:

- There are 24 Wavelet transform statistic based features. That is 16.5%.
- There are 24 Fourier transform based statistic features. That is 16.5%.
- There are 42 Fourier transform based band powers features. That is 28.8%.
- There are 20 Time-domain features selected by Josef Rieger. That is 13.7%.
- There are 6 Hjorth time domain features. That is 4.1%.
- There are 30 others time domain features. That is 20.5%.

	Significance in	
Feature group	**extracted features**	**selected features**
Wavelet transform	16.5%	26.1%
Fourier transform statistics	16.5%	24.8%
Fourier transform based band power	28.8%	15.5%
Time-domain features by Josef Rieger	13.7%	19.3%
Time-domain Hjorth's features	4.1%	6%
Other time-domain features	20.5%	8.1%

Table 9.2: Results on feature groups significance for stage "Breathing Irregularly"

Now the conclusion for this experiment should be, that for distinguishing "Breathing Irregularly" class from the others, the Wavelet transform based statistics features wins. But they are closely followed by the Fourier transform based statistic features. Time-domain features by J. Rieger and Hjorth features also gained a little.

9.6.4 Class: Breathing Regularly

Features are listed is tables C.7 on page 101, C.8 on page 102, C.9 on page 103. I have done statistic on selected features by all methods. I have counted types of features and I found following facts:

In following list are numbers of selected features in each feature group:

- There are 44 Wavelet transform statistic based features. That is 23.8%.
- There are 52 Fourier transform based statistic features. That is 28.1%.
- There are 38 Fourier transform based band powers features. That is 20.5%.
- There are 28 Time-domain features selected by Josef Rieger. That is 15.1%.
- There are 9 Hjorth time domain features. That is 4.9%.
- There are 14 others time domain features. That is 7.5%.

In this list is the same statistics on all extracted features I got following numbers:

- There are 24 Wavelet transform statistic based features. That is 16.5%.
- There are 24 Fourier transform based statistic features. That is 16.5%.
- There are 42 Fourier transform based band powers features. That is 28.8%.
- There are 20 Time-domain features selected by Josef Rieger. That is 13.7%.
- There are 6 Hjorth time domain features. That is 4.1%.
- There are 30 others time domain features. That is 20.5%.

	Significance in	
Feature group	**extracted features**	**selected features**
Wavelet transform	16.5%	23.8%
Fourier transform statistics	16.5%	28.1%
Fourier transform based band power	28.8%	20.5%
Time-domain features by Josef Rieger	13.7%	15.1%
Time-domain Hjorth's features	4.1%	4.9%
Other time-domain features	20.5%	7.5%

Table 9.3: Results on feature groups significance for stage ”Breathing Regularly”

Now the conclusion for this experiment should be, that for distinguishing ”Breathing Regularly” class from the others, the Fourier transform based statistic features are most valuable. Then again are the Wavelet transform based statistics features. Time-domain features by J. Rieger and Hjorth features also gained a little.

9.6.5 Class: Snoring

Features are listed is tables C.10 on page 104, C.11 on page 105, C.12 on page 106. I have done statistic on selected features by all methods. I have counted types of features and I found following facts:

In following list are numbers of selected features in each feature group:

- There are 40 Wavelet transform statistic based features. That is 24.1%.
- There are 44 Fourier transform based statistic features. That is 26.5%.
- There are 29 Fourier transform based band powers features. That is 17.5%.
- There are 34 Time-domain features selected by Josef Rieger. That is 20.5%.
- There are 9 Hjorth time domain features. That is 5.4%.

- There are 10 others time domain features. That is 6%.

In this list is the same statistics on all extracted features I got following numbers:

- There are 24 Wavelet transform statistic based features. That is 16.5%.
- There are 24 Fourier transform based statistic features. That is 16.5%.
- There are 42 Fourier transform based band powers features. That is 28.8%.
- There are 20 Time-domain features selected by Josef Rieger. That is 13.7%.
- There are 6 Hjorth time domain features. That is 4.1%.
- There are 30 others time domain features. That is 20.5%.

	Significance in	
Feature group	**extracted features**	**selected features**
Wavelet transform	16.5%	24.1%
Fourier transform statistics	16.5%	26.5%
Fourier transform based band power	28.8%	17.5%
Time-domain features by Josef Rieger	13.7%	20.5%
Time-domain Hjorth's features	4.1%	5.4%
Other time-domain features	20.5%	6%

Table 9.4: Results on feature groups significance for stage "Snoring"

The conclusion for this feature selection experiment is, that for distinguishing "Snoring" class from the others, the Fourier transform based statistic features are most significance. Then as usual are the Wavelet transform based statistics features. Time-domain features by J. Rieger and Hjorth features also gained a little.

9.6.6 Feature Selection Conclusions

Now on the end of feature selection experiment I can make following conclusions on significance of sets of features. For better insight to my conclusions see the table 9.5. The most significant set of features for distinguishing breathing classes from each other are statistic parameters for the Fourier transform coefficients and the Continuous Wavelet transform coefficients. More the features selected by Josef Rieger in his Diploma thesis are also usefull, but far less than both above mentioned. Next Hjorth parameters seems on the edge between usefull and useless features. Power of frequency bands seems to be quite useless features and the Time domain parameters and function interpolations features seems to be complete losers and brings no special information on distinguishing the classes.

	Significance in					
Feature group	FE	Not Br	Br Ir	Br Reg	Snore	Avg
Wavelet transform	16.5%	25.9%	26.1%	23.8%	24.1%	25.0%
Fourier transform statistics	16.5%	29.3%	24.8%	28.1%	26.5%	27.2%
Fourier transform based band power	28.8%	19.5%	15.5%	20.5%	17.5%	18.3%
Time-domain features by Josef Rieger	13.7%	14.6%	19.3%	15.1%	20.5%	17.3%
Time-domain Hjorth's features	4.1%	3.7%	6%	4.9%	5.4%	5%
Other time-domain features	20.5%	5.5%	8.1%	7.5%	6%	6.8%

Table 9.5: Results on feature groups significance for all stages – abbreviations meaning is in the text on page 46

The abbreviations in table 9.5 means followings :

- FE – Feature extraction.

- Not Br – Class "Not Breathing".
- Br Ir – Class "Breathing Irregularly".
- Br Reg – Class "Breathing Regularly".
- Snore – Class "Snoring".
- Avg – Average over all breathing classes.

9.7 Classificator Building and Their Results

After clustering I will turn my attention to building the classification models. Again I will use already well known data mining tools – the Weka and GAME. As I wrote above I will try to distinguish the one class from the others.

Now I will, at last, turn my attention to the classification. As many times before I will again use both the Weka and the GAME tools to build the classificator.

The GAME builds the neural network with use of genetic algorithms with different types of neurons as basic stones to build the network from. (See section A and appropriate sections in theoretical part for details.) I will allow neurons with several types of transfer functions (*linear transfer function, sigmoid function, polynomial function, exponential function and fraction of two polynomial function*). I will leave out neurons, which have a small back-propagation networks inside them. On one hand they can improve accuracy of the classifier, but on the other hand they takes a long time to learn properly. All genetic algorithms implemented in the GAME will be available for building the network.

The Weka offers much higher number of classification methods. I have chosen five methods which each represents one approach to building the classifier. They are the J48Tree, the DecisionTable, the SimpleLogistic function, the NaiveBayes method and the BayesNet. More I have decided to use K-Means clustering algorithm for classification. (See section A and appropriate sections in theoretical part for details.) All method settings in both Game and Weka were left on their default values.

So I have built above outlined classifiers and in following paragraphs I will show their results. I have built all classifiers for each set of features. I decided to do so because I wanted to found out if all methods are suitable for all sets of features or some feature set chosen by certain feature selection method is better for some classification method.

9.7.1 Classification Results – Class "Not Breathing"

	Bayes Net	Decision table	J48 Tree	Naive Bayes	K-Means	Simple Logistic	GAME
All features	17,96	7,54	7,54	32,57	49,42	6,13	6,5
CfsSubset BestFirst	9,38	7,28	7,12	24,24	36,8	7,07	5,97
CfsSubset Genetic	13,46	8,01	7,43	30,3	45,5	7,12	7,54
χ^2 Ranker	13,19	6,7	6,96	28,01	40	6,91	6,13
GainRatio Ranker	12,72	7,33	6,49	24,08	40,9	5,97	6
Game	17,49	7,23	7,38	23,5	42,82	5,92	7,17
InfoGain Ranker	13,98	6,91	6,86	28,95	41,25	6,86	6,23

Table 9.6: Error rate achieved by classificators for class "Not Breathing"

As you may see in the table 9.6 (and also on the graph 9.8), it is clear that class "Not Breathing" is very easily distinguished from the others. The best classifiers are the GAME, Decision Table and Simple Logistic. Other classifiers are too simple to handle the task with sufficient accuracy. In case of Simple K-Means algorithm the error is caused by fact that class "Not Breathing" and other classes do not group into two crisp clusters but both classes are mixed together.

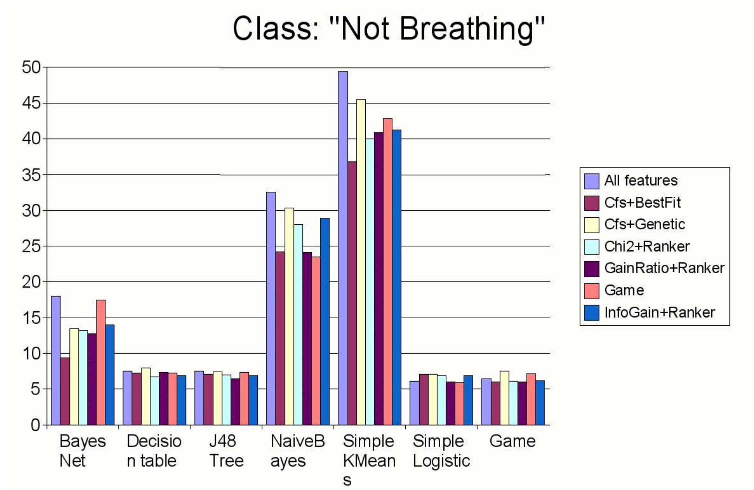

Figure 9.8: Error rate achieved by classificators for class "Not Breathing"

The best accuracy achieved by classifiers (Simple Logistic in this case) is 94% and this is excellent result. Remaining 6% of missclassified instances may be caused by one of following reasons:

- Error in class assignment – these instances were misclassified in process of preparing the learning set and the classifier set them correctly.
- Bordering instances – this includes instances which in time-domain contains two classes – in other words one part of the instance contains "Not Breathing" class and the other contains another class.

To illustrate misclassified instances I have done two screenshots 9.9 and 9.10 from my CATool. On each screenshot you may see a part of time domain signal. The blue signal comes from microphone and the black is from air flow detector and measures amount of air breathed in and out in a certain period of time. The dark and light bar above the signals is a classification by two classifiers (the above two "rows") and the bottom row is my classification. Currently selected instance is marked by red colour of the time domain signal and green bars on the top and the bottom of classification bar. The dark box in the bar means that in the instance the patient is breathing (is in other state than "Not Breathing") and the light bar means that the patient is "Not Breathing".

On figure 9.9 you may see an example of first situation mentioned above – when I misclassified the instance during preparation of learning set. On the other figure 9.10 you may see an example of error caused by border of two breathing states in the signal window.

From the result is also clear that feature selection did not improved accuracy too much. In case of GAME, Decision Table and Simple Logistic classifiers it is not big surprise because they do feature selection themself (because of nature of the classifier as the GAME or Simple Logistic or it explicitly uses another method).

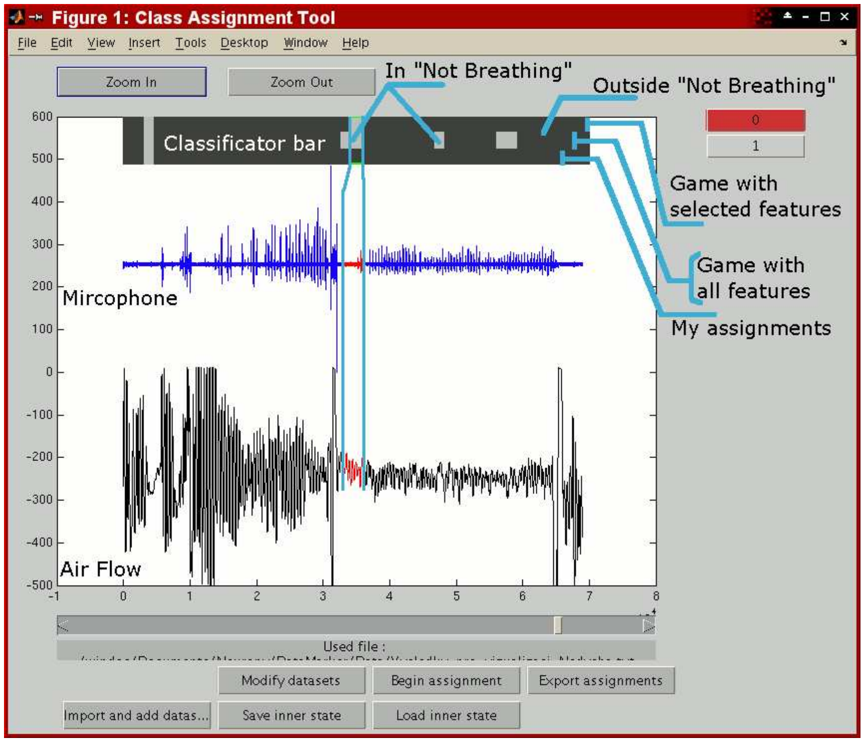

Figure 9.9: Example of error – instance misclassified in preprocessing

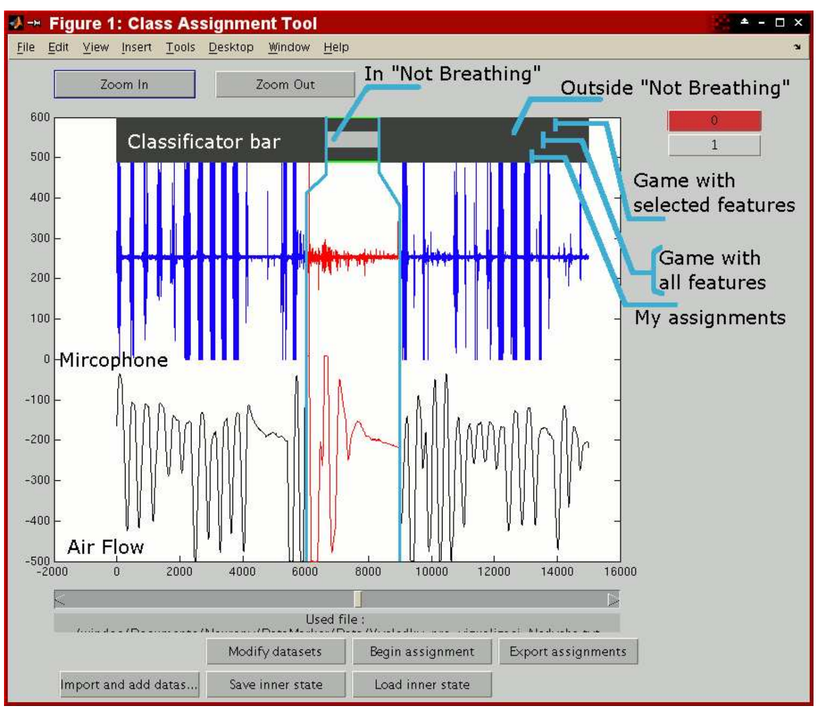

Figure 9.10: Example of error – example of border instance

9.7.2 Classification Results – Class "Breathing Irregularly"

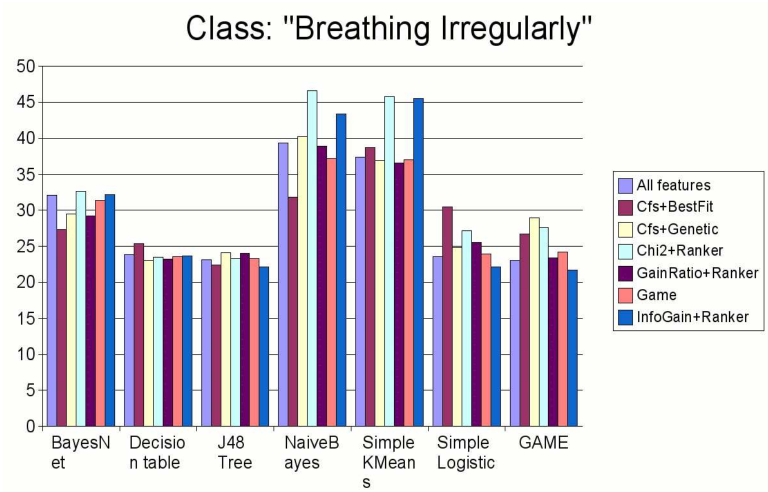

Figure 9.11: Error rate achieved by classificators for class "Breathing Irregularly"

	Bayes Net	Decision table	J48 Tree	Naive Bayes	K-Means	Simple Logistic	GAME
All features	32.1 %	23,82 %	23,13 %	39,37 %	37,33 %	23,56 %	23 %
CfsSubset BestFirst	27,33 %	25,39 %	22,46 %	31,78 %	38,71 %	30,52 %	26,7 %
CfsSubset Genetic	29,52 %	23,04 %	24,08 %	40,26 %	36,91 %	24,82 %	28,95 %
χ^2 Ranker	32,62 %	23,46 %	23,29 %	46,6 %	45,81 %	27,17 %	27,6 %
GainRatio Ranker	29,26 %	23,19 %	24,03 %	38,9 %	36,54 %	25,55 %	23,44 %
Game	31,41 %	23,56 %	23,29 %	37,17 %	37,01 %	23,92 %	24,21 %
InfoGain Ranker	32,15 %	23,66 %	22,15 %	43,35 %	45,49 %	22,15 %	21,72 %

Table 9.7: Error rate achieved by classificators for class "Breathing Irregularly"

Opposite to "Not Breathing" class, the class "Breathing Irregularly" is recognised very poorly. According to table 9.7 or figure 9.11 the lowest classification error rate achieved for this class was 21.72%. My opinion about this bad result is that this class is vaguely defined and is often indistinguishable from the other classes. It may be that I have badly classified too many instances and classifiers are not capable to learn characteristic patterns of the class.

To illustrate misclassified instances I have done screenshots 9.12 and 9.13 from CATool. On each screenshot you may see a part of time domain signal. The blue signal comes from microphone and the black is from air flow detector and measures amount of air breathed in and out in a certain period of time. Please note that, for this class, the "Microphone" signal is useless.

On the first figure 9.12 you may see example of badly learned classificator for distinguishing class "Breathing Irregularly". The shape of the Air-Flow signal is obviously irregular (and so it meet my criterion to assign "Breathing Irregularly" class and yet the classifier do not decided so.

On the figure 9.13, there is another example – the shape of Air-Flow signal is almost regularly and was classified by me as something else than "Breathing Irregularly" (most probable as "Breathing Regularly" or "Snoring") and the classificator marks it as "Breathing Irregularly". This may be my mistake during class assignment process or classificator's error. I am no sure which case is more probable, but it seems to me that the second (classificator error) because the shape does not seems to me more regular than signal on figure 9.12.

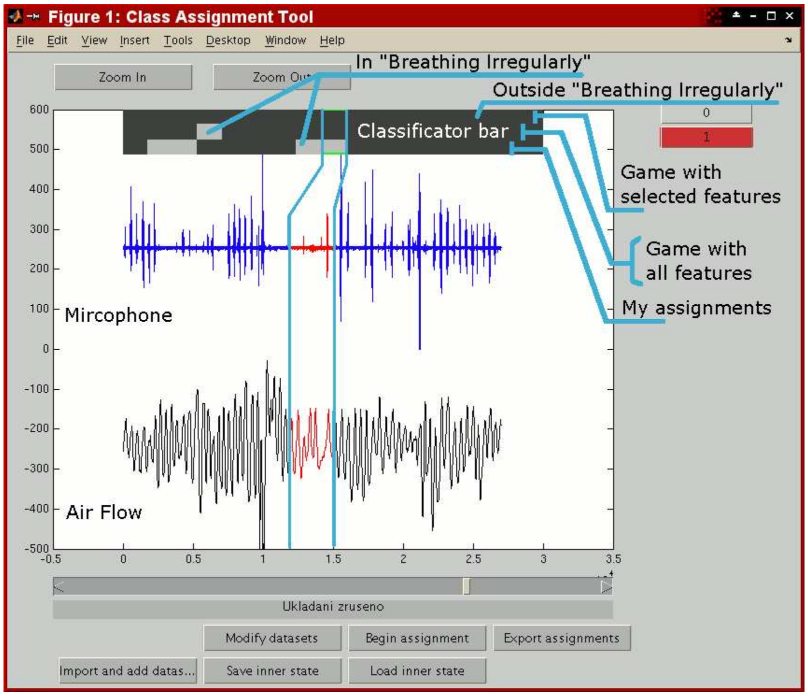

Figure 9.12: Example of error – example of bad learned classificator

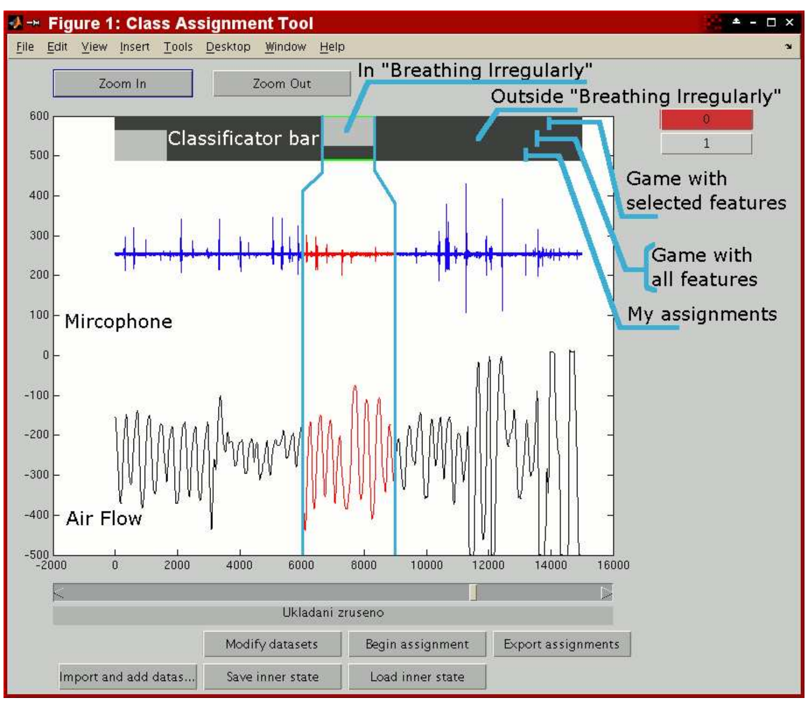

Figure 9.13: Example of error – instance misclassified in preprocessing or classificator error.

9.7.3 Classification Results – Class "Breathing Regularly"

	Bayes Net	Decision table	J48 Tree	Naive Bayes	K-Means	Simple Logistic	GAME
All features	19,16 %	11,52 %	12,46 %	31,26 %	49,58 %	9,37 %	10 %
CfsSubset BestFirst	16,23 %	11,15 %	10,94 %	38,48 %	39,69 %	9,9 %	10 %
CfsSubset Genetic	14,4 %	10,68 %	11,83 %	27,8 %	48,38 %	10,37 %	12,04 %
χ^2 Ranker	15,97 %	11,26 %	12,72 %	22,77 %	42,98 %	10,31 %	11,73 %
GainRatio Ranker	16,54 %	11,15 %	11,41 %	24,76 %	40,87 %	10,37 %	10,5 %
Game	22,25 %	12,36 %	12,14 %	31,51 %	44,65 %	10,89 %	11,1 %
InfoGain Ranker	16,64 %	11,2 %	12,15 %	22,2 %	41,41 %	9,95 %	11,52 %

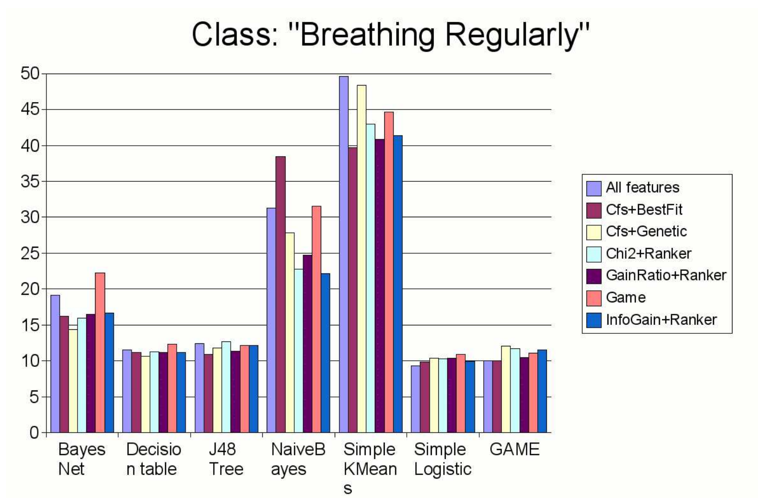

Figure 9.14: Error rate achieved by classificators for class "Breathing Regularly"

Now back to more accurate classifiers. The "Breathing Regularly" class is distinguished with 9.37 % error and it is good accuracy. This class is well separated from other classes. Misclassified instances are caused, as in above cases, by misclassification during preprocessing, bordering instances and also by some badly learned cases.

To illustrate misclassified instances I have done screenshots 9.15 and 9.16 from CATool. On each screenshot you may see a part of time domain signal. The blue signal comes from microphone and the black is from air flow detector and measures amount of air breathed in and out in a certain period of time. Please note that, for this class, the Microphone signal is quite useless.

On the first figure 9.15 you may see an example of bordering class which is hard to decide. One part of "Air-Flow" signal is obviously regular and the other is not. And here is an example how each classifier take different decision taking care about another part of instance.

On the second figure 9.16 is example of poorly learned classifier – the shape of the "Air-Flow" signal is obviously regular and one classifier (GAME with all features – the middle row in classificator bar) took bad decision. Because it is clear to see, that "Air-Flow" signal is very close to regular shape. May be that it did not select usefull features properly and it got confused with some useless features.

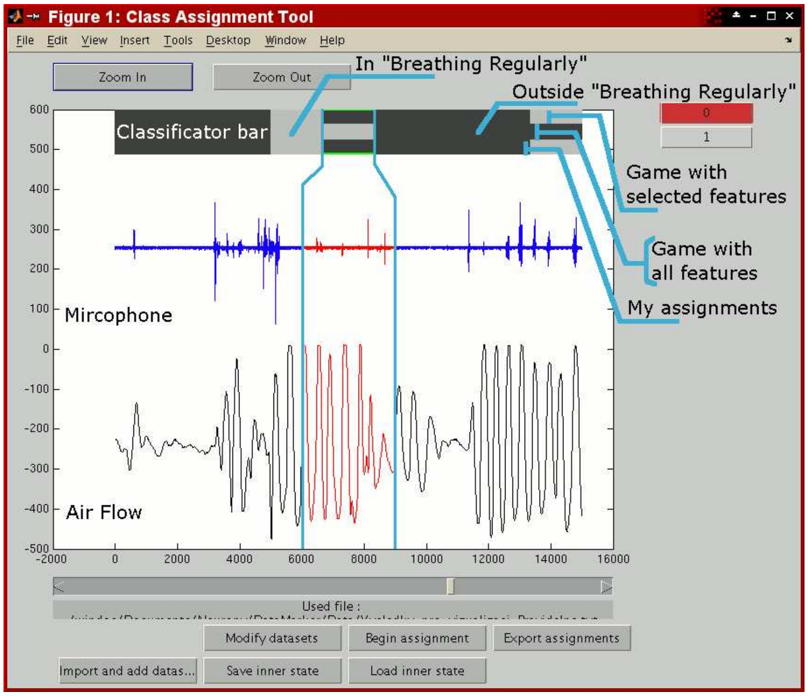

Figure 9.15: Example of error – error caused by border instance

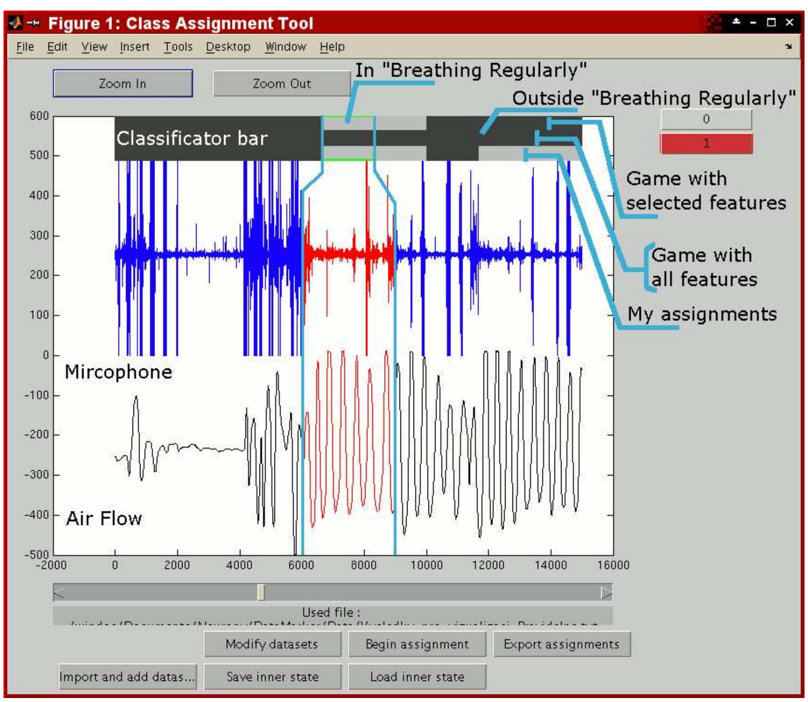

Figure 9.16: Example of error – poorly learned classifier – the middle row, GAME with all features

9.7.4 Classification Results – Class ”Snoring”

	Bayes Net	Decision table	J48 Tree	Naive Bayes	K-Means	Simple Logistic	GAME
All features	14,76 %	13,35 %	13,77 %	14,66 %	21,1 %	12,04 %	11,88 %
CfsSubset BestFirst	17,38 %	14,61 %	14,76 %	23,4 %	34,6 %	13,51 %	12,98 %
CfsSubset Genetic	14,87 %	14,97 %	14,56 %	18,12 %	21,83 %	13,25 %	12,36 %
χ^2 Ranker	21,2 %	13,72 %	16,34 %	43,09 %	37,75 %	14,03 %	13,87 %
GainRation Ranker	14,24 %	15,08 %	15,39 %	13,87 %	24,6 %	13,66 %	12,67 %
Game	14,35 %	15,13 %	14,35 %	13,87 %	16,49 %	12,25 %	12,46 %
InfoGain Ranker	21,36 %	13,92 %	14,5 %	43,09 %	37,23 %	14,29 %	14,4 %

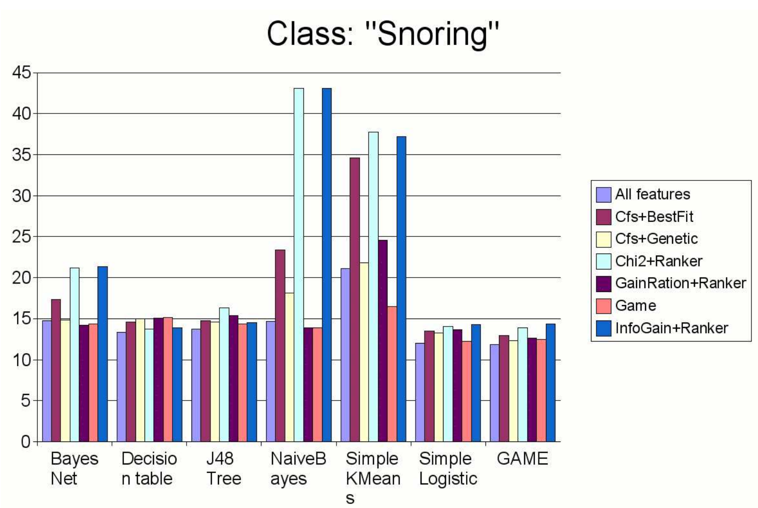

Figure 9.17: Error rate achieved by classificators for class ”Snore”

The last class is ”Snoring”. This class is classified with good error rate of 11.88 %. The reasons for misclassified instances are again the same as for above classes – errors in preprocessing, bordering instances and badly learned classificator.

As above to illustrate misclassified instances I have done two screenshots 9.18 and 9.19 from CATool. On each screenshot you may see a part of time domain signal. The blue signal comes from microphone and the black is from air flow detector and measures amount of air breathed in and out in a certain period of time. Please note that, for this class, the Microphone signal is more significant than Air-Flow signal.

On the first figure 9.18 is a little questionable situation. There are few peaks in Microphone signal which shows snoring, but also there are some not snoring parts. When I assigned classes to the signal in preprocessing I had assigned something else than class ”Snoring”. But classifiers found patterns enough to justify this instance as ”Snoring”. And as I think about it, it comes to me now that they are right and the patient is snoring constantly.

More on the figure 9.19 is another disputable instance. In the ”Microphone” signal is some activity which is not as big as on previous picture but has strength enough to, according to my opinion, mark it as ”Snoring”. But both classifiers disagree. Same situation is few instances on the right. This may be caused by badly learned classificator or somewhere else in the signal similar situation occurs and is not marked as snoring. And classificators learned to distinguish this situation as something else and applied it here also.

In the end it is possible to say, that classes ”Snoring”, ”Breathing Regularly” and ”Not Breathing”

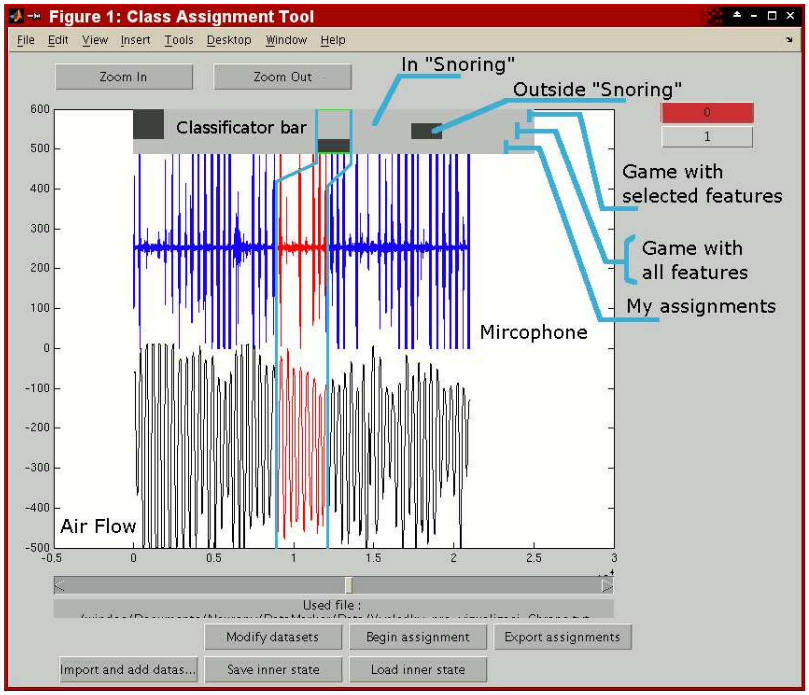

Figure 9.18: Example of error – error in class assignment during preprocessing

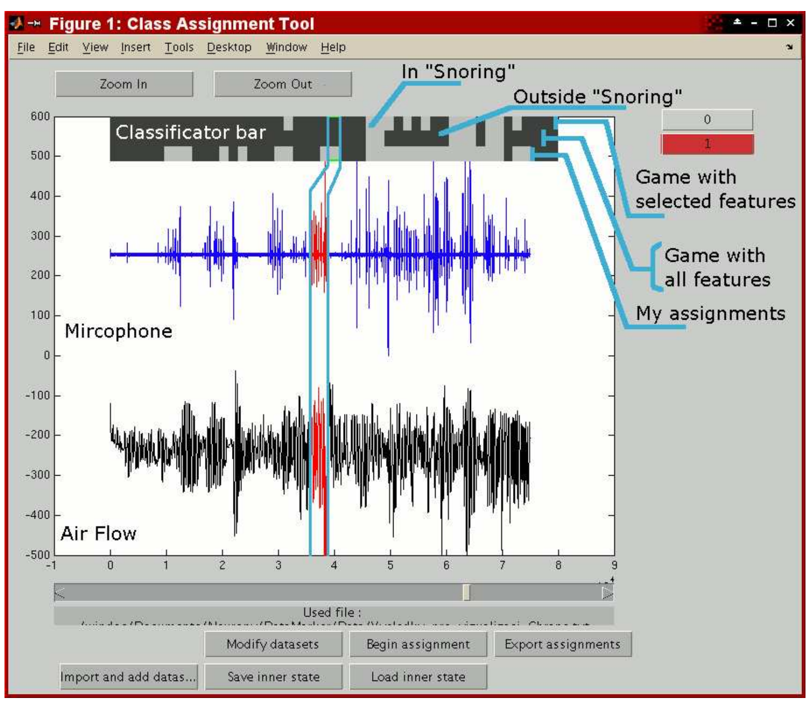

Figure 9.19: Example of error – bad learned classifiers (according to my opinion)

are recognised with good accuracy, but class "Breathing Irregularly" is recognised very poorly. In table 9.10 on page 57 you may see best results achieved.

9.7.5 Results Discussion and Confusion Matrix

To get a little better insight I created a single classificator which classifies all breathing classes at once. This classificator is the Simple Logistic and it achieves 25.86% error. Its goal was to produce so called confusion matrix which you may see in table 9.8. This confusion matrix shows how instances were classified and into which classes they were assigned during preprocessing.

	Classified as			
Class assigned in preprocessing	Not Breathing	Breathing Irregularly	Breathing Regularly	Snoring
Not Breathing	175	62	0	3
Breathing Irregularly	44	474	64	91
Breathing Regularly	0	69	219	27
Snoring	4	87	22	488

Table 9.8: Confusion table for Simple Logistic classificator classifying all classes at once

In this table you may see, that "Not Breathing" class is easily recognised – the only other class which holds large group of missclassified instances is "Breathing Irregularly" class. Please note that "Not Brathing" and "Breathing Regularly" are not confused at all. The worst recognisable class – "Breathing Irregularly" – is confused with all other classes in great numbers. A large number of misclassification will be caused be my mistakes in preprocessing. More errors are caused by bad learned classifier because of poor class definition. Although other two classes ("Snoring" and "Breathing Regularly") seems in this table to be recognised with same accuracy as "Breathing Irregularly" prior classifiers show far better result. But from this table you may see that class "Snoring" is mainly confused with "Breathing Irregularly" and sometimes with "Breathing Regularly". The other class, "Breathing Regularly" is mainly confused with "Breathing Irregularly" and a little with class "Snoring".

I have build several classifiers with satisfiable results – with one exception for "Breathing Irregularly" class and later in this section I discused these results.

9.8 Evaluation of Proposed Concept – Conclusions

In this chapter I tried my concept (see chapter 7) on simple and easy to understand data. I have gathered some more experience with concept methods and data handling.

I have done class assignment and feature extraction.

As a next step I did feature selection with six different methods and I got result on significance of different types of features (feature groups) for distinguishing each breathing class. Here I present the final percentage split between all types of features. Numbers in table 9.9 presents average for all breathing classes.

	Significance in	
Feature group	Feature extraction	Average
Wavelet transform statistic	16.5%	25.0%
Fourier transform statistics	16.5%	27.2%
Fourier transform based band power	28.8%	18.3%
Time-domain features by Josef Rieger	13.7%	17.3%
Time-domain Hjorth's features	4.1%	5%
Other time-domain features	20.5%	6.8%

Table 9.9: Results on feature groups significance for all stages (In percents of all features)

From this table (9.9) it is possible to do following conclusions:

- Wavelet transform statistic – this feature group is the second most usefull. The wavelet transform reacts on shapes and peaks in signals which describes the breathing classes quite well. Some future experiment could study how different mother wavelets affects capability of this features to describe the classes.
- Fourier transform statistics – this is the most significant feature group. This is not a big surprise, because frequency envelope naturally changes in different classes and also its statistic changes.
- Fourier transform based band power – this type of features does not contain any useful information. It is not much suprise because frequency bands are set up for another task. And it is possible that if the bands were set up different ways this group could got higher significance. I think that, at least, for distinguishing "Snoring" class it could be very useful.
- Time-domain features by Josef Rieger – these features are on the edge. Their description of data is not excellent but is sufficient to distinguish the classes.
- Time-domain Hjorth's features – these features were a little experiment and they prove to be quite useful. They are on the same level as above Rieger's features.
- Other time-domain features – these feature group contains all others extracted features – mainly approximation features. They proved as complete losers which do not describe signal at all.

I have build several classifiers for each breathing state and for all sets of selected features. In table 9.10 I present best results achieved for each breathing class. More detailed results you may find in section 9.7.

	Not Breathing	Breathing Irregularly	Breathing Regularly	Snoring
Classifier name	Simple Logistic	Simple Logistic	Simple Logistic	GAME
Feature selection method	GainRatio Ranker	InfoGain Ranker	All features	All features
Error Rate	5.92 %	21.72 %	9.37 %	11.88 %
Accuracy	94.08 %	78.28 %	90.63 %	88.12 %

Table 9.10: Best results achieved by classifiers for each class (Error rate and Accuracy are in percentage of misclassified or correctly classified instances.)

In general the best classifier for this task seems to be the Simple Logistic from the Weka closely followed by the GAME.

In general it is possible to say that feature selection did not improve classification accuracy much. Some classifiers are just too simple to distinguish the classes with sufficient accuracy and so they gets high error on full data as well as on selected features. The other classifiers do the feature selection themselves. They do the feature selection by explicit call of feature selection method or do it by nature of the method. And when I done a feature selection prior to building such classifies I ease it work but I may not improve its accuracy. I would improve its accuracy if the classifier stops in some local minimum for all features and prior feature selection remove this minimum. But this situation did not occurred in this case.

Now I turn my attention to classification accuracy. As you may have seen in table 9.10 the accuracy achieved for "Not Breathing" class 94 % is excellent. Also accuracy for classes "Snoring" and "Breathing Regularly" is good. But in case of "Breathing Irregularly" class the accuracy of 77.85% is quite poor. Reasons for such results are discused above in section 9.7 (Classification Results) and here is the brief summary:

- "Not Breathing" – this class is easy to recognise. The few misclassification are caused by errors during class assigning and instances which contains border between two neighbouring classes.
- "Breathing Irregularly" – this is the worst class. Many instances are confused with other classes (mainly "Snoring" and "Breathing Regularly"). The confusion with other classes is because of class assigning errors, similarity of the classes and badly learned classificators.

- "Breathing Regularly" – this class is in the middle (with the following class). Misclassification are caused by same reasons again (class assigning errors, similarity of the classes and bordering instances).
- "Snoring" – this class is the same as the above "Breathing Regularly" class.

If you are interested in reasons of misclassification please see section 9.7.

Based on the experiences gathered in this chapter I decided to modify a little my concept.

1. I will balance the instances prior to building classifiers. This is discused later in following section, but in general it is about changing size of set of positive and negative instances to the same number of examples.
2. I will leave out Simple K-Means classificator. It is too simple to produce any useful information.

As a final summary of evaluation of proposed concept I will type down a list of completed tasks with references to the chapters:

- Data Acquisition – section 9.3 on page 40.
- Breathing class assignment – section 9.4 on page 40.
- Feature extraction –section 9.5 on page 41.
- Feature ranking and Feature selection – section 9.6 on page 41.
- Classificators Building – section 9.7 on page 47.

Chapter 10

Application of the Concept on MIT Sleep Data

10.1 Description

Now I will come to the main experiment. Here I will use the concept introduced in the chapter 7.

The data I will process in this chapter comes for internet PhysioBank. The PhysioBank is located at http://www.physionet.org/physiobank/ and the Sleep EEG data can be downloaded from http://www.physionet.org/physiobank/database/sleep-edf/.

After first look to the data I decided to chose two process two recordings. This is because some sleep stages are rare in all recordings and so I used two recordings to increase number of these rare stages. Because both of these recordings comes from one source, so it is possible just to merge the recordings without any preprocessing. The recordings I will use are named `st7022j0.rec` and `st7121j0.rec` . In EEG PhysioBank repository is also sleep stages assignments of each signal in separate files with same name and extension `.hyp` instead of `.rec` .

Again I will try to build classifiers which will try to distinguish one sleep stage from others.

10.2 Experiment Setup

Experiment setup is almost the same as in previous chapter. The figure of actual setup is on th picture 10.1 below.

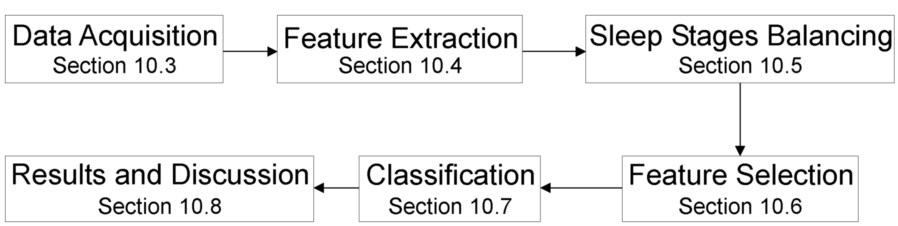

Figure 10.1: Experimental setup for Sleep EEG experiment

10.3 Data Acquisition

Again the first step in the experiment will be to acquire the data. As in previous case again I got the Sleep data in format called European Data Format. Its definition can be found in [13]. The task so was to import the breathing data from this EDF to my Matlab processing program – CATool (see A for details). Part of this program of mine is an import routine which reads data in EDF to CATool. (Example screenshot 9.2 can be found on page 40.)

The most of acquisition process runs separately for each file (remind that I use two pairs of recordings and scorings). The acquisition process consists of following steps:

1. Import each file separately to CATool.

2. Save each file in plain text format for easier manipulation in following step.

3. Merge appropriate files. (both recordings and both scorings).

4. In merged scoring file – replace numerical values, which presents sleep stages, with their string equivalents.

5. Now again import both merged files and the data acquisition is complete.

10.4 Features Extraction

When I acquired the data I split the signal into instances. Each instance, according to the obtained scoring and common practice, consists of 30 seconds of the signal and the signal's features. These features will be the same as in chapter "Evaluation of Proposed Concept" and are listed in appendix B or in shortened form in chapter 8.

For feature extraction process I will again use my CATool and picture that illustrates the feature extraction may be found on figure 9.7 on page 43.

10.5 Sleep Stages Balancing

I put in this step because of big differences in numbers of instances of each class. sleep stage number 4 and REM[1] stage is very common, presented in about 43 % respectively 23 % instances. On the other hand Stage 1 and Awake state are very rare presented in about 5 % each. In case of this rare classes some classifiers may completely ignore them and still it will have good accuracy. In case of Awake state or Sleep Stage 1 the classifier which ignores these stages will have 95 % accuracy. This high accuracy is excellent result. This great result will have only one disadvantage – we never got patient awake or in Stage 1 (falling asleep) and this may be a big problem when there is a need to detect these states.

To address this problem I will do a thing which I will call balancing. The target of the balancing is to make sure that number of positive and negative instances[2] will be nearly equal.

To ease this task I wrote a small script in Matlab which automate the task for me. Also I have balanced training sets for all sleep stages even if it was not necessary. I will use these balanced datasets from now on.

10.6 Feature Ranking, Feature Selection and Its Results

10.6.1 Introduction

As I wrote above in chapter 6 reducing the dimension of input data is very helpfull for correct data classification. It improves the speed of classification, reduces learning time and even can improve accuracy of the classification, because it leaves out unimportant or even confusing features. I also want to find which features are more important for distinguishing breathing classes. Let me recall that I will try to build classifiers which will distinguish one particular class from the others.

I will use same feature selection methods as in previous chapter. See section 9.6 for details. As in previous chapter I will do simple statistic for each sleep stage. But in addition I will be also interested in which electrodes supply useful features.

Because of too many features and too little instances for some Sleep stages I decided to select features in two steps. In the first step I will divide extracted features into two sets for sleep stage 3 and 4 and into four sets for sleep stage 1 and Wake. Then I will select representative features among these sets. And then in the second step I will select the most useful set of features among the features selected in the first step. The feature selection method will be always the same as in the first step.

[1]Abbreviation from Rapid Eye Movement

[2]Please keep in mind that I will build classifiers which will separate one particular sleep stage from others

10.6.2 Sleep Stage Wake

Features are listed is tables D.1 on page 108. With this and the following sleep stage is a small problem because feature ranking methods implemented in Weka did not provided any features. They simply outputed all features or none. Why is this I am unable to say. The ratio between number of instances and number of features is almost the same as in other Sleep stages.

So the numbers presented here are based only on GAME selected features. This is not completely right but at least it gives a small insight how the things can be.

I have done statistic on selected features by all methods. I have counted types of features and I found following facts:

In selected features for sleep stage Wake following numbers of features are present:

- There are 3 Wavelet transform statistic based features selected. That is 7%.
- There are 14 Fourier transform based statistic features selected. That is 34%.
- There are 10 Fourier transform based band powers features selected. That is 24%.
- There are 6 Time-domain features selected by Josef Rieger selected. That is 14%.
- There are 2 Hjorth time domain features selected. That is 5%.
- There are 5 others time domain features selected. That is 9%.

I have also counted electrodes from which came the signal presented in selected features I will got following numbers:

- There 5 selected features that have origin in EEG Fpz-Cz electrode. That is 9%.
- There 9 selected features that have origin in EEG Pz-Oz electrode. That is 22%.
- There 6 selected features that have origin in EOG horizontal electrode. That is 9%.
- There 21 selected features that have origin in EMG submental electrode. That is 51%.

When I did same statistics on all extracted features I got following numbers:

- There are 48 Wavelet transform statistic based features. That is 16.5%.
- There are 48 Fourier transform based statistic features. That is 16.5%.
- There are 84 Fourier transform based band powers features. That is 28.8%.
- There are 40 Time-domain features selected by Josef Rieger. That is 13.7%.
- There are 12 Hjorth time-domain features. That is 4.1%.
- There are 60 others time-domain features. That is 20.5%.

	Significance in	
Feature group	**extracted features**	**selected features**
Wavelet transform	16.5%	7%
Fourier transform statistics	16.5%	34%
Fourier transform based band power	28.8%	24%
Time-domain features by Josef Rieger	13.7%	14%
Time-domain Hjorth's features	4.1%	5 %
Other time-domain features	20.5%	9%

Table 10.1: Results on feature groups significance for sleep stage Wake

Now the feature set which gain higher percentage in selected features than it was its percentage a priory in extracted features, the feature set is important for the correct classification of the sleep stage. At least the feature selection methods think so. The same logic is for electrodes.

	Electrode significance in	
Electrode name	**extracted features**	**selected features**
EEG Fpz-Cz	25%	9%
EEG Pz-Oz	25%	22%
EOG Horizontal	25%	9%
EMG Submental	25%	51%

Table 10.2: Results on electrode significance for sleep stage Wake

Now it is possible to say, that the most useful features are the statistic properties of the Fourier transform coefficients. Their presence is almost doubled, which is mark of really high significance. The features based on frequency band powers, Hjorths's parameters and Josef Rieger's time-domain features hold their significance and the others are loosing their positions. In case of other time-domain features it is not a great suprise. But big suprise pops out in case of wavelet transform features. These parameters proved quite useful in previous chapter and also holds good significance in following Sleep stages. This may be caused by specific shapes of the signal which the wavelet transform features were unable to describe. The second reason may be, as I wrote above, data are only from one feature selection method and the GAME network by chance selected other features.

To electrode significance – the most significant electrode is EMG[3] submental electrode. This electrode looks after your muscle activity. The other electrodes are loosing. As I wrote above, this may be caused by chance, but it seems to me very natural that if the patient is awake he (or she) moves a lot and so he (or she) generates a lot of muscle activity which does not occur in other sleep stages. If I write in in different words it should look this way – he (or she) generates specific patterns in EMG electrode which differs from EMG activity in other sleep stages. Please note the warning about generalising too much these conclusions.

10.6.3 Sleep Stage 1

Features are listed is tables D.2 on page 109. Again I must warn the reader not to take results presented in this section to seriously. Feature ranking methods implemented in Weka did not provided any features. They simply outputed all features or none. So the numbers presented here are based only on GAME selected features. This is not completely right but at least it gives a small insight how the things can be.

I have done statistic on selected features by all methods. I have counted types of features and I found following facts:

In selected features for sleep stage Wake following numbers of features are present:

- There are 9 Wavelet transform statistic based features selected. That is 18.3%.
- There are 15 Fourier transform based statistic features selected. That is 30.6%.
- There are 12 Fourier transform based band powers features selected. That is 24.5%.
- There are 4 Time-domain features selected by Josef Rieger selected. That is 8.1%.
- There are 1 Hjorth time domain features selected. That is 2%.
- There are 8 others time domain features selected. That is 16.3%.

I have also counted electrodes from which came the signal presented in selected features I will got following numbers:

- There 15 selected features that have origin in EEG Fpz-Cz electrode. That is 30.6%.
- There 8 selected features that have origin in EEG Pz-Oz electrode. That is 16.3%.

[3] This stands for Electromyography

- There 15 selected features that have origin in EOG horizontal electrode. That is 30.6%.
- There 11 selected features that have origin in EMG submental electrode. That is 22.4%.

When I did same statistics on all extracted features I got following numbers:

- There are 48 Wavelet transform statistic based features. That is 16.5%.
- There are 48 Fourier transform based statistic features. That is 16.5%.
- There are 84 Fourier transform based band powers features. That is 28.8%.
- There are 40 Time-domain features selected by Josef Rieger. That is 13.7%.
- There are 12 Hjorth time-domain features. That is 4.1%.
- There are 60 others time-domain features. That is 20.5%.

	Significance in	
Feature group	**extracted features**	**selected features**
Wavelet transform	16.5%	18.3%
Fourier transform statistics	16.5%	30.6%
Fourier transform based band power	28.8%	24%
Time-domain features by Josef Rieger	13.7%	8.1%
Time-domain Hjorth's features	4.1%	2%
Other time-domain features	20.5%	8%

Table 10.3: Results on feature groups significance for sleep stage 1

	Electrode significance in	
Electrode name	**extracted features**	**selected features**
EEG Fpz-Cz	25%	30.6%
EEG Pz-Oz	25%	16.3%
EOG Horizontal	25%	30.6%
EMG Submental	25%	22.4%

Table 10.4: Results on electrode significance for sleep stage 1

Now the feature set which gain higher percentage in selected features than it was its apriory percentage in extracted features, the feature set is important for the correct classification of the sleep stage. At least the feature selection methods think so. The same logic is for electrodes.

For sleep stage 1 the Fourier transform statistic is the most useful again. Again the frequency band powers holds their significance and now they are accompanied with wavelet statistic features. All other are losing. These results are much closer to what I expect than results in previous sleep stage. But again please note that results are gathered only from one feature selection method and results may be erroneous.

As for the electrodes the situation here is not so clear as in previous case. The electrodes EEG Fpz-Cz and EOG horizontal gained a little more significance. The EMG submental lost a little and the EEG Pz-Oz is the least significant. Here comes the point where I am unable to precisely comment what's happening – I am not educated in this field. But I assume that as the patient falls asleep the EEG frequency lowers. This fact occurs on the EEG Fpz-Cz electrode. And at the same time his eyes moves around. This may be blinking or some kind automatic movements or may be he (or she) just lies open eyed on the bed and tries to fall asleep.

10.6.4 Sleep Stage 2

Features are listed is tables D.3 on page 111, D.4 on page 112, D.5 on page 114. Now I came to better Sleep stages where I obtained all 6 sets of selected features.

I have done statistic on selected features by all methods. I have counted types of features and I found following facts:

In selected features for sleep stage Wake following numbers of features are present:

- There are 63 Wavelet transform statistic based features selected. That is 23.9%.
- There are 39 Fourier transform based statistic features selected. That is 14.8%.
- There are 76 Fourier transform based band powers features selected. That is 28.9%.
- There are 33 Time-domain features selected by Josef Rieger selected. That is 12.5%.
- There are 12 Hjorth time domain features selected. That is 4.5%.
- There are 40 others time domain features selected. That is 15.2%.

I have also counted electrodes from which came the signal presented in selected features I will got following numbers:

- There 151 selected features that have origin in EEG Fpz-Cz electrode. That is 57.4%.
- There 48 selected features that have origin in EEG Pz-Oz electrode. That is 18.3%.
- There 48 selected features that have origin in EOG horizontal electrode. That is 18.3%.
- There 16 selected features that have origin in EMG submental electrode. That is 6%.

When I did same statistics on all extracted features I got following numbers:

- There are 48 Wavelet transform statistic based features. That is 16.5%.
- There are 48 Fourier transform based statistic features. That is 16.5%.
- There are 84 Fourier transform based band powers features. That is 28.8%.
- There are 40 Time-domain features selected by Josef Rieger. That is 13.7%.
- There are 12 Hjorth time-domain features. That is 4.1%.
- There are 60 others time-domain features. That is 20.5%.

	Significance in	
Feature group	**extracted features**	**selected features**
Wavelet transform	16.5%	23.9%
Fourier transform statistics	16.5%	14.8%
Fourier transform based band power	28.8%	28.9%
Time-domain features by Josef Rieger	13.7%	12.5%
Time-domain Hjorth's features	4.1%	4.5%
Other time-domain features	20.5%	15.2%

Table 10.5: Results on feature groups significance for sleep stage 2

Now the feature set which gain higher percentage in selected features than it was its percentage a priory in extracted features, the feature set is important for the correct classification of the sleep stage. At least the feature selection methods think so. The same logic is for electrodes.

In this sleep stage the wavelet transform statistic is winning. All other are losing slightly. These results are little unusual because this sleep stage does not have big winners and big losers but it have one small winner and a lot of yet smaller losers.

Electrode name	Electrode significance in	
	extracted features	**selected features**
EEG Fpz-Cz	25%	57.4%
EEG Pz-Oz	25%	18.3%
EOG Horizontal	25%	18.3%
EMG Submental	25%	6%

Table 10.6: Results on electrode significance for sleep stage 2

If I look on the electrodes there is much cleaner situation because the EEG Fpz-Cz electrode is supreme winner and the EMG submental electrode is a big looser. The other two electrods are losing but not so much. The reason probably is that the patterns generated in this sleep stage on EEG Fpz-Cz are much different in this stage and in other stages. Lets recall that sleep stage 2 is defined with presence of so called Sleep spindles and K-Complexes which occurs in EEG signal. But why it is only on EEG Fpz-Cz electrode I am unable to say.

In case of other electrodes patterns which occurs in their signals does not help to identify the stage too much.

10.6.5 Sleep Stage 3

Features are listed is tables D.6 on page 115, D.7 on page 116, D.8 on page 118.

I have done statistic on selected features by all methods. I have counted types of features and I found following facts:

In selected features for sleep stage Wake following numbers of features are present:

- There are 77 Wavelet transform statistic based features selected. That is 24.5%.
- There are 53 Fourier transform based statistic features selected. That is 16.9%.
- There are 85 Fourier transform based band powers features selected. That is 27%.
- There are 42 Time-domain features selected by Josef Rieger selected. That is 13.3%.
- There are 24 Hjorth time domain features selected. That is 7.6%.
- There are 33 others time domain features selected. That is 10.5%.

I have also counted electrodes from which came the signal presented in selected features I will got following numbers:

- There 115 selected features that have origin in EEG Fpz-Cz electrode. That is 36.6%.
- There 13 selected features that have origin in EEG Pz-Oz electrode. That is 4%.
- There 63 selected features that have origin in EOG horizontal electrode. That is 20%.
- There 123 selected features that have origin in EMG submental electrode. That is 39.1%.

When I did same statistics on all extracted features I got following numbers:

- There are 48 Wavelet transform statistic based features. That is 16.5%.
- There are 48 Fourier transform based statistic features. That is 16.5%.
- There are 84 Fourier transform based band powers features. That is 28.8%.
- There are 40 Time-domain features selected by Josef Rieger. That is 13.7%.
- There are 12 Hjorth time-domain features. That is 4.1%.
- There are 60 others time-domain features. That is 20.5%.

	Significance in	
Feature group	**extracted features**	**selected features**
Wavelet transform	16.5%	24.5%
Fourier transform statistics	16.5%	16.9%
Fourier transform based band power	28.8%	27%
Time-domain features by Josef Rieger	13.7%	13.3%
Time-domain Hjorth's features	4.1%	7.6%
Other time-domain features	20.5%	10.5%

Table 10.7: Results on feature groups significance for sleep stage 3

	Electrode significance in	
Electrode name	**extracted features**	**selected features**
EEG Fpz-Cz	25%	36.6%
EEG Pz-Oz	25%	4%
EOG Horizontal	25%	20%
EMG Submental	25%	39.1%

Table 10.8: Results on electrode significance for sleep stage 3

Now the feature set which gain higher percentage in selected features than it was its percentage a priory in extracted features, the feature set is important for the correct classification of the sleep stage. At least the feature selection methods think so. The same logic is for electrodes.

The situation is almost the same as in previous stage. The only difference is that Hjorth's parameters significance boosts to almost double value. Otherwise the situation is the same – wavelet features gained and the others features lose.

In field of electrodes the EMG submental and EEG Fpz-Cz wins. In case of EMG electrode it is a little suprise because the sleep stage 3 is defined only through the EEG frequency properties. In this case some uncommon patterns occurs on EMG electrode which looks after the muscle activity, which I did not expect. It is also possible that there is no muscle activity in stage 3 which is also valuable pattern. As I mentioned above the stage 3 is defined by frequency powers in EEG signals so it is no a big surprise that EEG Fpz-Cz electrode gains so much. The bad result of EEG Pz-Oz electrode is a constant fact which I can not now interpret. It is maybe the noise or some other facts which makes this electrode useless. It is also possible that the electrode is just placed over a area of the brain where nothing interesting happens in this sleep stage.

10.6.6 Sleep Stage 4

Features are listed is tables D.9 on page 119, D.10 on page 120, D.11 on page 121.

I have done statistic on selected features by all methods. I have counted types of features and I found following facts:

In selected features for sleep stage Wake following numbers of features are present:

- There are 66 Wavelet transform statistic based features selected. That is 25%.
- There are 47 Fourier transform based statistic features selected. That is 17.8%.
- There are 68 Fourier transform based band powers features selected. That is 25.7%.
- There are 33 Time-domain features selected by Josef Rieger selected. That is 12.5%.
- There are 20 Hjorth time domain features selected. That is 7.5%.
- There are 30 others time domain features selected. That is 11.3%.

I have also counted electrodes from which came the signal presented in selected features I will got following numbers:

- There 116 selected features that have origin in EEG Fpz-Cz electrode. That is 44%.
- There 8 selected features that have origin in EEG Pz-Oz electrode. That is 3%.
- There 19 selected features that have origin in EOG horizontal electrode. That is 7.2%.
- There 121 selected features that have origin in EMG submental electrode. That is 45.8%.

When I did same statistics on all extracted features I got following numbers:

- There are 48 Wavelet transform statistic based features. That is 16.5%.
- There are 48 Fourier transform based statistic features. That is 16.5%.
- There are 84 Fourier transform based band powers features. That is 28.8%.
- There are 40 Time-domain features selected by Josef Rieger. That is 13.7%.
- There are 12 Hjorth time-domain features. That is 4.1%.
- There are 60 others time-domain features. That is 20.5%.

	Significance in	
Feature group	**extracted features**	**selected features**
Wavelet transform	16.5%	25%
Fourier transform statistics	16.5%	17.8%
Fourier transform based band power	28.8%	25.7%
Time-domain features by Josef Rieger	13.7%	12.5%
Time-domain Hjorth's features	4.1%	7.5%
Other time-domain features	20.5%	11.3%

Table 10.9: Results on feature groups significance for sleep stage 4

	Electrode significance in	
Electrode name	**extracted features**	**selected features**
EEG Fpz-Cz	25%	44%
EEG Pz-Oz	25%	3%
EOG Horizontal	25%	7.2%
EMG Submental	25%	45.8%

Table 10.10: Results on electrode significance for sleep stage 4

Now the feature set which gain higher percentage in selected features than it was its percentage a priory in extracted features, the feature set is important for the correct classification of the sleep stage. At least the feature selection methods think so. The same logic is for electrodes.

The most significant features are the wavelet based, as usual, accompanied with Hjorth's parameters which again boosts to double significance. The features which holds their significance are Fourier coefficient statistic, Frequency bands, Rieger's features. The other time-domain features are losing, as usual.

The statements I quoted in previous stage is true here again. Only change is that EEG Fpz-Cz and EMG submental wins more and EOG horizontal sinks lower in significance. In case of EMG electrode so high significance is again a little suprise because the sleep stage 4 is defined only through the EEG frequency properties. In this case again some uncommon patterns occurs on EMG electrode, which make it useful. It is also possible that there is no muscle activity in stage 4 which is also valuable pattern. As I mentioned in introduction the stage 4 is defined by frequency powers in EEG signals so it is no a big surprise that EEG Fpz-Cz electrode gains so much. The bad result of EEG Pz-Oz electrode is a constant fact which I can not now interpret. It is maybe the noise or some other facts which makes this electrode useless. It is also possible that the electrode is just placed over a area of the brain where nothing interesting happens in this sleep stage.

10.6.7 Sleep Stage REM

Features are listed is tables D.12 on page 123, D.13 on page 124, D.14 on page 125.

I have done statistic on selected features by all methods. I have counted types of features and I found following facts:

In selected features for sleep stage Wake following numbers of features are present:

- There are 66 Wavelet transform statistic based features selected. That is 21.8%.
- There are 47 Fourier transform based statistic features selected. That is 15.5%.
- There are 79 Fourier transform based band powers features selected. That is 26.1%.
- There are 50 Time-domain features selected by Josef Rieger selected. That is 16.5%.
- There are 18 Hjorth time domain features selected. That is 6%.
- There are 42 others time domain features selected. That is 13.9%.

I have also counted electrodes from which came the signal presented in selected features I will got following numbers:

- There 180 selected features that have origin in EEG Fpz-Cz electrode. That is 59.6%.
- There 63 selected features that have origin in EEG Pz-Oz electrode. That is 20.8%.
- There 45 selected features that have origin in EOG horizontal electrode. That is 15%.
- There 14 selected features that have origin in EMG submental electrode. That is 4.5%.

When I did same statistics on all extracted features I got following numbers:

- There are 48 Wavelet transform statistic based features. That is 16.5%.
- There are 48 Fourier transform based statistic features. That is 16.5%.
- There are 84 Fourier transform based band powers features. That is 28.8%.
- There are 40 Time-domain features selected by Josef Rieger. That is 13.7%.
- There are 12 Hjorth time-domain features. That is 4.1%.
- There are 60 others time-domain features. That is 20.5%.

	Significance in	
Feature group	**extracted features**	**selected features**
Wavelet transform	16.5%	21.8%
Fourier transform statistics	16.5%	15.5%
Fourier transform based band power	28.8%	26.1%
Time-domain features by Josef Rieger	13.7%	16.7%
Time-domain Hjorth's features	4.1%	6%
Other time-domain features	20.5%	13.9%

Table 10.11: Results on feature groups significance for sleep stage REM

Now the feature set which gain higher percentage in selected features than it was its percentage a priory in extracted features, the feature set is important for the correct classification of the sleep stage. At least the feature selection methods think so. The same logic is for electrodes. The percentage significance of electrodes in extracted features is not presented in any table because it is trivial – from all electrodes the equal number of features was extracted and so each electrode has 25 % significance in extracted features.

Electrode name	Electrode significance in	
	extracted features	**selected features**
EEG Fpz-Cz	25%	59.6%
EEG Pz-Oz	25%	20.8%
EOG Horizontal	25%	15%
EMG Submental	25%	4.5%

Table 10.12: Results on electrode significance for sleep stage REM

For this stage the feature significance is as usual – the wavelet transform features wins, other time-domain features loses and all other are as significant as in feature extraction. So nothing unexpected happend here.

When I look on electrodes the EEG Fpz-Cz electrode again wins with more than two times higher significance than in the beginning. The EEG Pz-Oz at last got some significance, although it is still losing. This supports my opinion presented in previous sleep stages, that the electrode is just placed over a area of the brain where nothing interesting happens in those sleep stages. Unexpected result, for me, is low significance of EOG horizontal electrode. This electrode looks after the rotation of patients eyes in horizontal pane. From the definition of the REM sleep stage I expected that this electrode will play far more significant role in distinguishing this sleep stage. In other words that it will contain patterns which will easily identify the stage.

10.6.8 Feature Selection Conclusions

Now on the end of feature selection experiment I can make following conclusions on significance of sets of features. For better insight to my conclusions see the table 10.13. The most significant set of features for distinguishing breathing classes from each other are statistic parameters for the Fourier transform coefficients and the Continuous Wavelet transform coefficients. More the features selected by Josef Rieger in his Diploma thesis are also usefull, but far less than both above mentioned. Next Hjorth parameters seems on the edge between usefull and useless features. Power of frequency bands seems to be quite useless features and the Time domain parameters and function interpolations features seems to be complete losers and brings no special information on distinguishing the classes.

Feature group	All features	Significance in sleep stage					
		Wake	1	2	3	4	REM
Wavelet transform	16.5%	7%	18.3%	23.9%	24.5%	25%	21.8%
Fourier transform statistics	16.5%	34%	30.%	14.8%	16.9%	17.8%	15.5%
Fourier transform based band power	28.8%	24%	24%	28.9%	27%	25.7%	26.%
Time-domain features by Josef Rieger	13.7%	14%	8.1%	12.5%	13.3%	12.5%	16.7%
Time-domain Hjorth's features	4.1%	5%	2%	4.5%	7.6%	7.5%	6%
Other time-domain features	20.5%	9%	8%	15.2%	10.5%	11.3%	13.9%

Table 10.13: Feature groups significance for all sleep stages

Electrode	All features	Significance in sleep stage					
		Wake	1	2	3	4	REM
EEG Fpz-Cz	25%	9%	30.6%	57.4%	36.6%	44%	59.6%
EEG Pz-Oz	25%	22%	16.3%	18.3%	4%	3%	20.8%
EOG horizontal	25%	9%	30.6%	18.3%	20%	7.2%	15%
EMG submental	25%	51%	22.4%	6%	39.1%	45.8%	4.5%

Table 10.14: Significance of electrodes for all sleep stages

10.7 Classifiers Building and Their Results

Here I will present results of builded classifiers. As I wrote when I introduced my concept, I will use following classifiers: Bayes Net, Naive Bayes, Decision Table, J48 Tree and Simple Logistic from Weka and GAME. As in previous chapter I will build each classifier for each set of features selected by each feature selection method. These sets were obtained in section 10.6 and are fully listed below in appendixes. As I wrote many times before, I will try to distinguish each single class from the others.

For a little more details on this topic see section 9.7.

Now lets come to the results of classiers. I will use following classifiers: Bayes Net, Decision Table, J48 Tree, Naive Bayes and Simple Logistic from Weka and the last classifier will be GAME. I will teach each classifier on each set of selected features to determine which set and classifier is the best.

While I was processing the results of classifiers I found out a problem which makes a result discussion harder. When I did a balancing of the instances to contain same number of positive and negative instances, I did not record which instances I pick for further processing. Because of random nature of my balancing algorithm I am unable to reconstruct time domain signals for each instance. For positive instances it is not a problem – I took all positive instances, but for negative instances this problem occurs. The solution might be to show you values of features for each electrode, but I do not believe they will show anything useful.

10.7.1 Patient Wake – Results

	Bayes Net	Decision table	J48 Tree	Naive Bayes	Simple Logistic	GAME
All features	41,24%	35,4%	36,43%	41,58%	36,08%	19%
CfsSubset BestFirst	N/A	N/A	N/A	N/A	N/A	N/A
CfsSubset Genetic	N/A	N/A	N/A	N/A	N/A	N/A
χ^2 Ranker	N/A	N/A	N/A	N/A	N/A	N/A
GainRatio Ranker	N/A	N/A	N/A	N/A	N/A	N/A
InfoGain Ranker	N/A	N/A	N/A	N/A	N/A	N/A
Game	32,21%	29,96%	31,09%	44,57%	24,34%	21%

Table 10.15: Error rate achieved by classificators for stage Wake

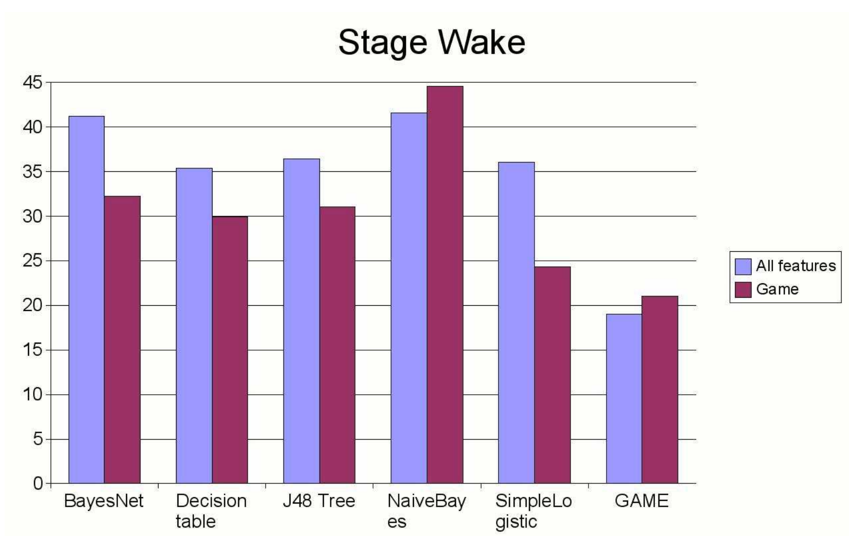

Figure 10.2: Error rate achieved by classificators for Sleep Stage Wake

This sleep stage is the second easiest for classification. My opinion is that this stage contains a lot of specific activity, for example on EMG electrode, which does not occur in other stages. But I can not prove this statement because I am unable find which instances in datasets belong to specific part of the signal.

The best classifier seems to be the GAME. It achieves 19% error rate. In context of other stages, it is very good result.

The results of Simple logistic classifiers illustrates the well known problem in data mining – it is so called curse of dimensionality. When build over all available features, the learning process fails and the classifier achieves 36% error rate. When we supply subset of features its ability to learn correctly increases and it achieves 24% error rate. The same thing occurs for J48 decision tree, Decision table and Bayes net.

10.7.2 Sleep Stage 1 – Results

	Bayes Net	Decision table	J48 Tree	Naive Bayes	Simple Logistic	GAME
All features	40,63%	40,18%	43,3%	47,77%	42,41%	24%
CfsSubset BestFirst	N/A	N/A	N/A	N/A	N/A	N/A
CfsSubset Genetic	N/A	N/A	N/A	N/A	N/A	N/A
χ^2 Ranker	N/A	N/A	N/A	N/A	N/A	N/A
GainRatio Ranker	N/A	N/A	N/A	N/A	N/A	N/A
InfoGain Ranker	N/A	N/A	N/A	N/A	N/A	N/A
Game	43,75%	42,86%	42,41%	40,18%	31,7%	21%

Table 10.16: Error rate achieved by classificators for sleep stage 1

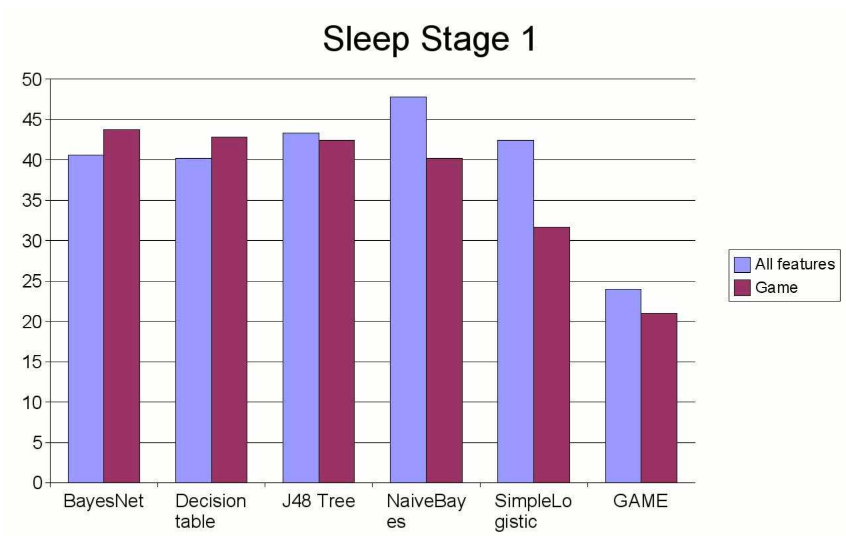

Figure 10.3: Error rate achieved by classificators for Sleep Stage 1

This sleep stage is also ease to classify. The best classifier seems to be the GAME. It achieves 21% error rate with GAME selected features. The curse of dimensionality again takes place in simple logistic classifier. Other classifiers are too simple to recognise this class.

10.7.3 Sleep Stage 2 – Results

This class is the most difficult to recognise. The best classifier is the GAME with features produced by CfsSubset + BestFirst feature selection method. It achieves 33% error rate, which is very poor result. This may be caused be several things – poor classification of the learning set and consequent badly learned classifier, or it is may be too close to other sleep stages. All these statement are just my deduction supported by none exact fact.

	Bayes Net	Decision table	J48 Tree	Naive Bayes	Simple Logistic	GAME
All features	38,94%	36,24%	42,22%	47,67%	39,95%	40%
CfsSubset BestFirst	34,88%	36,38%	35,25%	44,69	39,75%	33%
CfsSubset Genetic	40,6%	33,63%	35,21%	44,4	42,75%	41%
χ^2 Ranker	38,58%	34,71%	37,39%	44,63%	39,83%	39%
GainRatio Ranker	41,84%	35,03%	35,67%	43,96%	40,3%	35%
InfoGain Ranker	39,7%	34,33%	36,77%	44,38%	40,45%	40%
Game	41,57%	40,82%	40,51%	42,45%	38,95%	35%

Table 10.17: Error rate achieved by classificators for sleep stage 2

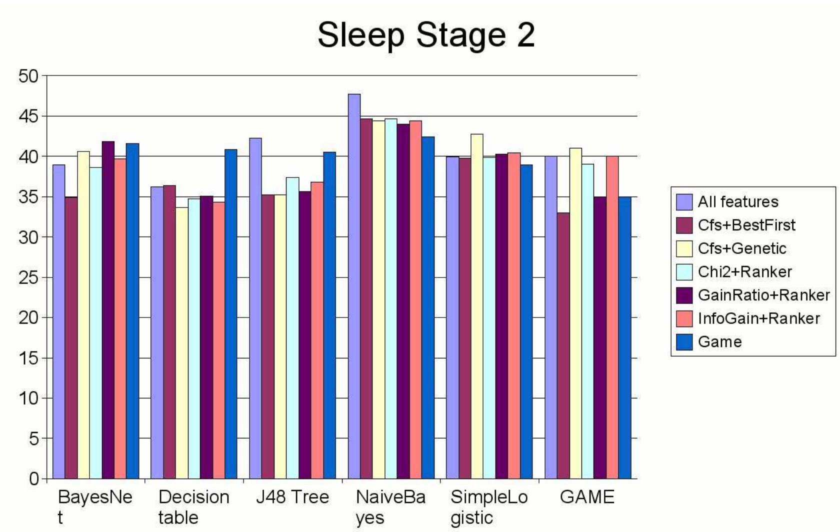

Figure 10.4: Error rate achieved by classificators for Sleep Stage 2

10.7.4 Sleep Stage 3 – Results

	Bayes Net	Decision table	J48 Tree	Naive Bayes	Simple Logistic	GAME
All features	34,65%	37,76%	40,25%	37,55%	32,16%	23%
CfsSubset BestFirst	35,14%	35,94%	39,96%	40,16%	40,36%	36%
CfsSubset Genetic	36,31%	31,74%	35,48%	34,65%	32,57%	30%
χ^2 Ranker	36,31%	33,61%	37,14%	37,76%	33,2%	27%
GainRatio Ranker	39,45%	38,86%	38,07%	42,41%	36,09%	28%
InfoGain Ranker	39,59%	41,43%	36,53%	40,82%	35,31%	35%
Game	37,67%	36,69%	39,84%	39,65%	36,29%	27%

Table 10.18: Error rate achieved by classificators for sleep stage 3

This stage is better than previous. The best classifier is again the GAME with 23% error rate. The set of features the best classifier operates on contains all features extracted. This can be explained by the fact than the genetic algorithm inside the GAME does feature selection by itself and by chance it found the optimum subset of features in this case.

As you may see on the figure 10.5 the CfsSubset + BestFirst feature selection method produces a bad subset features, at least for the GAME and simple logistic classifiers.

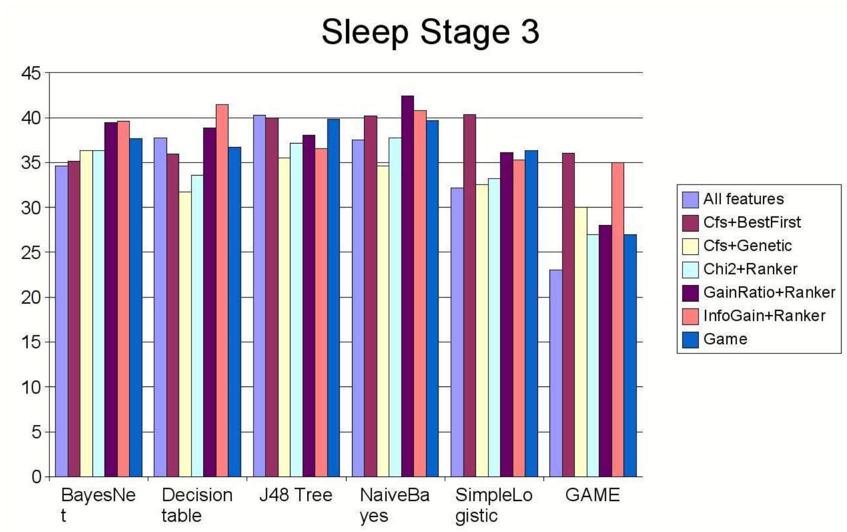

Figure 10.5: Error rate achieved by classificators for Sleep Stage 3

10.7.5 Sleep Stage 4 – Results

	Bayes Net	Decision table	J48 Tree	Naive Bayes	Simple Logistic	GAME
All features	31,86%	28,86%	33,07%	39,68%	26,45%	19%
CfsSubset BestFirst	26,45%	30,86%	28,26%	31,66%	26,65%	24%
CfsSubset Genetic	26,21%	24,74%	25,79%	24,74%	27,88%	20%
χ^2 Ranker	28,42%	26,95%	22,74%	36,84%	25,89%	22%
GainRatio Ranker	29,65%	26,99%	26,58%	37,22%	23,72%	18%
InfoGain Ranker	27,84%	28,95%	29,18%	34,52%	26,73%	20%
Game	24,53%	25,79%	26,21%	31,24%	22,01%	16%

Table 10.19: Error rate achieved by classificators for sleep stage 4

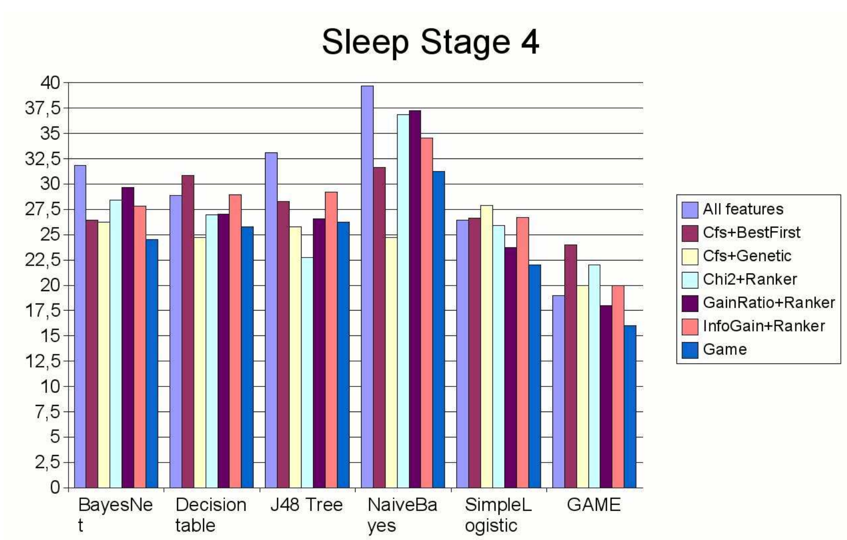

Figure 10.6: Error rate achieved by classificators for Sleep Stage 4

This is the best classified sleep stage. The best classifier is as usual the GAME with the GAME selected features. It achieves 16% error rate.

Again I am unable to say the exact reason, but my assumptions are the this stage contains some specific patterns which makes the classification easier, but which they are I can not say.

10.7.6 Sleep Stage REM – Results

	Bayes Net	Decision table	J48 Tree	Naive Bayes	Simple Logistic	GAME
All features	30,51%	28,85%	30,95%	47,14%	28,3%	26%
CfsSubset BestFirst	25,46%	29%	26,83%	40,3%	29,91	24%
CfsSubset Genetic	34,42%	29,7%	30,55%	34,42%	27,27	23%
χ^2 Ranker	35,99%	33,14%	30,76%	46,2%	29,22%	26%
GainRatio Ranker	35,1%	29,92%	31,21%	42,99%	30,51%	32%
InfoGain Ranker	34,63%	30,85%	30,16%	46,56%	31,42%	25%
Game	29,91%	32,42%	29,57%	35,84%	28,54%	25%

Table 10.20: Error rate achieved by classificators for sleep stage REM

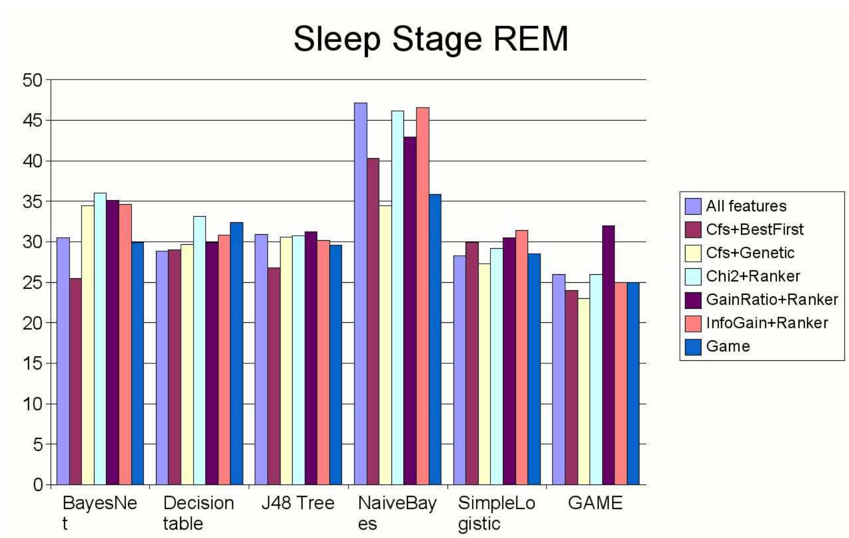

Figure 10.7: Error rate achieved by classificators for Sleep Stage REM

This stage is also easy to recognise. The best classifier is the GAME with CfsSubset + BestFirst selected features.

I thought that sleep stage ought be the one easiest to classify. It should contain very specific patterns on EOG electrode (the one which record patient's eye movements) and this patterns should be easily recognisable. But it is not so. Why is it I discuss later in section 10.8 when I try to find general reasons, why the sleep stages are not so easy to classify.

10.8 Sleep Stages Results and Conclusions

The table 10.21 you can see the best results for each sleep stage. The best classifier row is little boring because it contains only one name – GAME. Only Simple Logistic sometimes competes with it, but never wins. All other classification methods are too simple to handle such complex data.

As you also may see in the table, no feature selection method is superior to others. The features selected by GAME and again supplied to GAME seems to be useful. That usefulness is quite clear

	Results for sleep stage					
	Wake	1	2	3	4	REM
Best classifier name	GAME	GAME	GAME	GAME	GAME	GAME
Best feature set	All features	Game selected	CfsSubset BestFirst	All features	Game selected	CfsSubset Genetic
Accuracy	81 %	79 %	67 %	77 %	84 %	77 %
Error rate	19 %	21 %	33 %	23 %	16 %	23 %

Table 10.21: Best results achieved by classifiers for each class (Error rate and Accuracy are in percentage of misclassified or correctly classified instances.)

because the GAME holds same rules in both runs (feature selection and classification) and so the best features selected in first run are also the best in the second run over selected features. Also all features extracted from signal celebrates a little victory. The reason is that GAME have a feature selection algorithm inside by nature of the method. And so GAME takes advantage of all extracted features and do a feature selection again and it find again useful features. The other two sets of selected features seems to be successful by chance when they contain the most useful features or the GAME optimisation methods fails to find the best optimum for other sets.

As I wrote in introduction to classification above, I am unable to create pictures of misclassified instances as in Breathing data classification (section 9.7). To compensate this fact I will produce another pictures. These will not contain the classifiers output, but only expert classification. On the other hand I think that this expert classification is enough to identify the most common reasons for misclassification.

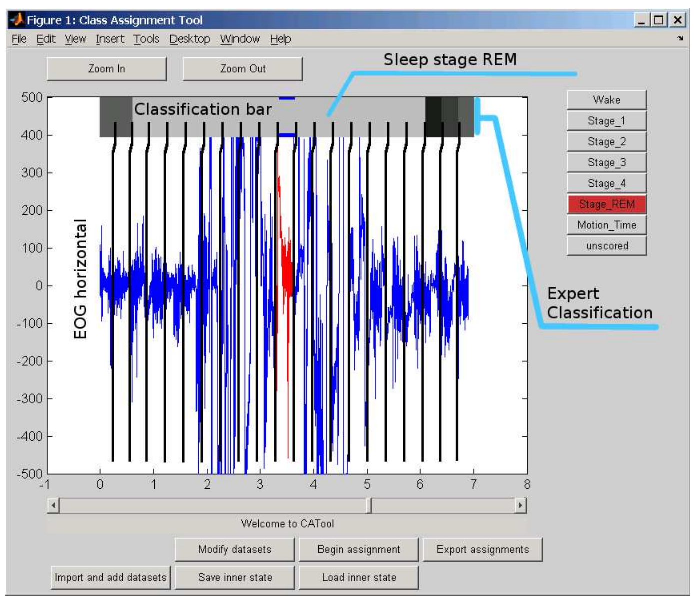

Figure 10.8: Example of confusing instances in electrode EOG for sleep stage REM

On picture 10.8 you may see an example, in my opinion, badly assigned sleep stage. On the picture the sleep stage REM is presented. And the signal is in time domain and comes from one electrode, EOG horizontal. Lets recall that if the patient is in sleep stage REM, his eyes should move fast. According to Sleep scoring manual [20] this is very important condition.

Now if you look at that figure (10.8), you may see regions, on each side of the central area, which are also marked as Stage REM even the signal shape is completely different from the the central area. On the other hand the signal shape is similar to regions which are on the left and right edge of the screen and are marked as other Sleep stages. This fact naturally confuses all classifiers and also may explain quite low significance (15 %) of EOG electrode in selected features for this sleep stage.

According to my opinion the figure 10.9 show another difficult situation for any classificator. If you look on instances next to active instance (the one marked with red colour) look exactly the same as

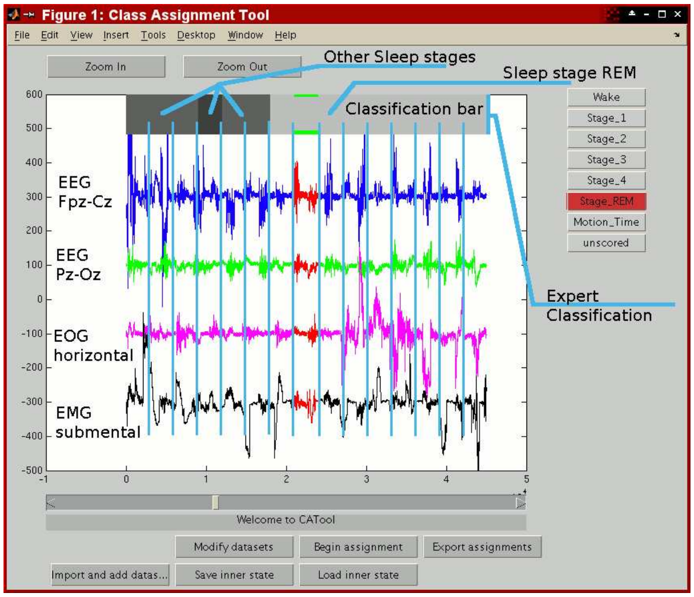

Figure 10.9: Example of confusing instances for sleep stage REM

instances marked as "Other sleep stages". This almost certainly results in missclassification.

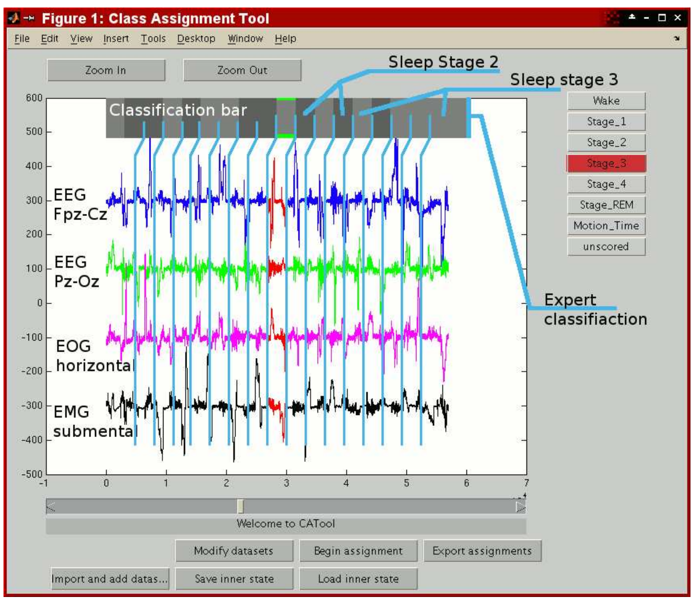

Figure 10.10: Example of confusing instances for sleep stage 2 and 3

The last example on figure 10.10 shows yet another confusing situation. The sleep stage oscillates between stage 2 and 3 even when signal show same shape in all electrodes. This sequence almost certainly results in high number of errors in classification.

In general it is possible to say that, in compare to results in previous chapter, the classification results are quite poor. The best classification accuracy of 84 % is in compare to 94 % from previous section poor result. But if you look at figures 10.9, 10.10 or 10.8 you may more cleanly understand why is it so. There is a lot of places which I believe were misclassified by expert. And this fact naturally causes problems. One problem is that correctly learned classificator classifies such instance to a different (correct) sleep stage and this fact causes an error. The second problem is that if classifier is learned with too many misclassified instances than it can not work properly and classifies by chance. This is in my opinion problem with sleep stage 2 which is correctly classified only with 77% accuracy. The other reason for poor classification results might be that the problem is just to complex to be sufficiently solved by used classificators. But I think that this reason is minor and valid only for unsuccessful classificators.

After above discusion I think that there are two ways to improve the classificators accuracy: more precise sleep stages assigning in learning set and the second is to supply the classificators with informations about the history and the future. I believe that these two steps would greatly improve classificators accuracy.

Comparison results with other works

Now I will compare my results with some other works (in table 10.22). You need to know that in these works different data are processed. And each uses own evaluation methods. Some uses training and testing data, some uses only training data to evaluate the models, etc... . But it gives an idea how accurate the models are.

	My results	Oropesa, Cycon [17]	Mutapcic, Shimayama [18]	Koprinska, Pfurtscheller [19]
Wake	19%	26%	20%	–
Sleep stage 1	21%	25%	23.3%	cca 25%
Sleep stage 2	33%	9.5%	11.5%	cca 15%
Sleep stage 3	23%	–	11.3%	cca 18%
Sleep stage 4	16%	–	–	cca 18%
Sleep stage REM	23%	35%	17.4%	cca 12%

Table 10.22: Comparison my error rate with other works. Empty spaces in the table means that the work did not concern on that sleep stage. Note that in work [19] the results are presented only in form of charts and there are no exact values.

As you may see in the table 10.22 my results are in the middle. There are several ways to improve my results. The ways in my opinion leads in two directions – the first is to precisely reassign sleep stages to the signal. The second is to provide the history to the classificator.

As a final summary of sleep stages classification I will type down a list of completed tasks with references to the chapters:

- Data Acquisition – section 10.3 on page 59.
- Feature extraction –section 10.4 on page 60.
- Class balancing –section 10.5 on page 60.
- Feature ranking and Feature selection – section 10.6 on page 60.
- Classificators Building – section 10.7 on page 70.

Chapter 11

Future work

As in the most cases a lot of questions and research directions still remains open. Here I present only the most interesting research directions.

First look at figure 11.1. You may see that sleep stages can be quite easily distinguished and recognised. And so some classifier should achieve very good results. In following list are ideas what to do and how to improve thinks.

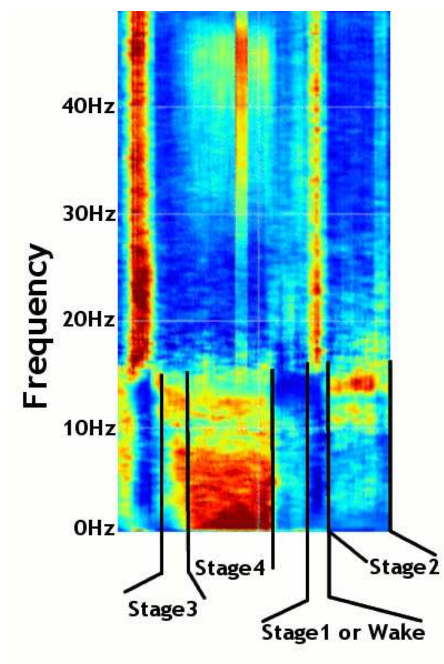

Figure 11.1: Sleep recording in frequency domain with sleep stages

- These ideas were mentioned above – to introduce history to the classifier and to precisely reassign sleep stages to meet the conditions introduced in Sleep scoring manual [20].

- If you again look at the figure 11.1 the natural conclusion ought to be that frequency band powers should give enough informations to distinguish each sleep stage. The one way to compute these powers is to use regular windowed Fourier transform. The other way is to use Welch method to estimates the power spectral density. This Welch method produces much smoother results than regular Fourier transform and so it seems to be better for classification. In fact the picture 11.1 was produced using this method.

- In the Sleep scoring manual [20] the Sleep stages are also defined with use of some shapes which occurs in the EEG (K-complexes, Sleep spindles). So the other research direction should be to identify these shapes in the signal. This field is addressed with many works.

- After the testing in real world I would like to bundle all methods used in my concept into one application and deploy it in Sleep laboratory in Faculty of Medicine in Hradec Králové.

Chapter 12

Conclusion

In my work I tried to find out which signal features and which electrodes are significant for correct Sleep stages classification. Prior to classification or feature extraction, the data preprocessing must be done. To support the data preprocessing I wrote a tool (program) in Matlab. This program proved to be useful and was successfully used by several students in their semestral works.

To systematise step, I need to do to obtain results, I created a concept inspired by CRISP Data mining industry standard. This concept proved to be successful and its results are comparable or even better than other works (see table 10.22). My results, concept and preprocessing program will be tested in Faculty of medicine in Hradec Králové.

Results I find out in my work also gives an answer I ask in abstract of this work:

In general the most significant features are (in no special order) Wavelet transform statistic coefficients and Fourier transform statistic coefficients. Although The Fourier transform based band powers features are used in many works it shows no superiority to above the two mentioned.

Features that are significant for EEG sleep stages classification may be not significant for other tasks – in other words the set of optimal features is problem depend. My concept can be used to select and evaluate significant features for any time-series data set.

My most significant theoretical results of my work are in following list

- The significant electrodes for sleep stage classification varies from stage to stage. Some surprise comes out in sleep stage REM. See section 10.6 for details.
- In general the Sleep EEG stages classification is not too accurate, because of complexity of the data and errors in learning set. See section 10.7 for details.
- The best classifier proved to be the GAME followed by the Simple logistic. The other classifiers were unable to handle such complex data. For details see section 10.7.
- The classification accuracy vary from sleep stage to stage. The best recognised stage is stage 4 with 84% accuracy. The worst is stage 2 with 67% accuracy. Reasons for this results are discused in section 10.7.
- I suggest some ideas how to increase the accuracy of the Sleep stages classification and some other research directions. See chapter 11 for details.
- On the other hand my Sleep stages classification results are as good as other works. See table 10.22 on page 77 for comparison.

Appendix A

Brief Overview of Used Programs

In this chapter I will shortly describe most of the software tools used in my work.

Game

The acronym GAME stands for Group of Adaptive Models Evolution. It is developed by Ing. Pavel Kordik on Department of Computer Science and Engineering. This tool can build powerfull models for classification, provides many feature ranking method and allows many more interesting things.

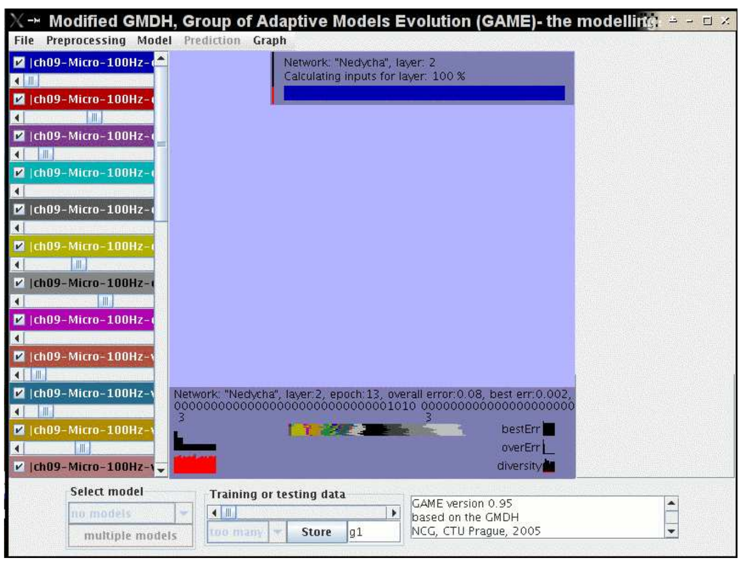

Figure A.1: GAME Screenshot

This tool may be downloaded from web page http://neuron.felk.cvut.cz/game.

Weka

Weka is quite powerful data mining tool. It implements many different methods for feature selection, classification, clustering, etc. It also provides basic preprocessing methods but I will not use them.

I will use this tool to build classifiers to check my features and also I will use it for feature selection.

This tool is mainly developed at Waikato university on New Zealand under the GNU GLP licence. So it may be freely downloaded from http://www.cs.waikato.ac.nz/∼ml/weka/.

Matlab

Matlab is commercial software by Mathworks. It is very powerfull, fast and complex tool for numerical computations. It may be extended by so called toolboxes. Many of these are inside the Matlab installa-

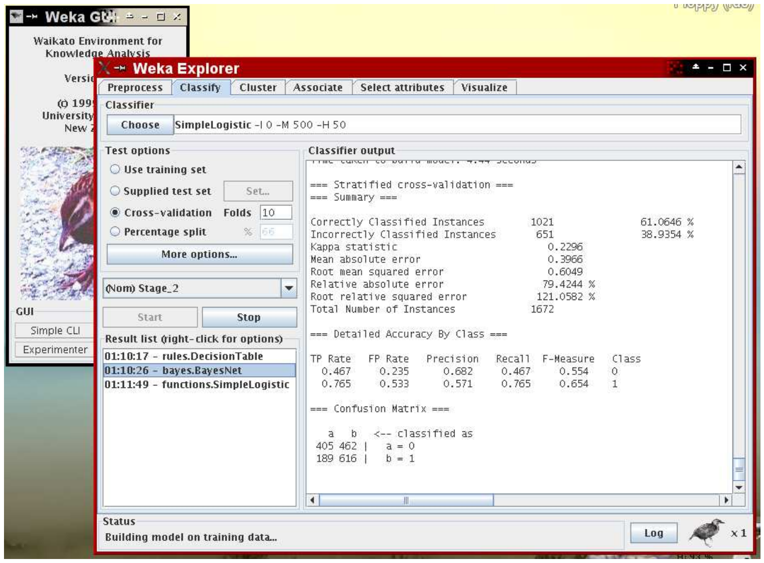

Figure A.2: Weka Screenshot

tion and others may be downloaded from various places on the internet. I have used many in my work like Statistical toolbox, Signal processing toolbox, etc.

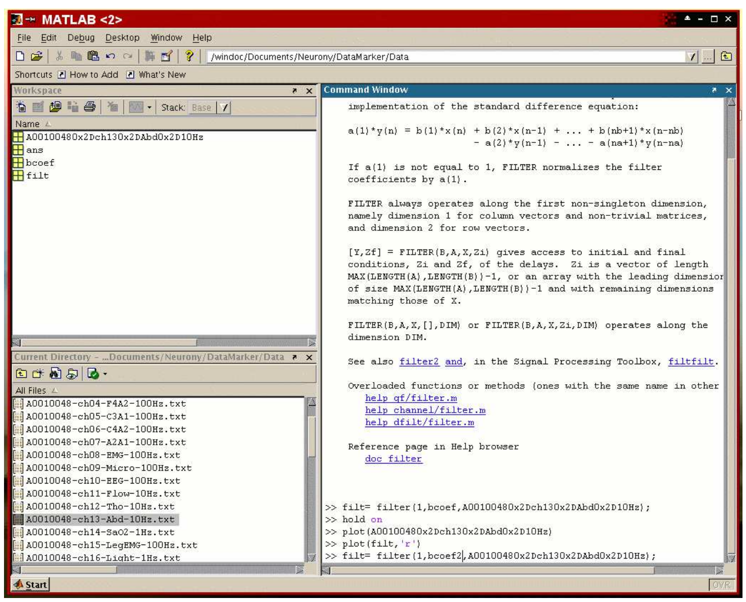

Figure A.3: Matlab screenshot

The CTU have obtained a Matlab education licence. Students and teachers may download it from http://nss.cvut.cz. The official site is http://www.mathworks.com. On the official site is also great support and documentation section.

Class Assignment Tool (CATool)

This is a tool I developed in Matlab to simplify my work with time data. It has nothing to do with softcomputing methods described above, but focus on earlier phase of data mining – preprocessing and feature extraction. I have done all tasks related to preprocessing in this tool.

It can import all data formats used in my work (plain text and European Data Format widely used

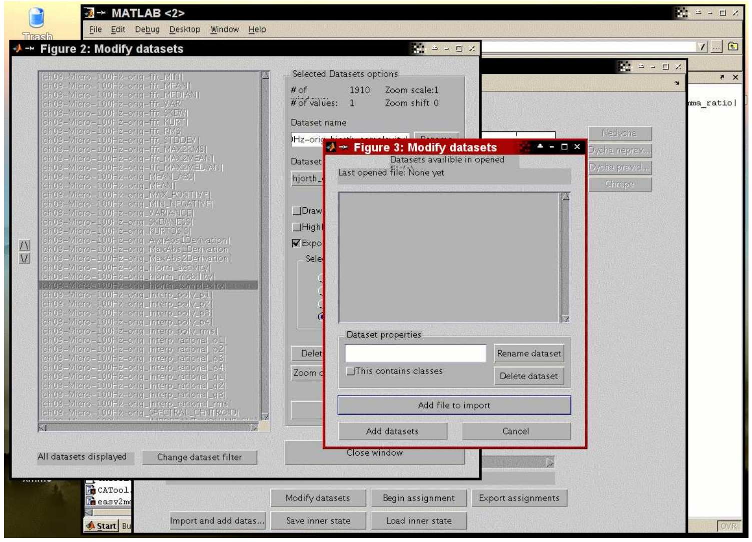

Figure A.4: CATool screenshot

for exchange the EEG data) and can export data files ready to use in both Weka (A) and GAME (A). I have also developed many plug-ins for this tools which calculates all wanted features for me (for list of features see 8).

This tool may be downloaded from http://neuron.felk.cvut.cz/∼cepekm1/CATool/index.html

Appendix B

Complete list of features

B.1 Note on Symbols

In following sections I will use some common symbols. Here I will describe its meaning.

N	Number of samples in currently processed window.
x	all samples in time domain window.
x'	first derivation of signal x.
x''	second derivation of signal x.
x_i	i-th sample in time domain signal.
$CWT(x)_i$	i-th value in continuous wavelet transformation of signal x.
$FFT(x)_i$	i-th value in Fourier transformation of signal x.

B.2 Features in original work

Here is list of features used by Ing. Josef Rieger in his diploma thesis [4].

Average absolute amplitude

$$A_{mean\ abs} = \frac{1}{N}\sum_{j=1}^{N}|x_j|$$

Average amplitude

$$A_{mean} = \frac{1}{N}\sum_{j=1}^{N}x_j$$

Maximum positive amplitude

$$A_{max+} = \max_{j=1,\cdots,N} x_i$$

Maximum negative amplitude

$$A_{max-} = -\max_{j=1,\cdots,N}(-x_j)$$

Variance

$$varX = \frac{1}{N}\sum_{j=1}^{N}(x_j - \bar{x})^2$$

Skewness

$$\gamma_1 = \frac{\mu_3}{\sigma^3} = \frac{\mu_3}{(\mu_2)^{\frac{3}{2}}}$$

Kurtosis

$$\gamma_2 = \frac{\mu_4}{\sigma^4} - 3 = \frac{\mu_4}{(\mu_2)^2} - 3$$

Average absolute first derivation

$$D^{(1)}_{mean\ abs} = \frac{1}{T_{vz}(N-1)} \sum_{j=2}^{N} |x_j - x_{j-1}|$$

Absolute maximum of first derivation

$$D^{(1)}_{max\ abs} = \frac{1}{T_{vz}} \max_{j=2,\cdots,N} |x_j - x_{j-1}|$$

Absolute maximum of second derivation

$$D^{(2)}_{max\ abs} = \frac{1}{T_{vz}} \max_{j=2,\cdots,N} |x'_j - x'_{j-1}|$$

B.3 New Tested Features

B.3.1 Continuous Wavelet Transformation Statistics

These features are selected statistics properties of countinuous wavelet transform.

CWT_mean

$$CWT_{mean}(x) = \frac{1}{N} \sum_{j=1}^{N} CWT(x)_j$$

CWT_std_dev

$$CWT_{std\_dev}(x) = \sigma_{CWT} = \sqrt{\frac{1}{N-1} \sum_{j=1}^{N} \left(CWT(x)_j - \overline{CWT(x)} \right)^2}$$

CWT_variance

$$CWT_{var}(x) = \frac{1}{N-1} \sum_{j=1}^{N} \left(CWT(x)_j - \overline{CWT(x)} \right)^2$$

CWT_median

$$\mathcal{P}\left(\mathcal{X} \geq CWT_{median}\right) = \int_{-\infty}^{CWT_{median}} f(x)dx = 0.5$$

CWT_max

$$CWT_{max}(x) = \max_{j=1\cdots N} (CWT(x)_j)$$

CWT_min

$$CWT_{min}(x) = \min_{j=1\cdots N} (CWT(x)_j)$$

CWT_rms

$$CWT_{RMS}(x) = \frac{1}{N} \sum_{j=1}^{N} (CWT(x)_j^2)$$

CWT_skewness

$$\gamma_1 = \frac{\mu_3}{\sigma^3} = \frac{\mu_3}{(\mu_2)^{\frac{3}{2}}}$$

CWT_kurtosis

$$\gamma_2 = \frac{\mu_4}{\sigma^4} - 3 = \frac{\mu_4}{(\mu_2)^2} - 3$$

CWT_max2rms

$$CWT_{max2rms}(x) = \frac{CWT_{max}(x)}{CWT_{rms}(x)}$$

CWT_max2mean

$$CWT_{max2mean}(x) = \frac{CWT_{max}(x)}{CWT_{mean}(x)}$$

CWT_max2median

$$CWT_{max2median}(x) = \frac{CWT_{max}(x)}{CWT_{median}(x)}$$

B.3.2 Fourier Coefitients Statistics

These features are selected statistics features of fourier transform.

FFT_mean

$$FFT_{mean}(x) = \frac{1}{N}\sum_{j=1}^{N} FFT(x)_j$$

FFT_std_dev

$$FFT_{std\_dev}(x) = \sigma_{FFT} = \sqrt{\frac{1}{N-1}\sum_{j=1}^{N}\left(FFT(x)_j - \overline{FFT(x)}\right)^2}$$

FFT_variance

$$FFT_{var}(x) = \frac{1}{N-1}\sum_{j=1}^{N}\left(FFT(x)_j - \overline{FFT(x)}\right)^2$$

FFT_median

$$\mathcal{P}\left(\mathcal{X} \geq FFT_{median}\right) = \int_{-\infty}^{FFT_{median}} f(x)dx = 0.5$$

FFT_max

$$FFT_{max}(x) = \max_{j=1\cdots N}(FFT(x)_j)$$

FFT_min

$$FFT_{min}(x) = \min_{j=1\cdots N}(FFT(x)_j)$$

FFT_rms

$$FFT_{RMS}(x) = \frac{1}{N}\sum_{j=1}^{N}(FFT(x)_j^2)$$

FFT_skewness

$$\gamma_1 = \frac{\mu_3}{\sigma^3} = \frac{\mu_3}{(\mu_2)^{\frac{3}{2}}}$$

FFT_kurtosis

$$\gamma_2 = \frac{\mu_4}{\sigma^4} - 3 = \frac{\mu_4}{(\mu_2)^2} - 3$$

FFT_max2rms

$$FFT_{max2rms}(x) = \frac{FFT_{max}(x)}{FFT_{rms}(x)}$$

FFT_max2mean

$$FFT_{max2mean}(x) = \frac{FFT_{max}(x)}{FFT_{mean}(x)}$$

FFT_max2median

$$FFT_{max2median}(x) = \frac{FFT_{max}(x)}{FFT_{median}(x)}$$

B.3.3 Fourier Frequency Band Spectrum Energy

These features represents band powers of selected frequency ranges. The frequency ranges corresponding to the EEG frequency band division are specified in the beginning of my work.

FFT_band_0-2

$$FFT_{band\_0-2}(x) = \sum_{j \in 0...2\ Hz\ FFT\ coefitients} |fft(x)|$$

FFT_band_2-4

$$FFT_{band\_2-4}(x) = \sum_{j \in 2...4\ Hz\ FFT\ coefitients} |fft(x)|$$

FFT_band_4-6

$$FFT_{band\_4-6}(x) = \sum_{j \in 4...6\ Hz\ FFT\ coefitients} |fft(x)|$$

FFT_band_6-8

$$FFT_{band\_6-8}(x) = \sum_{j \in 6...8\ Hz\ FFT\ coefitients} |fft(x)|$$

FFT_band_8-10

$$FFT_{band\_8-10}(x) = \sum_{j \in 8...10\ Hz\ FFT\ coefitients} |fft(x)|$$

FFT_band_10-13

$$FFT_{band\_10-13}(x) = \sum_{j \in 10...13\ Hz\ FFT\ coefitients} |fft(x)|$$

FFT_band_13-19

$$FFT_{band\_13-19}(x) = \sum_{j \in 13...19\ Hz\ FFT\ coefitients} |fft(x)|$$

FFT_band_19-24

$$FFT_{band\_19-24}(x) = \sum_{j \in 19...24\ Hz\ FFT\ coefitients} |fft(x)|$$

FFT_band_24-30

$$FFT_{band\_24-30}(x) = \sum_{j \in 24...30\ Hz\ FFT\ coefitients} |fft(x)|$$

FFT_band_30-50

$$FFT_{band\_30-50}(x) = \sum_{j \in 2...4\ Hz\ FFT\ coefitients} |fft(x)|$$

FFT_banddelta

$$FFT_{band\_delta}(x) = \sum_{j \in delta\ Hz\ FFT\ coefitients} |fft(x)|$$

FFT_bandtheta

$$FFT_{band\_delta}(x) = \sum_{j \in delta\ Hz\ FFT\ coefitients} |fft(x)|$$

FFT_bandalpha

$$FFT_{band\_aplha}(x) = \sum_{j \in alpha\ Hz\ FFT\ coefitients} |fft(x)|$$

FFT_bandbeta

$$FFT_{band\_beta}(x) = \sum_{j \in beta\ Hz\ FFT\ coefitients} |fft(x)|$$

FFT_bandgamma

$$FFT_{band\_gamma}(x) = \sum_{j \in gamma\ Hz\ FFT\ coefitients} |fft(x)|$$

FFT_band_delta2theta_ratio

$$FFT_{band\_delta2theta\_ratio} = \frac{FFT_{band\_delta}(x)}{FFT_{band\_theta}(x)}$$

FFT_band_delta2alpha_ratio

$$FFT_{band\_delta2theta\_ratio} = \frac{FFT_{band\_delta}(x)}{FFT_{band\_alpha}(x)}$$

FFT_band_delta2beta_ratio

$$FFT_{band\_delta2theta\_ratio} = \frac{FFT_{band\_delta}(x)}{FFT_{band\_beta}(x)}$$

FFT_band_delta2gamma_ratio

$$FFT_{band\_delta2theta\_ratio} = \frac{FFT_{band\_delta}(x)}{FFT_{band\_gamma}(x)}$$

FFT_band_theta2alpha_ratio

$$FFT_{band\_delta2theta\_ratio} = \frac{FFT_{band\_theta}(x)}{FFT_{band\_alpha}(x)}$$

FFT_band_theta2beta_ratio

$$FFT_{band\_delta2theta\_ratio} = \frac{FFT_{band\_theta}(x)}{FFT_{band\_beta}(x)}$$

FFT_band_theta2gamma_ratio

$$FFT_{band\_delta2theta\_ratio} = \frac{FFT_{band\_theta}(x)}{FFT_{band\_gamma}(x)}$$

FFT_band_alpha2beta_ratio

$$FFT_{band\_delta2theta\_ratio} = \frac{FFT_{band\_alpha}(x)}{FFT_{band\_beta}(x)}$$

FFT_band_alpha2gamma_ratio

$$FFT_{band\_delta2theta\_ratio} = \frac{FFT_{band\_alpha}(x)}{FFT_{band\_gamma}(x)}$$

FFT_band_beta2gamma_ratio

$$FFT_{band\_delta2theta\_ratio} = \frac{FFT_{band\_beta}(x)}{FFT_{band\_gamma}(x)}$$

B.3.4 Spectral Centroid

This term means the frequency which divides a spectrum power into two equal halves.

spectralcentroidfreq

$$\sum_{i \in 0 \ldots Spectral\_centroid\ Hz\ FT\ coefitients} |fft(x)| = \sum_{i \in Spectral\_centroid \ldots f_max\ Hz\ FT\ coefitients} |fft(x)|$$

B.3.5 Time Domain Features – Ratio Values of Time Signal

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1} N(x_i^2)}$$

Max2rms_ratio

$$max2rms\_ratio = \frac{\max(x)}{RMS}$$

Max2mean_ratio

$$max2rms\_ratio = \frac{\max(x)}{mean(x)}$$

Max2median_ratio

$$max2rms\_ratio = \frac{\max(x)}{median(x)}$$

B.3.6 Hjorth Parameters

hjorth_activity

$$activity = m_0 = \sigma_x^2$$

hjorth_mobility

$$mobility = \frac{\sigma_{x'}}{\sigma_x}$$

hjorth_complexity

$$complexity = \frac{\sigma_{x''}/\sigma_{x'}}{\sigma_{x'}/\sigma_x}$$

B.3.7 Important Volume Changes Detector

See 3.5 for details.

number_of_changes

$$number\_of\_changes = number\_of\_generated\_events$$

B.3.8 Polynomial Interpolation Parameters

Let P be a polynomial of 3th degree. $P(x) = \alpha_1 x^3 + \alpha_2 x^2 + \alpha_3 x + \alpha_4$. The polynomial will be fitted to the signal and α_1, α_2, α_3, α_4 will act as features this way.

interp_poly_p1

$$interp\_poly\_p1 = \alpha_1$$

interp_poly_p2

$$interp\_poly\_p2 = \alpha_2$$

interp_poly_p3

$$interp\_poly\_p3 = \alpha_3$$

interp_poly_p4

$$interp\_poly\_p4 = \alpha_4$$

interp_poly_rms

$$interp\_poly\_rms = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - P(x_i))^2}$$

B.3.9 Rational Interpolation Parameters

Let R be a fraction of two polynomials of 3th degree. $R(x) = \frac{\alpha_1 x^3 + \alpha_2 x^2 + \alpha_3 x + \alpha_4}{x^3 + \beta_1 x^2 + \beta_2 x + \beta_3}$. The fraction of polynomials will be fitted to the signal and α_1, α_2, α_3, α_4, β_1, β_2, β_3 will act as features this way.

interp_rational_p1

$$interp\_rational\_p1 = \alpha_1$$

interp_rational_p2

$$interp\_rational\_p2 = \alpha_2$$

interp_rational_p3

$$interp\_rational\_p3 = \alpha_3$$

interp_rational_p4

$$interp\_rational\_p4 = \alpha_4$$

interp_rational_q1

$$interp\_rational\_q1 = \beta_1$$

interp_rational_q2

$$interp\_rational\_q2 = \beta_2$$

interp_rational_q3

$$interp\_rational\_q3 = \beta_3$$

interp_poly_rms

$$interp\_poly\_rms = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - R(x_i))^2}$$

Appendix C

Evaluation of Proposed Concept – List of Selected Features

C.1 Class: Not Breathing

CfsSubsetEvaluator and BestFirst	CfsSubsetEvaluator and GeneticSearch
ch09-Micro-100Hz-orig-cwt_SKEW	ch09-Micro-100Hz-orig-cwt_MEAN
ch09-Micro-100Hz-orig_hjorth_complexity	ch09-Micro-100Hz-orig-fft_MAX
ch09-Micro-100Hz-wined_FFT_band_delta	ch09-Micro-100Hz-orig-fft_SKEW
ch11-Flow-100Hz-orig-cwt_MIN	ch09-Micro-100Hz-orig-fft_KURT
ch11-Flow-100Hz-orig-cwt_MEDIAN	ch09-Micro-100Hz-orig_MEAN
ch11-Flow-100Hz-orig-cwt_SKEW	ch09-Micro-100Hz-orig_VARIANCE
ch11-Flow-100Hz-orig-cwt_KURT	ch09-Micro-100Hz-orig_KURTOSIS
ch11-Flow-100Hz-orig-cwt_MAX2MEAN	ch09-Micro-100Hz-orig_AvgAbs1Derivation
ch11-Flow-100Hz-orig-cwt_MAX2MEDIAN	ch09-Micro-100Hz-orig_MaxAbs1Derivation
ch11-Flow-100Hz-orig-fft_SKEW	ch09-Micro-100Hz-orig_hjorth_mobility
ch11-Flow-100Hz-orig-fft_RMS	ch09-Micro-100Hz-orig_IMPORTANT...
ch11-Flow-100Hz-orig-fft_MAX2MEAN	ch09-Micro-100Hz-wined_13-19-Band
ch11-Flow-100Hz-orig_SKEWNESS	ch09-Micro-100Hz-wined_FFT_band_delta2theta...
ch11-Flow-100Hz-orig_KURTOSIS	ch09-Micro-100Hz-wined_FFT_band_delta2gamma...
ch11-Flow-100Hz-orig_MaxAbs2Derivation	ch09-Micro-100Hz-wined_FFT_band_theta2alpha...
ch11-Flow-100Hz-orig_hjorth_mobility	ch11-Flow-100Hz-orig-cwt_MIN
	ch11-Flow-100Hz-orig-cwt_MEAN
	ch11-Flow-100Hz-orig-cwt_SKEW
	ch11-Flow-100Hz-orig-cwt_MAX2RMS
	ch11-Flow-100Hz-orig-cwt_MAX2MEAN
	ch11-Flow-100Hz-orig-fft_MAX
	ch11-Flow-100Hz-orig-fft_VAR
	ch11-Flow-100Hz-orig-fft_SKEW
	ch11-Flow-100Hz-orig-fft_RMS
	ch11-Flow-100Hz-orig-fft_MAX2RMS
	ch11-Flow-100Hz-orig_MEAN
	ch11-Flow-100Hz-orig_VARIANCE
	ch11-Flow-100Hz-orig_KURTOSIS

Table C.1: Selected features – Class: Not Breathing – CfsSubsetEvaluator and BestFirst, CfsSubsetEvaluator and GeneticSearch

Continued on next page

GAME selected features	Ranker and ChiSquaredEvaluator
GAME selected features	**Ranker and ChiSquaredEvaluator**
ch09-Micro-100Hz-orig_MaxAbs2Derivation	ch11-Flow-100Hz-orig-cwt_MEDIAN
ch09-Micro-100Hz-orig_SPECTRAL...	ch11-Flow-100Hz-orig_KURTOSIS
ch09-Micro-100Hz-orig-cwt_MAX2MEAN	ch11-Flow-100Hz-orig-cwt_MAX2MEAN
ch09-Micro-100Hz-orig-cwt_MEDIAN	ch11-Flow-100Hz-orig-cwt_SKEW
ch09-Micro-100Hz-orig-fft_MAX2MEAN	ch11-Flow-100Hz-orig-cwt_KURT
ch09-Micro-100Hz-orig-fft_MAX2MEDIAN	ch11-Flow-100Hz-orig-cwt_MEAN
ch09-Micro-100Hz-orig-fft_MAX2RMS	ch11-Flow-100Hz-orig-cwt_MAX2RMS
ch09-Micro-100Hz-orig-fft_MEDIAN	ch11-Flow-100Hz-orig_AvgAbs1Derivation
ch09-Micro-100Hz-wined_FFT_band_beta	ch11-Flow-100Hz-orig-fft_MAX2MEAN
ch09-Micro-100Hz-wined_FFT_band_delta	ch11-Flow-100Hz-orig-cwt_MAX2MEDIAN
ch09-Micro-100Hz-wined_FFT_band_delta2alpha...	ch11-Flow-100Hz-orig-fft_MEAN
ch09-Micro-100Hz-wined_FFT_band_delta2beta...	ch11-Flow-100Hz-orig-fft_SKEW
ch09-Micro-100Hz-wined_FFT_band_delta2g...	ch11-Flow-100Hz-orig-fft_KURT
ch09-Micro-100Hz-wined_FFT_band_delta2theta...	ch11-Flow-100Hz-orig-fft_MAX2RMS
ch11-Flow-100Hz-orig_KURTOSIS	ch11-Flow-100Hz-wined_FFT_band_delta
ch11-Flow-100Hz-orig_MIN_NEGATIVE	ch11-Flow-100Hz-orig_SKEWNESS
ch11-Flow-100Hz-orig-cwt_MAX2RMS	ch11-Flow-100Hz-orig_interp_poly_rms
ch11-Flow-100Hz-orig-cwt_MEDIAN	ch09-Micro-100Hz-wined_FFT_band_delta
ch11-Flow-100Hz-orig-cwt_RMS	ch11-Flow-100Hz-orig-fft_RMS
ch11-Flow-100Hz-orig-fft_KURT	ch11-Flow-100Hz-orig-fft_VAR
ch11-Flow-100Hz-orig-fft_MAX2MEAN	ch11-Flow-100Hz-orig-fft_STDDEV
ch11-Flow-100Hz-orig-fft_MAX2RMS	ch11-Flow-100Hz-orig_hjorth_complexity
ch11-Flow-100Hz-orig-fft_MEAN	ch09-Micro-100Hz-wined_FFT_band_beta
ch11-Flow-100Hz-orig-fft_SKEW	ch09-Micro-100Hz-wined_13-19-Band
ch11-Flow-100Hz-orig-fft_STDDEV	ch09-Micro-100Hz-wined_FFT_band_delta2beta...
ch11-Flow-100Hz-orig-fft_VAR	ch09-Micro-100Hz-wined_FFT_band_alpha
ch11-Flow-100Hz-wined_19-24-Band	ch09-Micro-100Hz-orig-fft_MEAN
ch11-Flow-100Hz-wined_FFT_band_delta	ch09-Micro-100Hz-orig-fft_MAX2MEAN
ch11-Flow-100Hz-wined_FFT_band_delta2beta...	ch09-Micro-100Hz-wined_FFT_band_theta
	ch09-Micro-100Hz-wined_19-24-Band

Table C.2: Selected features – Class: Not Breathing – GAME selected features, Ranker and ChiSquaredEvaluator

Ranker and GainRatio	Ranker and InfoGain
ch11-Flow-100Hz-orig-cwt_KURT	ch11-Flow-100Hz-orig-cwt_MEDIAN
ch11-Flow-100Hz-orig-cwt_MEDIAN	ch11-Flow-100Hz-orig-cwt_MAX2MEAN
ch11-Flow-100Hz-orig_KURTOSIS	ch11-Flow-100Hz-orig_KURTOSIS
ch11-Flow-100Hz-orig_MaxAbs2Derivation	ch11-Flow-100Hz-orig-cwt_SKEW
ch11-Flow-100Hz-orig-cwt_SKEW	ch11-Flow-100Hz-orig-cwt_MAX2MEDIAN
ch11-Flow-100Hz-orig-cwt_MAX2MEAN	ch11-Flow-100Hz-orig-cwt_MAX2RMS
ch11-Flow-100Hz-orig-fft_SKEW	ch11-Flow-100Hz-orig-cwt_MEAN
ch11-Flow-100Hz-orig-fft_KURT	ch11-Flow-100Hz-orig_AvgAbs1Derivation
ch11-Flow-100Hz-orig-cwt_MAX2MEDIAN	ch11-Flow-100Hz-orig-fft_MAX2MEAN
ch11-Flow-100Hz-orig-cwt_MAX2RMS	ch11-Flow-100Hz-orig-cwt_KURT
ch11-Flow-100Hz-orig-fft_MAX2MEAN	ch11-Flow-100Hz-orig-fft_SKEW
ch11-Flow-100Hz-orig_SKEWNESS	ch11-Flow-100Hz-orig-fft_MEAN
ch11-Flow-100Hz-orig-fft_MAX2RMS	ch11-Flow-100Hz-orig-fft_KURT
ch11-Flow-100Hz-orig-fft_MEAN	ch11-Flow-100Hz-orig-fft_MAX2RMS
ch11-Flow-100Hz-wined_FFT_band_delta	ch11-Flow-100Hz-wined_FFT_band_delta
ch11-Flow-100Hz-orig-cwt_MIN	ch11-Flow-100Hz-orig_interp_poly_rms
ch11-Flow-100Hz-orig-cwt_MEAN	ch09-Micro-100Hz-wined_FFT_band_delta

Continued on next page

Ranker and GainRatio	**Ranker and InfoGain**
ch11-Flow-100Hz-orig_AvgAbs1Derivation	ch11-Flow-100Hz-orig-fft_RMS
ch11-Flow-100Hz-orig_interp_poly_rms	ch11-Flow-100Hz-orig-fft_STDDEV
ch11-Flow-100Hz-orig_hjorth_mobility	ch11-Flow-100Hz-orig-fft_VAR
ch09-Micro-100Hz-wined_FFT_band_delta2theta...	ch11-Flow-100Hz-orig_VARIANCE
ch09-Micro-100Hz-wined_FFT_band_delta	ch11-Flow-100Hz-orig_hjorth_activity
ch09-Micro-100Hz-wined_FFT_band_theta	ch09-Micro-100Hz-wined_FFT_band_beta
ch09-Micro-100Hz-wined_FFT_band_beta	ch09-Micro-100Hz-wined_13-19-Band
ch11-Flow-100Hz-orig-fft_RMS	ch11-Flow-100Hz-orig_SKEWNESS
ch11-Flow-100Hz-orig-fft_STDDEV	ch09-Micro-100Hz-wined_FFT_band_delta2beta...
ch11-Flow-100Hz-orig-fft_VAR	ch09-Micro-100Hz-orig-fft_MAX2MEAN
ch09-Micro-100Hz-wined_FFT_band_delta2beta...	ch09-Micro-100Hz-orig-fft_MEAN
ch09-Micro-100Hz-wined_13-19-Band	ch09-Micro-100Hz-wined_19-24-Band
ch09-Micro-100Hz-wined_FFT_band_alpha	ch09-Micro-100Hz-orig-fft_MEDIAN

Table C.3: Selected features – Class: Not Breathing – Ranker and GainRatio, Ranker and CInfoGain

C.2 Class: Breathing Irregulary

CfsSubsetEvaluator and BestFirst	**CfsSubsetEvaluator and GeneticSearch**
ch09-Micro-100Hz-orig-cwt_MAX2MEDIAN	ch09-Micro-100Hz-orig-cwt_MAX2MEAN
ch09-Micro-100Hz-wined_FFT_band_theta	ch09-Micro-100Hz-orig-fft_SKEW
ch11-Flow-100Hz-orig-cwt_MEDIAN	ch09-Micro-100Hz-orig-fft_STDDEV
ch11-Flow-100Hz-orig-cwt_MAX2RMS	ch09-Micro-100Hz-orig_VARIANCE
ch11-Flow-100Hz-orig-fft_KURT	ch09-Micro-100Hz-orig_AvgAbs1Derivation
ch11-Flow-100Hz-orig_SKEWNESS	ch09-Micro-100Hz-orig_MaxAbs1Derivation
ch11-Flow-100Hz-orig_KURTOSIS	ch09-Micro-100Hz-orig_MaxAbs2Derivation
ch11-Flow-100Hz-orig_AvgAbs1Derivation	ch09-Micro-100Hz-orig_hjorth_mobility
	ch09-Micro-100Hz-orig_interp_poly_rms
	ch09-Micro-100Hz-orig_SPECTRAL...
	ch09-Micro-100Hz-wined_FFT_band_theta
	ch09-Micro-100Hz-wined_FFT_band_delta2theta...
	ch09-Micro-100Hz-wined_FFT_band_delta2alpha...
	ch09-Micro-100Hz-wined_FFT_band_delta2beta...
	ch09-Micro-100Hz-wined_FFT_band_delta2g...
	ch11-Flow-100Hz-orig-cwt_MAX
	ch11-Flow-100Hz-orig-cwt_MIN
	ch11-Flow-100Hz-orig-cwt_MEAN
	ch11-Flow-100Hz-orig-cwt_MAX2MEAN
	ch11-Flow-100Hz-orig-fft_MAX
	ch11-Flow-100Hz-orig-fft_MIN
	ch11-Flow-100Hz-orig-fft_KURT
	ch11-Flow-100Hz-orig-fft_RMS
	ch11-Flow-100Hz-orig-fft_MAX2MEDIAN
	ch11-Flow-100Hz-orig_VARIANCE
	ch11-Flow-100Hz-orig_KURTOSIS
	ch11-Flow-100Hz-orig_AvgAbs1Derivation
	ch11-Flow-100Hz-orig_MaxAbs1Derivation
	ch11-Flow-100Hz-orig_MaxAbs2Derivation
	ch11-Flow-100Hz-orig_hjorth_mobility
	ch11-Flow-100Hz-wined_FFT_band_delta2alpha...
	ch11-Flow-100Hz-wined_FFT_band_theta2alpha...

Table C.4: Selected features – Class: Breathing Irregulary – CfsSubsetEvaluator and BestFirst, CfsSubsetEvaluator and GeneticSearch

GAME selected features	**Ranker and ChiSquaredEvaluator**
ch09-Micro-100Hz-orig_AvgAbs1Derivation	ch11-Flow-100Hz-orig_KURTOSIS
ch09-Micro-100Hz-orig-cwt_MEAN	ch11-Flow-100Hz-orig_AvgAbs1Derivation
ch09-Micro-100Hz-orig-cwt_RMS	ch11-Flow-100Hz-orig-cwt_MEAN
ch09-Micro-100Hz-orig_interp_poly_rms	ch11-Flow-100Hz-orig-cwt_MEDIAN
ch09-Micro-100Hz-orig-cwt_STDDEV	ch11-Flow-100Hz-orig_hjorth_complexity
ch09-Micro-100Hz-wined_FFT_band_gamma	ch11-Flow-100Hz-orig_interp_poly_rms
ch09-Micro-100Hz-orig-fft_MAX2MEAN	ch11-Flow-100Hz-orig-cwt_RMS
ch09-Micro-100Hz-wined_FFT_band_delta2theta...	ch11-Flow-100Hz-orig_VARIANCE
ch09-Micro-100Hz-wined_FFT_band_delta2alpha...	ch11-Flow-100Hz-orig_hjorth_activity
ch09-Micro-100Hz-wined_10-13-Band	ch11-Flow-100Hz-orig-fft_MAX2RMS
ch11-Flow-100Hz-orig-cwt_MEDIAN	ch11-Flow-100Hz-orig-fft_KURT
ch09-Micro-100Hz-wined_30-50-Band	ch11-Flow-100Hz-orig-fft_SKEW
ch11-Flow-100Hz-orig_hjorth_mobility	ch11-Flow-100Hz-orig_interp_rational_rms

Continued on next page

GAME selected features	Ranker and ChiSquaredEvaluator
ch11-Flow-100Hz-orig-cwt_MEAN	ch11-Flow-100Hz-orig_SKEWNESS
ch11-Flow-100Hz-orig-cwt_MIN	ch11-Flow-100Hz-orig-cwt_MAX2RMS
ch09-Micro-100Hz-wined_FFT_band_delta2g...	ch11-Flow-100Hz-orig-cwt_SKEW
ch09-Micro-100Hz-orig_IMPORTANT...	ch11-Flow-100Hz-orig-cwt_VAR
ch09-Micro-100Hz-orig-fft_MAX2MEDIAN	ch11-Flow-100Hz-orig-cwt_STDDEV
ch09-Micro-100Hz-orig_SPECTRAL...	ch11-Flow-100Hz-orig-fft_MAX2MEAN
ch09-Micro-100Hz-orig-cwt_SKEW	ch11-Flow-100Hz-orig-fft_MEAN
ch09-Micro-100Hz-orig_VARIANCE	ch11-Flow-100Hz-orig-cwt_MAX2MEAN
ch11-Flow-100Hz-orig_interp_poly_rms	ch11-Flow-100Hz-wined_FFT_band_delta
ch11-Flow-100Hz-orig_AvgAbs1Derivation	ch11-Flow-100Hz-orig-fft_RMS
ch09-Micro-100Hz-wined_FFT_band_theta	ch11-Flow-100Hz-orig-fft_VAR
ch09-Micro-100Hz-orig_MaxAbs1Derivation	ch11-Flow-100Hz-orig-fft_STDDEV
ch09-Micro-100Hz-orig-fft_MEAN	ch11-Flow-100Hz-orig-cwt_MAX
ch09-Micro-100Hz-wined_FFT_band_delta2beta...	ch11-Flow-100Hz-orig-cwt_KURT
ch09-Micro-100Hz-wined_FFT_band_alpha	ch11-Flow-100Hz-orig_MaxAbs1Derivation
ch11-Flow-100Hz-wined_FFT_band_delta2beta...	ch11-Flow-100Hz-orig-cwt_MAX2MEDIAN
ch09-Micro-100Hz-orig_hjorth_activity	
ch11-Flow-100Hz-orig-fft_MEAN	
ch11-Flow-100Hz-orig_KURTOSIS	
ch11-Flow-100Hz-orig-fft_MAX2MEAN	
ch11-Flow-100Hz-orig_interp_poly_p4	
ch09-Micro-100Hz-orig_MEAN	

Table C.5: Selected features – Class: Breathing Irregulary – GAME selected features, Ranker and ChiSquaredEvaluator

Ranker and GainRatio	Ranker and InfoGain
ch11-Flow-100Hz-orig_KURTOSIS	ch11-Flow-100Hz-orig_KURTOSIS
ch11-Flow-100Hz-orig_AvgAbs1Derivation	ch11-Flow-100Hz-orig_AvgAbs1Derivation
ch11-Flow-100Hz-orig-cwt_MEAN	ch11-Flow-100Hz-orig-cwt_MEAN
ch11-Flow-100Hz-orig-cwt_MEDIAN	ch11-Flow-100Hz-orig-cwt_MEDIAN
ch09-Micro-100Hz-wined_FFT_band_theta	ch11-Flow-100Hz-orig_interp_poly_rms
ch11-Flow-100Hz-orig-fft_KURT	ch11-Flow-100Hz-orig_hjorth_complexity
ch11-Flow-100Hz-orig-fft_MAX2RMS	ch11-Flow-100Hz-orig-cwt_RMS
ch11-Flow-100Hz-orig_interp_poly_rms	ch11-Flow-100Hz-orig-fft_SKEW
ch11-Flow-100Hz-orig_VARIANCE	ch11-Flow-100Hz-orig-fft_KURT
ch11-Flow-100Hz-orig_hjorth_activity	ch11-Flow-100Hz-orig_VARIANCE
ch11-Flow-100Hz-orig-fft_SKEW	ch11-Flow-100Hz-orig_hjorth_activity
ch09-Micro-100Hz-wined_FFT_band_delta2g...	ch11-Flow-100Hz-orig-fft_MAX2RMS
ch09-Micro-100Hz-orig-fft_MEDIAN	ch11-Flow-100Hz-orig_SKEWNESS
ch09-Micro-100Hz-orig-fft_MAX2MEDIAN	ch11-Flow-100Hz-orig_interp_rational_rms
ch09-Micro-100Hz-orig-fft_RMS	ch11-Flow-100Hz-orig-fft_MAX2MEAN
ch09-Micro-100Hz-orig_AvgAbs1Derivation	ch11-Flow-100Hz-orig-fft_MEAN
ch11-Flow-100Hz-orig_SKEWNESS	ch11-Flow-100Hz-orig-cwt_MAX2RMS
ch09-Micro-100Hz-orig-cwt_MEAN	ch11-Flow-100Hz-wined_FFT_band_delta
ch09-Micro-100Hz-wined_FFT_band_delta2theta...	ch11-Flow-100Hz-orig-cwt_VAR
ch09-Micro-100Hz-orig-fft_MAX2MEAN	ch11-Flow-100Hz-orig-cwt_STDDEV
ch09-Micro-100Hz-orig-fft_MEAN	ch11-Flow-100Hz-orig-cwt_SKEW
ch09-Micro-100Hz-orig-fft_MAX2RMS	ch11-Flow-100Hz-orig-cwt_MAX2MEAN
ch11-Flow-100Hz-orig_hjorth_complexity	ch11-Flow-100Hz-orig-cwt_MAX
ch09-Micro-100Hz-wined_FFT_band_delta	ch11-Flow-100Hz-orig-fft_RMS
ch09-Micro-100Hz-orig-fft_SKEW	ch11-Flow-100Hz-orig-fft_VAR
ch09-Micro-100Hz-orig-fft_KURT	ch11-Flow-100Hz-orig-fft_STDDEV
ch09-Micro-100Hz-orig-cwt_STDDEV	ch11-Flow-100Hz-orig_MaxAbs1Derivation

Continued on next page

Ranker and GainRatio	**Ranker and InfoGain**
ch09-Micro-100Hz-orig-cwt_VAR	ch11-Flow-100Hz-orig-cwt_KURT
ch09-Micro-100Hz-wined_24-30-Band	ch09-Micro-100Hz-orig-cwt_RMS

Table C.6: Selected features – Class: Breathing Irregulary – Ranker and GainRatio, Ranker and CInfoGain

C.3 Class: Breathing Regulary

CfsSubsetEvaluator and BestFirst	**CfsSubsetEvaluator and GeneticSearch**
ch09-Micro-100Hz-orig-cwt_SKEW	ch09-Micro-100Hz-orig-cwt_MEAN
ch09-Micro-100Hz-orig_hjorth_complexity	ch09-Micro-100Hz-orig-fft_MAX
ch09-Micro-100Hz-wined_FFT_band_delta	ch09-Micro-100Hz-orig-fft_SKEW
ch11-Flow-100Hz-orig-cwt_MIN	ch09-Micro-100Hz-orig-fft_KURT
ch11-Flow-100Hz-orig-cwt_MEDIAN	ch09-Micro-100Hz-orig_MEAN
ch11-Flow-100Hz-orig-cwt_SKEW	ch09-Micro-100Hz-orig_VARIANCE
ch11-Flow-100Hz-orig-cwt_KURT	ch09-Micro-100Hz-orig_KURTOSIS
ch11-Flow-100Hz-orig-cwt_MAX2MEAN	ch09-Micro-100Hz-orig_AvgAbs1Derivation
ch11-Flow-100Hz-orig-cwt_MAX2MEDIAN	ch09-Micro-100Hz-orig_MaxAbs1Derivation
ch11-Flow-100Hz-orig-fft_SKEW	ch09-Micro-100Hz-orig_hjorth_mobility
ch11-Flow-100Hz-orig-fft_RMS	ch09-Micro-100Hz-orig_IMPORTANT...
ch11-Flow-100Hz-orig-fft_MAX2MEAN	ch09-Micro-100Hz-wined_13-19-Band
ch11-Flow-100Hz-orig_SKEWNESS	ch09-Micro-100Hz-wined_FFT_band_delta2theta...
ch11-Flow-100Hz-orig_KURTOSIS	ch09-Micro-100Hz-wined_FFT_band_delta2g...
ch11-Flow-100Hz-orig_MaxAbs2Derivation	ch09-Micro-100Hz-wined_FFT_band_theta2alpha...
ch11-Flow-100Hz-orig_hjorth_mobility	ch11-Flow-100Hz-orig-cwt_MIN
	ch11-Flow-100Hz-orig-cwt_MEAN
	ch11-Flow-100Hz-orig-cwt_SKEW
	ch11-Flow-100Hz-orig-cwt_MAX2RMS
	ch11-Flow-100Hz-orig-cwt_MAX2MEAN
	ch11-Flow-100Hz-orig-fft_MAX
	ch11-Flow-100Hz-orig-fft_VAR
	ch11-Flow-100Hz-orig-fft_SKEW
	ch11-Flow-100Hz-orig-fft_RMS
	ch11-Flow-100Hz-orig-fft_MAX2RMS
	ch11-Flow-100Hz-orig_MEAN
	ch11-Flow-100Hz-orig_VARIANCE
	ch11-Flow-100Hz-orig_KURTOSIS
	ch11-Flow-100Hz-orig_hjorth_activity
	ch11-Flow-100Hz-orig_interp_poly_p4
	ch11-Flow-100Hz-orig_interp_poly_rms
	ch11-Flow-100Hz-orig_interp_rational_q3
	ch11-Flow-100Hz-wined_FFT_band_delta2alpha...
	ch11-Flow-100Hz-wined_FFT_band_alpha2g...

Table C.7: Selected features – Class: Breathing Regulary – CfsSubsetEvaluator and BestFirst, CfsSubsetEvaluator and GeneticSearch

GAME selected features	**Ranker and ChiSquaredEvaluator**
ch11-Flow-100Hz-orig-cwt_MEAN	ch11-Flow-100Hz-orig-cwt_MEDIAN
ch11-Flow-100Hz-orig-cwt_MEDIAN	ch11-Flow-100Hz-orig_KURTOSIS
ch11-Flow-100Hz-orig_AvgAbs1Derivation	ch11-Flow-100Hz-orig-cwt_MAX2MEAN
ch11-Flow-100Hz-orig-fft_MAX2RMS	ch11-Flow-100Hz-orig-cwt_SKEW
ch11-Flow-100Hz-orig_interp_poly_rms	ch11-Flow-100Hz-orig-cwt_KURT
ch11-Flow-100Hz-orig-fft_KURT	ch11-Flow-100Hz-orig-cwt_MEAN
ch09-Micro-100Hz-orig-fft_MEDIAN	ch11-Flow-100Hz-orig-cwt_MAX2RMS
ch09-Micro-100Hz-orig-fft_MAX2MEAN	ch11-Flow-100Hz-orig_AvgAbs1Derivation
ch09-Micro-100Hz-orig-fft_MEAN	ch11-Flow-100Hz-orig-fft_MAX2MEAN
ch09-Micro-100Hz-wined_FFT_band_gamma	ch11-Flow-100Hz-orig-cwt_MAX2MEDIAN
ch11-Flow-100Hz-orig_VARIANCE	ch11-Flow-100Hz-orig-fft_MEAN

Continued on next page

GAME selected features

ch09-Micro-100Hz-orig_IMPORTANT...
ch09-Micro-100Hz-wined_30-50-Band
ch11-Flow-100Hz-orig_hjorth_activity
ch11-Flow-100Hz-orig-cwt_MAX2RMS
ch09-Micro-100Hz-orig-cwt_RMS
ch09-Micro-100Hz-wined_24-30-Band
ch11-Flow-100Hz-orig-cwt_RMS
ch09-Micro-100Hz-orig-cwt_STDDEV
ch09-Micro-100Hz-orig_AvgAbs1Derivation
ch09-Micro-100Hz-wined_FFT_band_beta
ch11-Flow-100Hz-orig-fft_MAX2MEAN
ch11-Flow-100Hz-orig-fft_MEAN
ch11-Flow-100Hz-orig-fft_SKEW
ch09-Micro-100Hz-orig_interp_poly_rms
ch11-Flow-100Hz-orig-fft_STDDEV
ch09-Micro-100Hz-wined_19-24-Band
ch09-Micro-100Hz-orig-cwt_VAR
ch11-Flow-100Hz-orig-fft_RMS
ch09-Micro-100Hz-wined_FFT_band_delta
ch09-Micro-100Hz-orig_MAX_POSITIVE
ch11-Flow-100Hz-orig-fft_VAR
ch11-Flow-100Hz-orig_MIN_NEGATIVE
ch09-Micro-100Hz-orig_hjorth_activity
ch09-Micro-100Hz-wined_FFT_band_alpha
ch09-Micro-100Hz-wined_10-13-Band
ch11-Flow-100Hz-orig_MaxAbs1Derivation
ch11-Flow-100Hz-wined_FFT_band_delta
ch09-Micro-100Hz-orig_VARIANCE
ch09-Micro-100Hz-orig-fft_MAX2RMS
ch09-Micro-100Hz-wined_FFT_band_theta
ch11-Flow-100Hz-orig_interp_rational_rms
ch11-Flow-100Hz-orig_IMPORTANT...
ch09-Micro-100Hz-wined_FFT_band_delta2theta...
ch09-Micro-100Hz-orig-cwt_MAX2MEDIAN
ch11-Flow-100Hz-orig_MaxAbs2Derivation
ch11-Flow-100Hz-orig-cwt_STDDEV
ch09-Micro-100Hz-orig-fft_SKEW

Ranker and ChiSquaredEvaluator

ch11-Flow-100Hz-orig-fft_SKEW
ch11-Flow-100Hz-orig-fft_KURT
ch11-Flow-100Hz-orig-fft_MAX2RMS
ch11-Flow-100Hz-wined_FFT_band_delta
ch11-Flow-100Hz-orig_SKEWNESS
ch11-Flow-100Hz-orig_interp_poly_rms
ch09-Micro-100Hz-wined_FFT_band_delta
ch11-Flow-100Hz-orig-fft_RMS
ch11-Flow-100Hz-orig-fft_VAR
ch11-Flow-100Hz-orig-fft_STDDEV
ch11-Flow-100Hz-orig_hjorth_complexity
ch09-Micro-100Hz-wined_FFT_band_beta
ch09-Micro-100Hz-wined_13-19-Band
ch09-Micro-100Hz-wined_FFT_band_delta2beta...
ch09-Micro-100Hz-wined_FFT_band_alpha
ch09-Micro-100Hz-orig-fft_MEAN
ch09-Micro-100Hz-orig-fft_MAX2MEAN
ch09-Micro-100Hz-wined_FFT_band_theta

Table C.8: Selected features – Class: Breathing Regulary – GAME selected features, Ranker and ChiSquaredEvaluator

Ranker and GainRatio

ch11-Flow-100Hz-orig-cwt_KURT
ch11-Flow-100Hz-orig-cwt_MEDIAN
ch11-Flow-100Hz-orig_KURTOSIS
ch11-Flow-100Hz-orig_MaxAbs2Derivation
ch11-Flow-100Hz-orig-cwt_SKEW
ch11-Flow-100Hz-orig-cwt_MAX2MEAN
ch11-Flow-100Hz-orig-fft_SKEW
ch11-Flow-100Hz-orig-fft_KURT
ch11-Flow-100Hz-orig-cwt_MAX2MEDIAN
ch11-Flow-100Hz-orig-cwt_MAX2RMS
ch11-Flow-100Hz-orig-fft_MAX2MEAN
ch11-Flow-100Hz-orig_SKEWNESS

Ranker and InfoGain

ch11-Flow-100Hz-orig-cwt_MEDIAN
ch11-Flow-100Hz-orig-cwt_MAX2MEAN
ch11-Flow-100Hz-orig_KURTOSIS
ch11-Flow-100Hz-orig-cwt_SKEW
ch11-Flow-100Hz-orig-cwt_MAX2MEDIAN
ch11-Flow-100Hz-orig-cwt_MAX2RMS
ch11-Flow-100Hz-orig-cwt_MEAN
ch11-Flow-100Hz-orig_AvgAbs1Derivation
ch11-Flow-100Hz-orig-fft_MAX2MEAN
ch11-Flow-100Hz-orig-cwt_KURT
ch11-Flow-100Hz-orig-fft_SKEW
ch11-Flow-100Hz-orig-fft_MEAN

Continued on next page

Ranker and GainRatio	**Ranker and InfoGain**
ch11-Flow-100Hz-orig-fft_MAX2RMS	ch11-Flow-100Hz-orig-fft_KURT
ch11-Flow-100Hz-orig-fft_MEAN	ch11-Flow-100Hz-orig-fft_MAX2RMS
ch11-Flow-100Hz-wined_FFT_band_delta	ch11-Flow-100Hz-wined_FFT_band_delta
ch11-Flow-100Hz-orig-cwt_MIN	ch11-Flow-100Hz-orig_interp_poly_rms
ch11-Flow-100Hz-orig-cwt_MEAN	ch09-Micro-100Hz-wined_FFT_band_delta
ch11-Flow-100Hz-orig_AvgAbs1Derivation	ch11-Flow-100Hz-orig-fft_RMS
ch11-Flow-100Hz-orig_interp_poly_rms	ch11-Flow-100Hz-orig-fft_STDDEV
ch11-Flow-100Hz-orig_hjorth_mobility	ch11-Flow-100Hz-orig-fft_VAR
ch09-Micro-100Hz-wined_FFT_band_delta2theta...	ch11-Flow-100Hz-orig_VARIANCE
ch09-Micro-100Hz-wined_FFT_band_delta	ch11-Flow-100Hz-orig_hjorth_activity
ch09-Micro-100Hz-wined_FFT_band_theta	ch09-Micro-100Hz-wined_FFT_band_beta
ch09-Micro-100Hz-wined_FFT_band_beta	ch09-Micro-100Hz-wined_13-19-Band
ch11-Flow-100Hz-orig-fft_RMS	ch11-Flow-100Hz-orig_SKEWNESS
ch11-Flow-100Hz-orig-fft_STDDEV	ch09-Micro-100Hz-wined_FFT_band_delta2beta...
ch11-Flow-100Hz-orig-fft_VAR	ch09-Micro-100Hz-orig-fft_MAX2MEAN
ch09-Micro-100Hz-wined_FFT_band_delta2beta...	ch09-Micro-100Hz-orig-fft_MEAN
ch09-Micro-100Hz-wined_13-19-Band	ch09-Micro-100Hz-wined_19-24-Band

Table C.9: Selected features – Class: Breathing Regulary – Ranker and GainRatio, Ranker and CInfoGain

C.4 Class: Snoring

CfsSubsetEvaluator and BestFirst	**CfsSubsetEvaluator and GeneticSearch**
ch09-Micro-100Hz-orig-cwt_MAX2MEDIAN	ch09-Micro-100Hz-orig-cwt_MAX2MEAN
ch09-Micro-100Hz-wined_FFT_band_theta	ch09-Micro-100Hz-orig-fft_SKEW
ch11-Flow-100Hz-orig-cwt_MEDIAN	ch09-Micro-100Hz-orig-fft_STDDEV
ch11-Flow-100Hz-orig-cwt_MAX2RMS	ch09-Micro-100Hz-orig_VARIANCE
ch11-Flow-100Hz-orig-fft_KURT	ch09-Micro-100Hz-orig_AvgAbs1Derivation
ch11-Flow-100Hz-orig_SKEWNESS	ch09-Micro-100Hz-orig_MaxAbs1Derivation
ch11-Flow-100Hz-orig_KURTOSIS	ch09-Micro-100Hz-orig_MaxAbs2Derivation
ch11-Flow-100Hz-orig_AvgAbs1Derivation	ch09-Micro-100Hz-orig_hjorth_mobility
	ch09-Micro-100Hz-orig_interp_poly_rms
	ch09-Micro-100Hz-orig_SPECTRAL...
	ch09-Micro-100Hz-wined_FFT_band_theta
	ch09-Micro-100Hz-wined_FFT_band_delta2theta...
	ch09-Micro-100Hz-wined_FFT_band_delta2alpha...
	ch09-Micro-100Hz-wined_FFT_band_delta2beta...
	ch09-Micro-100Hz-wined_FFT_band_delta2g...
	ch11-Flow-100Hz-orig-cwt_MAX
	ch11-Flow-100Hz-orig-cwt_MIN
	ch11-Flow-100Hz-orig-cwt_MEAN
	ch11-Flow-100Hz-orig-cwt_MAX2MEAN
	ch11-Flow-100Hz-orig-fft_MAX
	ch11-Flow-100Hz-orig-fft_MIN
	ch11-Flow-100Hz-orig-fft_KURT
	ch11-Flow-100Hz-orig-fft_RMS
	ch11-Flow-100Hz-orig-fft_MAX2MEDIAN
	ch11-Flow-100Hz-orig_VARIANCE
	ch11-Flow-100Hz-orig_KURTOSIS
	ch11-Flow-100Hz-orig_AvgAbs1Derivation
	ch11-Flow-100Hz-orig_MaxAbs1Derivation
	ch11-Flow-100Hz-orig_MaxAbs2Derivation
	ch11-Flow-100Hz-orig_hjorth_mobility
	ch11-Flow-100Hz-wined_FFT_band_delta2alpha...
	ch11-Flow-100Hz-wined_FFT_band_theta2alpha...

Table C.10: Selected features – Class: Snoring – CfsSubsetEvaluator and BestFirst, CfsSubsetEvaluator and GeneticSearch

GAME selected features	**Ranker and ChiSquaredEvaluator**
ch09-Micro-100Hz-orig-fft_MEAN	ch11-Flow-100Hz-orig_KURTOSIS
ch09-Micro-100Hz-orig-fft_MEDIAN	ch11-Flow-100Hz-orig_AvgAbs1Derivation
ch09-Micro-100Hz-wined_FFT_band_beta	ch11-Flow-100Hz-orig-cwt_MEAN
ch09-Micro-100Hz-wined_10-13-Band	ch11-Flow-100Hz-orig-cwt_MEDIAN
ch09-Micro-100Hz-wined_FFT_band_gamma	ch11-Flow-100Hz-orig_hjorth_complexity
ch09-Micro-100Hz-wined_30-50-Band	ch11-Flow-100Hz-orig_interp_poly_rms
ch09-Micro-100Hz-wined_FFT_band_delta	ch11-Flow-100Hz-orig-cwt_RMS
ch09-Micro-100Hz-wined_FFT_band_alpha	ch11-Flow-100Hz-orig_VARIANCE
ch09-Micro-100Hz-orig-cwt_SKEW	ch11-Flow-100Hz-orig_hjorth_activity
ch09-Micro-100Hz-orig-cwt_MEDIAN	ch11-Flow-100Hz-orig-fft_MAX2RMS
ch09-Micro-100Hz-wined_24-30-Band	ch11-Flow-100Hz-orig-fft_KURT
ch09-Micro-100Hz-wined_13-19-Band	ch11-Flow-100Hz-orig-fft_SKEW
ch09-Micro-100Hz-wined_FFT_band_theta	ch11-Flow-100Hz-orig_interp_rational_rms
ch09-Micro-100Hz-orig-fft_MAX2MEDIAN	ch11-Flow-100Hz-orig_SKEWNESS

Continued on next page

GAME selected features	Ranker and ChiSquaredEvaluator
ch09-Micro-100Hz-orig_SKEWNESS	ch11-Flow-100Hz-orig-cwt_MAX2RMS
ch11-Flow-100Hz-orig_MaxAbs2Derivation	ch11-Flow-100Hz-orig-cwt_SKEW
ch09-Micro-100Hz-orig_SPECTRAL...	ch11-Flow-100Hz-orig-cwt_VAR
ch09-Micro-100Hz-orig_MaxAbs1Derivation	ch11-Flow-100Hz-orig-cwt_STDDEV
ch09-Micro-100Hz-orig-cwt_MEAN	ch11-Flow-100Hz-orig-fft_MAX2MEAN
ch09-Micro-100Hz-orig-fft_MAX2MEAN	ch11-Flow-100Hz-orig-fft_MEAN
ch09-Micro-100Hz-orig_KURTOSIS	ch11-Flow-100Hz-orig-cwt_MAX2MEAN
ch09-Micro-100Hz-orig-cwt_MAX2MEAN	ch11-Flow-100Hz-wined_FFT_band_delta
ch09-Micro-100Hz-orig_interp_rational_rms	ch11-Flow-100Hz-orig-fft_RMS
ch11-Flow-100Hz-orig_hjorth_mobility	ch11-Flow-100Hz-orig-fft_VAR
ch11-Flow-100Hz-wined_30-50-Band	ch11-Flow-100Hz-orig-fft_STDDEV
ch11-Flow-100Hz-orig_SKEWNESS	ch11-Flow-100Hz-orig-cwt_MAX
ch09-Micro-100Hz-orig_MEAN	ch11-Flow-100Hz-orig-cwt_KURT
ch09-Micro-100Hz-orig_interp_poly_rms	ch11-Flow-100Hz-orig_MaxAbs1Derivation
ch09-Micro-100Hz-wined_FFT_band_delta2theta...	ch11-Flow-100Hz-orig-cwt_MAX2MEDIAN
ch09-Micro-100Hz-orig-fft_MAX	
ch11-Flow-100Hz-orig-fft_MEDIAN	
ch09-Micro-100Hz-orig-fft_MAX2RMS	
ch09-Micro-100Hz-wined_8-10-Band	
ch09-Micro-100Hz-wined_FFT_band_delta2alpha...	
ch09-Micro-100Hz-orig-fft_RMS	
ch11-Flow-100Hz-orig-fft_MAX2MEDIAN	
ch09-Micro-100Hz-orig-cwt_KURT	
ch11-Flow-100Hz-orig_KURTOSIS	
ch11-Flow-100Hz-wined_FFT_band_gamma	

Table C.11: Selected features – Class: Snoring – GAME selected features, Ranker and ChiSquaredEvaluator

Ranker and GainRatio	Ranker and InfoGain
ch11-Flow-100Hz-orig_KURTOSIS	ch11-Flow-100Hz-orig_KURTOSIS
ch11-Flow-100Hz-orig_AvgAbs1Derivation	ch11-Flow-100Hz-orig_AvgAbs1Derivation
ch11-Flow-100Hz-orig-cwt_MEAN	ch11-Flow-100Hz-orig-cwt_MEAN
ch11-Flow-100Hz-orig-cwt_MEDIAN	ch11-Flow-100Hz-orig-cwt_MEDIAN
ch09-Micro-100Hz-wined_FFT_band_theta	ch11-Flow-100Hz-orig_interp_poly_rms
ch11-Flow-100Hz-orig-fft_KURT	ch11-Flow-100Hz-orig_hjorth_complexity
ch11-Flow-100Hz-orig-fft_MAX2RMS	ch11-Flow-100Hz-orig-cwt_RMS
ch11-Flow-100Hz-orig_interp_poly_rms	ch11-Flow-100Hz-orig-fft_SKEW
ch11-Flow-100Hz-orig_VARIANCE	ch11-Flow-100Hz-orig-fft_KURT
ch11-Flow-100Hz-orig_hjorth_activity	ch11-Flow-100Hz-orig_VARIANCE
ch11-Flow-100Hz-orig-fft_SKEW	ch11-Flow-100Hz-orig_hjorth_activity
ch09-Micro-100Hz-wined_FFT_band_delta2g...	ch11-Flow-100Hz-orig-fft_MAX2RMS
ch09-Micro-100Hz-orig-fft_MEDIAN	ch11-Flow-100Hz-orig_SKEWNESS
ch09-Micro-100Hz-orig-fft_MAX2MEDIAN	ch11-Flow-100Hz-orig_interp_rational_rms
ch09-Micro-100Hz-orig-fft_RMS	ch11-Flow-100Hz-orig-fft_MAX2MEAN
ch09-Micro-100Hz-orig_AvgAbs1Derivation	ch11-Flow-100Hz-orig-fft_MEAN
ch11-Flow-100Hz-orig_SKEWNESS	ch11-Flow-100Hz-orig-cwt_MAX2RMS
ch09-Micro-100Hz-orig-cwt_MEAN	ch11-Flow-100Hz-wined_FFT_band_delta
ch09-Micro-100Hz-wined_FFT_band_delta2theta...	ch11-Flow-100Hz-orig-cwt_VAR
ch09-Micro-100Hz-orig-fft_MAX2MEAN	ch11-Flow-100Hz-orig-cwt_STDDEV
ch09-Micro-100Hz-orig-fft_MEAN	ch11-Flow-100Hz-orig-cwt_SKEW
ch09-Micro-100Hz-orig-fft_MAX2RMS	ch11-Flow-100Hz-orig-cwt_MAX2MEAN
ch11-Flow-100Hz-orig_hjorth_complexity	ch11-Flow-100Hz-orig-cwt_MAX
ch09-Micro-100Hz-wined_FFT_band_delta	ch11-Flow-100Hz-orig-fft_RMS

Continued on next page

Ranker and GainRatio	**Ranker and InfoGain**
ch09-Micro-100Hz-orig-fft_SKEW	ch11-Flow-100Hz-orig-fft_VAR
ch09-Micro-100Hz-orig-fft_KURT	ch11-Flow-100Hz-orig-fft_STDDEV
ch09-Micro-100Hz-orig-cwt_STDDEV	ch11-Flow-100Hz-orig_MaxAbs1Derivation
ch09-Micro-100Hz-orig-cwt_VAR	ch11-Flow-100Hz-orig-cwt_KURT
ch09-Micro-100Hz-wined_24-30-Band	ch09-Micro-100Hz-orig-cwt_RMS

Table C.12: Selected features – Class: Snoring – Ranker and GainRatio, Ranker and CInfoGain

Appendix D

Sleep EEG Data – List of Selected Features

D.1 Sleep Stage Wake

Please note that Wake did not select any features.

GAME selected features
EOG_horizontal_cwt_MEAN
EMG_submental_fft_SKEW
EOG_horizontal_cwt_MEDIAN
EMG_submental_fft_MAX2RMS
EMG_submental_FFT_band_alpha...
EEG_Pz-Oz_24-30-Band
EMG_submental_fft_KURT
EMG_submental_MaxAbs2Derivation
EEG_Pz-Oz_fft_MAX2RMS
EOG_horizontal_IMPORTANT_VOLU...
EMG_submental_fft_MAX
EOG_horizontal_AvgAbs1Derivation
EMG_submental_cwt_VAR
EMG_submental_interp_poly_p4
EEG_Pz-Oz_13-19-Band
EMG_submental_MEAN_ABS
EEG_Pz-Oz_fft_SKEW
EEG_Pz-Oz_KURTOSIS
EEG_Pz-Oz_fft_RMS
EMG_submental_24-30-Band
EMG_submental_30-50-Band
EMG_submental_fft_STDDEV
EOG_horizontal-windowed_FFT_band_gamma
EEG_Fpz-Cz_interp_rational_rms
EEG_Fpz-Cz_fft_STDDEV
EEG_Pz-Oz_FFT_band_alpha...
EOG_horizontal_fft_MAX2RMS
EEG_Fpz-Cz_fft_RMS
EEG_Fpz-Cz_fft_MEAN
EMG_submental_fft_MIN
EEG_Pz-Oz_cwt_MEAN
EEG_Pz-Oz_FFT_band_theta...
EMG_submental_fft_VAR
EMG_submental_FFT_band_alpha...

Continued on next page

GAME selected features
EMG_submental_hjorth_mobility
EMG_submental_FFT_band_delta
EMG_submental_KURTOSIS
EMG_submental_hjorth_complexity
EEG_Fpz-Cz_MIN_NEGATIVE
EMG_submental_cwt_MEAN
EMG_submental_interp_poly_p3

Table D.1: Selected features – Wake – GAME selected features

D.2 Sleep Stage 1

Please note that Wake did not select any features.

GAME selected features
EEG_Fpz-Cz_fft_MAX
EEG_Pz-Oz_FFT_band_theta...
EEG_Fpz-Cz_MEAN_ABS
EOG_horizontal_MIN_NEGATIVE
EOG_horizontal-windowed_FFT_band_gamma
EEG_Fpz-Cz_cwt_SKEW
EMG_submental_fft_SKEW
EOG_horizontal_cwt_RMS
EEG_Pz-Oz_cwt_MEDIAN
EMG_submental_fft_KURT
EEG_Fpz-Cz_fft_MAX2MEAN
EEG_Fpz-Cz_IMPORTANT_VOLU...
EOG_horizontal-windowed_30-50-Band
EEG_Pz-Oz_FFT_band_alpha...
EOG_horizontal_cwt_STDDEV
EEG_Pz-Oz_FFT_band_theta...
EMG_submental_fft_MAX2RMS
EOG_horizontal_interp_poly_p2
EEG_Fpz-Cz_cwt_MEDIAN
EMG_submental_MEAN_ABS
EEG_Fpz-Cz_fft_STDDEV
EEG_Fpz-Cz_FFT_band_theta...
EMG_submental_fft_MAX2MEAN
EOG_horizontal_MaxAbs2Derivation
EMG_submental_MEAN
EEG_Fpz-Cz_interp_poly_p4
EOG_horizontal-windowed_24-30-Band
EOG_horizontal_cwt_MAX
EMG_submental_FFT_band_delta
EOG_horizontal-windowed_FFT_band_delta
EMG_submental_SKEWNESS
EEG_Pz-Oz_FFT_band_alpha...
EEG_Pz-Oz_hjorth_complexity
EMG_submental_fft_RMS
EEG_Fpz-Cz_MIN_NEGATIVE
EEG_Pz-Oz_fft_MEDIAN
EMG_submental_fft_MAX
EEG_Fpz-Cz_fft_SKEW
EMG_submental_fft_VAR
EOG_horizontal_fft_KURT
EOG_horizontal_cwt_KURT
EEG_Fpz-Cz_fft_RMS
EEG_Fpz-Cz_interp_rational_rms
EEG_Fpz-Cz_MEAN
EEG_Fpz-Cz_cwt_MAX2RMS
EOG_horizontal-windowed_FFT_band_theta...
EOG_horizontal_fft_MEDIAN
EOG_horizontal_cwt_MAX2MEDIAN
EEG_Pz-Oz_24-30-Band

Table D.2: Selected features – Sleep Stage 1 – GAME selected features

D.3 Sleep Stage 2

CfsSubsetEvaluator and BestFirst	CfsSubsetEvaluator and GeneticSearch
EEG_Fpz-Cz_fft_VAR	EEG_Fpz-Cz_cwt_MAX
EEG_Fpz-Cz_fft_MAX2MEAN	EEG_Fpz-Cz_cwt_MEDIAN
EEG_Fpz-Cz_fft_MAX2MEDIAN	EEG_Fpz-Cz_cwt_SKEW
EEG_Fpz-Cz_FFT_band_delta...	EEG_Fpz-Cz_cwt_STDDEV
EEG_Fpz-Cz_FFT_band_theta...	EEG_Fpz-Cz_cwt_MAX2RMS
EEG_Fpz-Cz_MIN_NEGATIVE	EEG_Fpz-Cz_fft_MAX
EEG_Fpz-Cz_hjorth_complexity	EEG_Fpz-Cz_fft_STDDEV
EEG_Fpz-Cz_IMPORTANT_VOLU...	EEG_Fpz-Cz_8-10-Band
EEG_Pz-Oz_cwt_MEAN	EEG_Fpz-Cz_10-13-Band
EEG_Pz-Oz_cwt_MEDIAN	EEG_Fpz-Cz_13-19-Band
EEG_Pz-Oz_cwt_RMS	EEG_Fpz-Cz_19-24-Band
EEG_Pz-Oz_19-24-Band	EEG_Fpz-Cz_24-30-Band
EEG_Pz-Oz_FFT_band_delta...	EEG_Fpz-Cz_FFT_band_theta...
EEG_Pz-Oz_FFT_band_theta...	EEG_Fpz-Cz_FFT_band_theta...
EEG_Pz-Oz_AvgAbs1Derivation	EEG_Fpz-Cz_FFT_band_alpha...
EOG_horizontal_cwt_MAX2RMS	EEG_Fpz-Cz_MEAN_ABS
EOG_horizontal-windowed_FFT_band_theta...	EEG_Fpz-Cz_MEAN
EOG_horizontal-windowed_FFT_band_alpha...	EEG_Fpz-Cz_MIN_NEGATIVE
EOG_horizontal_MEAN_ABS	EEG_Fpz-Cz_VARIANCE
EOG_horizontal_MaxAbs1Derivation	EEG_Fpz-Cz_AvgAbs1Derivation
EOG_horizontal_interp_poly_p1	EEG_Fpz-Cz_interp_rational_rms
EOG_horizontal_interp_poly_p4	EEG_Fpz-Cz_IMPORTANT_VOLU...
EOG_horizontal_interp_rational_rms	EEG_Pz-Oz_cwt_MEAN
EMG_submental_cwt_MEAN	EEG_Pz-Oz_cwt_VAR
EMG_submental_cwt_MEDIAN	EEG_Pz-Oz_fft_MAX
	EEG_Pz-Oz_FFT_band_gamma
	EEG_Pz-Oz_MAX_POSITIVE
	EEG_Pz-Oz_AvgAbs1Derivation
	EEG_Pz-Oz_interp_poly_p4
	EEG_Pz-Oz_interp_rational_p2
	EOG_horizontal_cwt_MAX2MEDIAN
	EOG_horizontal_fft_MIN
	EOG_horizontal-windowed_FFT_band_beta
	EOG_horizontal-windowed_FFT_band_theta...
	EOG_horizontal_AvgAbs1Derivation
	EOG_horizontal_interp_poly_p4
	EOG_horizontal_interp_poly_rms
	EOG_horizontal_interp_rational_p4
	EOG_horizontal_interp_rational_q2
	EOG_horizontal_interp_rational_q3
	EMG_submental_cwt_MAX
	EMG_submental_cwt_MEAN
	EMG_submental_fft_MEAN
	EMG_submental_fft_RMS
	EMG_submental_AvgAbs1Derivation
	EMG_submental_interp_rational_p2

Table D.3: Selected features – Sleep stage 2 – CfsSubsetEvaluator and BestFirst, CfsSubsetEvaluator and GeneticSearch

GAME selected features	**Ranker and ChiSquaredEvaluator**
EEG_Fpz-Cz_hjorth_mobility	EEG_Fpz-Cz_cwt_MAX
EOG_horizontal_MEAN_ABS	EEG_Fpz-Cz_cwt_MEAN
EEG_Fpz-Cz_cwt_MEDIAN	EEG_Fpz-Cz_cwt_MEDIAN
EOG_horizontal_interp_poly_rms	EEG_Fpz-Cz_cwt_VAR
EEG_Fpz-Cz_SPECTRAL_CENTROID	EEG_Fpz-Cz_cwt_KURT
EEG_Fpz-Cz_IMPORTANT_VOLU...	EEG_Fpz-Cz_cwt_RMS
EEG_Fpz-Cz_cwt_SKEW	EEG_Fpz-Cz_cwt_STDDEV
EOG_horizontal-windowed_FFT_band_delta...	EEG_Fpz-Cz_cwt_MAX2RMS
EMG_submental_MEAN	EEG_Fpz-Cz_cwt_MAX2MEDIAN
EOG_horizontal-windowed_FFT_band_theta...	EEG_Fpz-Cz_fft_MAX
EOG_horizontal-windowed_FFT_band_alpha...	EEG_Fpz-Cz_fft_MEDIAN
EMG_submental_fft_RMS	EEG_Fpz-Cz_fft_SKEW
EOG_horizontal-windowed_FFT_band_delta	EEG_Fpz-Cz_fft_MAX2MEAN
EEG_Pz-Oz_SPECTRAL_CENTROID	EEG_Fpz-Cz_fft_MAX2MEDIAN
EOG_horizontal_fft_SKEW	EEG_Fpz-Cz_19-24-Band
EOG_horizontal_MaxAbs2Derivation	EEG_Fpz-Cz_24-30-Band
EOG_horizontal_interp_poly_p2	EEG_Fpz-Cz_30-50-Band
EOG_horizontal_cwt_KURT	EEG_Fpz-Cz_FFT_band_beta
EOG_horizontal_interp_rational_rms	EEG_Fpz-Cz_FFT_band_gamma
EEG_Pz-Oz_FFT_band_theta...	EEG_Fpz-Cz_FFT_band_delta...
EMG_submental_cwt_RMS	EEG_Fpz-Cz_FFT_band_delta...
EOG_horizontal_fft_KURT	EEG_Fpz-Cz_FFT_band_delta...
EEG_Pz-Oz_FFT_band_theta...	EEG_Fpz-Cz_FFT_band_delta...
EMG_submental_cwt_MAX2RMS	EEG_Fpz-Cz_FFT_band_theta...
EEG_Pz-Oz_IMPORTANT_VOLU...	EEG_Fpz-Cz_FFT_band_theta...
EOG_horizontal-windowed_8-10-Band	EEG_Fpz-Cz_MEAN
EEG_Fpz-Cz_cwt_MEAN	EEG_Fpz-Cz_MAX_POSITIVE
EEG_Fpz-Cz_hjorth_complexity	EEG_Fpz-Cz_MIN_NEGATIVE
EEG_Fpz-Cz_FFT_band_theta...	EEG_Fpz-Cz_AvgAbs1Derivation
EMG_submental_KURTOSIS	EEG_Fpz-Cz_MaxAbs1Derivation
EMG_submental_fft_KURT	EEG_Fpz-Cz_hjorth_mobility
EMG_submental_fft_MEDIAN	EEG_Fpz-Cz_hjorth_complexity
	EEG_Fpz-Cz_SPECTRAL_CENTROID
	EEG_Fpz-Cz_IMPORTANT_VOLU...
	EEG_Pz-Oz_cwt_MEAN
	EEG_Pz-Oz_cwt_MEDIAN
	EEG_Pz-Oz_cwt_VAR
	EEG_Pz-Oz_cwt_RMS
	EEG_Pz-Oz_cwt_STDDEV
	EEG_Pz-Oz_fft_MEDIAN
	EEG_Pz-Oz_19-24-Band
	EEG_Pz-Oz_24-30-Band
	EEG_Pz-Oz_30-50-Band
	EEG_Pz-Oz_FFT_band_gamma
	EEG_Pz-Oz_AvgAbs1Derivation
	EEG_Pz-Oz_IMPORTANT_VOLU...
	EOG_horizontal_MEAN_ABS
	EOG_horizontal_VARIANCE
	EOG_horizontal_hjorth_activity
	EOG_horizontal_interp_poly_rms

Table D.4: Selected features – Sleep Stage 2 – GAME selected features, Ranker and ChiSquaredEvaluator

Ranker and GainRatio

EEG_Fpz-Cz_cwt_MEAN
EEG_Fpz-Cz_cwt_MEDIAN
EEG_Fpz-Cz_cwt_VAR
EEG_Fpz-Cz_cwt_SKEW
EEG_Fpz-Cz_cwt_KURT
EEG_Fpz-Cz_cwt_RMS
EEG_Fpz-Cz_cwt_STDDEV
EEG_Fpz-Cz_cwt_MAX2MEDIAN
EEG_Fpz-Cz_fft_MAX
EEG_Fpz-Cz_fft_MIN
EEG_Fpz-Cz_fft_MEDIAN
EEG_Fpz-Cz_fft_VAR
EEG_Fpz-Cz_fft_SKEW
EEG_Fpz-Cz_fft_KURT
EEG_Fpz-Cz_fft_RMS
EEG_Fpz-Cz_fft_STDDEV
EEG_Fpz-Cz_fft_MAX2MEAN
EEG_Fpz-Cz_fft_MAX2MEDIAN
EEG_Fpz-Cz_19-24-Band
EEG_Fpz-Cz_24-30-Band
EEG_Fpz-Cz_30-50-Band
EEG_Fpz-Cz_FFT_band_beta
EEG_Fpz-Cz_FFT_band_gamma
EEG_Fpz-Cz_FFT_band_delta...
EEG_Fpz-Cz_FFT_band_delta...
EEG_Fpz-Cz_FFT_band_delta...
EEG_Fpz-Cz_FFT_band_delta...
EEG_Fpz-Cz_FFT_band_theta...
EEG_Fpz-Cz_FFT_band_theta...
EEG_Fpz-Cz_FFT_band_theta...
EEG_Fpz-Cz_FFT_band_alpha...
EEG_Fpz-Cz_FFT_band_alpha...
EEG_Fpz-Cz_FFT_band_beta...
EEG_Fpz-Cz_MEAN_ABS
EEG_Fpz-Cz_AvgAbs1Derivation
EEG_Fpz-Cz_hjorth_mobility
EEG_Fpz-Cz_hjorth_complexity
EEG_Fpz-Cz_SPECTRAL_CENTROID
EEG_Fpz-Cz_IMPORTANT_VOLU...
EEG_Pz-Oz_cwt_MEAN
EEG_Pz-Oz_cwt_MEDIAN
EEG_Pz-Oz_cwt_RMS
EEG_Pz-Oz_FFT_band_theta...
EEG_Pz-Oz_AvgAbs1Derivation
EOG_horizontal_fft_MAX
EOG_horizontal_MIN_NEGATIVE
EOG_horizontal_MaxAbs1Derivation
EOG_horizontal_interp_poly_p1
EOG_horizontal_interp_poly_p3
EOG_horizontal_interp_poly_p4

Ranker and InfoGain

EEG_Fpz-Cz_cwt_MAX
EEG_Fpz-Cz_cwt_MEAN
EEG_Fpz-Cz_cwt_MEDIAN
EEG_Fpz-Cz_cwt_VAR
EEG_Fpz-Cz_cwt_SKEW
EEG_Fpz-Cz_cwt_KURT
EEG_Fpz-Cz_cwt_RMS
EEG_Fpz-Cz_cwt_STDDEV
EEG_Fpz-Cz_cwt_MAX2RMS
EEG_Fpz-Cz_cwt_MAX2MEDIAN
EEG_Fpz-Cz_fft_MAX
EEG_Fpz-Cz_fft_MEDIAN
EEG_Fpz-Cz_fft_SKEW
EEG_Fpz-Cz_fft_MAX2MEAN
EEG_Fpz-Cz_fft_MAX2MEDIAN
EEG_Fpz-Cz_10-13-Band
EEG_Fpz-Cz_19-24-Band
EEG_Fpz-Cz_24-30-Band
EEG_Fpz-Cz_30-50-Band
EEG_Fpz-Cz_FFT_band_alpha
EEG_Fpz-Cz_FFT_band_beta
EEG_Fpz-Cz_FFT_band_gamma
EEG_Fpz-Cz_FFT_band_delta...
EEG_Fpz-Cz_FFT_band_delta...
EEG_Fpz-Cz_FFT_band_delta...
EEG_Fpz-Cz_FFT_band_delta...
EEG_Fpz-Cz_FFT_band_theta...
EEG_Fpz-Cz_FFT_band_theta...
EEG_Fpz-Cz_FFT_band_alpha...
EEG_Fpz-Cz_FFT_band_alpha...
EEG_Fpz-Cz_MEAN_ABS
EEG_Fpz-Cz_MEAN
EEG_Fpz-Cz_MAX_POSITIVE
EEG_Fpz-Cz_MIN_NEGATIVE
EEG_Fpz-Cz_AvgAbs1Derivation
EEG_Fpz-Cz_MaxAbs1Derivation
EEG_Fpz-Cz_hjorth_mobility
EEG_Fpz-Cz_hjorth_complexity
EEG_Fpz-Cz_SPECTRAL_CENTROID
EEG_Fpz-Cz_IMPORTANT_VOLU...
EEG_Pz-Oz_cwt_MEAN
EEG_Pz-Oz_cwt_MEDIAN
EEG_Pz-Oz_cwt_VAR
EEG_Pz-Oz_cwt_RMS
EEG_Pz-Oz_cwt_STDDEV
EEG_Pz-Oz_fft_MEDIAN
EEG_Pz-Oz_19-24-Band
EEG_Pz-Oz_24-30-Band
EEG_Pz-Oz_30-50-Band
EEG_Pz-Oz_FFT_band_gamma
EEG_Pz-Oz_AvgAbs1Derivation
EEG_Pz-Oz_IMPORTANT_VOLU...
EOG_horizontal_fft_RMS
EOG_horizontal_fft_STDDEV
EOG_horizontal_MEAN_ABS

Continued on next page

Ranker and GainRatio	**Ranker and InfoGain**
	EOG_horizontal_VARIANCE
	EOG_horizontal_hjorth_activity
	EOG_horizontal_hjorth_complexity
	EOG_horizontal_interp_poly_rms
	EMG_submental_cwt_MEDIAN

Table D.5: Selected features – Sleep Stage 2 – Ranker and GainRatio, Ranker and InfoGain

D.4 Sleep Stage 3

CfsSubsetEvaluator and BestFirst	**CfsSubsetEvaluator and GeneticSearch**
EEG_Fpz-Cz_cwt_MEDIAN	EEG_Fpz-Cz_cwt_MEAN
EEG_Fpz-Cz_cwt_RMS	EEG_Fpz-Cz_cwt_KURT
EEG_Fpz-Cz_hjorth_mobility	EEG_Fpz-Cz_FFT_band_delta...
EOG_horizontal_cwt_SKEW	EEG_Fpz-Cz_hjorth_mobility
EOG_horizontal_cwt_MAX2RMS	EEG_Fpz-Cz_IMPORTANT_VOLU...
EOG_horizontal-windowed_19-24-Band	EOG_horizontal_cwt_SKEW
EOG_horizontal-windowed_30-50-Band	EOG_horizontal_cwt_MAX2RMS
EOG_horizontal_VARIANCE	EOG_horizontal_fft_MEDIAN
EOG_horizontal_hjorth_mobility	EOG_horizontal-windowed_FFT_band_gamma
EMG_submental_fft_MAX2RMS	EOG_horizontal_hjorth_complexity
EMG_submental_MAX_POSITIVE	EMG_submental_fft_KURT
EMG_submental_VARIANCE	EMG_submental_fft_MAX2MEDIAN
	EMG_submental_VARIANCE
	EMG_submental_interp_poly_p4
	EMG_submental_interp_rational_rms

Table D.6: Selected features – Sleep stage 3 – CfsSubsetEvaluator and BestFirst, CfsSubsetEvaluator and GeneticSearch

GAME selected features	**Ranker and ChiSquaredEvaluator**
EEG_Fpz-Cz_cwt_MEAN	EEG_Fpz-Cz_cwt_MIN
EEG_Fpz-Cz_cwt_MEDIAN	EEG_Fpz-Cz_cwt_MEAN
EEG_Fpz-Cz_cwt_SKEW	EEG_Fpz-Cz_cwt_MEDIAN
EEG_Fpz-Cz_cwt_KURT	EEG_Fpz-Cz_cwt_VAR
EEG_Fpz-Cz_cwt_STDDEV	EEG_Fpz-Cz_cwt_SKEW
EEG_Fpz-Cz_fft_MAX2RMS	EEG_Fpz-Cz_cwt_KURT
EEG_Fpz-Cz_24-30-Band	EEG_Fpz-Cz_cwt_RMS
EEG_Fpz-Cz_FFT_band_alpha	EEG_Fpz-Cz_cwt_STDDEV
EEG_Fpz-Cz_FFT_band_alpha...	EEG_Fpz-Cz_fft_MEDIAN
EEG_Fpz-Cz_FFT_band_beta...	EEG_Fpz-Cz_fft_MAX2MEAN
EEG_Fpz-Cz_AvgAbs1Derivation	EEG_Fpz-Cz_fft_MAX2MEDIAN
EEG_Fpz-Cz_hjorth_mobility	EEG_Fpz-Cz_10-13-Band
EEG_Fpz-Cz_SPECTRAL_CENTROID	EEG_Fpz-Cz_19-24-Band
EEG_Fpz-Cz_IMPORTANT_VOLU...	EEG_Fpz-Cz_24-30-Band
EEG_Pz-Oz_cwt_MEAN	EEG_Fpz-Cz_30-50-Band
EEG_Pz-Oz_fft_MAX	EEG_Fpz-Cz_FFT_band_alpha
EEG_Pz-Oz_fft_VAR	EEG_Fpz-Cz_FFT_band_beta
EEG_Pz-Oz_fft_KURT	EEG_Fpz-Cz_FFT_band_gamma
EEG_Pz-Oz_fft_MAX2RMS	EEG_Fpz-Cz_FFT_band_delta...
EEG_Pz-Oz_13-19-Band	EEG_Fpz-Cz_FFT_band_delta...
EEG_Pz-Oz_24-30-Band	EEG_Fpz-Cz_FFT_band_delta...
EEG_Pz-Oz_FFT_band_beta	EEG_Fpz-Cz_FFT_band_theta...
EEG_Pz-Oz_FFT_band_gamma	EEG_Fpz-Cz_FFT_band_theta...
EEG_Pz-Oz_FFT_band_theta...	EEG_Fpz-Cz_FFT_band_theta...
EEG_Pz-Oz_FFT_band_theta...	EEG_Fpz-Cz_MEAN_ABS
EEG_Pz-Oz_FFT_band_theta...	EEG_Fpz-Cz_KURTOSIS
EEG_Pz-Oz_SKEWNESS	EEG_Fpz-Cz_AvgAbs1Derivation
EOG_horizontal_cwt_SKEW	EEG_Fpz-Cz_hjorth_mobility
EOG_horizontal_cwt_KURT	EEG_Fpz-Cz_hjorth_complexity
EOG_horizontal_cwt_MAX2MEAN	EEG_Fpz-Cz_SPECTRAL_CENTROID
EOG_horizontal_cwt_MAX2MEDIAN	EEG_Fpz-Cz_IMPORTANT_VOLU...

Continued on next page

GAME selected features	**Ranker and ChiSquaredEvaluator**
EOG_horizontal-windowed_19-24-Band	EOG_horizontal_cwt_MAX
EOG_horizontal-windowed_FFT_band_theta...	EOG_horizontal_cwt_SKEW
EOG_horizontal-windowed_FFT_band_alpha...	EOG_horizontal_cwt_KURT
EOG_horizontal-windowed_FFT_band_alpha...	EOG_horizontal_cwt_MAX2RMS
EOG_horizontal-windowed_FFT_band_beta...	EOG_horizontal_cwt_MAX2MEAN
EOG_horizontal_hjorth_complexity	EOG_horizontal_fft_MEDIAN
EMG_submental_cwt_MAX2RMS	EOG_horizontal-windowed_19-24-Band
EMG_submental_cwt_MAX2MEAN	EOG_horizontal-windowed_30-50-Band
EMG_submental_fft_VAR	EOG_horizontal-windowed_FFT_band_gamma
EMG_submental_fft_KURT	EOG_horizontal_VARIANCE
EMG_submental_fft_RMS	EOG_horizontal_hjorth_activity
EMG_submental_fft_STDDEV	EOG_horizontal_hjorth_mobility
EMG_submental_fft_MAX2MEAN	EOG_horizontal_hjorth_complexity
EMG_submental_FFT_band_alpha...	EOG_horizontal_interp_rational_rms
EMG_submental_FFT_band_alpha...	EMG_submental_cwt_MEAN
EMG_submental_MEAN	EMG_submental_cwt_MEDIAN
EMG_submental_VARIANCE	EMG_submental_cwt_SKEW
EMG_submental_KURTOSIS	EMG_submental_cwt_KURT
EMG_submental_interp_poly_p4	EMG_submental_cwt_MAX2MEAN
	EMG_submental_cwt_MAX2MEDIAN
	EMG_submental_fft_MAX
	EMG_submental_fft_VAR
	EMG_submental_fft_SKEW
	EMG_submental_fft_KURT
	EMG_submental_fft_RMS
	EMG_submental_fft_STDDEV
	EMG_submental_fft_MAX2RMS
	EMG_submental_fft_MAX2MEAN
	EMG_submental_fft_MAX2MEDIAN
	EMG_submental_FFT_band_delta
	EMG_submental_FFT_band_delta...
	EMG_submental_FFT_band_delta...
	EMG_submental_FFT_band_delta...
	EMG_submental_FFT_band_delta...
	EMG_submental_MEAN_ABS
	EMG_submental_MEAN
	EMG_submental_MAX_POSITIVE
	EMG_submental_VARIANCE
	EMG_submental_SKEWNESS
	EMG_submental_KURTOSIS
	EMG_submental_AvgAbs1Derivation
	EMG_submental_hjorth_activity
	EMG_submental_interp_poly_p1
	EMG_submental_interp_poly_p4
	EMG_submental_interp_rational_p3
	EMG_submental_interp_rational_rms
	EMG_submental_SPECTRAL_CENTROID
	EMG_submental_IMPORTANT_VOLU...

Table D.7: Selected features – Sleep Stage 3 – GAME selected features, Ranker and ChiSquaredEvaluator

Ranker and GainRatio	**Ranker and InfoGain**
EEG_Fpz-Cz_cwt_MIN	EEG_Fpz-Cz_cwt_MIN

Continued on next page

Ranker and GainRatio
EEG_Fpz-Cz_cwt_MEAN
EEG_Fpz-Cz_cwt_MEDIAN
EEG_Fpz-Cz_cwt_VAR
EEG_Fpz-Cz_cwt_SKEW
EEG_Fpz-Cz_cwt_KURT
EEG_Fpz-Cz_cwt_RMS
EEG_Fpz-Cz_cwt_STDDEV
EEG_Fpz-Cz_fft_MEDIAN
EEG_Fpz-Cz_fft_MAX2MEAN
EEG_Fpz-Cz_fft_MAX2MEDIAN
EEG_Fpz-Cz_10-13-Band
EEG_Fpz-Cz_19-24-Band
EEG_Fpz-Cz_24-30-Band
EEG_Fpz-Cz_30-50-Band
EEG_Fpz-Cz_FFT_band_alpha
EEG_Fpz-Cz_FFT_band_beta
EEG_Fpz-Cz_FFT_band_gamma
EEG_Fpz-Cz_FFT_band_delta...
EEG_Fpz-Cz_FFT_band_delta...
EEG_Fpz-Cz_FFT_band_delta...
EEG_Fpz-Cz_FFT_band_theta...
EEG_Fpz-Cz_FFT_band_theta...
EEG_Fpz-Cz_FFT_band_theta...
EEG_Fpz-Cz_MEAN_ABS
EEG_Fpz-Cz_KURTOSIS
EEG_Fpz-Cz_AvgAbs1Derivation
EEG_Fpz-Cz_hjorth_mobility
EEG_Fpz-Cz_hjorth_complexity
EEG_Fpz-Cz_SPECTRAL_CENTROID
EEG_Fpz-Cz_IMPORTANT_VOLU...
EOG_horizontal_cwt_MAX
EOG_horizontal_cwt_SKEW
EOG_horizontal_cwt_KURT
EOG_horizontal_cwt_MAX2RMS
EOG_horizontal_cwt_MAX2MEAN
EOG_horizontal_fft_MEDIAN
EOG_horizontal-windowed_19-24-Band
EOG_horizontal-windowed_30-50-Band
EOG_horizontal-windowed_FFT_band_gamma
EOG_horizontal_VARIANCE
EOG_horizontal_hjorth_activity
EOG_horizontal_hjorth_mobility
EOG_horizontal_hjorth_complexity
EOG_horizontal_interp_rational_rms
EMG_submental_cwt_MEAN
EMG_submental_cwt_MEDIAN
EMG_submental_cwt_SKEW
EMG_submental_cwt_KURT
EMG_submental_cwt_MAX2MEAN
EMG_submental_cwt_MAX2MEDIAN
EMG_submental_fft_MAX
EMG_submental_fft_VAR
EMG_submental_fft_SKEW
EMG_submental_fft_KURT
EMG_submental_fft_RMS

Ranker and InfoGain
EEG_Fpz-Cz_cwt_MEAN
EEG_Fpz-Cz_cwt_MEDIAN
EEG_Fpz-Cz_cwt_VAR
EEG_Fpz-Cz_cwt_SKEW
EEG_Fpz-Cz_cwt_KURT
EEG_Fpz-Cz_cwt_RMS
EEG_Fpz-Cz_cwt_STDDEV
EEG_Fpz-Cz_fft_MEDIAN
EEG_Fpz-Cz_fft_MAX2MEAN
EEG_Fpz-Cz_fft_MAX2MEDIAN
EEG_Fpz-Cz_10-13-Band
EEG_Fpz-Cz_19-24-Band
EEG_Fpz-Cz_24-30-Band
EEG_Fpz-Cz_30-50-Band
EEG_Fpz-Cz_FFT_band_alpha
EEG_Fpz-Cz_FFT_band_beta
EEG_Fpz-Cz_FFT_band_gamma
EEG_Fpz-Cz_FFT_band_delta...
EEG_Fpz-Cz_FFT_band_delta...
EEG_Fpz-Cz_FFT_band_delta...
EEG_Fpz-Cz_FFT_band_theta...
EEG_Fpz-Cz_FFT_band_theta...
EEG_Fpz-Cz_FFT_band_theta...
EEG_Fpz-Cz_MEAN_ABS
EEG_Fpz-Cz_KURTOSIS
EEG_Fpz-Cz_AvgAbs1Derivation
EEG_Fpz-Cz_hjorth_mobility
EEG_Fpz-Cz_hjorth_complexity
EEG_Fpz-Cz_SPECTRAL_CENTROID
EEG_Fpz-Cz_IMPORTANT_VOLU...
EOG_horizontal_cwt_MAX
EOG_horizontal_cwt_SKEW
EOG_horizontal_cwt_KURT
EOG_horizontal_cwt_MAX2RMS
EOG_horizontal_cwt_MAX2MEAN
EOG_horizontal_fft_MEDIAN
EOG_horizontal-windowed_19-24-Band
EOG_horizontal-windowed_30-50-Band
EOG_horizontal-windowed_FFT_band_gamma
EOG_horizontal_VARIANCE
EOG_horizontal_hjorth_activity
EOG_horizontal_hjorth_mobility
EOG_horizontal_hjorth_complexity
EOG_horizontal_interp_rational_rms
EMG_submental_cwt_MEAN
EMG_submental_cwt_MEDIAN
EMG_submental_cwt_SKEW
EMG_submental_cwt_KURT
EMG_submental_cwt_MAX2MEAN
EMG_submental_cwt_MAX2MEDIAN
EMG_submental_fft_MAX
EMG_submental_fft_VAR
EMG_submental_fft_SKEW
EMG_submental_fft_KURT
EMG_submental_fft_RMS

Continued on next page

Ranker and GainRatio	Ranker and InfoGain
EMG_submental_fft_STDDEV	EMG_submental_fft_STDDEV
EMG_submental_fft_MAX2RMS	EMG_submental_fft_MAX2RMS
EMG_submental_fft_MAX2MEAN	EMG_submental_fft_MAX2MEAN
EMG_submental_fft_MAX2MEDIAN	EMG_submental_fft_MAX2MEDIAN
EMG_submental_FFT_band_delta	EMG_submental_FFT_band_delta
EMG_submental_FFT_band_delta...	EMG_submental_FFT_band_delta...
EMG_submental_FFT_band_delta...	EMG_submental_FFT_band_delta...
EMG_submental_FFT_band_delta...	EMG_submental_FFT_band_delta...
EMG_submental_FFT_band_delta...	EMG_submental_FFT_band_delta...
EMG_submental_MEAN_ABS	EMG_submental_MEAN_ABS
EMG_submental_MEAN	EMG_submental_MEAN
EMG_submental_MAX_POSITIVE	EMG_submental_MAX_POSITIVE
EMG_submental_VARIANCE	EMG_submental_VARIANCE
EMG_submental_SKEWNESS	EMG_submental_SKEWNESS
EMG_submental_KURTOSIS	EMG_submental_KURTOSIS
EMG_submental_AvgAbs1Derivation	EMG_submental_AvgAbs1Derivation
EMG_submental_hjorth_activity	EMG_submental_hjorth_activity
EMG_submental_interp_poly_p1	EMG_submental_interp_poly_p1
EMG_submental_interp_poly_p4	EMG_submental_interp_poly_p4
EMG_submental_interp_rational_p3	EMG_submental_interp_rational_p3
EMG_submental_interp_rational_rms	EMG_submental_interp_rational_rms
EMG_submental_SPECTRAL_CENTROID	EMG_submental_SPECTRAL_CENTROID
EMG_submental_IMPORTANT_VOLU...	EMG_submental_IMPORTANT_VOLU...

Table D.8: Selected features – Sleep Stage 3 – Ranker and GainRatio, Ranker and InfoGain – please note, that sets of selected features are the same for both methods

D.5 Sleep Stage 4

CfsSubsetEvaluator and BestFirst	**CfsSubsetEvaluator and GeneticSearch**
EEG_Fpz-Cz_24-30-Band	EEG_Fpz-Cz_cwt_SKEW
EEG_Fpz-Cz_IMPORTANT_VOLU...	EEG_Fpz-Cz_fft_MIN
EEG_Pz-Oz_hjorth_mobility	EEG_Fpz-Cz_fft_MEAN
EMG_submental_cwt_MEDIAN	EEG_Fpz-Cz_fft_MAX2MEDIAN
EMG_submental_cwt_MAX2RMS	EEG_Fpz-Cz_10-13-Band
EMG_submental_FFT_band_theta...	EEG_Fpz-Cz_24-30-Band
EMG_submental_FFT_band_alpha...	EEG_Fpz-Cz_FFT_band_theta
EMG_submental_SKEWNESS	EEG_Fpz-Cz_FFT_band_delta...
EMG_submental_AvgAbs1Derivation	EEG_Fpz-Cz_FFT_band_alpha...
EMG_submental_hjorth_mobility	EEG_Fpz-Cz_interp_poly_rms
EMG_submental_hjorth_complexity	EOG_horizontal_cwt_MAX2RMS
EMG_submental_IMPORTANT_VOLU...	EMG_submental_cwt_MAX
	EMG_submental_cwt_MEAN
	EMG_submental_cwt_MEDIAN
	EMG_submental_cwt_SKEW
	EMG_submental_fft_RMS
	EMG_submental_fft_MAX2MEDIAN
	EMG_submental_FFT_band_gamma
	EMG_submental_FFT_band_delta...
	EMG_submental_MEAN
	EMG_submental_hjorth_activity
	EMG_submental_hjorth_mobility
	EMG_submental_interp_rational_rms

Table D.9: Selected features – Sleep stage 4 – CfsSubsetEvaluator and BestFirst, CfsSubsetEvaluator and GeneticSearch

GAME selected features	**Ranker and ChiSquaredEvaluator**
EEG_Fpz-Cz_cwt_MEAN	EEG_Fpz-Cz_cwt_MEAN
EEG_Fpz-Cz_cwt_MEDIAN	EEG_Fpz-Cz_cwt_MEDIAN
EEG_Fpz-Cz_cwt_VAR	EEG_Fpz-Cz_cwt_VAR
EEG_Fpz-Cz_cwt_SKEW	EEG_Fpz-Cz_cwt_RMS
EEG_Fpz-Cz_fft_MAX	EEG_Fpz-Cz_cwt_STDDEV
EEG_Fpz-Cz_fft_SKEW	EEG_Fpz-Cz_cwt_MAX2MEDIAN
EEG_Fpz-Cz_fft_KURT	EEG_Fpz-Cz_fft_MEDIAN
EEG_Fpz-Cz_fft_RMS	EEG_Fpz-Cz_fft_SKEW
EEG_Fpz-Cz_fft_MAX2RMS	EEG_Fpz-Cz_fft_KURT
EEG_Fpz-Cz_fft_MAX2MEAN	EEG_Fpz-Cz_fft_MAX2RMS
EEG_Fpz-Cz_FFT_band_theta	EEG_Fpz-Cz_fft_MAX2MEAN
EEG_Fpz-Cz_SKEWNESS	EEG_Fpz-Cz_fft_MAX2MEDIAN
EEG_Fpz-Cz_AvgAbs1Derivation	EEG_Fpz-Cz_8-10-Band
EEG_Fpz-Cz_SPECTRAL_CENTROID	EEG_Fpz-Cz_10-13-Band
EEG_Fpz-Cz_IMPORTANT_VOLU...	EEG_Fpz-Cz_13-19-Band
EEG_Pz-Oz_cwt_MEDIAN	EEG_Fpz-Cz_24-30-Band
EEG_Pz-Oz_fft_SKEW	EEG_Fpz-Cz_FFT_band_alpha
EEG_Pz-Oz_interp_poly_p3	EEG_Fpz-Cz_FFT_band_beta
EEG_Pz-Oz_interp_poly_p4	EEG_Fpz-Cz_FFT_band_delta...
EOG_horizontal_cwt_MAX2MEAN	EEG_Fpz-Cz_FFT_band_delta...
EOG_horizontal-windowed_10-13-Band	EEG_Fpz-Cz_FFT_band_delta...
EOG_horizontal-windowed_FFT_band_alpha	EEG_Fpz-Cz_MEAN
EOG_horizontal-windowed_FFT_band_delta...	EEG_Fpz-Cz_MAX_POSITIVE

Continued on next page

GAME selected features	**Ranker and ChiSquaredEvaluator**
EOG_horizontal-windowed_FFT_band_theta...	EEG_Fpz-Cz_VARIANCE
EOG_horizontal_hjorth_mobility	EEG_Fpz-Cz_SKEWNESS
EOG_horizontal_hjorth_complexity	EEG_Fpz-Cz_AvgAbs1Derivation
EOG_horizontal_IMPORTANT_VOLU...	EEG_Fpz-Cz_hjorth_activity
EMG_submental_cwt_MEDIAN	EEG_Fpz-Cz_interp_poly_rms
EMG_submental_cwt_SKEW	EEG_Fpz-Cz_SPECTRAL_CENTROID
EMG_submental_cwt_KURT	EEG_Fpz-Cz_IMPORTANT_VOLU...
EMG_submental_cwt_MAX2RMS	EEG_Pz-Oz_cwt_MEDIAN
EMG_submental_cwt_MAX2MEAN	EOG_horizontal_cwt_MEDIAN
EMG_submental_cwt_MAX2MEDIAN	EOG_horizontal_AvgAbs1Derivation
EMG_submental_fft_VAR	EMG_submental_cwt_MEAN
EMG_submental_fft_SKEW	EMG_submental_cwt_MEDIAN
EMG_submental_fft_KURT	EMG_submental_cwt_SKEW
EMG_submental_fft_MAX2MEAN	EMG_submental_cwt_KURT
EMG_submental_FFT_band_theta...	EMG_submental_cwt_MAX2RMS
EMG_submental_MEAN_ABS	EMG_submental_cwt_MAX2MEAN
EMG_submental_AvgAbs1Derivation	EMG_submental_cwt_MAX2MEDIAN
EMG_submental_hjorth_mobility	EMG_submental_fft_MAX
EMG_submental_hjorth_complexity	EMG_submental_fft_VAR
EMG_submental_SPECTRAL_CENTROID	EMG_submental_fft_SKEW
EMG_submental_IMPORTANT_VOLU...	EMG_submental_fft_KURT
	EMG_submental_fft_MAX2MEAN
	EMG_submental_fft_MAX2MEDIAN
	EMG_submental_FFT_band_delta...
	EMG_submental_FFT_band_delta...
	EMG_submental_FFT_band_delta...
	EMG_submental_FFT_band_theta...
	EMG_submental_FFT_band_alpha...
	EMG_submental_MEAN_ABS
	EMG_submental_VARIANCE
	EMG_submental_AvgAbs1Derivation
	EMG_submental_MaxAbs1Derivation
	EMG_submental_hjorth_activity
	EMG_submental_hjorth_mobility
	EMG_submental_hjorth_complexity
	EMG_submental_interp_poly_rms
	EMG_submental_interp_rational_rms
	EMG_submental_SPECTRAL_CENTROID
	EMG_submental_IMPORTANT_VOLU...

Table D.10: Selected features – Sleep Stage 4 – GAME selected features, Ranker and ChiSquaredEvaluator

Ranker and GainRatio	**Ranker and InfoGain**
EEG_Fpz-Cz_cwt_MAX	EEG_Fpz-Cz_cwt_MEAN
EEG_Fpz-Cz_cwt_MEAN	EEG_Fpz-Cz_cwt_MEDIAN
EEG_Fpz-Cz_cwt_MEDIAN	EEG_Fpz-Cz_cwt_VAR
EEG_Fpz-Cz_cwt_VAR	EEG_Fpz-Cz_cwt_RMS
EEG_Fpz-Cz_cwt_RMS	EEG_Fpz-Cz_cwt_STDDEV
EEG_Fpz-Cz_cwt_STDDEV	EEG_Fpz-Cz_cwt_MAX2MEDIAN
EEG_Fpz-Cz_cwt_MAX2MEDIAN	EEG_Fpz-Cz_fft_MIN
EEG_Fpz-Cz_fft_MIN	EEG_Fpz-Cz_fft_MEDIAN
EEG_Fpz-Cz_fft_MEAN	EEG_Fpz-Cz_fft_SKEW
EEG_Fpz-Cz_fft_MEDIAN	EEG_Fpz-Cz_fft_KURT

Continued on next page

Ranker and GainRatio	**Ranker and InfoGain**
EEG_Fpz-Cz_fft_SKEW	EEG_Fpz-Cz_fft_MAX2RMS
EEG_Fpz-Cz_fft_KURT	EEG_Fpz-Cz_fft_MAX2MEAN
EEG_Fpz-Cz_8-10-Band	EEG_Fpz-Cz_fft_MAX2MEDIAN
EEG_Fpz-Cz_10-13-Band	EEG_Fpz-Cz_8-10-Band
EEG_Fpz-Cz_13-19-Band	EEG_Fpz-Cz_10-13-Band
EEG_Fpz-Cz_19-24-Band	EEG_Fpz-Cz_13-19-Band
EEG_Fpz-Cz_24-30-Band	EEG_Fpz-Cz_19-24-Band
EEG_Fpz-Cz_30-50-Band	EEG_Fpz-Cz_24-30-Band
EEG_Fpz-Cz_FFT_band_delta	EEG_Fpz-Cz_30-50-Band
EEG_Fpz-Cz_FFT_band_theta	EEG_Fpz-Cz_FFT_band_alpha
EEG_Fpz-Cz_FFT_band_alpha	EEG_Fpz-Cz_FFT_band_beta
EEG_Fpz-Cz_FFT_band_gamma	EEG_Fpz-Cz_FFT_band_gamma
EEG_Fpz-Cz_FFT_band_theta...	EEG_Fpz-Cz_FFT_band_delta...
EEG_Fpz-Cz_FFT_band_alpha...	EEG_Fpz-Cz_MEAN
EEG_Fpz-Cz_FFT_band_beta...	EEG_Fpz-Cz_MAX_POSITIVE
EEG_Fpz-Cz_MAX_POSITIVE	EEG_Fpz-Cz_AvgAbs1Derivation
EEG_Fpz-Cz_MIN_NEGATIVE	EEG_Fpz-Cz_interp_poly_rms
EEG_Fpz-Cz_AvgAbs1Derivation	EEG_Fpz-Cz_SPECTRAL_CENTROID
EEG_Fpz-Cz_hjorth_mobility	EEG_Fpz-Cz_IMPORTANT_VOLU...
EEG_Fpz-Cz_IMPORTANT_VOLU...	EEG_Pz-Oz_cwt_MEDIAN
EEG_Pz-Oz_FFT_band_delta...	EOG_horizontal_cwt_MEDIAN
EOG_horizontal_cwt_MIN	EOG_horizontal_AvgAbs1Derivation
EOG_horizontal_cwt_MAX2RMS	EMG_submental_cwt_MEAN
EOG_horizontal-windowed_FFT_band_delta...	EMG_submental_cwt_MEDIAN
EOG_horizontal-windowed_FFT_band_delta...	EMG_submental_cwt_SKEW
EOG_horizontal-windowed_FFT_band_theta...	EMG_submental_cwt_KURT
EOG_horizontal_hjorth_mobility	EMG_submental_cwt_MAX2RMS
EMG_submental_cwt_MEAN	EMG_submental_cwt_MAX2MEAN
EMG_submental_cwt_MEDIAN	EMG_submental_cwt_MAX2MEDIAN
EMG_submental_cwt_SKEW	EMG_submental_fft_MAX
EMG_submental_cwt_KURT	EMG_submental_fft_SKEW
EMG_submental_cwt_MAX2RMS	EMG_submental_fft_KURT
EMG_submental_cwt_MAX2MEAN	EMG_submental_fft_MAX2MEAN
EMG_submental_cwt_MAX2MEDIAN	EMG_submental_fft_MAX2MEDIAN
EMG_submental_fft_MAX2RMS	EMG_submental_FFT_band_delta...
EMG_submental_fft_MAX2MEDIAN	EMG_submental_FFT_band_delta...
EMG_submental_FFT_band_delta...	EMG_submental_FFT_band_delta...
EMG_submental_FFT_band_delta...	EMG_submental_FFT_band_theta...
EMG_submental_FFT_band_delta...	EMG_submental_FFT_band_alpha...
EMG_submental_FFT_band_delta...	EMG_submental_MEAN_ABS
EMG_submental_FFT_band_theta...	EMG_submental_VARIANCE
EMG_submental_FFT_band_theta...	EMG_submental_AvgAbs1Derivation
EMG_submental_MEAN	EMG_submental_MaxAbs1Derivation
EMG_submental_MAX_POSITIVE	EMG_submental_MaxAbs2Derivation
EMG_submental_SKEWNESS	EMG_submental_hjorth_activity
EMG_submental_AvgAbs1Derivation	EMG_submental_hjorth_mobility
EMG_submental_hjorth_mobility	EMG_submental_hjorth_complexity
EMG_submental_hjorth_complexity	EMG_submental_interp_poly_rms
EMG_submental_interp_poly_p4	EMG_submental_interp_rational_rms
EMG_submental_interp_rational_rms	EMG_submental_SPECTRAL_CENTROID
EMG_submental_SPECTRAL_CENTROID	EMG_submental_IMPORTANT_VOLU...
EMG_submental_IMPORTANT_VOLU...	

Table D.11: Selected features – Sleep Stage 4 – Ranker and GainRatio, Ranker and InfoGain

D.6 Sleep Stage REM

CfsSubsetEvaluator and BestFirst

EEG_Fpz-Cz_cwt_SKEW
EEG_Fpz-Cz_cwt_RMS
EEG_Fpz-Cz_fft_SKEW
EEG_Fpz-Cz_fft_KURT
EEG_Fpz-Cz_FFT_band_delta
EEG_Fpz-Cz_FFT_band_theta
EEG_Fpz-Cz_MEAN_ABS
EEG_Fpz-Cz_interp_poly_rms
EEG_Pz-Oz_cwt_MEAN
EEG_Pz-Oz_cwt_MEDIAN
EEG_Pz-Oz_cwt_MAX2MEDIAN
EEG_Pz-Oz_10-13-Band
EEG_Pz-Oz_FFT_band_delta...
EEG_Pz-Oz_MEAN_ABS
EEG_Pz-Oz_AvgAbs1Derivation
EEG_Pz-Oz_interp_poly_rms
EEG_Pz-Oz_IMPORTANT_VOLU...
EOG_horizontal_cwt_MEDIAN
EOG_horizontal-windowed_8-10-Band
EOG_horizontal-windowed_30-50-Band
EOG_horizontal_MEAN
EOG_horizontal_VARIANCE
EOG_horizontal_MaxAbs2Derivation
EMG_submental_cwt_MEDIAN

CfsSubsetEvaluator and GeneticSearch

EEG_Fpz-Cz_cwt_MEAN
EEG_Fpz-Cz_cwt_MEDIAN
EEG_Fpz-Cz_cwt_STDDEV
EEG_Fpz-Cz_cwt_MAX2RMS
EEG_Fpz-Cz_cwt_MAX2MEDIAN
EEG_Fpz-Cz_fft_MAX
EEG_Fpz-Cz_fft_MEDIAN
EEG_Fpz-Cz_fft_VAR
EEG_Fpz-Cz_fft_SKEW
EEG_Fpz-Cz_fft_RMS
EEG_Fpz-Cz_fft_STDDEV
EEG_Fpz-Cz_fft_MAX2RMS
EEG_Fpz-Cz_fft_MAX2MEAN
EEG_Fpz-Cz_fft_MAX2MEDIAN
EEG_Fpz-Cz_19-24-Band
EEG_Fpz-Cz_24-30-Band
EEG_Fpz-Cz_30-50-Band
EEG_Fpz-Cz_FFT_band_theta
EEG_Fpz-Cz_FFT_band_delta...
EEG_Fpz-Cz_FFT_band_delta...
EEG_Fpz-Cz_FFT_band_delta...
EEG_Fpz-Cz_FFT_band_delta...
EEG_Fpz-Cz_FFT_band_theta...
EEG_Fpz-Cz_FFT_band_alpha...
EEG_Fpz-Cz_FFT_band_alpha...
EEG_Fpz-Cz_VARIANCE
EEG_Fpz-Cz_KURTOSIS
EEG_Fpz-Cz_AvgAbs1Derivation
EEG_Fpz-Cz_hjorth_mobility
EEG_Fpz-Cz_hjorth_complexity
EEG_Fpz-Cz_interp_poly_p2
EEG_Fpz-Cz_interp_poly_p4
EEG_Fpz-Cz_interp_rational_p4
EEG_Fpz-Cz_interp_rational_q3
EEG_Fpz-Cz_SPECTRAL_CENTROID
EEG_Fpz-Cz_IMPORTANT_VOLU...
EEG_Pz-Oz_cwt_MEDIAN
EEG_Pz-Oz_cwt_SKEW
EEG_Pz-Oz_fft_MEDIAN
EEG_Pz-Oz_FFT_band_delta
EEG_Pz-Oz_FFT_band_gamma
EEG_Pz-Oz_VARIANCE
EEG_Pz-Oz_AvgAbs1Derivation
EEG_Pz-Oz_hjorth_mobility
EEG_Pz-Oz_interp_poly_p1
EEG_Pz-Oz_interp_poly_p4
EEG_Pz-Oz_IMPORTANT_VOLU...
EOG_horizontal_cwt_MEDIAN
EOG_horizontal_cwt_VAR
EOG_horizontal_cwt_KURT
EOG_horizontal_cwt_MAX2MEDIAN
EOG_horizontal_fft_MIN
EOG_horizontal_fft_MEDIAN

Continued on next page

CfsSubsetEvaluator and BestFirst	**CfsSubsetEvaluator and GeneticSearch**
	EOG_horizontal-windowed_8-10-Band
	EOG_horizontal-windowed_FFT_band_beta
	EOG_horizontal-windowed_FFT_band_gamma
	EOG_horizontal-windowed_FFT_band_delta...
	EOG_horizontal-windowed_FFT_band_delta...
	EOG_horizontal-windowed_FFT_band_theta...
	EOG_horizontal-windowed_FFT_band_beta...
	EOG_horizontal_MaxAbs1Derivation
	EOG_horizontal_hjorth_activity
	EOG_horizontal_interp_poly_p3
	EOG_horizontal_interp_poly_rms
	EMG_submental_cwt_MAX
	EMG_submental_fft_MEAN
	EMG_submental_fft_MEDIAN
	EMG_submental_fft_VAR
	EMG_submental_fft_RMS
	EMG_submental_24-30-Band
	EMG_submental_FFT_band_theta...
	EMG_submental_FFT_band_theta...
	EMG_submental_VARIANCE
	EMG_submental_interp_poly_p3
	EMG_submental_interp_poly_rms
	EMG_submental_interp_rational_rms
	EMG_submental_IMPORTANT_VOLU...

Table D.12: Selected features – Sleep stage REM – CfsSubsetEvaluator and BestFirst, CfsSubsetEvaluator and GeneticSearch

GAME selected features	**Ranker and ChiSquaredEvaluator**
EEG_Fpz-Cz_cwt_SKEW	EEG_Fpz-Cz_cwt_MEAN
EEG_Fpz-Cz_cwt_KURT	EEG_Fpz-Cz_cwt_MEDIAN
EEG_Fpz-Cz_fft_SKEW	EEG_Fpz-Cz_cwt_VAR
EEG_Fpz-Cz_fft_KURT	EEG_Fpz-Cz_cwt_SKEW
EEG_Fpz-Cz_fft_MAX2RMS	EEG_Fpz-Cz_cwt_KURT
EEG_Fpz-Cz_FFT_band_theta	EEG_Fpz-Cz_cwt_RMS
EEG_Fpz-Cz_SPECTRAL_CENTROID	EEG_Fpz-Cz_cwt_STDDEV
EEG_Fpz-Cz_IMPORTANT_VOLU...	EEG_Fpz-Cz_cwt_MAX2RMS
EEG_Pz-Oz_cwt_MEAN	EEG_Fpz-Cz_cwt_MAX2MEAN
EEG_Pz-Oz_cwt_MEDIAN	EEG_Fpz-Cz_cwt_MAX2MEDIAN
EEG_Pz-Oz_AvgAbs1Derivation	EEG_Fpz-Cz_fft_MAX
EEG_Pz-Oz_MaxAbs2Derivation	EEG_Fpz-Cz_fft_MEAN
EEG_Pz-Oz_IMPORTANT_VOLU...	EEG_Fpz-Cz_fft_VAR
EOG_horizontal_cwt_MAX	EEG_Fpz-Cz_fft_SKEW
EOG_horizontal-windowed_FFT_band_delta...	EEG_Fpz-Cz_fft_KURT
EOG_horizontal-windowed_FFT_band_alpha...	EEG_Fpz-Cz_fft_RMS
EOG_horizontal-windowed_FFT_band_alpha...	EEG_Fpz-Cz_fft_STDDEV
EOG_horizontal_MEAN_ABS	EEG_Fpz-Cz_fft_MAX2RMS
EOG_horizontal_hjorth_complexity	EEG_Fpz-Cz_fft_MAX2MEAN
	EEG_Fpz-Cz_fft_MAX2MEDIAN
	EEG_Fpz-Cz_8-10-Band
	EEG_Fpz-Cz_10-13-Band
	EEG_Fpz-Cz_13-19-Band
	EEG_Fpz-Cz_FFT_band_theta
	EEG_Fpz-Cz_FFT_band_alpha

Continued on next page

GAME selected features	Ranker and ChiSquaredEvaluator
	EEG_Fpz-Cz_FFT_band_delta...
	EEG_Fpz-Cz_FFT_band_delta...
	EEG_Fpz-Cz_FFT_band_delta...
	EEG_Fpz-Cz_FFT_band_delta...
	EEG_Fpz-Cz_FFT_band_theta...
	EEG_Fpz-Cz_FFT_band_alpha...
	EEG_Fpz-Cz_MEAN_ABS
	EEG_Fpz-Cz_MEAN
	EEG_Fpz-Cz_MAX_POSITIVE
	EEG_Fpz-Cz_MIN_NEGATIVE
	EEG_Fpz-Cz_VARIANCE
	EEG_Fpz-Cz_SKEWNESS
	EEG_Fpz-Cz_AvgAbs1Derivation
	EEG_Fpz-Cz_hjorth_activity
	EEG_Fpz-Cz_hjorth_mobility
	EEG_Fpz-Cz_hjorth_complexity
	EEG_Fpz-Cz_interp_poly_p4
	EEG_Fpz-Cz_interp_poly_rms
	EEG_Fpz-Cz_interp_rational_rms
	EEG_Fpz-Cz_SPECTRAL_CENTROID
	EEG_Fpz-Cz_IMPORTANT_VOLU...
	EEG_Pz-Oz_cwt_MEAN
	EEG_Pz-Oz_cwt_MEDIAN
	EEG_Pz-Oz_cwt_RMS
	EEG_Pz-Oz_MEAN_ABS
	EEG_Pz-Oz_VARIANCE
	EEG_Pz-Oz_AvgAbs1Derivation
	EEG_Pz-Oz_hjorth_activity
	EEG_Pz-Oz_interp_rational_rms
	EEG_Pz-Oz_IMPORTANT_VOLU...
	EOG_horizontal-windowed_30-50-Band
	EOG_horizontal-windowed_FFT_band_gamma
	EOG_horizontal_MEAN_ABS
	EOG_horizontal_VARIANCE
	EOG_horizontal_MaxAbs2Derivation
	EOG_horizontal_hjorth_activity

Table D.13: Selected features – Sleep Stage REM – GAME selected features, Ranker and ChiSquaredEvaluator

Ranker and GainRatio	Ranker and InfoGain
EEG_Fpz-Cz_cwt_MAX	EEG_Fpz-Cz_cwt_MEAN
EEG_Fpz-Cz_cwt_MEAN	EEG_Fpz-Cz_cwt_MEDIAN
EEG_Fpz-Cz_cwt_MEDIAN	EEG_Fpz-Cz_cwt_VAR
EEG_Fpz-Cz_cwt_VAR	EEG_Fpz-Cz_cwt_SKEW
EEG_Fpz-Cz_cwt_SKEW	EEG_Fpz-Cz_cwt_KURT
EEG_Fpz-Cz_cwt_RMS	EEG_Fpz-Cz_cwt_RMS
EEG_Fpz-Cz_cwt_STDDEV	EEG_Fpz-Cz_cwt_STDDEV
EEG_Fpz-Cz_cwt_MAX2MEAN	EEG_Fpz-Cz_cwt_MAX2RMS
EEG_Fpz-Cz_cwt_MAX2MEDIAN	EEG_Fpz-Cz_cwt_MAX2MEAN
EEG_Fpz-Cz_fft_MAX	EEG_Fpz-Cz_cwt_MAX2MEDIAN
EEG_Fpz-Cz_fft_SKEW	EEG_Fpz-Cz_fft_MAX
EEG_Fpz-Cz_fft_KURT	EEG_Fpz-Cz_fft_MEAN
EEG_Fpz-Cz_fft_MAX2RMS	EEG_Fpz-Cz_fft_VAR

Continued on next page

Ranker and GainRatio	**Ranker and InfoGain**
EEG_Fpz-Cz_fft_MAX2MEAN	EEG_Fpz-Cz_fft_SKEW
EEG_Fpz-Cz_fft_MAX2MEDIAN	EEG_Fpz-Cz_fft_KURT
EEG_Fpz-Cz_8-10-Band	EEG_Fpz-Cz_fft_RMS
EEG_Fpz-Cz_10-13-Band	EEG_Fpz-Cz_fft_STDDEV
EEG_Fpz-Cz_13-19-Band	EEG_Fpz-Cz_fft_MAX2RMS
EEG_Fpz-Cz_FFT_band_theta	EEG_Fpz-Cz_fft_MAX2MEAN
EEG_Fpz-Cz_FFT_band_alpha	EEG_Fpz-Cz_fft_MAX2MEDIAN
EEG_Fpz-Cz_FFT_band_delta...	EEG_Fpz-Cz_8-10-Band
EEG_Fpz-Cz_FFT_band_delta...	EEG_Fpz-Cz_10-13-Band
EEG_Fpz-Cz_FFT_band_delta...	EEG_Fpz-Cz_13-19-Band
EEG_Fpz-Cz_FFT_band_delta...	EEG_Fpz-Cz_FFT_band_theta
EEG_Fpz-Cz_FFT_band_theta...	EEG_Fpz-Cz_FFT_band_alpha
EEG_Fpz-Cz_MEAN_ABS	EEG_Fpz-Cz_FFT_band_delta...
EEG_Fpz-Cz_MAX_POSITIVE	EEG_Fpz-Cz_FFT_band_delta...
EEG_Fpz-Cz_MIN_NEGATIVE	EEG_Fpz-Cz_FFT_band_delta...
EEG_Fpz-Cz_VARIANCE	EEG_Fpz-Cz_FFT_band_delta...
EEG_Fpz-Cz_AvgAbs1Derivation	EEG_Fpz-Cz_FFT_band_theta...
EEG_Fpz-Cz_MaxAbs1Derivation	EEG_Fpz-Cz_FFT_band_alpha...
EEG_Fpz-Cz_hjorth_activity	EEG_Fpz-Cz_MEAN_ABS
EEG_Fpz-Cz_hjorth_mobility	EEG_Fpz-Cz_MEAN
EEG_Fpz-Cz_interp_poly_rms	EEG_Fpz-Cz_MAX_POSITIVE
EEG_Fpz-Cz_SPECTRAL_CENTROID	EEG_Fpz-Cz_MIN_NEGATIVE
EEG_Fpz-Cz_IMPORTANT_VOLU...	EEG_Fpz-Cz_VARIANCE
EEG_Pz-Oz_cwt_MIN	EEG_Fpz-Cz_SKEWNESS
EEG_Pz-Oz_cwt_MEAN	EEG_Fpz-Cz_AvgAbs1Derivation
EEG_Pz-Oz_cwt_MEDIAN	EEG_Fpz-Cz_hjorth_activity
EEG_Pz-Oz_cwt_RMS	EEG_Fpz-Cz_hjorth_mobility
EEG_Pz-Oz_cwt_MAX2MEAN	EEG_Fpz-Cz_hjorth_complexity
EEG_Pz-Oz_cwt_MAX2MEDIAN	EEG_Fpz-Cz_interp_poly_p4
EEG_Pz-Oz_13-19-Band	EEG_Fpz-Cz_interp_poly_rms
EEG_Pz-Oz_30-50-Band	EEG_Fpz-Cz_interp_rational_rms
EEG_Pz-Oz_FFT_band_alpha	EEG_Fpz-Cz_SPECTRAL_CENTROID
EEG_Pz-Oz_FFT_band_gamma	EEG_Fpz-Cz_IMPORTANT_VOLU...
EEG_Pz-Oz_FFT_band_theta...	EEG_Pz-Oz_cwt_MEAN
EEG_Pz-Oz_FFT_band_theta...	EEG_Pz-Oz_cwt_MEDIAN
EEG_Pz-Oz_FFT_band_alpha...	EEG_Pz-Oz_cwt_RMS
EEG_Pz-Oz_FFT_band_alpha...	EEG_Pz-Oz_MEAN_ABS
EEG_Pz-Oz_MEAN_ABS	EEG_Pz-Oz_VARIANCE
EEG_Pz-Oz_VARIANCE	EEG_Pz-Oz_AvgAbs1Derivation
EEG_Pz-Oz_AvgAbs1Derivation	EEG_Pz-Oz_hjorth_activity
EEG_Pz-Oz_hjorth_activity	EEG_Pz-Oz_interp_poly_rms
EEG_Pz-Oz_interp_poly_rms	EEG_Pz-Oz_IMPORTANT_VOLU...
EEG_Pz-Oz_IMPORTANT_VOLU...	EOG_horizontal-windowed_30-50-Band
EOG_horizontal_cwt_MEAN	EOG_horizontal-windowed_FFT_band_gamma
EOG_horizontal-windowed_8-10-Band	EOG_horizontal_VARIANCE
EOG_horizontal-windowed_FFT_band_alpha	EOG_horizontal_MaxAbs2Derivation
EOG_horizontal_VARIANCE	EOG_horizontal_hjorth_activity
EOG_horizontal_AvgAbs1Derivation	

Table D.14: Selected features – Sleep Stage REM – Ranker and GainRatio, Ranker and InfoGain

Appendix E

Enclosed CD-ROM content

Directory	**Descrition**
`./Diploma thesis - text`	Contains PDF version of my thesis
`./Diploma thesis - text/Src`	Contains all source files necessary for recreation of my thesis in LaTeX.
`./Used software`	Contains all software which I used to go through the concept. If contains Weka, GAME, my preprocessing tool – CATool and few othe scripts which I wrote to ease the task.
`./Data/Raw`	Contains data I processed in format they came to me.
`./Data/CATool`	Contains data I processed in format which fits the CATool.
`./Results/Sleep EEG Data`	Contains all results I gained from Sleep data.
`./Results/Snoring data`	Contains all results I gained from Breathing data.

Table E.1: Enclosed CD-ROM content

Bibliography

[1] Daubechies, I. Ten lectures on Wavelets. CBMS-NSF Regional Conference Series In Applied Mathematics, Vol. 61. Society for Industrial and Applied Mathematics. 1992.

[2] Polikar, R. The Wavelet Tutorial. http://users.rowan.edu/p̃olikar/WAVELETS/WTtutorial.html

[3] Mathworks, http://www.mathworks.com/access/helpdesk/help/toolbox/wavelet/ch01_28a.html

[4] Reiger, J. Diploma thesis. 2004.

[5] MacQueen, J.B. Some Methods for Classification and Analysis of Multivariate Observations. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, p. 281-297. 1967.

[6] Mařík V., Štěpánková O., Lažanský J. Umělá inteligence 4. Academia. 2003.

[7] Mařík V., Štěpánková O., Lažanský J. Umělá inteligence 1. Academia. 1993.

[8] Šnorek M. doc. Ing. CSc. Neuronové sítě a neuropočítače. Vydavatelství ČVUT. 2002.

[9] Kordík P. Ing. Dizertační práce.

[10] Pyle Dorian. Data preparation for data mining. Morgan Kaufmann publishers, inc. 1999.

[11] Duda Richard O., Hart Peter E., Stork David G. . Pattern Classification. Wiley-Interscience. Second edition, 2001.

[12] Hjorth Bo. EEG Analysis based on time domain domain properties. Electroencephalography and Clinical Neurophysiology, 1970, 29: 306-310.

[13] EDF definition. http://www.hsr.nl/edf/

[14] Sporka Adam. Signal Segmentation for Speech And Non-Speech User Interface Control. Postgraduate Neural Network Coursework Report, 2005.

[15] Rogalewitz V. Pravděpodobnost a statistika pro inženýry. Vydavatelstvý CVUT. First edition, 2000.

[16] Tadao Hori, Yoshio Sugita, Einosuke Koga, Shuichiro Shirakwa, Katuhiro Inouem Sunao Uchida, Hiroo Kuwahara, Masako Kousaka, Toshinori Kobayashi, Yoichi Tsuji, Masayosi Terashima, Kazuhiko Fukuda, Noriko Fukuda. Proposed supplements and amendments to *'A manual of Standardized Terminology, Technigues and Scoring System of Sleep Stages of Human Objects'*, the Rechtchaffen & Kales (1968) standard. Psychiatry and Clinical Neurosciences. 55:305-310. 2001.

[17] Oropesa Edgar, Cycon Hans, Jobert Marc. Sleep Stage Classification using Wavelet Transform and Neural Netowrk. TR-99-008. 1999.

[18] Mutapcic Almir, Shimayama Toshihide. Automatic sleep stage classification using frequency analysis of EEG signals. EE373A,B Final Project Report.

[19] Koprinska Irena, Pfurtscheller Gert, Flotzinger Doris. Sleep classification in infants by decision tree-based neural networks. Artificial Intelligence in Medicine 8 (I 9961387-401).

[20] Rechtschaffen A., and Kales A.. A Manual of Standardized Terminology, Technique and Scoring System for Sleep Stages of Human Subjects. Public Health Service, U.S. Government Printing Office, Washington, DC. 1968.

[21] John George H., Langley Pat. Estimating continuous distribution in Bayesian classifiers. Proceedings of the Eleventh Conference on Uncertainty in Artifitial Inteligence. Morgan Kaufmann Publishers, San Mateo. 1995.

[22] Kordík P. Selecting Subset of Relevant Variables by Means of Niching Genetic Algorithm. Poster 2004. CTU Prague 2004.

[23] Muller J. A. Lemke F. Self-Organising Data Mining. Berlin. ISBN 3-89811-861-4. 2000.

[24] Koller D., Sahami M. Towards optimal Feature Selection. Machine Learning 13. 1996.

[25] Šima, J., & Neruda R. Teoretické otázky neuronových sítí.

[26] Piramuthu S. Evaluating feature selection methods for learning in data mining applications. European Journal of Operational Research 156. page 483-494. 2004.

[27] Witten I., Frank E. Data Mining Practical Machine Learning Tools and Techniques. Second Edition. Elsevier. ISBN: 0-12-088407-0. 2005.

[28] Press W., Teukolsky S., Vetterling W., Flannery B. Numerical Recipes in C, The Art of Scientific Computing. Second Edition. CAMBRIDGE UNIVERSITY PRESS. 2002.

[29] Ralston A., Wilf H.S. Mathematical Methods for Digital Computers. New York: Wiley. 1960.

[30] Goldberg, D. E. Genetic algorithms in search, optimization, and machine learning. Reading, MA: Addison-Wesley. 1989.

[31] Müller, J.-A., F. Lemke. Self-Organizing Data Mining. Libri, Hamburg 2000. ISBN 3-89811-861-4.